MONETARY-UNIT SAMPLING: AN INVESTIGATION

by

Jane M. Horgan, M.A., M.Sc.

A Dissertation Presented in Fulfilment

of the

Requirements for the Degree

of

Doctor of Philosophy

1994.

VOLUME 1

City University Business School, London

February 1994

TABLE OF CONTENTS

Volume 1

TABLE OF CONTENTS

Volume 11

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge the support and direction of my supervisor Professor Georges Selim of London City University.

The advice and suggestions of Professor Anthony Walsh and Professor Michael Ryan of Dublin City University were invaluable.

I would also like to thank the library staff for their infinite patience in dealing with my requests, particulary Ms Helen Fallon.

Monetary-Unit Sampling: An Investigation

Jane M. Horgan

ABSTRACT

This study examined the performance of six monetary-unit sampling methods in substantive auditing under various population conditions. The investigation involved both a theoretical point estimator analysis and an empirical study of upper bound estimates of the total error amount using the Stringer, Cell and Moment bounds with two methods of error assignment, three nominal confidence levels and three sample sizes.

For a range of thirty audit populations simulated from real accounting populations of debtors from commercial entities in the Public Sector in Ireland, it was found that the differential effects of simple random, systematic, cell and sieve sampling were independent of the bound used, the error assignment method and the nominal confidence level. The reliability and tightness of the bounds were found to be similar for all selection methods but differences in the precision of the bounds did exist. In terms of the precision of the estimates, systematic and cell sampling favoured populations with large line items and sieve sampling favoured populations with large line items and low error rates.

Lahiri sampling, a selection method not used in auditing previously and proposed in this study as a practical alternative to simple random sampling was not found to be significantly different from simple random sampling with respect to any of the performance measures.

Stabilised sieve sampling, a new monetary-unit sampling method developed in this study as an alternative to simple random sampling and sieve sampling, was found to be reliable for the range of audit populations on which it was tested. In populations with small line items, stabilised sieve sampling tended to have a tightness similar to that of simple random sampling and sieve sampling for any given error rate, taint size, sample size and bound and in populations with large line items, stabilised sieve sampling was more conservative than simple random sampling and sieve sampling but the differences were not significant in any case. It was more precise than simple random sampling and its precision was similar to that of sieve sampling in most cases. As stabilised sieve sampling overcomes the primary disadvantage of sieve sampling by returning a constant sample size of monetary units, it was concluded that it may be a useful alternative to simple random sampling and sieve sampling in real substantive auditing environments.

KEY TO ABBREVIATIONS

AICPA	American Institute of Certified Public Accountants
ANOVA	Analysis of Variance
AON	'All-or-nothing' Method of Error Assignment
APB	Auditing Practices Board
APC	Auditing Practices Committee
ASB	Auditing Standards Board
CAG	Comptroller and Auditor General
CAV	Combined Attributes and Variables
CICA	Canadian Institute of Chartered Accountants
DEFF	Design Effect
ICAEW	Institute of Chartered Accountants of England and Wales
ICAS	Institute of Chartered Accountants of Scotland
IIA	Institute of Internal Auditors
MPU	Mean-per-unit
MUS	Monetary-Unit Sampling
PPS	Probability Proportional to Size
SAMSIZE	Sample Size
SRS	Simple Random Sampling of Monetary Units
SYS	Systematic Sampling of Monetary Units
UK	United Kingdom
USA	United States of America

Chapter 1

An Introduction to the Background of the Study

1.1 Introduction

Accounting populations consist of a compilation of items (e.g. creditors, debtors) put together over a specific period. The populations sometimes contain a large number of items amounting to balances totalling millions of pounds and the auditor has the responsibility of attesting to their truth and fairness. A detailed examination of all these accounts is not always practicable and it is often necessary to rely on statistical sampling to select a subset in order to estimate the characteristics of the whole population.

This study explores some of the problems facing the auditor when using statistical sampling. Specifically, it investigates how existing monetary-unit sampling methods perform in obtaining estimates of the total error amount in accounting populations of debtors. New monetary-unit sampling methods are developed for application in substantive testing of debtors and are tested on data obtained from commercial entities in the Public Sector in Ireland.

In the remainder of this chapter;

The nature of auditing and the audit process is discussed (1.2)

Sampling of accounting populations is explained (1.3)

Statistical sampling in auditing is discussed (1.4)

The development of monetary-unit sampling is outlined (1.5)

The need for the research is summarised (1.6)

The scope of the research is detailed (1.7)

The objectives of the study are stated (1.8)

The limitations of the study are explained (1.9)

An overview of the remaining chapters is provided (1.10)

1.2 The Nature of Auditing and the Audit Process

The American Accounting Association's Committee on Basic Auditing Concepts (1973, p2) defines auditing as:

> 'a systematic process of objectively obtaining and evaluating evidence regarding assertions about economic actions and events to ascertain the degree of correspondence between these assertions and established criteria and communicating the results to interested users.'

and the explanatory forward of the APC's Auditing Standards and Guidelines defines an audit as:

> 'an independent examination of, and expression of opinion on, the financial statements of an enterprise by an appointed auditor in pursuance of that appointment and in compliance with any relevant statutory obligations' (APC, 1989).

Auditing may be classified in a number of ways and two classifications which are particularly relevant to this study are (i) Financial and operational auditing and (ii) External and internal auditing. These are defined below.

(i) Financial and Operational Auditing

A financial audit is:

> ' a systematic examination of financial statements, records and related operations to determine adherence to generally accepted accounting principles, management policies, or stated requirements'(Schlosser, 1971, pp 1-4).

Operational auditing is defined by the AICPA's Special Committee on Operational and Management Auditing as

> 'a systematic review of an organisation's activities (or a stipulated segment of them) in relation to specified objectives for the purposes of assessing performance, identifying opportunities for improvement, and developing recommendations for improvement or further action'(AICPA,1973).

(ii) External and Internal Auditing.

External audits are carried out by personnel who are not employees of the organisations being audited. Usually, external audits are carried out by firms of public accountants who offer their services on a contractual basis. The majority of audits performed by public accountants are financial. The most common financial audit is an examination of the financial statements for the purpose of forming an opinion of their truth and fairness in conformity with generally accepted accounting principles (Taylor and Glezen, 1994, p3).

Internal auditing is described by the Institute of Internal Auditors as 'an independent appraisal function established within an organisation to examine and evaluate its activities as a service to the organisation' (IIA, 1992). The objective of an internal audit is to assist members of the organisation, including those in management and on the board, in the effective discharge of their responsibilities. To this end, the scope of internal auditing should encompass 'the examination and evaluation of the adequacy and the effectiveness of the organizations's system of internal controls and the quality of performance in carrying out assigned responsibilities' (IIA, 1992).

As will be discussed further in 1.2.4, this study is concerned with external financial auditing of commercial entities in the Irish Public Sector by government auditors. Leslie (1975) noted that

> 'Government auditors are frequently called upon to carry out many tests of accounting populations similar to those undertaken by external and internal auditors in the Private Sector.'

To develop an opinion on the financial statements 'the auditor must gather and evaluate many different types of information both financial and non-financial' (Mock and Watkins, 1982). This gathering and evaluation activity is called the audit process and the information obtained is called the audit evidence.

1.2.1 Conceptual Framework for Auditing

Authors have sought to identify a conceptual framework for the audit process as a means of determining the criteria necessary for the adequate performance of the audit function. The early development of a conceptual framework consisted of a 'postulate-based' approach. Mautz and Sharaf (1961, Chap. 3) proposed 'tentative' postulates and these are listed below:

(i) Financial statements and financial data are verifiable.

(ii) There is no necessary conflict of interest between the auditor and the management of the enterprise under audit.

(iii) The financial statements and other information submitted for verification are free from collusive and other unusual irregularities.

(iv) The existence of a satisfactory system of internal control eliminated the probability of irregularities.

(v) Consistent application of generally accepted principles of accounting results in the fair presentation of the financial position and the results of operations.

(vi) In the absence of clear evidence to the contrary, what has held true in the past for the enterprise under examination will hold true in the future.

(vii) When examining financial data for the purpose of expressing an independent opinion thereon, the auditor acts exclusively in the capacity of an auditor.

(viii) The professional status of the independent auditor imposes commensurate professional obligations.

Many authors see some of these postulates as untenable and have striven to interpret them less rigidly than stated. For example, regarding the verifiability of the audit process, Higson (1987) explains that the word 'verifiable' was not taken to mean beyond all doubt, but verification is a process that 'carries one to a position of confidence about any given proposition'. Hamilton (1978) suggests that the conceptual model advanced by Mautz and Sharaf (1961) fails to provide a foundation because it has neither yielded any testable results nor led to a process of theory refinement and development.

The Mautz and Sharaf postulates deal mainly with whether an audit is feasible and what the scope of such an audit should be. They say very little about whether an audit is in fact necessary. The postulates approach was developed further by Lee (1982) who classified postulates under three headings, (i) justifying postulates, (ii) behavioral postulates and (iii) functional postulates.

Justifying Postulates:

Justifying postulates attempt to justify the audit on the basis of the credibility of the accounts. They are:

(i) Statutory accounts in general have insufficient credibility to be used confidently by the shareholders.

(ii) The enhancement of credibility is the most important requirement of the statutory audit.

(iii) The statutory audit provides the best means of enhancing the credibility of the accounts.

(iv) The credibility of accounts can be enhanced or verified by the statutory audit.

(v) The shareholders are unable personally to satisfy themselves as to the credibility of the information in the accounts.

Behaviourial Postulates

Behavioral postulates refer to the qualities of the auditor. They are:

(i) The audit is not impeded by the existence of conflict between the auditor and management.

(ii) The law does not restrict the auditor.

(iii) The auditor is independent, both mentally and physically.

(iv) The auditor is sufficiently skilled to undertake the audit.

(v) The auditor is accountable for the quality of the work and opinion.

Functional Postulates

Functional postulates are primarily concerned with the existence of evidence and the interface of auditing with financial accounting. They are:

(i) There is sufficient reliable evidence available, and in an appropriate form, to enable the auditor to carry out an audit within reasonable time at a reasonable cost.

(ii) The accounts are free from major fraud and error.

(iii) There exist generally recognised accounting concepts and bases which, when used properly and consistently, result in fair presentation.

Sherer and Kent (1988, p19) state that Lee's postulates succeed in fulfilling an important purpose of an audit which is to enhance the credibility of the financial statements in the minds of the users. They also maintain that the categorisation of Lee 'forms a rational and comprehensible basis on which to base an examination of auditing theory.'

However, the 'postulates-based' approach is not without its critics mainly because it focuses more on the circumstances in which an audit may be successfully performed rather than on the rationale for the auditing. Also, authors disagree as to whether there is in fact an adequate theoretical foundation to auditing. On the one hand, Robertson (1984) maintains that there is, and suggests that the Mautz and Sharaf theory, although imperfect, can in fact serve auditors quite well. Robertson also believes that no alternative theory structure exists in such breadth and depth. On the other hand, Gwilliam (1987) does not believe that a theoretical foundation for auditing exists. He contends that

> 'setting out the basic assumptions on which the majority of audits are conducted... may well be of value in bringing forward certain fundamental issues and in clarifying thinking about assumptions that lie behind much of present day auditing' (Gwilliam, 1987,p47).

But he does not believe that it is possible to build a universal theory of how auditing should be carried out, by arguing from 'first principles'. Gwilliam (1987, p49) argues that a greater insight may be gained by viewing auditing as an economic activity in which benefits may be purchased (e.g., improved management performance) in exchange for costs (e.g., time and resources). He maintains that currently an agreed economic framework for auditing does not exist and states that

> ' what theoretical and historical evidence there is suggests that external auditing does provide a valuable monitoring device between management and shareholders and also between shareholders and bond-holders' (Gwilliam, 1987, p58).

1.2.3 Auditing Standard and Guidelines

In the conduct of an audit, auditors are expected to comply with the auditing standards published by the Auditing Practices Board (APB) (formally the Auditing Practices Committee (APC)). This consists of representatives of five of the professional accountancy bodies whose members are recognised as qualified to act as auditors in Ireland or the UK, i.e., The Institute of Chartered Accounts of England and Wales, The Institute of Chartered Accountants of Scotland, The Institute of Chartered Accountants of Ireland, The Association of Certified Accountants and the Institute of Certified Public Accountants. In 1980, the APC issued its definitive statement on auditing standards and guidelines (APC, 1980a) and since then additions have been made and some of the initial standards and guidelines have been revised. In the explanatory forward of the auditing standards and guidelines (APC,1989), auditing standards are defined as

> 'Basic principles and practices which members of the accountancy bodies are expected to follow in the conduct of an audit.'

Failure to observe auditing standards may result in disciplinary action imposed by the accountancy bodies (APC, 1989).

Auditing Guidelines, also issued by the APC, give the auditor guidance on what procedures may be applied and how to apply auditing standards. They are meant to be persuasive rather than prescriptive but 'should normally be followed' (APC, 1989).

The original standards and guidelines were written primarily in the context of limited company audits. Since 1980, many of the original guidelines have been revised and reissued. The revised guidelines extend to Public Sector auditing. The guideline on Public Sector auditing issued in July 1987 (APC, 1987a) states that auditing standards issued before June 1985 are applicable to the Public Sector. All auditing standards and guidelines issued since June 1985 apply to the audit of the financial statements of Public Sector bodies unless otherwise stated in the individual pronouncements (APC,1987a).

1.2.4 Background to Auditing in the Public Sector

In Ireland, the office of the Comptroller and Auditor General audits all government departments and some Public Sector commercial entities as listed in the Comptroller and Auditor General (Amendment) Act (1993). Other Public Sector commercial entities (e.g. Aer Lingus and Bord na Mona) are audited, with the consent of relevant government ministers, by commercial auditors from the private sector. This situation is also encountered in other countries. In the UK, the Comptroller and Auditor-

General's office is responsible for auditing of some 500 government entities, but the nationalised industries (e.g. The Coal Board) are audited by private sector auditing firms. In the USA, the office of the Comptroller General is responsible for the audit of all Federal departments and agencies but directly audits only 20 of some 50 government companies (Hardman, 1991). The auditors of the Comptroller and Auditor General's office are employed by the state. They function as external auditors when undertaking a financial audit of Public Sector commercial entities or other recipients of Government funds. The Comptroller and Auditor General's office examines the financial statements of the commercial entities for which it is responsible, in order to form conclusions about their financial positions and to assess the financial accuracy and regularity of the accounts. The audit work culminates in a report which contains the Comptroller and Auditor General's opinion on the financial statements of the entity in a given year.

1.2.5 Stages in the Audit Process

The guideline on operational standards (APC, 1988) states that a typical audit should cover:

1. Planning, controlling and recording: The auditor should adequately plan, control and record all the work needed for the audit;

2. Accounting systems: The auditor should ascertain the enterprise's system of recording and processing transactions and assess its adequacy as a basis for the preparation of financial statements;

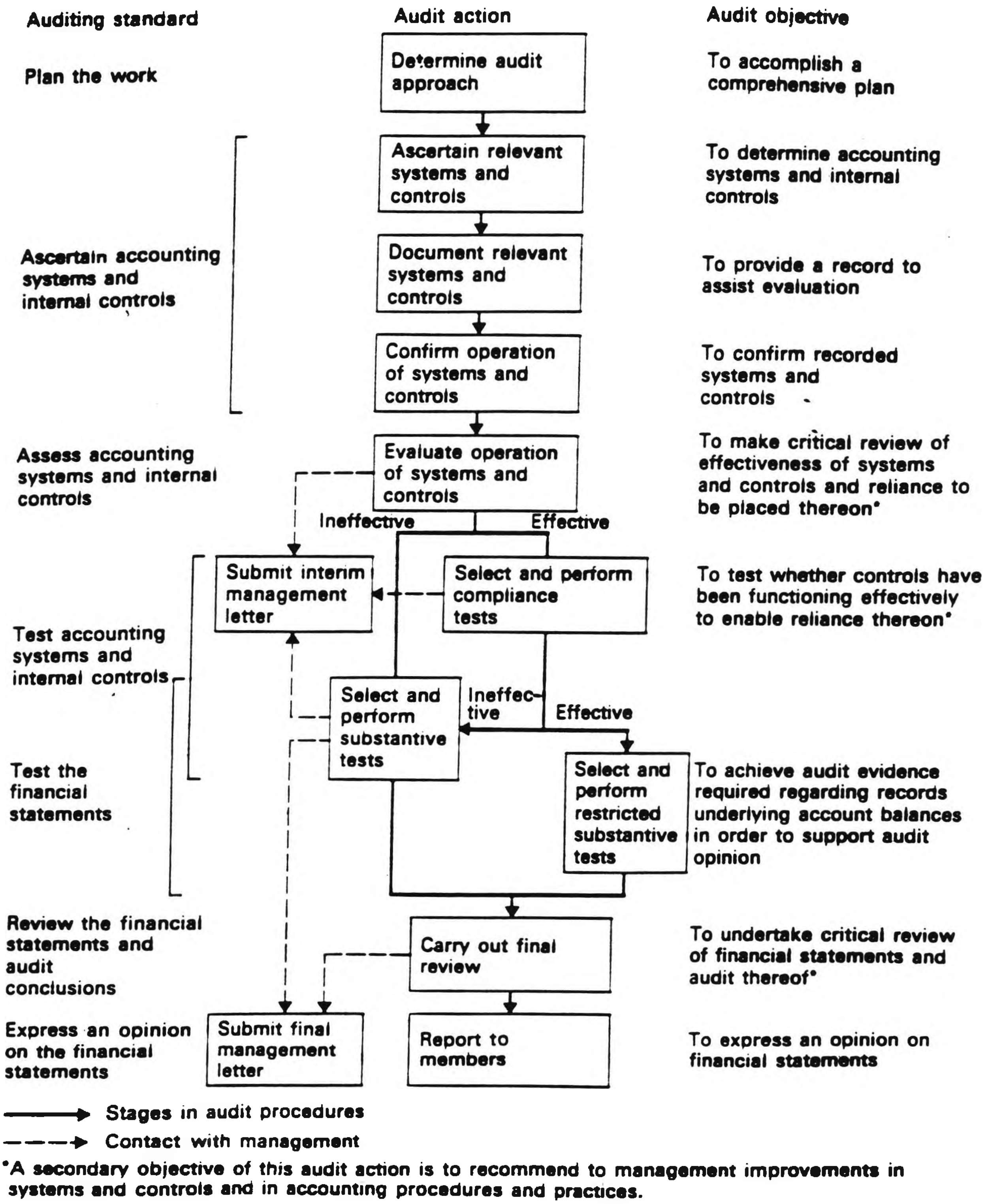
3. Audit evidence: The auditor should obtain relevant and reliable audit evidence sufficient to draw reasonable conclusions therefrom;

4. Internal Controls: If the auditor wishes to place reliance on any internal controls, then these controls should be ascertained and evaluated and compliance tests should be performed on their operation;

5. Review of financial statements: The auditor should carry out a review of the financial statements as is sufficient, in conjunction with the conclusions drawn from the other audit evidence obtained, to give a reasonable basis for the opinion on the financial statements.

A diagrammatic representation of the stages of an audit and their objectives is given in Figure 1.1 and from this, it can be seen that audit evidence is obtained by carrying out audit tests which may be classified as compliance or substantive according to their principal purpose.

Figure 1.1 The General Audit



(Source: dePaula and Attwood, 1982)

1.2.5.1 Compliance Testing

The purpose of tests of compliance is to provide reasonable assurance that the accounting control procedures are being applied as described and are complying with the stated policies, plans, laws and regulations. Internal controls are a set of procedures that are designed to minimise the chance of errors in the operation of the accounting system. The APC's auditing guideline on Internal Control (APC,1980c) defines the internal control system as

> 'the whole system of controls, financial and otherwise, established by the management to carry on the business of the company in an orderly and efficient manner, ensure adherence to management policies, safeguard the assets and secure, as far as possible, the accuracy and reliability of its records.'

And COSO (1992) describes internal control as a process

> 'effected by an entity's board of directors, management and other personnel, designed to provide reasonable assurance regarding the achievement of objectives in the following categories: Effectiveness and efficiency of organisations; Reliability of financial reporting; Compliance with applicable laws and regulations.'

Compliance tests are designed to establish to what extent the controls can be relied on to detect material error and whether the internal controls were operating effectively throughout the period being audited. The APC's operational standard (APC, 1980b) suggests that testing internal control is optional and 'need only be evaluated and tested if the auditor is seeking to place reliance on them.'

1.2.5.2 Substantive Procedures

The purpose of substantive procedures is

> 'to provide audit evidence as to the completeness, accuracy and validity of the information contained in the accounting records or in the financial statements (APC, 1980d).'

Substantive procedures consist of an examination of individual transactions in the accounts (substantive tests) and other procedures of a more general nature (analytical review).

1.2.5.2.1 Analytical Review

The analytical review examines the financial statements as a whole and reviews them for credibility, feasibility and consistency (Shaw, 1980, p61). The guideline on analytical review (APC, 1988) states that analytical review procedures can be carried out at the planning stage, the testing stage or at the financial statements review. Analytical review procedures listed in this guideline include:

(i) analysing relationships between items of financial data (e.g. sales and costs of sales), or between financial and non-financial data (e.g. payroll costs and number of staff);

(ii) comparing actual data with predictions derived from other analysis;

(iii) comparing data for latest period with corresponding data for:
earlier periods;
comparable enterprises;
industry averages;

(iv) investigating unexpected variations;

(v) obtaining and substantiating explanations to variances;

(vi) evaluating the analysis in the light of other evidence.

1.2.5.2.2 Substantive Tests

The purpose of substantive tests is to draw conclusions about the materiality of the error amounts in the accounts. Substantive testing involves detailed examination of the monetary value of the account balances to determine their accuracy. The extent and nature of substantive testing, depends upon the decisions taken about the effectiveness of the system of internal control (Shaw, 1980, p61). The auditing guideline on audit evidence (APC, 1980d) states that;

> 'the auditor may rely on appropriate evidence by substantive testing to form his opinion, provided that sufficient of such evidence is obtained. Alternatively, he may be able to obtain assurance from the presence of a reliable system of internal control, and thereby reduce the extent of substantive testing.'

Many government audits deal only with compliance issues and are concerned with whether a reporting entity has utilised its resources according to statutes and specific legislative appropriations (Banker, Cooper, and Potter, 1992). However, the office of the CAG carries our an annual audit of all its commercial entities and this audit always involves some substantive testing as noted in the Audit System Manual (CAG, 1992) that 'a minimum of substantive testing is necessary in all cases'.

1.3 Sampling of Accounting Populations

Accounting populations consist of a compilation of items, (e.g. debtors, creditors) put together over a specific period. Internal controls are invoked in an attempt to ensure that the items are entered correctly. The populations sometimes consist of a large number of items amounting to balances totalling millions of pounds and a detailed examination of all accounts is not always practicable. Consequently, sampling is often used when carrying out compliance or substantive tests. Sampling is a process of selecting a subset of a population of items for the purpose of making inferences to the whole population. The APC issued a draft guideline on audit sampling in 1987 (APC, 1987b) and this was followed in 1993 by yet another draft issued by the APB (1993) (formally the APC). In the new draft, audit sampling is defined as

> 'the application of audit procedures to less than 100% of the items within an account balance or class of transactions to enable the auditor to obtain and evaluate evidence about some characteristic of the items selected in order to form or assist in forming a conclusion concerning the population which makes up the account balance or class of transactions.' (APB, 1993)

To date this draft has not been finalised and its status is that of an exposure draft. In the US however, the Auditing Standards Board (ASB) of the American Institute of Certified Public Accountants (AICPA) published its standard on audit sampling in 1981 (AICPA, 1981). This established a framework for auditors in planning and evaluating audit samples. The AICPA has also published an audit sampling guideline (AICPA,1983), to help auditors in implementing the concepts of the audit sampling standard.

Adams (1989) is critical of the progress of the APC (now the APB) and at its tardiness in issuing finalised guidelines. He states

> 'We are still waiting for the final product of the sampling exercise, although our North American friends have already made significant strides in this direction.'

Sampling of some sort or another is used in many audit programmes as noted by Taylor (1985).

> 'A fundamental element of any audit programme will be the selection of transactions to be tested as a sample of all available transactions.'

Generally, sampling in auditing is either judgemental or statistical and the professional bodies allow for either selection method (see for example, APB ,1993; AICPA, 1983; AICPA, 1981).

1.3.1 Judgemental Sampling

Judgemental sampling is a selection process where the auditor decides which items should be audited. It involves a subjective selection of items for testing and a subjective evaluation of the results. Judgemental sampling is accepted by the accounting professions as a means of gathering evidence concerning the truth and fairness of the financial statements. In the current draft guideline on sampling issued by the Auditing Practices Board, it is stated that judgemental sampling 'is an acceptable method of selection provided the auditor is satisfied that the sample is not unrepresentative of the entire population' (APB, 1993).

It could be contended that the reliability of the sample results obtained using judgemental sampling cannot be estimated because the probability of selection of the individual line items cannot be ascertained. Vance (1950, p2) maintains that the sampling method is not scientific and enunciates a number of subjective influences that may affect the conclusions based on judgemental sampling and render them inconsistent and unreliable:

(i) Differences in individual auditor's ability, knowledge, experience and prejudices.

(ii) Pressure on the auditor to reduce the client's cost of the audit.

(iii) Auditor's state of physical and mental health.

Taylor (1985) criticises judgmental sampling because it 'relies on intuition and non-quantitative methods in the evaluation process'. It has also been criticised on the basis that the extent of audit testing is not consistent between auditors or across audits. Sneed (1979) found that different audit firms demonstrated significantly different degrees of conservatism with regard to sample size in judgemental sampling. In an investigation on audit testing, the Canadian Institute of Chartered Accountants (CICA, 1980) found that a wide variation existed between audits and auditors in the amount of auditing done using judgemental sampling.

1.3.2 Statistical Sampling

Statistical sampling involves the random selection of a number of items for inspection and is endorsed by the accountancy bodies (APB, 1993; AICPA, 1981). In statistical sampling, each item has a calculable chance of being selected and therefore the reliability of the sample results can be estimated.

A commonly held misconception about statistical sampling is that it removes the need for the use of professional judgement. While it is true that statistical sampling uses statistical methods to determine the sample size and to select and evaluate audit samples, it is the responsibility of the auditor to consider and specify in advance, factors such as materiality, the expected error rate or amount, the risk of over-reliance or the risk of incorrect acceptance, audit risk, inherent risk, control risk, standard deviation and population size, before the sample size can be determined (AICPA, 1983).

McRae (1982, p14) asserts that

> 'Statistical sampling allows an auditor's judgement to be concentrated on those areas of the audit where it is most needed'

Taylor (1985) also described how the auditor uses his/her professional judgement in choosing a statistical sample.

> 'statistical sampling allows the quantification of key factors and the risk of errors. This is not to suggest that statistical sampling methods remove the need for professional judgement, but rather that they allow elements of the evaluation process to be quantified, measured and controlled. As a result statistical methods may give greater assistance to the auditor in the control and direction of audit work.'

Statistical sampling has won increasing acceptance in the auditing profession as a means of efficiently and effectively gathering evidence concerning the fairness of the client's financial statements (Carpenter and Dirsmith, 1993). Studies have shown that the use of statistical firms in the UK has increased in recent years. McRae (1982, p178) found that only 11.8% of 136 medium-sized firms in the UK used some form of statistical sampling but Abdul-Hamid (1993) found that 43% of 61 medium-sized accounting firms sampled used some form of statistical sampling in their audits. It should be noted also that a study carried out by Higson (1987) suggests that the use of statistical sampling in auditing is on the decrease among the big six firms in the UK.

In the opinion of Arkin (1984) statistical sampling is one of the most valuable audit tools and cites the main advantages as:

(i) The sample result is objective and defensible. Nearly all phases of the statistical process are based on demonstrable statistics principles;

(ii) The method provides a means of advance estimation of sample size on an objective basis. The sample size is no longer determined by traditional methods of guesswork; it is determined by statistical method;

(iii) The method provides an estimate of error. When probability sampling is used, the results may be validated in terms of how far the sample projection might deviate from the value that could be obtained by a 100% check;

(iv) Statistical samples may be combined and evaluated, even though accomplished by different auditors. That the entire test operation has an objective and scientific basis makes it possible for different auditors to participate independently in the same test and for the results to be combined as though accomplished by one auditor;

(v) Objective evaluation of test results is possible. Thus, all auditors performing this audit would be able to reach the same conclusion about the numerical extent of error in the population. While the impact of these errors might be interpreted differently, there can be no question as to the facts obtained, since the method of determining their frequency in the population is objective.

McRae (1982, p15) states that statistical sampling ensures that the sampling process is being handled in a logical, economical and consistent way. Colbert (1991) maintains that statistical sampling is to be preferred over judgemental sampling when an objective measure of risk is needed and an bound estimate of the monetary amount of error is desired. As such it seems an ideal tool for substantive testing.

1.3.3 Judgemental versus Statistical Evidence in a Court of Law

Auditors disagree on which sampling method is more defensible in court. Those favouring statistical sampling maintain that such sample testing would carry greater evidential weight in a court of law and that 'conclusions drawn from statistical sampling are more defensible in court because the risk of error in the population is objectively determined' (Colbert,1991). Copeland and Englebrecht (1975) maintain that statistical sampling results provide the auditor with data readily defensible in court because

> 'it gives the court quantitative standards to measure quantitative results, and the probability that deviations from the universe are not included in the results have been mathematically determined.'

On the other hand, auditors favouring a non-statistical approach believe that

> 'the use of professional judgement is a better defence--say in court-- than a statistical measure of risk. They would prefer to have expert witnesses explain how critical professional judgement is on an audit than a statistician explain that there is a known chance, say 5 or 10 percent, that the auditor's conclusion was incorrect.' (Colbert, 1991)

McRae (1982, p328) points out that overall there appears to have been very little discussion of the use of sample evidence in UK courts and states that there is no evidence to suggest that inferences from a statistically validated audit sample carries more evidential weight in a court of law than inferences and opinions based on a purely judgement audit sample. In the US however, there has been a greater legal interest in sampling techniques in general. Although no specific case in which the extent of audit sampling was a major issue has been brought before the courts, McRae (1982, p328) asserts that there was some evidence from the US to suggest that the law may be beginning to weigh up the adequacy of audit samples, which could lead towards statistical evaluation of audit sample size.

1.4 Statistical Sampling in Auditing

Sampling is used in both compliance and substantive testing and is described in numerous textbooks in auditing (see, for example, Arens and Loebbecke, 1981; Arkin, 1984; Guy, Carmichael and Whittington 1994; McRae, 1974; Roberts, 1986).

1.4.1 Statistical Sampling in Compliance Testing

Compliance testing is typically concerned with qualitative characteristics or attributes and statistical sampling is used to estimate the proportion of violations associated with a particular set of controls. For example, purchase orders may need to be authorised and compliance testing might estimate the proportion of times that they have not been authorised. Tests of compliance have normally been designed so as to provide information as to the rate of error in terms of control failure rather than to enable direct extrapolation in terms of monetary errors in the financial statements. There are a number of well known statistical techniques which have been utilised for compliance test purposes. They include:

(i) Estimation Sampling: With estimation sampling, a random sample of items of a specified size is selected and the proportion of errors or the average error amount is estimated to establish if it is less than some acceptable level. This is the most widely used statistical approach to compliance testing.

(ii) Acceptance Sampling: Acceptance sampling is a technique which enables the auditor to reject or accept the population under certain conditions. A sample of a given size is drawn and if more than a certain amount of errors is found, the population is accepted, otherwise it is rejected. The auditor using acceptance sampling seeks to balance out the risks of rejecting 'satisfactory' populations (and thereby frequently involving further audit costs) and of accepting 'unsatisfactory' populations (and thereby exposing the auditor to the potential risk of giving an inaccurate clean audit opinion). More efficient forms of acceptance sampling exist which involve two stage, multiple stage and sequential sampling plans. These methods are discussed in McRae (1974, pp24-34). Multistage sampling plans reduce the average sample size but the gains may be outweighed by the operational problems and costs incurred by the need for multiple evaluations.

(iii) Discovery Sampling: Discovery sampling is a sampling plan which selects a sample of a given size, accepts the population if the sample is error free, and rejects the population if it contains at least one error. With discovery sampling, the auditor may not be interested in determining how many errors there are in the population. Where there is a possibility of avoidance of the internal control system, it may be sufficient to disclose one example to precipitate further action or investigation.

The theoretical properties of these sampling plans are well known and have been applied in other fields, notably the area of industrial quality control (see Duncan, 1986). Their use in auditing was initially suggested by Vance (1947) and later by Vance and Neter (1956) and Arkin (1961). The application of statistical sampling plans and inference procedures to compliance testing presents few difficulties and does not involve any special problems (Neter, 1986, Neter and Godfrey, 1988). It will not be dealt with in this study.

1.4.2 Statistical Sampling in Substantive Testing

In substantive testing, statistical sampling is used to obtain monetary estimates of the total error amount or confidence limits for the total error amount in a particular accounting population. The objective is to obtain reliable confidence limits, (i.e. confidence limits with actual confidence levels never less than their nominal levels) which are not conservative (i.e, the estimate of the total error amount should not be very much greater than the true error amount) with sample sizes that are not too large for practical audit applications. Neter (1986) points out that serious problems are encountered when applying statistical sampling methods to sampling account balances and Neter and Godfrey (1988) state that the application of statistical sampling and inferential procedures to substantive testing presents some challenging problems. The following section discusses the statistical sampling problem in substantive testing.

1.4.2.1 The Sampling Problem in Substantive Testing

At first glance, it would appear that the sampling and estimation procedures for obtaining estimates of the total error amount in substantive testing are straightforward. The problem appears to be well defined. The population is finite, containing N accounts say. The book values B_1, B_2 ... B_N are known. The true values $A_1, A_2 \ldots A_N$ and the errors $E_i = B_i - A_i$ are unknown. Since N may to too large to verify the accuracy of the complete population, it is necessary to obtain a sample of accounts in order to estimate the total error amount. It would seem appropriate to use the set of book values as the auxiliary variable and to stratify the population by book value size and select a stratified sample using optimum allocation. Auxiliary estimators using the book values as the auxiliary variable, and the central limit theorem might then be used to estimate the total error amount and to obtain confidence limits. This is the classical sample survey approach as outlined in standard texts (see, for example, Cochran, 1977; Kish, 1965; and Moser and Kalton, 1971).

Both simple random sampling and stratified sampling of line items have been used in substantive testing. Stratification is a process of dividing a population into subgroups each of which is a set of sampling units with similar characteristics. Stratification of accounting populations is usually based on the recorded book value amount of the line items and a sample is selected independently from each stratum. In the draft guideline on audit sampling issued by the APB (1993), stratification is advocated as an acceptable sampling method on the basis that it enables the auditor

> 'to direct audit efforts towards the items which, for example, contain the greatest potential monetary error. For example, the auditors may direct attention to larger value items for accounts receivable to detect major overstatement errors.'

Roberts (1978, Chap. 6) gives a detail account of the application of stratified sampling methods in auditing.

Two major problems are encountered when the classical sampling and estimation approach is applied to auditing. First, accounting populations often have very low error rates and consequently the selected sample may yield zero errors and hence fail to give any information on the population total error amount. For example, when the error rate in the population is

.01, the probabilities that sample random samples of sizes 30, 60 and 100 will contain no errors are 0.74, 0.55 and 0.38 respectively. When this situation occurs, the population error amount would be estimated at zero if classical estimation procedures are used and confidence limits for the total error amount cannot be obtained.

The second problem pertains to the unreliability of confidence intervals, i.e., confidence intervals with actual confidence less than the nominal. Often, the average line item error amount (mean-per-unit) is used as an estimate of the total error amount and the central limit theorem is applied to obtain the confidence limits. Numerous studies have shown that the mean-per-unit estimator with simple random sampling leads to unreliable confidence intervals when the populations have low error rates and when the line items are highly skewed (see for example, Kaplan, 1973; Neter and Loebbecke, 1975). Studies have also shown that the confidence intervals are unreliable with stratified sampling for most sample sizes used in auditing. Menzefricke and Smieliauskas (1987b) examined the mean-per-unit stratified estimator and found that the sample size required to obtain reliable confidence intervals may be too large for practical applications. Dunmore (1986) confirmed that large sample sizes within each stratum are necessary for the confidence level to be sufficiently close to the nominal.

Auxiliary estimators have been proposed to overcome the unreliability problem (see Cochran, 1977 for a description of auxiliary estimators). Kaplan (1973) examined the performance of the ratio, regression and difference auxiliary estimators. Both conceptually and through simulation studies, with sample sizes ranging from twenty five to two hundred, Kaplan demonstrated that the use of the t-statistic for statistical inference is inadequate and leads to unreliable confidence intervals. He pointed out that since there is often a high correlation between the estimator and the estimate of the standard error, the numerator and the denominator of the t-statistic will not be independent and hence the t-statistic is invalid. Neter and Loebbecke (1975) obtained similar results in an empirical study using actual accounting populations. They found that with highly skewed, low error rate populations, the achieved confidence levels with auxiliary estimators are far below the nominal levels. Beck (1980) examined the regression estimator using samples of size 600 and found that even with this large sample, the regression estimator cannot be relied upon to provide a confidence interval with actual confidence close to the nominal. Frost and Tamura (1986) proved that the skewness of the distribution of auxiliary estimators largely accounted for their failure to yield reliable confidence intervals in statistical auditing. Neter and Loebbecke (1977) noted that the low error rates in accounting populations lead to downward bias in the estimates of the standard error and consequent unreliable confidence limits.

The jackknife technique has been used to improve the estimate of the standard error (see Efron, 1979; Miller,1974; Mosteller, 1971 for a description of the jackknife technique) and hence to provide reliable confidence intervals. Empirical evidence has shown that this has not succeeded in its objective. Frost and Tamura (1982) found that the unreliability problem is not completely solved when the jackknife technique is used to reduce the downward bias in the standard error of the estimator. They showed that although the jackknifed ratio estimator provides higher reliability than the ratio estimator when the error rate is high and involves both understatement and overstatement errors, however, neither the ordinary ratio estimator nor the jackknifed ratio estimator are reliable in accounting populations with low error rates or in situations where the errors are all overstatements.

Authors are practically unanimous in their verdict that conventional sampling methods and estimation procedures are not appropriate to substantive testing. Kaplan (1973) concluded from his study on auxiliary estimators that 'entirely new approaches may be required for statistical sampling in auditing.' Smith (1979) deduced that the conventional approach based on confidence intervals does not answer the auditors questions. Gwilliam (1987, p244) noted that even though there has been an extended debate on the properties and suitability of the classical methods of statistical sampling, it would not appear that these methods enjoy more than limited use by auditing firms for the purpose of the substantive testing of value.

A vast amount of research has gone into developing statistical techniques suitable for substantive testing. Sampling methods and estimation procedures which do not rely on large sample normal theory have been developed to give more reliable bound estimates of the total error amount. A non-classical procedure which has gained the most significant acceptance among the major accounting firms for the purpose of regular audit use is monetary-unit sampling (Gwilliam, 1987, p245). The following discussion provides the historical development of monetary-unit sampling methods pertaining to substantive testing.

1.5 The Development of Monetary-Unit Sampling

The term 'monetary-unit sampling', is often used to denote the sample selection procedure and the evaluation procedure. Each of these is discussed below.

1.5.1 Monetary-Unit Sample Selection

Monetary-unit sample selection views the population, not as a population of accounts of different sizes, but as a population of monetary units. The size of the population is taken to be the total number of monetary units in all the accounts and each monetary unit is selected with epsem probability i.e., each monetary-unit has an equal chance of selection. Monetary-unit sample selection gives each line item a probability of selection proportional to its stated monetary value. Probability proportional to size selection was originally developed in survey sampling theory by Hansen and Hurwitz in 1943 for selecting clusters of unequal size. Deming (1960, pp110-185) was the first to put forward the idea of monetary-unit sampling. He suggested that an individual dollar of investment could be considered to be the sampling unit but he did not elaborate further on the idea of monetary-unit sampling.

The first known published research on monetary-unit sampling applied to accounting populations was done by Van Heerden (1961). He defined the sampling unit as a guilder, a monetary unit. He suggested that an account balance or line item could be thought of as a collection of monetary units, some of which were 100 percent correct and some were 100 percent in error, i.e., the all-or-nothing method of allocating error amounts to monetary units. The incorrect units were considered to be the last monetary units appearing in the line item. Van Heerden suggested that a simple random sample of monetary units be taken from the total balance.

In 1968, the Canadian Institute of Chartered Accountants (CICA) commissioned a study of statistical sampling as it is applied to auditing in an attempt to promote a better understanding of statistical sampling among auditors. The study was published in 1972 by Meikle. It included a description of a monetary-unit sampling plan, referred to as Cumulative-Monetary-Amounts sampling. This sampling method had been used for some years previously by Haskins and Sells (later reorganised as Deloitte, Haskins and Sells) and described in the Haskins and Sells instruction manual on audit sampling (Haskins and Sells, 1970).

In 1973, Anderson and Teitlebaum provided the first complete description of a monetary-unit sampling method which they called Unrestricted Dollar-Unit sampling. This was equivalent to the Cumulative-Monetary-Amounts sampling described by Meikle (1972). In this study, Unrestricted Dollar-Unit sampling will be referred to as simple random sampling of monetary units. Simple random sampling of monetary units selects a simple random sample of monetary units from a list of cumulated book values. Anderson and Teitlebaum (1973) also proposed systematic sampling, a sampling method which divides the total book value amount into **n** (the sample size) equi-sized sections and selects a sample of monetary units systematically after a random start in the first section. Anderson (1973) discussed the use of statistical sampling in auditing and gave a brief description of monetary-unit sampling. Leslie (1975) provided a non-technical explanation of monetary-unit sampling with particular emphasis on its use by government auditors in the International Journal of Government Auditing.

In a paper presented at the National Meeting of the American Statistical Association in 1973, Teitlebaum presented a more formal definition of monetary-unit sampling and introduced monetary-unit cell selection. Cell sampling is similar to systematic sampling in that it divides the total book value amount into **n** (the sample size) equi-sized sections or cells and selects an element at random from each cell. In cell sampling, however, an independent random selection is made within each cell. Leslie, Teitlebaum and Anderson (1979, p110) described the cell selection method in detail and recommended that cell selection should be used whenever monetary-unit sampling is being considered and they suggested that cell sampling is preferable to systematic sampling 'because it is not significantly harder to do and because it avoids any doubt as to rigorousness'. Goodfellow, Loebbecke and Neter (1974) discussed the basic concepts of monetary-unit sampling.

In the Netherlands, a sample selection method called sieve sampling was proposed by Rietveld (1978,79) as a practical alternative to simple random, cell and systematic sampling. Instead of looking at the population as a collection of monetary units, the sieve selection method looks at the population as a collection of items, each of which has a probability proportional to its monetary value of being selected. The mathematical background of the selection method is given by Gill (1983). Driessen (1986) showed that the sieve method may, under certain conditions, be validly used for two- or three-stage sampling schemes, even though the statistical evaluation is based on simple random sampling of monetary units.

1.5.2 Monetary-Unit Sample Evaluation

Research on monetary-unit sampling and non-classical estimation procedures has concentrated on trying to obtain bounds for the total error amount which are reliable but not too conservative. Numerous bound estimates of the total error amount have been developed.

In 1963, Stringer proposed a heuristic procedure for estimating the total error amount based on the Poisson distribution and calculated the error in each monetary unit as a proportion of the error in the associated item. This bound is referred to as the Stringer bound and is widely used in practice (Felix, Leslie and Neter, 1982). Stringer (1963) used stratification and line item selection when choosing the items to evaluate the bound. Anderson and Teitlebaum (1973) suggested that monetary-unit sampling be used with the Stringer estimation procedure and claimed that the upper bound estimate gives a conservative upper error limit for any population. There is no formal proof of this assertion but all empirical evidence tend to support their claim. Simulation studies have shown that although the bound always achieves the specified confidence, it is conservative in the sense that the estimate of the error amount obtained from this bound is usually far in excess of the true error amount (see, for example, Felix and Kinney 1982; Leitch, Neter, Plante and Sinha,1982; Neter and Loebbecke, 1975; Reneau, 1978).

Teitlebaum (1973) introduced the Cell bound for estimating the total error amount. Leslie, Teitlebaum and Anderson (1979) defined the Cell bound in detail and suggested that it be used 'when it is necessary to eliminate the conservatism present in regular DUS evaluations.' While this bound was developed for cell sampling, its computational form does not restrict its use to cell sampling alone (Wurst, Neter and Godfrey, 1989b). Fienberg, Neter and Leitch (1977) introduced the multinominal bound, a bound for estimating the total error amount based on the multinominal distribution. It was found to be computationally tedious when the number of errors in the sample was large. Leitch, Neter, Planta and Sinha (1982) modified the bound to reduce the number of computations required under the original model.

Garstka and Ohlson (1979) developed a bound based on an unbiased point estimator of the total error amount. It is similar to the central limit theorem approach in that the upper bound is defined by the point estimator plus the standard error multiplied by an appropriate constant. The multiple is a function of both the sample size and the number of errors found in the sample. Garstka and Ohlson (1979) showed that the multiple is greater than the corresponding normal coefficient when the number of errors in the sample is low. The main advantage of the Garstka-Ohlson bound is its ability to obtain bounds with greater reliability than the central limit theorem when the number of errors in the sample is low. However, like the central limit theorem, it gives no information on error bounds when the number of errors in the sample is zero. Tamura (1985) showed that the Garstka-Ohlson bound is unreliable in populations with small proportionate errors.

The Moment bound was proposed by Dworkin and Grimlund (1984) as an alternative to the Stringer bound. They tested the bound under a wide range of test conditions and found that it is reliable under most of the test conditions. Menzefricke and Smieliauskas (1987b) noted that the Moment bound 'is more comparable to complex bounds and its robustness is supported by considerable empirical evidence'. It has recently been adopted by Arthur Anderson as a replacement for the Stringer bound (Felix, Grimlund, Koster and Roussey, 1990).

1.5.2.1 Modified Bounds

The non-classical approach appears to have overcome the two major problems of the classical approach (i.e., it provides an estimate of the total error amount when there are no errors found in the sample and it provides reliable confidence intervals). However, most empirical research on monetary-unit sampling indicates that the bounds are conservative (i.e., the estimate of the total error amount is usually far in excess of the true error amount). Conservative estimates lead to rejection of acceptable populations and increases the cost of the audit unnecessarily. In an attempt to obtain estimates which are less conservative, Smith (1979) suggested a modification of the Stringer bound by projecting sample results into the unsampled population. In a simulation study comparing the modified and unmodified Stringer bounds, he observed that although modification did lead to tighter upper error limits, the limits were still conservative. Phillips (1985) found that the modified Stringer bound is tighter than the unmodified bound and that the mean coverage of the modified bound is nearly always equal to the unmodified bound. Wurst, Neter and Godfrey (1991) examined the modified Stringer and Cell bounds with simple random, systematic and cell selection for samples of sizes 65, 150 and 300. They observed that while modification reduced the risk of incorrect rejection for samples of sizes 150 and 300, the risks of incorrect rejection was not reduced for samples of size 65. Since a sample size of 65 may be closest to the sizes used in practice, the findings indicate that modification may not be of practical use to the auditor.

Modification has been considered by MacGuidwin,Roberts and Shedd (1982) but not pursued on the basis that it resulted in biased estimates of the total error amount. Leslie and Andersley (1982) discouraged the use of bound modification, pointing out its dangers thus;

> 'If errors are concentrated in smaller book values, the bound will be understated too frequently. If errors are concentrated in larger population items, the bound will usually be too high. It is only when the errors are evenly distributed across the entire range of population values that the results are acceptable - a condition never known to the auditor. Johnson, Neter and Leitch (1981) found larger items have a larger probability of containing an error and this approach would usually be unduly conservative for most acceptable bounds. In addition, the point estimates produced by this bound are biased.

1.5.2.2 Bayesian Bounds

Another approach suggested for reducing the conservativeness of the estimates in monetary-unit sampling has been Bayesian estimation. Bayesian estimation consists of obtaining a probability distribution reflecting possible values of a parameter, called a prior distribution and combining this with the sample results to obtain a posterior distribution. The mathematical foundation of Bayesian methods and a description of its applications can be found in Winkler (1972). A Bayesian bound allows the auditor to systematically integrate evidence derived from sampling procedures with evidence derived from other sources and seems appropriate to the particular needs of auditing. In fact Smith (1979) maintained that the only satisfactory solution to the problem of estimating the total error amount in auditing is the Bayesian approach.

Numerous Bayesian bounds for estimating total error amount have been developed for use with monetary-unit sampling. Cox and Snell (1979) derived a Bayesian upper bound for the total overstatement amount. Much work has been done on the Cox and Snell bound to obtain prior distributions that will produce reliable bounds which are not conservative (see for example Godfrey and Neter 1984; Neter and Godfrey, 1985; Phillips, 1985; Tsui, Matsumura and Tsui, 1985). McCray (1984) introduced a Bayesian bound for the total error amount which assumed a discrete prior distribution. Dworkin and Grimlund (1986) found that the Moment bound provided narrower confidence intervals than the McCray bound in populations with low error rates (less than or equal to 6%) and that the McCray bound was not as conservative as the Moment bound in populations with high error rates (greater than or equal to 75%). Both bounds provided comparable results in the intervening cases but the Moment bound had less computational requirements.

The application of the Bayesian methodology to the audit situation has been reported by some authors. McRae (1982, p160) reports that a UK firm, Thomsom McLintock had developed a Bayesian sampling procedure in auditing. Kirtland and Holstrum (1984) state that Deloitte, Haskins and Sells are using a semi-Bayesian approach in the MUS system and Abdolmohammadi (1987) reports that Touch Ross provide guidelines in their audit sampling for assessing inherent assurance based on the analysis of the factors that contribute to the likelihood of material error.

However, there is very little evidence that auditors are actually using Bayesian methods in a rigorous manner. The reason for this appears to be the difficulty of estimating prior probability distributions, together with an inherent lack of confidence in the prior probability distributions. Godfrey and Andrews (1982) claim that 'the requirement that prior beliefs be quantified is the probably main reason Bayesian methods have not been more widely adopted by practitioners'.

1.5.2.3 Theoretical Approach To Evaluation of Monetary-Unit Sampling

Most of the bounds used with monetary-unit selection are heuristic in nature (a notable exception is the Stringer/Cell bound with the AON error assignment using simple random, cell and sieve sampling (see section 2.8)). Attempts to analyse them mathematically have been limited in their success. Testing the reliability of the heuristic bounds has largely been in terms of simulation studies using accounting populations seeded with errors appropriate to what might occur in real life. It should be pointed out, however, that any conclusions drawn from simulation studies on specific audit populations are tentative and cannot immediately be generalised to all existing audit populations. But this does not appear to have worried practitioners, as McRae (1982, p253) explains that

> 'since the most commonly used versions of the MUS system appear to work as predicted under the simulated conditions, the absence of a full mathematical proof is not likely to deter an auditor from using it'

It must be admitted that the lack of a theoretical validation of the bound evaluation seems more likely to be of concern to the statistician than to the auditor. Recently, a rather novel attempt at theoretical validation of the heuristic bounds has been taken by Wurst, Neter and Godfrey (1989a, 1989b). They hypothesised that the accuracy of a point estimator of the total error amount may carry over to the accuracy of the bounds and hence a theoretical analysis of the point estimator may give some indication of the comparative accuracy of the bounds with different sampling methods. They derived and compared the precision of a point estimator of the total error amount using simple random, cell and sieve sampling (Wurst, Neter and Godfrey, 1989a). They showed that the comparative behaviour of the bound estimates of the total error amount with the different sampling methods is similar to the comparative behaviour of the point estimator for the total error amount (Wurst, Neter and Godfrey, 1989b). For example, they proved theoretically that a point estimator of the total error amount is more precise with sieve sampling than with simple random sampling for larger sample sizes but not as precise as cell sampling. In a simulation study, they found that the sample selection methods had the same effects on the Stringer and Cell bounds estimates of the total error amount.

While in no way can this be taken to be a theoretical validation of the heuristic bounds, it is an interesting approach to comparing the properties of the bounds for different sampling methods. This research intends to develop this approach further and to compare the precision of the point estimator with the precision of the bound estimators of the total error amount for other sampling methods.

1.5.2.4 The Use of Monetary-Unit Sampling

Monetary-Unit sampling has gained wide acceptance in the auditing profession in recent years. McRae (1974, p224) stated that the monetary-unit sampling system is the best so far devised for external auditing. Kaplan (1975) described monetary-unit sampling as 'one of the prime statistical sampling procedures available to auditors'. Leslie, Teitlebaum and Anderson (1979) noted that it has been extensively used by government auditors. McRae (1982, p178) found that in the UK, over 90% of firms using statistical sampling to evaluate the total error amount use some form of monetary-unit sampling. In a more recent study, Abdul-Hamid (1993) also found that, of those firms using statistical sampling in substantive testing, the monetary-unit sampling approach was used predominantly. Variants of MUS are currently being employed by Deloitte, Haskins and Sells, Arthur Young, Peat, Marwick, Mitchell and Touche Ross among others. Menzefricke and Smieliauskas (1987a) noted that virtually all large firms in the US are now adopting monetary-unit sampling, 'at least on an experimental basis'.

1.6 Need for the Research

Most of the empirical research on monetary-unit sampling method has been done in the US. Abdul-Hamid (1993) in comparing the UK and the US auditing environment pointed out that:

> 'Despite the obvious similarities between the USA and the UK auditing environments, there are important economic, legal and cultural differences which differentiate results obtained in the USA from those in the UK'

The same could be said of Ireland and therefore research which attempts to validate the US findings in an Irish environment would be useful. This study replicates some of the work done in the US and Canada by investigating the performance of existing monetary-unit sampling when applied to two Irish accounting populations.

In addition, this study addresses some of the issues raised by previous authors. In his research on sieve sampling Wurst (1990) called for future research on;

(i) the performance of the Moment bound using sieve sampling;

(ii) the comparative performance of sieve sampling and systematic sampling of monetary units.

Atkinson (1990) studied the performance of some monetary-unit sampling methods and also called for further research on the performance of the Moment bound with different sampling methods.

This study extends the work of Wurst (1990) and Atkinson (1990) by investigating the performance of sieve sampling using the Moment bound for estimating the total error amount and by carrying out a study of the comparative performance of sampling methods currently used in practice including a comparative investigation of systematic sampling and sieve sampling.

There are practical aspects of the monetary-unit sampling methods currently used in practice which may be of concern to the auditor. For example, simple random, systematic and cell sampling ignore the line item structure of the population when selecting the sample and treat the population as a collection of monetary units from which a random sample of monetary units is selected. Since only the line items containing the selected monetary units can be tested by the auditor, the selected monetary units must be traced back to their associated line items. Wurst, Neter and Godfrey (1989a) point out that this may cause implementation problems. Leslie, Teitlebaum and Anderson (1979) cite some of the practical disadvantages of simple random, of monetary units.

> 'The first is the minor nuisance of having to accumulate the book value totals. The second and more important is the need to know the total book value amount accurately *before* selection can begin'

The need to know the book value total in advance of sampling may impede the planning and implementation of the auditing process since as Leslie, Teitlebaum and Anderson (1979) point out that

> 'often the total book value amount is not known accurately during the planning stage, nor is it known for transaction streams prior to the end of the year'.

Clearly, a sample selection method which overcomes these difficulties would be of use to the auditor. The sieve sample selection method is a possibility (see 4.7). Sieve sampling uses the line item structure of the population when selecting the monetary units and does not require that the book amounts be accumulated or that the total book value be known in advance of sampling. However, a disadvantage of sieve sampling is that the sample size is not constant. It varies depending on the random numbers chosen when selecting the sample and this may be of serious concern to the auditor when the costs of carrying out the audit are being estimated prior to the audit.

New monetary-unit sample selection methods which preserve the advantages of sieve sampling while returning a fixed sample size may be of benefit to the auditor. Two such sampling methods are proposed in this study. A new sampling method, 'Stabilised Sieve Sampling' is defined, and a sampling method which has not been applied previously in auditing 'Lahiri Sampling' is introduced. Both sampling methods use the line item structure of the population when selecting a sample. They also return a constant sample size of monetary units.

To test the performance of the monetary-unit sampling methods, it is necessary to investigate the results obtained when the methods are applied to populations with book value and error characteristics similar to those found in real accounting populations. Previous research on the performance of the monetary-unit sampling methods have been tested on data derived mainly from accounting populations in the US (for example, Neter and Loebbecke, 1975; Johnson, Leitch and Neter, 1981). As far as the writer is aware, only two studies have been carried out by UK authors to investigate the performance of monetary-unit sampling. Smith (1979) used simulated accounting populations to compare the performance of upper bound estimates of the total error using monetary-unit sampling. Abdul-Hamid (1993) conducted a small simulation study to find out if monetary-unit sampling, for sample sizes usually used in the UK, is likely to pick up the degree of error that can be anticipated to exist in audited populations of accounting data. He based his study on a population from the Neter and Loebbecke US database (Neter and Loebbecke, 1975).

It is not unreasonable to assume that book values and error characteristics of Irish accounting populations may be similar to those in the US and hence to expect the comparative performance of existing monetary-unit sampling methods to be the same when used on Irish accounting populations and US accounting populations. However, as has been pointed out above, research which attempts to validate the US findings in an Irish environment would be useful. This study obtains data on the characteristics of the book values and the patterns of errors from two commercial entities of debtors in the Public Sector and uses these to test the performance of the sampling methods.

1.7 Scope of the Research

This study develops new monetary-unit sampling methods for use in substantive auditing of debtors and tests their performance on accounting populations of debtors from commercial entities in the Irish Public Sector. This research goes beyond prior audit sampling research in the following ways:

(i) It examines error characteristics from a sector that has not been previously investigated. This study will provide, for the first time, information on the characteristics of book values and error patterns of two populations of debtors in the Public Sector;

(ii) It extends the work done on sieve sampling by Wurst, Neter and Godfrey (1989a, 1989b) by investigating sieve sampling using the Moment bound. It also compares the performance of sieve sampling with systematic sampling;

(iii) It devises and applies monetary-unit sampling methods that have not been previously used in auditing.

1.8 Objectives of the Study

There are five major objectives. These are:

(i) To obtain information on the characteristics of book values and patterns of errors in two populations of debtors in the Public Sector.

(ii) To carry out a theoretical analysis of a point estimator of the total error amount for six monetary-unit sampling methods. Four methods are currently used in practice i.e., simple random, systematic, cell and sieve sampling), one 'Lahiri Sampling' which has not been previously applied in auditing (see also objective (iv)), and a new sampling method 'Stabilised Sieve Sampling' which has been developed in this study for use in substantive testing (see also objective (v).

(iii) To investigate the performance of the Stringer, Cell and Moment bounds for estimating the total error amount in substantive auditing using monetary-unit sampling methods currently used in practice.

(iv) To investigate Lahiri sampling as an alternative to simple random sampling of monetary units.

(v) To investigate stabilised sieve sampling as an alternative to sieve sampling and simple random sampling of monetary units.

A more detailed description of these objectives is given below.

1.8.1 Objective 1. Characteristics Debtors in the Public Sector

The first objective is to obtain information on the book values and the error characteristics of debtors in the Public Sector and to compare the structures of these with accounting populations already available.

1.8.2 Objective 2. Point Estimator Analysis

The second objective is to examine and compare the precision of a point estimator of the total error amount for all the monetary-unit sampling methods used in this study. Wurst, Neter and Godfrey (1989a) investigated the statistical properties of an unbiased point estimator of the total error amount under sieve sampling, simple random sampling and cell sampling. This study extends their work and derives the properties of a point estimator of the total error amount with Lahiri and stabilised sieve sampling and with simple random, cell, systematic and sieve sampling. The purpose is to establish whether the theoretical properties of the point estimator are consistent with the properties of the bounds observed using simulation.

1.8.3 Objective 3. Comparative Bound Performance with Sampling Methods Currently Used in Practice

In the bound analysis for sieve sampling, Wurst, Neter and Godfrey (1989b) confined their investigation to the Stringer and Cell bounds to compare sieve, cell and simple random sampling of monetary units. This investigation extends their work to include the Moment bound and systematic sampling. This study investigates the Stringer, Cell, and Moment bounds using four monetary-unit sampling methods currently used in practice, i.e. simple random, systematic, cell and sieve sampling of monetary units. The comparison is performed by means of a simulation study using two accounting populations of debtors from the Public Sector.

1.8.4 Objective 4. Comparison of the Bound Performance with Lahiri Sampling and Simple Random Sampling of Monetary Units

The fourth objective is to investigate Lahiri sampling as an alternate to simple random sampling of monetary units. The two sampling methods are compared using the Stringer, Cell and Moment bounds to estimate the total error amount. The comparative performance of the sampling methods is tested on two accounting populations of debtors from the Public Sector.

1.8.5 Objective 5. Comparison of the Bound Performance with Stabilised Sieve Sampling and Sieve Sampling of Monetary Units

The fifth objective is to investigate stabilised sieve sampling as an alternate to sieve sampling and simple random sampling of monetary units. The sampling methods are compared using the Stringer, Cell and Moment bounds to estimate the total error amount. The comparative performance of the sampling methods is tested on two accounting populations of debtors from the Public Sector.

1.9 Limitations of the Study

This study is an investigation of the comparative performance of monetary-unit sampling methods and estimation procedures used in substantive testing and applied to accounting populations of debtors in the Public Sector. It does not deal with;

(i) Creditors and Stock

(ii) Understatements

(iii) Lower Bounds

(iv) All the populations of debtors in the Public Sector

(v) Wider issues of public accountability.

The details of the limitations are given below.

(i) Creditors and Stock

In monetary-unit sampling, the probability of selection of any line item is proportional to its reported value and not necessarily its actual value. Consequently, this selection method is predisposed to selection of overstatements as opposed to errors of understatement. Since errors in debtors are predominantly overstatements (see Chapter 3), monetary-unit sampling is suitable to selecting samples of debtors. On the other hand, errors in creditors are likely to be understatements (Johnson, Leitch and Neter, 1981) and therefore, monetary-unit sampling is not an appropriate selection method. Consequently, accounting populations of creditors are not included. The commercial entities studied are service enterprises and in these, stock is of negligible importance, hence accounting populations of stock are not investigated.

(ii) Understatements.

The study confines itself to overstatement errors because these errors are of primary interest to auditors (Wurst, Neter and Godfrey, 1989a). It does not consider understatements. While some studies have shown (see Chapter 3) that errors in debtors are predominantly overstatements, others have found that understatements may occur in debtors (see for example, Johnson, 1987; McRae, 1982, p70) but, as was pointed out in (i) above, monetary-unit sampling may not be the most effective type of sampling when understatement errors play a dominant role.

(iii) Lower Bounds

Because this study deals with debtors only and because overstatements and the estimation of the maximum overstatement is one of the main priorities in debtors accounts (Wurst, Neter and Godfrey, 1989a), the analysis concentrates on obtaining precise estimates of the total error amount and reliable upper bounds for the total error amount.

Estimation of lower bounds for the total error amount is not included. Some of the procedures used in the study may also be applied to estimating lower bounds for the total error amount but lower bound analysis is not specifically dealt with here.

(iv) The Populations

This study confines itself to two populations of commercial entities audited by the office of the Comptroller and Auditor General. The error modelling described in Chapter 3 attempts to recreate all possible situations for the populations studied. Generalisations to other populations must at most be tentative. In particular, Public Sector commercial entities audited by private sector auditing firms are not included and would be of interest in future research.

(v) Wider Issues of Public Accountability

Hardman (1991) noted that the CAG has a wider public duty, extending beyond the mere expression of an opinion on the financial statements, which encompasses the public accountability of the government to the legislature and ultimately to the electorate. This issue is deemed to be beyond the scope of this study.

1.10 Structure of the Thesis

The remainder of the thesis is structured as follows:

Chapter 2 gives an overview of the research methodology. It describes how the audit populations are generated. It details the sampling methods and defines the point estimator and the bounds used in the investigation. The measures for assessing the performance of the bounds are defined and the simulation procedure is outlined.

Chapter 3 details the statistical characteristics of the book values of the Public Sector debtors used in the study. It outlines the methodology used in designing the sampling plan to obtain the audit samples and describes the error patterns found in the audit samples. It also describes how the populations used in the simulation study are generated.

Chapter 4 details the monetary-unit sampling methods used in the study and derives the properties of each sampling method.

Chapter 5 examines the precision of a point estimator under the various sampling methods. In this chapter, an attempt is made to discover whether different sampling methods give more precise estimates of the total error amount for populations with different line item structures and error distributions.

Chapter 6 compares the performance of the four monetary-unit sampling methods currently used in practice, using the Stringer, Cell and Moment bounds to estimate the total error amount. A comparative investigation of the sampling methods is carried out by means of a large scale simulation study using the two actual accounting populations of debtors from the Public Sector.

Chapter 7 compares the performance of Lahiri and simple random sampling of monetary units using the Stringer, Cell and Moment bounds to obtain upper bound estimates of the total error amount.

Chapter 8 compares the performance of stabilised sieve and sieve sampling of monetary units using the Stringer, Cell and Moment bounds to obtain upper bound estimates of the total error amount. Simple random sampling of monetary units is used as a bench mark in the comparison of stabilised sieve and sieve sampling of monetary units.

Chapter 9 summarizes the results and provides suggestions for future research.

Chapter 2

The Methodology

2.1 Introduction

This chapter outlines the research methodology used to achieve the objectives stated in 1.8 and discusses issues arising from that methodology.

The study examines the effects of six sample selection methods on the behaviour of estimates of the total error amount in substantive testing. Accounting populations in the Public Sector are examined, and their error characteristics are determined by means of large scale investigative audits. Audit populations are created with different error rates and error amounts reflecting the error patterns found in the investigative audits. A theoretical analysis of a point estimator of the total error amount is carried out for each sampling method. The sampling methods are tested on the populations by means of a simulation study, using three upper bound estimates of the total error amount, three different sample sizes and three different confidence levels.

This chapter:

introduces;

(i) the monetary-unit sampling methods (2.2)

(ii) the populations of book values (2.3)

(iii) the investigative audits (2.4)

(iv) the error models (2.5)

(v) the simulation study (2.10)

(vi) the data analysis (2.12)

(These issues will be discussed in more detail in later chapters)

and discusses;

(vii) the error assignment methods (2.6)

(viii) the point estimator (2.7)

(ix) the upper bounds (2.8)

(x) the confidence levels (2.9)

(xi) the criteria for assessing the performance of the sampling methods (2.11)

2.2 The Sampling Methods

Monetary-unit sampling is a method of random sampling from accounting populations where the sampling unit is a monetary unit, e.g., Ir£1. The size of the population is defined as the total monetary amount of all the line items. In monetary-unit sampling, a line item of monetary value Ir£B_i say, is considered to consist of B_i monetary units, each having an equal chance of selection. This study examines six such monetary-unit sampling methods, four of which are currently used in substantive testing, one which has not been applied previously in auditing and a new sampling method developed in this study.

The six monetary-unit sampling methods are as follows:

(i) Simple random sampling of monetary units where each possible combination of monetary units has an equal chance of being selected;

(ii) Systematic sampling of monetary units where units are chosen systematically after a random start;

(iii) Cell sampling of monetary units where the population is divided into cells and an independent selection of one monetary unit is made from each cell;

(iv) Sieve sampling of monetary units where a random number is chosen for each line item and one monetary unit from each line item is considered for inclusion into the sample. It is selected depending on the random number chosen;

(v) Lahiri sampling of monetary units which selects a pair of random numbers, one to decide on a line item for consideration for inclusion into the sample and the other to determine what monetary unit, if any, should be selected from this line item. Sampling of pairs of random numbers continues until the required number of monetary units is selected;

(vi) Stabilised sieve sampling of monetary units which selects an initial sample by means of sieve sampling and increases or decreases the sample randomly so that the final number of monetary units in the sample is equal to the nominal sample size.

Simple random, systematic and cell sampling ignore the line item structure of the population when selecting the sample of monetary units and consider the population as a collection of monetary units from which a random sample is drawn. Lahiri, sieve and stabilised sieve sampling, on the other hand, use the line item structure of the population when selecting the monetary units. With Lahiri, sieve and stabilised sieve sampling, the line items are selected randomly and a monetary unit is chosen from each selected item. Sieve sampling was proposed as an alternative to simple random, systematic and cell sampling by Gill (1983). Lahiri sampling was suggested by Lahiri (1951) as a convenient method of selecting clusters of unequal size with probability proportional to size. It has not yet been used in statistical sampling in auditing. Stabilised sieve sampling of monetary units is a new sampling method developed in this study. The six sampling methods are defined in detail and their statistical properties are derived in Chapter 4.

2.3 The Populations of Book Values

The populations of book values on which the sampling methods are tested, were supplied by the office of the Comptroller and Auditor General. They consist of debtors accounts from two commercial entities audited by the office of the Comptroller and Auditor General. One population contains a relatively large number of small debtors and the second population contains a relatively small number of large debtors.

Population 1

Population 1 is composed of 3725 debtors accounts from a state scientific consultancy firm. It has a total book value amount of Ir£3,522,610. The firm is responsible for a number of national standards and also provides various technical services to industry.

Population 2

Population 2 consists of 662 debtors accounts from an industrial support body. It has a total book value of Ir£11,630,830. The firm provides grant aid to industries setting up in Ireland. It is also involved in renting factories and other premises to entrepreneurs.

Summary

Populations 1 contains relatively low valued line items while Population 2 contains relatively high valued line items. The median sizes of Population 1 is Ir£240 and the median size of the line items in Population 2 is Ir£3,740. These populations provide an excellent opportunity of investigating the sampling methods under different conditions. Detailed descriptions of the populations together with their distributional characteristics are given in Chapter 3.

2.4 The Investigative Audits

A large investigative audit was carried out on each of the populations to obtain information on error patterns and error amounts. The populations were stratified by book value size and disproportionate stratified random samples of line items were selected from each population. Stratification by book value size was used to investigate error characteristics in line items of differing sizes. The purpose was to obtain sufficient information on error patterns in order to model the errors in the populations. Stratification was not used in the subsequent sampling and estimation procedures. The sample designs used in the investigative audits and the characteristics of the error patterns are outlined in Chapter 3

2.5 The Error Models and Audit Populations

Audit populations were created from each of the accounting populations by seeding errors into the populations of book values with error rates and error sizes reflecting the patterns found in the data obtained from the investigative audits.

Line items in error found in the investigative audit were either (i) 100% in error or (ii) less than 100% in error. A mass of 100% errors was randomly seeded into each population reflecting the amount found in the investigative audit. The simulated errors with amounts less than 100% were determined using models derived from the data in the investigative audits.

In order to be able to investigate the sampling methods under a wide range of conditions, audit populations were created with lower and higher error rates and with lower and higher error amounts than those found in the investigative audits. In these cases, the relative error patterns across strata were maintained.

The audit populations created thus, are used in the simulation study described in 2.10 below, to compare the performance of the bounds under the differing sampling methods as stated in objectives 3, 4 and 5.

Details of the error models and the resultant audit populations are given in Chapter 3.

2.6 The Error Assignment Methods

Each of the sampling methods outlined in 2.2 selects monetary units rather than line items for inclusion in the sample. Prior to sampling, it is necessary to assign an error amount to each monetary unit. In the cases where the line items are less than 100% in error, the error amount is distributed among the constituent monetary units in each line item. Two methods of assigning errors in line items to monetary units are used in practice. They are the 'taint' and the 'all-or-nothing' methods of error assignment and they are described below.

2.6.1 The Taint Method of Error Assignment

In the taint method of assigning errors to monetary units, the errors are assumed to be distributed equally among all the monetary units in the line item. For example, if line item i is selected and an error amount $E_i = (B_i - A_i)$ is found, then each monetary unit is considered to have an error amount of E_i/B_i. E_i/B_i is referred to as the taint in item i. The number of errors found in the sample of **n** monetary units is the number of items in which a non-zero taint is found.

The sample elements obtained using the taint method of error assignment are denoted by,

$$t_1, t_2, \ldots t_n$$

where

$t_i = (B_i - A_i)/B_i$ if the selected monetary unit is in error

$= 0$ if the selected monetary unit is not in error

2.6.2 The All-Or-Nothing Method of Error Assignment

The 'all-or-nothing' (AON) method of assigning errors to monetary units considers the line items as a collection of monetary units some of which are 100% correct and some of which are 100% in error. The monetary units in error must be specified for each line item in advance of sampling. Van Heerden (1961), who developed this error assignment method, suggested that the incorrect units should be the last units in each line item. Gill (1983) adopts the convention of assigning the errors to the monetary units at the beginning of the line item. It is not important which convention one adopts. In this study, the errors are assigned to the units at the beginning of the line item in keeping with recent studies (e.g. Wurst, Neter and Godfrey 1989a and 1989b).

The sample elements obtained using the all-or-nothing error assignment method are denoted by,

$$u_1, u_2, \ldots\ldots u_n$$

where $u_i = 1$ if the selected monetary unit is in error

$= 0$ if the selected monetary unit is not in error

2.6.3 Discussion

Both error assignment methods are applicable to all the sampling methods. The taint method has gained wide acceptance by practising auditors and is the method used most often by auditors in the USA. The AON error assignment method has gained some acceptance in Europe (Wurst, Neter and Godfrey 1989a) but is not widely used in general. To illustrate the difference between the two methods of error assignment, consider a line item of stated value Ir£10 which is overstated by Ir£2. The error assigned to the monetary units under each error assignment method is illustrated in Table 2.1

Table 2.1 Error Assignment

Monetary Unit	Taint	AON
1	0.2	1.0
2	0.2	1.0
3	0.2	0.0
4	0.2	0.0
5	0.2	0.0
6	0.2	0.0
7	0.2	0.0
8	0.2	0.0
9	0.2	0.0
10	0.2	0.0

An advantage of the taint approach is that small error amounts on large line items have a high probability of being detected. i.e an increased visibility of small taintings. A disadvantage is that with the taint approach, the commonly used bounds are heuristic and lack mathematical justification.

The AON approach on the other hand is statistically valid. It can be shown that the Stringer and Cell bounds will always exceed the population total error amount with a probability greater than or equal to the nominal confidence level. The AON bound has however serious disadvantages which may be the reason why it is not widely used by practitioners. A bound has a greater variability when calculated with the AON method of error assignment than with the taint method. A second disadvantage is that the AON bound does not use all the information available. For example, a monetary unit from a line item in error may be classified as a zero error and hence information about the error amount in the line item is not used. Small tainted errors are more visible because they are embedded in large line items and with PPS sampling they have a large chance of selection with the taint error assignment. But as Leslie, Teitlebaum and Anderson (1979, p269) point out, the AON approach 'loses the benefit of the increased visibility of small taintings in the many audit applications where small taintings are typical'.

2.7 The Point Estimator of the Total Error Amount

A point estimate is a single value obtained from the data intended to represent the 'best estimate' of the unknown population value (i.e. the population parameter). A point estimator of the total used with probability proportional to size sampling was suggested by Horvitz and Thompson (1952) and is defined by

$$\hat{Y}_{HT} = \sum_{i=1}^{n} \frac{y_i}{\pi_i}$$

where y_i is the measurement for the i^{th} item and π_i is the probability that the i^{th} item is selected into the sample. This is the best known general estimator of the population total for unequal probability sampling (Cochran, 1977, p 259). It shall be shown in Chapter 4 that, when sampling of line items is done with replacement, the probability of selecting a particular item, item i say, in a sample of size **n** is nB_i/B. The form the Horvitz-Thompson estimator takes when estimating the total error amount

$$T = \sum_{i=1}^{N} (B_i - A_i)$$

is therefore

$$\hat{T} = \sum_{i=1}^{n} \frac{(B_i - A_i)}{\frac{nB_i}{B}} = \frac{B}{n}\sum_{i=1}^{n} \frac{(B_i - A_i)}{B_i}) = \frac{B}{n}\sum_{i=1}^{n} t_i$$

In this study, the point estimator of the total error amount is used with both the taint and AON error assignment methods. It is defined as

$$\hat{T} = \frac{B}{n} \sum_{i=1}^{n} t_i \text{ with the taint error asssignment}$$

$$= \frac{B}{n} \sum_{i=1}^{n} u_i \text{ with the AON assignment}$$

This estimator will be used when sampling is done with and without replacement.

In sieve sampling, where the achieved sample size $\mathbf{n_o}$ may not equal the nominal sample size, the summation is over the achieved sample size.
i.e.

$$\hat{T} = \frac{B}{n} \sum_{i=1}^{n_0} t_i \qquad \text{with taint error assignment}$$

$$= \frac{B}{n} \sum_{i=1}^{n_0} u_i \qquad \text{with AON error assignment}$$

In stabilised sieve sampling it is necessary to adjust this estimator to avoid bias. The adjustments are defined in 5.3.

The point estimator has been studied empirically by Kaplan (1973), Neter and Loebbecke (1975) and Duke, Neter and Leitch

(1982) using simple random sampling of monetary units with the taint error assignment method. Wurst, Neter and Godfrey (1989a) derived the statistical properties of the estimator with both assignment methods for simple random, cell and sieve sampling of monetary units. This study investigates the estimator with both error assignments for the six sampling methods outlined in 2.2.

The mean and variance of the estimator for each sampling method are derived in Chapter 5. The precision of the estimator for each sampling method is compared with the precision for simple random sampling of monetary units using the design effect suggested by Kish (1965, p258).

2.7.1 The Design Effect

The design effect is a measure of efficiency for comparing two sample designs. It is defined as follows;

> 'The *design effect* or *deff* is the ratio of the actual variance of a sample design to the variance of a simple random sample of the same number of elements' (Kish, 1965, p258).

The design effect has two primary uses:

(i) In appraising the efficiency of the sample design compared to simple random sampling;

This involves measuring the accuracy of a sampling method compared to simple random sampling. If deff is less than one, the sampling method is more precise than simple random sampling. If deff is greater than one, the accuracy of the sampling method under consideration is less than simple random sampling and if deff is equal to one the accuracy of the two sampling methods are the same. Simple random sampling acts as a useful basis of comparison for other methods of random sampling.

(ii) In sample size planning;

The design effect may be interpreted as the proportion increase or decrease in the sample size of simple random sampling to obtain the same precision as the sampling method under consideration. For example, if it is estimated that the design effect of a particular sampling method compared to simple random sampling is .8 with a sample of size n = 60, then a simple random sample of size n = 75 (i.e. 60/.8) is needed to give the same precision. Similarly, if the design effect is 1.2 based on a sample of size n = 60, simple random sampling will give the same precision with a sample of size n = 50 (i.e. 60/1.2).

Studies, notably those carried out by Kaplan (1973) and Neter and Loebbecke (1975) have shown that the confidence intervals obtained using the point estimator and the central limit theorem are unreliable (see Chapter 1). This study will not use the central limit theorem to obtain upper bounds. However, the theoretical properties of the point estimator are considered worthy of investigation because they may give some indication of how the heuristic upper bounds estimates of the total error amount may be expected to behave for the different sampling methods. Wurst, Neter and Godfrey (1989a, 1989b) investigated the performance of simple random, cell and sieve sampling using the Stringer and Cell bounds for estimating the total error amount and found that the sampling methods have the same effects on the bounds as on the point estimator.

The design effect is derived for systematic, cell, sieve, Lahiri and stabilised sieve sampling. For each sampling method, a comparative analysis is carried out between the statistical properties of this estimator and the properties of the simulated sampling distributions of upper bound estimates of the total error amount.

2.8 Upper Bound Estimates of the Total Error Amount

In an attempt to overcome the reliability problem of classical estimation procedures, confidence bounds have been developed for estimating the total error amount using the error taint and the error rate components in the sample data which do not depend on the assumptions of large sampling theory of classical statistics. Most of these bounds are heuristic. No theoretical proof of their validity exist and their performance can only be investigated empirically. This study looks at three such bounds. The notation assumes that **m** errors are found in a sample of size **n** and that the total book value amount is **B**.

2.8.1 The Stringer Bound

The Stringer bound (1963) is the most widely used non-classical procedure for estimating the total error amount (Felix, Leslie and Neter, 1982). An upper limit for the total error amount is obtained by combining the upper limit for the error rate obtained in the sample with the sample error amounts. It is defined below for the taint and AON methods of error assignment.

2.8.1.1 The Stringer Bound with the Taint Assignment

With the taint assignment method, and a confidence level of $1-\alpha$, the bound is defined as

$$STR_{taint,1-\alpha} = B.P_{1-\alpha}(n,\ 0) + B \sum_{i=1}^{m} [P_{1-\alpha}(n,\ i) - P_{1-\alpha}(n,\ i-1)]\ t_i$$

where $t_1 >= t_2 \ldots >= t_m$ are the **m** non-zero taints found in a sample of size **n**. $P_{(1-\alpha)}(n,i)$ is the upper $100(1-\alpha)$ percent confidence limit for a population proportion of errors when a random sample of **n** is selected and **i** errors are found in the sample $1 <= i <= m$.

The Stringer bound may be looked upon as a refinement of the bound obtained when all the taints are 100%. If $P_{(1-\alpha)}(n,m)$ is the upper limit for the proportion of errors, then $B.P_{(1-\alpha)}(n,m)$ is an upper limit for the error amount when all the taints are 1. When the taints are less than 1 this will overestimate the total error amount. Stringer (1963) suggested that the bound be adjusted downwards for each taint found in the sample. The Stringer bound adjustment for the ith largest taint is;

$$B.P_{1-\alpha}(n, m) - B.[(P_{1-\alpha}(n, i) - P_{1-\alpha}(n, i-1)] (1-t_i)$$

Similar adjustments are made for the other taints in the sample. The Stringer bound then becomes

$$STR_{taint,1-\alpha} = B.P_{1-\alpha}(n, m) - B\sum_{i=1}^{m} [P_{1-\alpha}(n;i) - P_{1-\alpha}(n, i-1)] (1-t_i)$$

which is equivalent to $STR_{taint,1-\alpha}$ above.

2.8.1.2 The Stringer Bound with the AON Error Assignment

With the AON error assignment method, the Stringer bound becomes

$$STR_{aon,1-\alpha} = B.P_{1-\alpha}(n, m^*)$$

where m^* is the number of AON errors found in the sample. This is clearly an upper confidence limit for the total error amount.

2.8.1.3 Discussion

Studies have shown that the Stringer bound always achieves a coverage larger than the nominal and that it is conservative in the sense that the estimate of error amount obtained by using this bound is usually far in excess of the true error amount (see 1.5).

2.8.2 The Cell Bound

The Cell bound was developed by Leslie, Teitlebaum and Anderson (1979, pp135-147) to provide a bound estimate of the total error amount which is not as conservative as the Stringer bound. It is obtained by means of an iterative procedure described below.

2.8.2.1 The Cell Bound with Taint Error Assignment

$$UEL_0 = P_{1-\alpha}(n, 0)$$

$$UEL_j = \max\left(UEL_{j-1} + \frac{t_j}{n},\ P_{1-\alpha}(n, j) * \sum_{i=1}^{j} \frac{t_i}{j} \right) \quad j = 1 \text{ to } m$$

$$Cell_{(taint, 1-\alpha)} = B * UEL_m$$

2.8.2.2 The Cell Bound with AON Error Assignment

With AON error assignment, the Cell bound reduces to

$$Cell_{(aon, 1-\alpha)} = B\, P_{1-\alpha}(n, m^*)$$

similar to the Stringer bound.

2.8.2.3 Discussion

Clearly, the Cell and the Stringer bounds with the 'all-or-nothing' assignment are equivalent and the bound is an upper confidence limit for the total error amount under simple random sampling of monetary units. Gill (1983) proved that it is also a upper confidence bound under cell sampling and sieve sampling. Gill (1983) showed that with this bound under sieve sampling, one may run a lower chance than the chosen risk α of rejecting an acceptable population but never a higher risk. The Cell (and Stringer) bound is included in the simulation study to examine its magnitude and its variance. It is calculated for all the sampling methods.

2.8.3 The Moment Bound

More recently a bound has been developed by Dworkin and Grimlund (1984, 1986) called the Moment bound. This bound assumes that the sampling distribution of the mean error can be represented by a three-parameter Gamma distribution. To estimate the parameters of the Gamma distribution, an analysis of the error rate and error taint components is carried out. The error rate is based on the binomial likelihood function with the sample size and the proportion of errors found in the sample taken to be the binomial parameters. The error taint distribution is based on the first three sample moments as determined from the taints observed in the sample. The error taint distribution also uses a hypothetical taint denoted by t^*. This is used like any other observation when calculating the sample moments of the error taint distribution. It ensures that the bound value is not zero when no errors occur in the sample. The three moments of the error rate distribution and the three moments of the taint distribution are combined to obtain a three moment representation of the sampling distribution of the mean error. The Moment bound is defined below for the taint and AON methods of error assignment.

2.8.3.1 The Moment Bound with the Taint Error Assignment

The mean error is assumed to have a Gamma distribution.

$$f(\bar{t}) = \frac{1}{D.\Gamma A} \cdot \left[\frac{\bar{t} - G)}{D}\right]^{A-1} \exp\left[-\frac{(\bar{t}-G)}{D}\right]$$

The parameters, A, D and G are estimated from the sample data by combining the first three moments of the sampling distribution of the error taint with the first three moments of the sampling distribution of the error rate as follows;

1 Average and Hypothetical Error Taint

$$\bar{t}_m = \sum_{i=1}^{m} \frac{t_i}{m}$$

$$t^* = .81\ [1 - .667\ \tanh(10\bar{t}_m)]\ [1 + .667\ \tanh\ (m/10)]$$

2 Error Taint Moments about Zero

$$TN_j = \frac{(t^*)^j + \sum_{i=1}^{m} t_i^j}{(m+1)} \qquad j = 1,2,3$$

3 Error Rate Moments about Zero

$RN_1 = (m+1)/(n+2)$

$RN_2 = [(m+2)/(n+3)]RN_1$

$RN_3 = [(m+3)/(n+4)]RN_2$

4. Mean Error Moments about Zero

$UN_1 = RN_1 * TN_1$

$UN_2 = [RN_1 * TN_2 + (n-1)RN_2 * TN^2{}_1]/n$

$UN_3 = [RN_1 * TN_3 + 3(n-1)RN_2 * TN_1 * TN_2 + (n-1)(n-2)RN_3 * TN^3{}_1]/n^2$

5 Mean Error Moments about the Mean

$UC_1 = UN_1$

$UC_2 = UN_2 - UN_1{}^2$

$UC_3 = UN_3 - 3UN_1 * UN_2 + 2UN^3{}_1$

6. The parameters of the Gamma distribution may be estimated from these moments as shown in Johnson and Kotz (1970) as follows;

$A = 4UC^3_2/UC^2_3$

$D = 0.5UC_3/UC_2$

$G = UC_1 - 2UC^2_2/UC_3$,

7. The 1-α Upper confidence bound for the Total Error Amount

The 100(1-α)% upper limit for the mean error is obtained by using the Wilson-Hilferty approximation (Wilson and Hilferty, 1931) to the Gamma distribution and can be written as;

$$CB_{1-\alpha} = G + D.A[1 + Z_{1-\alpha}/3\sqrt{A} - 1/9A]^3$$

where $z_{(1-\alpha)}$ is the 1-α percentile of the standardized normal distribution.

The 100(1 - α)% confidence bound for the total error amount is

$$MB_{(taint,1-\alpha)} = B * CB_{1-\alpha}$$

2.8.3.2. The Moment Bound with the AON Error Assignment

The Moment bound with the AON error assignment is similar to the Moment bound with taint assignment except that (1) and (2) are changed to take account of the different error assignment method.

They are:

1. $\bar{u}_{m^*} = \sum_{i=1}^{m^*} \frac{u_i}{m^*}, \quad u^* = [1 - .667 \tanh(100\bar{u}_{m^*}] * [1 + .667 \tanh(\frac{m}{10})]$

2.

$$TN_j = \frac{(u^*)^j + \sum_{i=1}^{m} u_i^j}{m + 1} \qquad j = 1,2,3$$

2.8.3.3. Discussion

The most distinctive feature of the Moment bound is the hypothetical error t^* or u^* which ensures that the bound has a non-zero value when the sample contains no errors. Its value has been selected so that the Moment bound is approximately the same as the Stringer bound when no errors occur in the sample. As the number of errors increases in the sample the relative impact of the hypothetical error taint decreases. Dworkin and Grimlund (1984) give the mathematical development of this bound.

2.9 The Confidence Levels

A reliable confidence interval of size $1 - \alpha$ has the property that it will include the true value $100(1 - \alpha)\%$ of the time, in repeated sampling. An upper confidence bound has the property that the true value is less than the upper bound $100*(1 - \alpha)\%$ of the time, in the long run. The upper confidence bound is of particular interest to the auditor in trying to assess the maximum amount of error present in a set of accounts. An auditor has to decide what value of α to choose before carrying out substantive testing. The chosen confidence level will depend on the circumstances of the audit. For example, the confidence level at the substantive testing stage will depend on the degree of confidence the auditor can derive from the previous stages of the audit process (Grimlund and Felix, 1987). Studies have shown that some bounds perform differently at different confidence levels (see for example Jenne, 1982; Grimlund and Felix, 1987; Chan and Smieliauskas, 1990). Wurst, Neter and Godfrey (1989b) stated that confidence levels of .85 and .95 are frequently used in audit practice and Grimlund and Felix (1987) suggested that a confidence level as low as .70 may be used in practice, when auditors develop their overall confidence using other sources of information. McRae(1982, p99) deduced that while the choice of a confidence interval is almost invariably a subjective estimate based on the auditor's prior knowledge about the quality of the accounting procedures in the particular audit, a degree of risk somewhere between 30% and 5% seems a reasonable range from which to choose. In this study, the bounds are calculated at three nominal confidence levels of .70, .85, .95.

2.10 The Simulation Study

In a situation where there is no clear-cut theoretical solution, one way of attempting to assess the properties of alternative procedures is by simulation. This is particularly true of auditing where the populations are finite and real (Smith, 1979). In this study, a large scale simulation study is carried out to investigate the performance of the sampling methods for each bound on the actual accounting populations. Samples are drawn from each audit population using the six sampling methods outlined in 2.2. For each sample design, one thousand replications are performed using sample sizes typically used in substantive testing.

In practice, the extent of substantive testing, depends upon the decisions taken about the effectiveness of the systems of internal control (Shaw, 1980, p61). The auditor may 'obtain assurance from the presence of reliable internal controls and thereby reduce the extent of substantive testing' (APC, 1980d). If the auditor is satisfied that the internal controls have operated satisfactorily throughout the period under review, then the level of substantive testing may be reduced (Coopers and Lybrand, 1985). If the overall level of assurance derived from compliance and analytical procedures is high, a smaller sample size will suffice for substantive testing. In a recent study of medium-sized accounting firms it was found that samples sizes of 25-100 were used in auditing in the UK (Abdul-Hamid, 1993). In this study, samples of sizes 30, 60 and 100 are drawn from each audit population using the six sampling methods outlined in 2.2.

Upper bound estimates of the total error amount are calculated using the Stringer, Cell, and Moment bounds with both error assignment methods at each confidence level. The computations are carried out on a VAX 6230 computer. The programmes are written in Fortran and the NAG (1988) and the IMSL (1987) libraries of subroutines are used for the statistical analysis. The programmes are listed in Appendix K.

2.11 Criteria for Assessing the Performance of the Sampling Methods

Royall (1970) stated that the most important criterion for evaluating theoretical calculations is their ability to predict the actual performance of sampling and estimation procedures in practical problems. In this study, the criteria used for assessing the performance of the sampling methods are (i) the reliability, (ii) the tightness and (iii) the precision of the bound estimates of the total error amount. For any real accounting population, the performance measures may differ for different bounds, for different sample sizes, and for different confidence levels. When reference is made to the performance of a sampling method, it is assumed that the bound, the confidence level and the sample size are known. When it is not clear what they are, they are specified.

2.11.1 Reliability

Reliability measures the coverage of the bounds. It refers to the proportion of the 1000 replications in which the value of bound is greater than or equal to the true error amount in the population. The observed coverage is an estimate of the true probability that the bound will be correct in repeated sampling of the same size from the same population, i.e., the confidence level. Since the bounds are heuristic, the actual confidence level may not be equal to the nominal confidence level and has to be ascertained by experiment. A sampling method is said to be reliable for a particular bound if the actual coverage obtained in the 1000 replications reaches the nominal confidence level. The coverage statistics are calculated for each sampling method, using each of the bounds. Numerous studies have used the coverage as a measure of the reliability of a sampling method (see, for example, Neter and Loebbecke, 1975; Reneau, 1978; Dworkin and Grimlund, 1984; Wurst, Neter and Godfrey, 1989b; Atkinson, 1990; Chan and Smieliauskas, 1990).

2.11.2 Tightness

Tightness refers to the difference between the bound value and the true error amount T. If the average size of the bound in repeated sampling is near the true error amount for a particular sampling method, the sampling method is said to be tight for that bound. If the average size of the bound is substantially greater than the true error amount for a sampling method, the sampling method is said to be conservative for that particular bound. A conservative sampling method will lead to too many rejections of populations on the basis of a material error when a material error does not exist.

Various measures of tightness have been used in previous research. Garstka (1977) presented both the number of upper bounds which exceeded the true error amount and the average amount by which it was exceeded. Reneau (1978) used the upper bound estimate expressed as a proportion of the total error amount to estimate the tightness. Wurst, Neter and Godfrey (1989b) used the mean bound as the measure of tightness. The measure of tightness used in this study is the difference of the mean bound and the total error amount in the population, expressed in units of the total error amount. It is defined as

$$\frac{\bar{T} - T}{T}$$

where $\bar{T}$ is the mean of the upper bound averaged over the 1000 replications. This measure indicates how close the mean bound is to the total error amount expressed in units of the total error amount.

2.11.3 Precision

Precision refers to the variability of the bound in repeated sampling. For any given tightness and reliability, a less variable bound is preferred to a more variable bound for helping to distinguish between acceptable and unacceptable total error amounts (Wurst, Neter and Godfrey, 1989b). In keeping with previous research (see, for example, Neter and Loebbecke, 1975; Reneau, 1978; Plante, Neter and Leitch,1985; Wurst, Neter and Godfrey,1989b), this study uses the standard deviation of the 1000 replications as a measure of the precision for each bound using each sampling method. In addition, the comparative variability of two sampling methods is measured in terms of the relative efficiency. It is the ratio of the variance of a particular bound using the two sampling methods. This concept is equivalent to the design effect, defined for the point estimator in 2.7.1.

2.12 The Data Analysis

Analysis of variance (ANOVA) models are constructed to test the performance of the sampling methods using different sample sizes and upper bound estimates of the total error amount for populations with differing error rates and taint sizes. Five factor analysis of variance models are constructed using the performance measures as the dependent variables and the sample size, error rate, taint size, bound and sampling method as the independent variables. Separate ANOVA models are derived for each performance measure, for each nominal confidence level and for each error assignment. The purpose of the models is to determine to what extent the performance measures are influenced by the different sampling methods and to what extent the sampling methods interact with other factors to affect the performance measures. T-, Tukey and Dunnett tests for comparisons of means are used to test for significant differences between the sampling methods with respect to coverage, tightness and precision. The comparative precision of each bound is also compared with the comparative precision of the point estimator for the different sampling methods. The SPSSx software package (SPSS, 1987) is used to estimate the ANOVA models. The details of the models and the results are given in Chapters 6, 7 and 8.

2.13 Chapter Summary

The research methodology employed to achieve the objectives has been outlined in this chapter. The study investigates the performance of six monetary-unit sampling methods in substantive auditing under various population conditions. Four of these are currently used in practice, (i.e., simple random, systematic, cell and sieve sampling), one 'Lahiri sampling' has not been previously applied in auditing and a new sampling method 'Stabilised Sieve sampling' will be developed in this study for use in substantive testing. The investigation involves both a theoretical point estimator analysis and an empirical study of upper bound estimates of the total error amount.

Two accounting populations of debtors from the Public Sector are examined, and their error characteristics are determined by means of large scale investigative audits. Thirty audit populations are created with different error rates and error amounts reflecting the error patterns found in the investigative audits and these are used to test the comparative performance of the sampling methods.

Samples of size 30, 60, 100 are drawn from each audit population using each sample selection method. One thousand replications are performed for each sample size and for each sampling plan. Upper bound estimates of the total error amount are calculated using the Stringer, Cell and Moment bounds with the taint and AON error assignment at the 95%, 85% and 70% confidence levels.

The criteria for assessing the performance of the sampling methods include reliability, tightness and precision of the upper bound estimates. Five factor analysis of variance models are constructed using the performance measures as the dependent variables and the sample size, error rate, taint size, bound and sampling method as the independent variables. The purpose of the models is to determine to what extent the performance measures are influenced by the different sampling methods and to what extent the sampling methods interact with other factors to affect the performance measures. The comparative precision of each bound is also compared with the comparative precision of the point estimator for the different sampling methods.

The following chapters will:

(i) detail the actual accounting populations, the investigative audits, the error models and the study populations (Chapter 3);

(ii) investigate the properties of the sampling methods (Chapter 4);

(iii) derive the analytical properties of the point estimator under each sampling method (Chapter 5);

(iv) analyse the results of the simulation study (Chapters 6, 7 and 8);

(v) provide a summary of the research, the conclusions drawn from the research and recommendations for future research (Chapter 9).

Chapter 3

The Data

3.1 Introduction

To assess the performance of the different sampling methods, it is necessary to investigate the results obtained when the methods are applied to a spectrum of accounting populations. These populations can be created in two main ways;

(i) Hypothetically, by postulating various population characteristics and building a model to capture the essential features of the populations. This approach can fail to represent the relationships that exist between different population parameters in real accounting populations. It may also fail to capture all the essential features of real populations.

(ii) Experimentally, by carrying out audits on a wide range of accounting populations. This method is not feasible due to cost.

The approach taken in this study is essentially a hybrid of the two methods above. An experimental study involving large audits of two accounting populations with different characteristics, is carried out and the relationships between different population parameters are identified for each population. Hypothetical populations are then created by choosing different values for population parameters while maintaining consistency with the relationships between parameters observed in the real populations. Neter and Loebbecke (1975) used a similar approach to generate populations for testing the performance of sampling methods in substantive auditing. They studied four populations and adjusted the error distributions in various ways to generate audit populations whose results were expected to be typical of real world audits. This study examines two populations of debtors in commercial entities in the Irish Public Sector and modifies the error distributions to generate audit populations in various plausible ways. These audit populations are then used to test the performance of the sampling methods.

This chapter describes the populations used, giving in particular:

(i) A summary of previous research on error characteristics in debtors accounts.

(ii) The detailed characteristics of the book values of two populations of debtors obtained from commercial entities in the Public Sector.

(iii) The investigative audits which were carried out on each population to obtain information on the error patterns.

(iv) The models used to create the audit populations for use in the empirical investigation of the sampling methods.

(v) The characteristics of the created audit populations.

3.2 Prior Research on Accounting Populations and Error Characteristics

The need for empirical study of error characteristics in accounting populations is well established (see, for example Garstka, 1977; McRae 1982; Jenne, 1986; Menzefricke and Smieliauskas, 1987a). Knight (1979) noted that many auditors choose the appropriate statistical sampling technique depending on the nature of the population being sampled and the prior knowledge of likely error rates. With increased knowledge of error distributions, auditors can more effectively and efficiently design audit sampling plans to detect errors. A limited number of databases of error characteristics in accounting populations are available and their characteristics are summarised in Table 3.1

Table 3.1 Available Databases of Accounting Errors

Author	Source of Data	Number of Populations.	Line Item Error Rate (Debtors)	Taint Size (exc. 100% taints)	Occurrence of 100% taints	Direction of Debtors Errors
Neter and Loebbecke (1975)	Freight and Manufacturing Companies (US)	4 (3 debtors)	5.7, 7.3, 28.6	0.5% - 50%	Mostly in low-valued items	All Overstatements
Ramage,Kriegar and Spero(1979)/ Johnson, Leitch and Neter (1981)	Peat, Marwick, Mitchell (US)	97 populations (55 debtors)	< 5% in 72% of cases. > 25% in 4% of cases.	>= 20%	Mostly in low-valued items	Mostly Overstatements.
McRae (1982)	Two Accounting Firms (UK)	76 populations (58 debtors)	< 5% in most cases. > 20% in 5% of cases.	< 10%	Large-taints in low valued items	59% Overstatement
Ham, Losell and Smieliauskas (1985)	Price Waterhouse, (US)	20 medium sized firms over 5 years (58 debtors)	< 5% in 50% of cases.			72% Overstatements
Johnson (1986)	Six UK National Accounting firms	55 manufacturing firms				A slight bias towards overstatements
Kreutzfelt and Wallace (1986)	Arthur Anderson (US)	260 engagements 258 (debtors)		2.36% on average		66% Overstatements

These data sets have been used by various researchers to generate populations which are considered to be typical of actual accounting populations. The generated populations are used to examine the performance of various sampling plans. The problem with this approach is that, when considering a sampling plan for populations unrelated to the generated ones, one can never be sure that the models used will represent the actual situation being studied. Certain patterns do emerge from these data sets which appear to be common to all the populations. For example,

(i) The line item error rate in debtors accounts is usually small.

(ii) The average taint size is small in most of the populations

(iii) Line items which are 100% in error occur predominantly in the low valued items

But differences also exist. One major difference which could have important implications for the choice of sampling plan concerns the incidence of understatement and overstatements. Errors in debtors from the US are mostly overstatements while data collected in the UK show only a slight bias towards overstatements. The direction of errors is an important factor when deciding on a sampling plan. For example, since monetary-unit sampling selects line items with probability proportional to their recorded value and not their actual value, this selection method may not be the most effective type of sampling when understatement errors play a dominant role.

As far as the author is aware, there is no data on book value and error characteristics available from Irish firms. Also there appears to be no data available from the Public Sector. In 1989, a Panel on Nonstandard Mixtures of Distributions set up by the Committee on Applied and Theoretical Sciences of the National Research Council Board of Mathematical Sciences, concluded that 'concerted efforts must be made to improve the statistical methodologies used in the Public Sector.' They also noted that there is no information available on Public Sector accounting populations. (Committee on Applied and Theoretical Research, 1989).

This study obtains data from two commercial entities of debtors in the Irish Public Sector and uses these to generate audit populations whose results are expected to be typical of such populations. The performance of the monetary-unit sampling methods are tested on the audit populations.

3.3 Characteristics of the Book Values of Two Irish Public Sector Debtors

Two populations are used as a basis for this study. They were obtained from commercial entities from the Public Sector in Ireland audited by the office of the Comptroller and Auditor General. Both populations consist of all positive balances of debtors. Line items having zero or negative book values were eliminated on the assumption that if these were not of negligible importance, the auditor would wish to audit them separately (Neter and Loebbecke, 1975). Population 1 consists of 3725 debtors from a state scientific consultancy firm and Population 2 consists of 662 debtors from a government industrial support body which provides grant aid to industry setting up in Ireland. Tables 3.2 to 3.5 contain frequency tables of the book values and the main descriptive parameters of the book value distribution for each population.

Table 3.2 Frequency Table of Book Values of Population 1

Amount (Irfs)	No. of Line Items	% Line Items	Monetary Amount (Irfs)	% Monetary Amount
0-500	2601	69.8	452,988	12.9
500-1000	523	14.0	375,908	10.7
1000-1500	168	4.5	210,834	6.0
1500-2000	95	2.6	167,998	4.8
2000-2500	67	1.8	151,015	4.3
2500-3000	56	1.5	153,488	4.4
3000-3500	34	0.9	110,320	3.1
3500-4000	25	0.7	94,899	2.7
4000-4500	19	0.5	81,041	2.3
4500-5000	23	0.6	110,327	3.1
5000-10000	74	2.0	512,392	14.5
10000-20000	21	0.6	289,187	8.2
20000-30000	6	0.2	152,001	4.3
>30000	13	0.3	660,212	18.7
Total	3725	100	3,522,610	100

Table 3.3 Book Value Parameters of Population 1

Total Book Value	Ir£3,522, 610.0
Mean	Ir£945.7
Standard Deviation	3,661.5
Skewness	14.00
Kurtosis	259.1
Minimum	Ir£2.0
First Quartile	Ir£88.0
Median	Ir£240.0
Second Quartile	Ir£646.0
Maximum	Ir£96,962.0
Number of Line Items	3725

Table 3.4 Frequency Table of Book Values of Population 2

Amount (Irfs)	No.of Line Items	% of Line Items	Monetary Amount (Irfs)	% Monetary Amount
0-500	130	19.6	19,884	0.2
500-1000	49	7.4	33,597	0.3
1000-2000	69	10.4	101,427	0.9
2000-3000	51	7.7	123,770	1.1
3000-4000	45	6.8	156,487	1.4
4000-5000	34	5.1	155,484	1.3
5000-6000	28	4.2	153,094	1.3
6000-7000	21	3.2	136,978	1.2
7000-8000	20	3.0	151,667	1.3
8000-9000	10	1.5	85,138	0.7
9000-10000	9	1.4	83,976	0.7
10000-15000	40	6.0	491,178	4.2
15000-30000	61	9.2	1,299,087	11.2
30000-60000	48	7.3	1,988,827	17.1
60000-100000	30	4.5	2,239,324	19.2
>100000	17	2.6	4,410,913	37.9
Total	662	100	11,630,831	100

Table 3.5 Book Value Parameters of Population 2

Total Book Value	Ir£11,630,831.0
Mean	Ir£17,569.2
Standard Deviation	54,687.8
Skewness	Ir£10.0
Kurtosis	130.1
Minimum	Ir£1.0
First Quartile	Ir£756.6
Median	Ir£3,740.7
Third Quartile	Ir£13,328.6
Maximum	Ir£880,918.2
Number of Line Items	662

3.3.1 Comparison of the Book Value Characteristics of the Two Populations.

The characteristics of the two populations of book values are similar to that found in other studies, in that they are both highly skewed and in each case a small proportion of the line items account for a large proportion of the book value total (e.g. Neter and Loebbecke, 1975). The total book value of the 13 largest accounts in Population 1 represents 18.7% of the total population amount and in Population 2, the largest 17 accounts represents 37.9% of the total book value amount.

The two populations also display contrasting features. Population 1 consists of a relatively large number of small debtors while Population 2 consists of a relatively small number of large debtors. The number of line items in Population 2 is less than one fifth that of Population 1. Also, the average book value amount in Population 2 is nearly twice that of Population 1 and the median book value amount in Population 2 is nearly sixteen times that of Population 1. Plante, Neter and Leitch (1985) noted that the variation in sampling method performance could be substantial when the line item sizes are large.

3.4 The Investigative Audits and the Error Patterns

Audits were carried out on each population to investigate the error distributions. Stratified random samples of line items with disproportionate allocation were selected for auditing from each population. Stratification by line item size was used to investigate error characteristics in line items of differing sizes. The purpose was to obtain sufficient information on error patterns to be able to model the errors in each population.

3.4.1 The Investigative Audit of Population 1

A stratified random sample of 213 line items was selected with disproportionate allocation for auditing from Population 1. The population was divided into four strata including a stratum containing the largest items which was subjected to 100% sampling. Table 3.6 shows the sample allocation among the strata in Population 1.

Table 3.6 Sample Selection from Population 1

Stratum	Book value	No. of Line Items	Sample Size
1	0-500	2601	80
2	500-5000	1010	100
3	5000-30000	101	20
4	>30000	13	13
Total		3725	213

3.4.1.1 Characteristics of Errors in Population 1

A total of sixteen errors was found in the audit of 213 line items from Population 1. Since the sample allocation among the strata was not proportionate to the strata size, it was necessary to use weighting to obtain the sample estimates of the population error rate and the standard deviation of the error rate. Formulae for the weighted estimates are given in Appendix A. The estimated population error rate is 0.055 with a standard error of .016. Table 3.7 gives the distribution of the error incidence across strata and the characteristics of the error taints.

Table 3.7 Sample Results from Population 1

Stratum	Sample Size	No. of Errors (100% taints)	Line Item Error Rate	Average Error Amount (Ir£s)	Ave. Taint (inc.100%)	Ave. Taint (exc.100%)
1	80	3 (3)	.0375	143.9	1	-
2	100	9 (4)	.09	1173	0.502	.10415
3	20	3 (0)	0.15	179	0.12	0.012
4	13	1	0.077	120	0.003	

With such a limited amount of information on the actual errors, it is not possible to make definitive extrapolations to the population. However, the following observations are made:

1. All errors are overstatements;

2. 7 of the 16 errors are 100% overstatements and all of these occur in the first two strata. All the errors in the first strata are 100% overstatement errors;

3. The taints tend to decrease with book value size. Taints, other than the 100% overstatements range from .09% to 28%.
4. If the top stratum is excluded, the error rate tends to increase with book value size;
5. The amount of error does not tend to increase with book value size.

3.4.1.2 Comparisons with Other Studies

The findings coincide with those of other studies in many respects. Johnson, Leitch and Neter (1981) found that 100% overstatement errors are frequently present to a large extent in debtors and that there is strong evidence to suggest that error rate increases as the size of the line item increases. Neter and Loebbecke (1975) found that small line items have larger taints than larger line items and Neter, Johnson and Leitch (1985) noted that in the majority of audits, the size of positive taints tends to vary inversely with book amounts.

3.4.2 The Investigative Audit of Population 2

A stratified random sample of 217 line items was selected with disproportionate allocation for auditing from Population 2. The population was divided into six strata including a stratum containing the largest items which was subjected to 100% sampling. Table 3.8 shows the sample allocation amongst strata for Population 2.

Table 3.8 Sample Selection from Population 2

Stratum	Book Value	Population Size	Sample Size
1	0-500	130	35
2	500-5000	248	60
3	5000-30000	189	60
4	30000-60000	48	25
5	60000-100000	30	20
6	>100000	17	17
Total		662	217

3.4.2.1 Characteristics of Errors of Population 2

A total of sixteen errors was found in the audit of 217 line items from Population 2. Again, since the sample allocation among the strata was not proportionate to the strata size it was necessary to use weighting to obtain the sample estimates of the population error rate and the standard deviation of the error rate. Formulae for the weighted estimates are given in Appendix A. The estimated population error rate is 0.081 with a standard error of .017. Table 3.9 gives the distribution of the error incidence across strata and the characteristics of the error taints.

Table 3.9 Sample Results from Population 2

Stratum	Sample Size	No. of Errors (100% taints)	Line Item Error Rate	Average Error Amount (Ir£s)	Ave. Taint (inc.100%)	Ave. Taint (exc.100%
1	35	2 (2)	.06	206	1	
2	60	8 (7)	.13	2917	0.99	0.93396
3	60	2 (1)	.03	5432	0.67	0.34671
4	25	1 (1)	.04	38,822	.988	
5	20	3 (1)	.15	58,889	0.89	0.8423
6	17	0	0	0	0	

Again, with such a limited amount of information on the actual errors, it is not possible to make definitive extrapolations to the population. However, the following observations are made:

1. All error are overstatements;
2. 75% of the errors are 100% overstatements. All the errors in the first strata are 100% overstatement errors;
3. The taints do not tend to decrease with book value size. Taints, other than the 100% overstatements, range from 35% to 98%;
4. The line item error rate is highest in the stratum with the largest book value items;
5. The amount of error increases with book value size.

3.4.2.2 Comparisons with Population 1 and with Previous Research

The error patterns in this population differ in many respects from the error patterns found in Population 1. The taints are larger and the proportion of 100% overstatement errors is greater in Population 2 than in Population 1. Also, in Population 1, the taints size tends to decrease with book value size and line item error rate increases with book value size (excluding the top stratum) but this is not the case in Population 2. However, these findings coincide with some aspects of the research carried out by Neter and Loebbecke (1975). They found that there is no difference in the error rate for large and small line items. Neter and Loebbecke (1975) also found that error amount tends to increase with the book value size.

The main difference between the error patterns found in Population 2 and the findings of other researchers is that the average taint size is substantially larger and the incidence of 100% errors is greater in Population 2 than in any of the previous populations studied .

3.5 The Audit Populations

In order to be able to investigate the sampling methods under a wide range of conditions, audit populations with different error rates and error amounts were created on the basis of the observed error patterns found in the two accounting populations described above. The error rates and error taint levels observed in the investigative audits were varied consistently to provide a range of audit populations with error patterns that might occur in real audit situations. Five different error rates and three different taint levels were used to create the audit populations giving fifteen audit populations for each of the two accounting populations. This approach is similar to that used by Neter and Loebbecke (1975) where study populations were constructed to exhibit the same error patterns but different error occurrence rates. In addition, a proportion of 100% taints was generated into each audit population in keeping with previous research which has shown that 100% taints are present in many real accounting populations (see, for example, Johnson, Leitch and Neter,1981).

3.5.1.1 The Error Rates and Error Taints of Audit Populations from Population 1

Audit populations were created for Population 1 with error patterns reflecting the patterns found in the sample data.

3.5.1 The Error Rates for Population 1

Errors were generated into Population 1 with five different error rates. Table 3.10 shows the error rate in each stratum for each overall error rate. In addition, the top stratum with an error rate of 0.08 was included.

Table 3.10 Line Item Error Rates in Population 1

Error Rate	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5
Stratum 1	.0125	.025	.0375	.075	.1125
Stratum 2	.03	.06	.09	.18	.27
Stratum 3	.05	.10	.15	.30	.45
Overall Error *	.019	.037	.055	.110	.165

* Weighted to take account of disproportionate allocation and corrected to include the top stratum

In this choice of error rate distribution, it is assumed that the relative error rates between strata remain the same as the overall error rate increases. The error rate 3 in Table 3.8 is the exact error rate found in the sample. The error rates 1 and 2 are one third and two thirds times the set of sample error rates respectively. The error rates 4 and 5 are twice and three times the set of sample error rates respectively. This set of error rates provides a range of audit populations with error patterns that might occur in real audit situations and enables the sampling methods to be tested under different conditions.

3.5.1.2 The Taint Distribution for Population 1

To obtain the taint distribution for Population 1, the error taint was modeled on the corresponding book value and the model was used to predict the error taint in the items containing errors. Some previous studies use theoretical probability distribution functions (e.g. Plante, Neter and Leitch, (1985) and Dworkin and Grimlund (1984) use the chi-squared distribution, and Wurst, Neter and Godfrey (1989b) use the exponential distribution) to seed errors into the book values. However, other studies (e.g. Neter and Loebbecke, 1975, and Duke (1980) and Duke, Neter and Leitch, 1982) fit a probability function to the sample data and use this to seed errors into the population and this research takes a similar approach. A simple linear regression model is fitted to the data obtained from the investigative audit. Estimates of the regression coefficients are based on the actual data obtained in the investigative audits. Line items which are not in error and items with 100% taints are not included in the taint modelling. The purpose of the regression analysis is to predict the taint value other than 1 or 0 from a particular book value in error. To improve the fit, it was necessary to transform both the independent and dependent variables and the following model was obtained;

$$\sqrt{taint} = \alpha + \beta\left(\frac{1}{bookvalue}\right) + \epsilon$$

The sample data were used to obtain estimates of the coefficients using least squares regression and the regression equation was found to be;

$$\sqrt{taint} = .09 + 220\frac{1}{bookvalue}$$

Table 3.11 gives the output from the regression analysis.

Table 3.11 Regression of Taint on Book Value in Population 1

Predictor	Coef	Stdev	t-ratio	p
Constant	0.08356	0.03395	2.49	0.04
$\frac{1}{bookvalue}$	219.69	42.38	5.18	0.00
	s = 0.0767		r^2(adj)=.764	

This model has an r-squared of .764, indicating that the model is able to take account of 76.4% of the total variation in the dependent variable. Examination of the residuals indicated that heteroscedasticity was not present. The model assumes simple random sampling of line items, rather than the disproportionate stratified sample design used here. However, Kish and Frankel (1974) demonstrated that the relative biases (ratio of the biases to parameter) in the estimates of the regression coefficients are small even for small stratified samples and Warren (1971) showed that selective sampling may lead to a precision greater than that of simple random sampling in the estimated regression coefficients.

Three taint sizes were used when generating the study populations, one exactly as found in the regression equation, a lower and a higher taint with the same pattern as found in the regression relationship but with the intercept decreased by one third and increased by three respectively. The lower and higher taints maintain the relationship of the regression in the sense that the correlation between the $\sqrt{t}$ and 1/bv remain the same but the intercept is increased and decreased to obtain lower and higher taints.

In addition, a proportion of 100% taints was generated into each study population, reflecting the proportions obtained in the investigative audit. All the errors in the first stratum were 100% overstatement errors, 50% in the second stratum were 100% overstatement errors and there were no 100% taints in the third stratum.

3.5.2 The Error Rates and Error Taints of the Audit Populations for Population 2

Audit populations were also created for Population 2 with error patterns reflecting the patterns found in the sample data.

3.5.2.1 The Error Rates for Population 2

Errors were generated into Population 2 with five different error rates. The number of strata used in the sample investigation was reduced to three in order to be able to establish the error pattern more clearly as suggested by Johnson, Leitch and Neter (1981) and Table 3.12 shows the five error rates. The error rate 3 in Table 3.12 is the exact error rate found in the sample. The error rates 1 and 2 are one third and two thirds times the set of sample error rates respectively. The error rates 4 and 5 are twice and three times the set of sample error rates respectively.

Table 3.12 Line Item Error Rates in Population 2

Error Rate Stratum	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5
1-5000	0.035	0.07	0.105	0.210	0.315
5000-60000	0.017	0.023	0.035	0.07	0.107
60000-10000	0.05	0.10	0.15	0.30	0.45
Overall Error Rate*	0.029	0.054	0.080	0.161	0.245

* weighted for disproportionate allocation and including the top stratum.

3.5.2.2 The Taint Distribution for Population 2

As there were only four taints found in the sample audit of Population 2 which were less than 1, it was not possible to model the taints from the sample data and it was therefore decided to use a theoretical distribution to generate the taint values. While empirical evidence indicates that no single assumption about the shape of the taint distribution will be appropriate in all audit situations, many studies have shown (for example, Johnson, Leitch and Neter, 1981) that most error taints follow a reversed J-shaped distribution with a mass at 1. Exponential distributions truncated at 1 have been used in many empirical studies to model the taints which are less than 1 (e.g. Wurst, Neter and Godfrey, 1989a; Leitch, Neter, Plante and Sinha, 1982; Peek, Neter and Warren, 1991). The exponential distributions truncated at 1 with mean taints of .94 for the first stratum, 0.35 for the second stratum and 0.84 for the third stratum were used to generate the taints which are less than 1. These taints represent the mean taints found in the sample data. Audit populations with lower and higher mean taint values were also created. The lower mean taints were one third and the higher mean taints were three times that of the mean taints found in the sample data. The formula for the general truncated exponential distribution is given in Appendix B. In addition, 100% taints were also generated into the audit populations reflecting the proportions found in the investigative audit. 90% of errors in the first stratum, 30% of the errors in the second stratum and 33.3% of the errors in the third stratum were 100% in error.

3.5.3 Creation of the Audit Populations.

In order to be able to test the sampling methods under a wide range of conditions, audit populations were created with a variety of error rates and error taints. This was done by utilizing the error pattern found in the investigative audits to assign errors at random to the line items in the population in order to achieve the specified error rate. For each population, the set of fifteen audit populations was created by generating errors into the population with the line item error rates and error taint distributions outlined above. The procedure used to generate errors into each population is detailed below.

(i) A proportion of line items corresponding to the highest error rate was selected at random from each stratum.

(ii) From each stratum, a proportion of line items corresponding to the first (lowest) error rate was selected from the line items selected in step 1.

(iii) From the selected line items, the number of line items which were 100% in error were allocated randomly.

(iv) The remainder of the line items were allocated a taint value according to the taint distributions described above. Three taint values were assigned to each line item in error. This gave three audit populations with the same line item error rate and with three different taint levels.

(v) For the each of the other error rates, steps (iii) to (iv) were repeated using the appropriate number of extra items selected in step 1 to obtain each of the remaining four line item error rates.

Thus, a set of 15 study populations with five different error rates and three different taint levels was created using this procedure for each of the two populations.

3.5.4 The Audit Populations and the Simulation Study.

Before implementing the simulation study, the high value items were eliminated from each population. Leslie, Teitlebaum and Anderson (1979) recommended that 'the top stratum cut off should generally be set equal to the sampling cell-width'. Also, in sieve sampling, line items which have book values greater than B/n, where n is the sample size are usually audited on a 100% basis (Wurst, Neter and Godfrey, 1989b). Therefore, line items which were greater than B/100 were not included in the investigation. (In the simulation study, 100 is the largest sample taken). Wurst, Neter and Godfrey (1989a) point out that auditors usually examine large items on a 100% basis regardless of the method of the sample selection because of the exposure to risk with these large line items. High value items are not included in the bound evaluation. This is consistent with current audit practice where

> 'high value items are aggregated separately... and reported separately. They are not projected by the sample onto the population but are added to the statistical projection after the errors in the lower stratum are assessed.' (CICA, 1990)

The population characteristics of the two accounting populations with the high value items excluded together with the set of fifteen audit populations associated with each accounting population are given in Appendix C. The set of audit populations generated for each population are similar in many respects (e.g. standard deviation, skewness and kurtosis) and differ mainly in their error amounts caused by the different error rates and taint sizes. When the high value items are eliminated, the total book value amounts for the two populations are similar but the line item sizes for Population 1 are smaller than that of Population 2. Population 1 consists of a relatively large number of small items and Population 2 consists of a relatively small number of large line items. This provides an excellent opportunity of comparing the sampling methods under different conditions.

3.6 Summary

This chapter describes the data collected and generated for use in this study.

Two populations of Public Sector debtors were investigated. Population 1 contains a relatively large number of small accounts while Population 2 contains a relatively small number of large accounts.

The error patterns in the two populations differ in many respects. The taints are larger and the proportion of 100% errors is greater in Population 2 than in Population 1. The taint size tends to decrease with book value size in Population 1 but this is not the case in Population 2.

Audit populations were created based on the two populations, with error patterns reflecting the patterns found in the investigative audits.

The difference in book value characteristics and error patterns found in the two populations provide an excellent opportunity of investigating the sampling methods under different conditions.

The main difference between the data obtained in this study and the McRae (1982, p178) and Johnson (1987) studies pertains to the incidence of overstatements errors. All the errors found in this study were overstatements while the McRae and Johnson studies found only a slight bias towards overstatements.

Chapter 4

The Monetary-Unit Sampling Methods

4.1 Introduction

This chapter provides a basis for the subsequent analysis of estimation procedures by deriving the properties of six monetary-unit sampling methods. It defines a new monetary-unit sampling method, 'Stabilised Sieve Sampling', and introduces one which has not been applied previously in auditing 'Lahiri Sampling'.

The methods are;

(i) simple random sampling

(ii) systematic sampling

(iii) cell sampling

(iv) Lahiri sampling

(v) sieve sampling

(vi) stabilised sieve sampling

Simple random sampling of monetary units is widely used in the States (Wurst, 1990) and systematic and cell sampling of monetary units are used in audit practice in Canada (CICA, 1990). These selection methods ignore the line item structure of the population when selecting the sample of monetary units and consider the population as a collection of monetary units from which a random sample is chosen. Since only the line items containing the selected monetary units can be tested by the auditor, the selected monetary units are traced back to their associated line items. The need to identify the line items selected for auditing may at times create some practical implementation problems (Wurst, Neter and Godfrey, 1989a).

Lahiri, sieve and stabilized sieve sampling, on the other hand,use the line item structure of the population when selecting the monetary units. In each of these sampling methods, the line items are selected randomly and a monetary unit is chosen from each selected item.

Lahiri sampling was proposed by Lahiri (1951) as a convenient method of selection for unequal size clusters in survey sampling. This sample selection method has never been used in auditing and is adapted to the auditing situation in this study.

Sieve sampling was proposed by Gill (1983) as an alternative to the sampling methods currently used in auditing. It has gained some acceptance in Europe (Wurst, Neter and Godfrey, 1989a). A disadvantage of sieve sampling, which may be an important consideration of the auditor, is that the sample size is not constant. It varies depending on the random numbers chosen when selecting the sample.

Stabilized sieve sampling is a new monetary-unit sampling method developed in this study which attempts to preserve the advantages of sieve sampling while returning a constant sample size. It is defined in detail in 4.3.

In the remainder of the chapter, the selection procedures are defined and the characteristics of each sampling method are derived. In subsequent chapters the effects of these sample selection methods on the behaviour of estimates of the total error amount are examined.

4.2 Notation

The notation used in this chapter is given in table 4.1.

Table 4.1 Notation

N	The number of line items in the population
B_i	The book value of the i^{th} line item
B_{max}	The book value of the largest line item
B	$\Sigma^N_{i=1} B_i$ is the total book value amount of the population.
n	The sample size
n_0	The sample size achieved using sieve sampling
π_i	The probability that the i^{th} line item is included in the sample.
(a,b)	Integers between a and b excluding a and b. i.e. the open interval.
(a,b]	Integers between a and b, including b but not including a. i.e. the half open interval.
[a,b]	Integers between a and b including both a and b. i.e. the closed interval
k	The sampling interval. $k = B/n$

4.3 Simple Random Sampling

Simple random sampling of monetary units (SRS) selects line items and monetary units as follows;

(i) The cumulative sum of the line items is formed and a range of numbers is assigned to each item.

(ii) A simple random sample of **n** numbers between 1 and B is chosen with replacement and the selected random numbers are sorted in order of magnitude.

(iii) The line item whose assigned range first exceeds the first random number is selected into the sample. Similarly for the second random number and so on until **n** line items are selected.

(iv) Within each selected line item, the monetary unit corresponding to the selected random number is chosen. If the same line item is selected more than once, then one monetary unit is selected from that line item for each time the line item is selected.

This method of selection is illustrated in table 4.2 using a hypothetical population of 6 items.

Table 4.2 Selection of Line Items with Simple Random Sampling

Item Number	Size B_i	Cumulative Sum	Assigned Range
1	30	30	1 - 30
2	24	54	31 -54
3	20	74	55 - 74
4	34	108	75 - 108
5	12	120	109 - 120
6	30	150	121 - 150

A random number is selected between 1 and 150. Suppose this is 69. This falls in the assigned range of item 3 and hence item 3 is selected. The sampling continues until **n** items are selected. This sampling method was developed originally by Hansen and Hurwitz (1943) for selecting unequal size clusters with probability proportional to size.

4.3.1 Properties of Simple Random Sampling of Monetary Units

The selection probabilities of line items and monetary units are derived in theorems 4.1 to 4.4.

Theorem 4.1

The probability of line item i being selected is B_i/B for each selection. i.e. SRS selects line items with probability proportional to size (monetary value).

Proof:

At each selection a random number, r say, is selected uniformly in (0,B]. The probability that line item i is selected is

$$P[\sum_{l=1}^{i-1} B_l < r \leq \sum_{l=1}^{i} B_l] = \frac{\sum_{l=1}^{i} B_l - \sum_{l=1}^{i-1} B_l}{B} = \frac{B_i}{B}$$

i.e. at each selection, each line item has a probability proportional to its size (PPS) of being selected.

Theorem 4.2

SRS is epsem for monetary units

Proof:

In SRS, **n** random numbers are chosen uniformly with replacement in (0, B].

The probability that a particular monetary unit is chosen in one selection is 1/B.

Therefore, the probability of selecting any monetary unit is the same, regardless in what line item it occurs.

i.e. SRS is epsem for monetary units

SRS views the population as a collection of B monetary units from which a simple random sample of **n** monetary units is selected. More than one monetary unit from the same line item may be selected into the sample. Thus with SRS, line items are selected with replacement, i.e. the number of line items selected for auditing may be less than the number of selected monetary units. The sample may contain more than one monetary unit from any line item. It would seem, at first glance, that sampling of line items without replacement would be a better procedure because it would return a sample of **n** monetary units from **n** distinct items. However, when SRS selection is done without replacement of line items, it is difficult to keep the line item selection probabilities proportional to size and sooner or later becomes impossible as the sample size increases. For example, if a sample of size 6 is selected without replacement, from the population described in Table 4.2 above, every item would be included with a probability of 1 irrespective of the sizes of the items.

The proportion of times that a line item will occur in a sample for SRS is derived in theorem 4.3 and the mean and variance of the number of times a particular line item will occur in any given sample are derived in theorem 4.4.

Theorem 4.3

When a sample of size **n** is selected with replacement, the probability that line item i is included at least once in the sample is

$$\pi_i = 1 - [1 - \frac{B_i}{B}]^n$$

Proof:

The probability that the ith item is included at least once in the sample is

1 - Prob(ith item is not included in any of the **n** selections)

The probability that line item is included in any one selection is B_i/B (theorem 4.1).

Therefore the probability that it is not included in any one selection is

$$1 - \frac{B_i}{B}$$

Hence, the probability that the ith item is not included in the **n** selections is

$$[1 - \frac{B_i}{B}]^n$$

Therefore, the probability that the ith line item is included at least once in a sample of size n is

$$1 - [1 - \frac{B_i}{B}]^n$$

This completes the proof.

Theorem 4.4

In simple random sampling with replacement, the mean and variance of the number of times line item i is selected in the sample in **n** selections is;

$$\mu_i = \frac{nB_i}{B}$$

$$\sigma_i^2 = \frac{nB_i}{B}(1 - \frac{B_i}{B})$$

Proof:

In theorem 4.1 it was shown that the probability that line item i is included in the sample, in any 1 selection is B_i/B.

Now, let k_i be the number of times that the i^{th} line item appears in a sample of size **n**, where k_i may have any integer value in [0,n].

Consider the joint frequency distribution of the k_i's for all N line items in the population. This is the multinominal distribution with a probability density of

$$P(k_1, k_2 \ldots\ldots k_N) = \frac{n!}{k_1! k_2! \ldots . k_N!} \left(\frac{B_1}{B}\right)^{k_1} \left(\frac{B_2}{B}\right)^{k_2} \ldots . \left(\frac{B_N}{B}\right)^{k_N}$$

$$where \sum_{i=1}^{N} k_i = n$$

The properties of multinominal distribution are derived in numerous texts (see for example Kendall, Stuart and Ord, 1987, p195). It can be shown that the marginal distribution of k_i is binomial with parameters n and B_i /B, 1 <=i <= n. Therefore, for each i, the mean and variance of k_i are

$$\mu_i = E(k_i) = \frac{nB_i}{B}$$

$$\sigma_i^2 = V(k_i) = \frac{nB_i}{B}\left(1 - \frac{B_i}{B}\right)$$

This completes the proof.

In sampling with replacement of line items, nB_i/B represents the average rate of occurrence of line item i in **n** selections.

Clearly, the mean occurrence rate is greater than one when $B_i > B/n$ and less than one when $B_i < B/n$. In this study, line items which are greater than the sampling interval $k = B/n$ are considered to be the 'high-valued' items and it is assumed that the auditor will examine these on a 100% basis. Hence, in the sampled population all the line items have book value amounts less than the sampling interval. So, even though SRS is sampling with replacement of line items, each line item does not occur more than once on average in samples of size **n**.

4.3.2 Summary

The fundamental characteristics of simple random sampling of monetary units are derived in theorems 4.1 to 4.4. The selection method chooses an epsem sample of monetary units. The line items are selected with probability proportional to size in any one selection but in a sample of size **n**, the line items may occur more than once. However, the average rate of occurrence of any line item in a sample of size **n** is less than 1 provided $B_i < B/n$ for all line items.

4.4 Systematic Sampling

A systematic sample of **n** monetary units drawn from a population of N line items and B monetary units is obtained as follows:

(i) The cumulative sum of the line items is formed and a range of numbers is assigned to each line item.

(ii) The population is divided into **n** subsections of width $k = B/n$.

(iii) A random integer is selected in [1,k], r say.

(iv) The monetary units are then selected in intervals of k starting at the r^{th} monetary unit.
i.e. the sample consists of the

$$r^{th},\ (r+k)^{th},\ (r+2k)^{th}, \ldots\ldots\ (r+(n-1)k)^{th}$$

monetary units.

Systematic sampling differs from SRS in that the monetary units are not selected independently at each selection. When the random number r is selected uniformly from [1,k], the sample is completely determined. Also, unlike SRS, systematic does not give all the possible sets of monetary units a chance of being selected. In fact, there are only k possible samples that can be selected under systematic sampling. Once r is selected, the sample elements are completely determined, and r can take on only k distinct values.

4.4.1 Properties of Systematic Sampling of Monetary Units

The selection probabilities of line items and monetary units are derived in theorems 4.5 to 4.7.

Theorem 4.5

Systematic sampling is sampling without replacement of line items provided $B_i < k$. i.e. in a sample of size **n**, the monetary units are selected from **n** distinct line items.

Proof:

Suppose B_i is selected in two distinct selections. This means that there exists two distinct integers j and h, j < h say, so that

$$\sum_{l=1}^{i-1} B_l < r + (j - 1)k \le \sum_{l=1}^{i} B_l$$

$$\sum_{l=1}^{i=1} B_l < r + (h - 1)k \le \sum_{l=1}^{i} B_l$$

Since j < h, then

$$\sum_{l=1}^{i-1} B_l < r + (j - 1)k < r + (h - 1)k \le \sum_{l=1}^{i} B_l$$

Now
$$[r + (h - 1)k] - [r + (j - 1)k] = (h - j)k > k$$

Therefore

$$\sum_{l=1}^{i} B_l - \sum_{l=1}^{i-1} B_l = B_i > k$$

which contradicts the assumption that $B_i < k$ for all i.

Therefore, it is not possible to select a line item twice.

i.e. systematic sampling is sampling without replacement of line items.

Theorem 4.6

Systematic sampling selects line items with probability proportional to size. Specifically, the probability of selecting line item i in a sample of size n is

$$\pi_i = \frac{nB_i}{B}$$

provided $B_i < B/n$.

Proof:

Let j be the largest integer so that

$$(j - 1)k \leq \sum_{l=1}^{i-1} B_l$$

if B_i is to be included in the sample, it will be the j^{th} item since r lies in the interval [1,k]. This implies that item i is selected as the jth item provided r is chosen so that

$$\sum_{l=1}^{i-1} B_l < r + (j - 1)k \leq \sum_{l=1}^{i} B_l$$

Then, the probability that line item i is selected into the sample is the probability that r is chosen so that

$$\pi_i = prob\left[\sum_{l=1}^{i-1} B_l < r + (j - 1)k \leq \sum_{l=1}^{i} B_l\right]$$

$$= Prob\left[\sum_{l=1}^{i-1} B_l - (j - 1)k < r \leq \sum_{l=1}^{i} B_l - (j - 1)k\right]$$

$$= \frac{\sum_{l=1}^{i} B_l - (j - 1)k - \sum_{l=1}^{i-1} B_i - (j - 1)k}{k}$$

$$= \frac{B_i}{k}$$

$$= \frac{n\,B_i}{B}$$

i.e. item i is chosen with probability proportional to size.

Theorem 4.7

Systematic sampling is epsem for monetary units.

Proof:

In systematic sampling, a random selection of one monetary unit in made from each interval.

Since each interval contains k monetary units, the probability of selecting a monetary unit from any interval is

$$\frac{1}{k} = \frac{n}{B}$$

i.e. systematic sampling is epsem for monetary units.

4.4.2 Summary

Theorems 4.5 to 4.7 derive the fundamental characteristics of systematic sampling of monetary units. The sampling method is epsem for monetary units. The line items are selected with probability proportional to size and sampling is done without replacement of line items provided all the line items are less than the sampling interval size k = B/n.

4.5 Cell Sampling

A cell sample of **n** monetary units from a population of **N** line items and **B** monetary units is obtained as follows;

(i) The cumulative sum of the line items is formed and a range of numbers is assigned to each line item

(ii) The population is divided into **n** subsections or cells, each of width $k = B/n$

(iii) One random number is selected independently from each cell as follows;

r_1 is chosen uniformly from $(0,k]$

r_2 is chosen uniformly from $(k,2k]$

r_3 is chosen uniformly from $(2k,3k]$

.

.

r_j is chosen uniformly from $((j-1)k,jk]$

.

.

r_n is chosen uniformly from $((n-1)k, nk]$

(iv) The monetary units corresponding to the random numbers are selected into the sample and the line items in which the monetary units occur are chosen for examination by the auditor.

Cell sampling is similar to systematic sampling in the sense that the population of size B monetary units is divided into $k = B/n$ intervals or cells and one element is selected from each cell. However, in cell sampling, an independent selection of a monetary unit is made within each interval whereas in systematic sampling, the units occur at the same relative position in each interval. This may lead to bias if there are regularities and patterns in the error patterns and cell sampling avoids this bias potential. Some of the advantages of cell sampling described by Leslie, Anderson and Teitlebaum (1979, p140) are;

(i) The method is simple.

(ii) It results in exactly the sample size desired.

(iii) It avoids any risk of bias associated with systematic sampling.

(iv) The sample is distributed evenly across the entire population.

4.5.1 Properties of Cell Sampling of Monetary units

The selection probabilities of cell sampling are derived in theorems 4.8 and 4.9

Theorem 4.8

The following line item selection probabilities prevail in cell sampling:

(a) Case 1. Line items contained completely in one cell.

If line items do not straddle two cells, cell sampling is sampling without replacement of line items and selects line items with probability proportional to size.

(b) Case 2. Line items not contained completely in one cell

If a line item straddles two cells, the selection probabilities of line items in each cell are proportional to the amount of the line items in each cell. The maximum number of times any line item is selected into a sample is two.

Proof:

Consider π_i the probability that the ith line item with book value B_i will be selected into the sample.

Case 1

Suppose item i is contained completely in the jth interval

$$((j-1)k,\ jk].$$

i.e.

$$\sum_{l=1}^{i-1} B_l \geq (j-1)k$$

and

$$\sum_{i=1}^{i} B_l \leq jk$$

Line item i will be selected at most once in a sample of size n. If a random number, r_j is selected uniformly in this interval, then line item i is selected as the jth item provided r_j is chosen in the interval $((j-1)k,\ jk]$ so that

$$\sum_{l=1}^{i-1} B_l < r_j \leq \sum_{l=1}^{i} B_l$$

Therefore,the probability of selecting item i as the jth sample item is

$$\pi_i = Prob\left(\sum_{l=1}^{i-1} B_l < r_j \le \sum_{l=1}^{i} B_l\right)$$

$$= \frac{\sum_{l=1}^{i} B_l - \sum_{l=1}^{i-1} B_l}{jk - (j-1)k}$$

$$= \frac{B_i}{jk - (j-1)k}$$

$$= \frac{B_i}{k}$$

$$= \frac{nB_i}{B}$$

i.e. line items are selected with probability proportional to size.

Case 2

Suppose item i straddles two cells, the first part in cell (j -1) and the second part in cell j say.

i.e.

$$(j-2)k < \sum_{l=1}^{i-1} B_l \le (j-1)k < \sum_{l=1}^{i} B_l \le jk$$

In this case, cell sampling is not sampling without replacement. Since an independent selection is made in each cell, item i may be selected more than once.

Line item i, straddling two cells, may be selected as the (j-1) item provided r_{j-1} is chosen so that

$$\sum_{l=1}^{i-1} B_l < r_{j-1} \leq (j-1)k$$

Now since r_{j-1} is chosen uniformly in the interval

$$((j-2)k,\ (j-1)k]$$

Then

$$Prob[\sum_{l=1}^{i-1} B_l < r_{j-1} \leq (j-1)k] = \frac{(j-1)k - \sum_{l=1}^{i-1} B_l}{(j-1)k-(j-2)k}$$

$$= \frac{B_i \cap cell(j-1)}{k}$$

$$= \frac{n(B_i \cap cell(j-1))}{B}$$

i.e. the probability that the line item i is selected in the $(j-1)^{th}$ cell is proportional to the amount of the book value of the line item contained in the $(j-1)^{th}$ cell.

Similarly the line item i will get selected as the jth item provided r_j is chosen so that

$$(j-1)k < r_j \leq \sum_{l=1}^{i} B_l$$

Since r_j is chosen uniformly in the interval

$$((j-1)k, jk]$$

Then

$$Prob[(j-1)k < r_j \leq \sum_{i=1}^{i} B_l] = \frac{\sum_{l=1}^{i} B_i - (j-1)k}{jk - (j-1)k}$$

$$= \frac{B_i \cap Cell(j)}{k}$$

$$= \frac{n[B_i \cap Cell(j)]}{B}$$

Therefore the selection probability of a line item in each cell is proportional to the amount of the line item in each cell.

Since $B_i < k$, for all i, then two is the maximum number of times that item i can be selected.

Theorem 4.9

Cell sampling is epsem for monetary units

Proof:

Case 1 Line item i is contained completely in the j^{th} cell

When line item i is selected as the jth unit and is included in the jth cell then the probability of selecting a particular monetary unit from line item i is

$$prob(item\ i\ is\ selected) * prob(\ one\ unit\ is\ selected)$$

$$= \frac{B_i}{k} * \frac{1}{B_i}$$

$$= \frac{nB_i}{B} \frac{1}{B_i}$$

$$= \frac{n}{B}$$

Case 2: Item i straddles two cells, cell (j -1) and cell (j) say

In the $(j - 1)^{th}$ cell, the monetary units will get selected with probability

$$\frac{(j - 1)k - \sum_{l=1}^{i-1} B_l}{k} * \frac{1}{(j - 1) - \sum_{l=1}^{i-1} B_l} = \frac{1}{k} = \frac{n}{B}$$

And in the jth cell, the monetary units will get selected with probability

$$\frac{\sum_{l=1}^{i} B_i - (j-1)k}{k} * \frac{1}{\sum_{l=1}^{i} B_l - (j-1)k} = \frac{1}{k} = \frac{n}{B}$$

Therefore cell sampling is epsem for monetary units.

4.5.2 Summary

Theorems 4.8 and 4.9 derive the characteristics of cell sampling of monetary units. Cell sampling is sampling without replacement of line items if line items do not straddle two cells. But when an item straddles two cells, cell sampling is not sampling without replacement of line items. It is possible for that item to be selected in each of the two cells and hence two monetary units will be selected from that line item. Two is the maximum number of times that any line item can be selected into a sample when $B_i < B/n$ for all line items. Cell sampling returns an epsem sample of monetary units.

4.6 Lahiri Sampling

Lahiri sampling was proposed by Lahiri (1951) as a convenient method of selection for unequal size clusters with probability proportional to size and is adapted to the auditing situation in this study.

A Lahiri sample of monetary units is selected as follows;

(i) A random integer is selected uniformly from [1,N], i say

(ii) Another random integer is selected uniformly in $[1, B_{max}]$, r_i say.

(iii) If r_i is less than or equal to B_i, then the line item i and the monetary unit corresponding to r_i are selected.

(iv) Steps (ii) and (iii) are repeated until **n** monetary units are selected.

4.6.1 Properties of Lahiri Sampling

The main characteristics of Lahiri sampling are derived in theorems 4.10 to 4.14.

Theorem 4.10

Lahiri sampling selects line items with probability proportional to size in each selection.

Proof:

The probability of selecting item i in any one selection is;

Prob (selection of i from [1,N]) * prob (r_i <= Bi from [1,B_{max}]

$$\frac{1}{N} * \frac{B_i}{B_{max}}$$

which is proportional to size of line item i.

Theorem 4.11

Lahiri sampling is epsem for monetary units.

Proof:

The probability of selecting a monetary unit in one draw can be looked upon as a two-stage process. The first stage is the selection of a line item and the second stage is the selection of a monetary unit within the selected line item.

For any line item, i say, the probability of selection of a monetary unit in any one selection is

Prob (selection of line item i) * Prob (selection of one monetary unit from line item i)

$$\frac{B_i}{N.B_{\max}}\,\frac{1}{B_i} = \frac{1}{N.B_{\max}}$$

which is independent of i. Therefore, each monetary unit has the same chance of being selected irrespective of what line item it belongs.

i.e. Lahiri sampling is epsem for monetary units.

Lahiri sampling may involve many rejections before a monetary unit is selected, especially if the largest line item is very much bigger than the smaller ones. The method involves fewest rejections when the line items do not differ too much in size. Theorems 4.12 derives the average number of trials necessary to obtain the desired sample size.

Theorem 4.12

The expected number of trials necessary to obtain **n** monetary units using Lahiri sampling is

$$\frac{nNB_{max}}{B}$$

Proof:

The probability of selecting line item i in one trial is

$$\frac{B_i}{NB_{max}}$$

Therefore, the probability of selecting any item in one trial is

$$\sum_{i=1}^{N} \frac{B_i}{NB_{max}} = \frac{B}{NB_{max}}$$

If x is the number of trials necessary to obtain **n** items, then x is a negative binomial distribution with probability density function;

$$f(x) = \binom{x-1}{n-1} \left(\frac{B}{NB_{max}}\right)^{n} \left(1 - \frac{B}{NB_{max}}\right)^{x-n} \qquad for\ x = n,\ n+1,\ n+2, ..$$

The properties of the negative binomial distribution are derived in numerous texts (see for example Kendall, Stuart and Ord, 1987).

The mean of x is

$$\mu = \frac{nNB_{max}}{B}$$

which is the average number of trials necessary to obtain a sample of size **n**.

Lahiri sampling is sampling with replacement of line items. Any line item can occur in any trial. Like simple random sampling of monetary units, it places no restriction on how often a line item may occur in the sample. The proportion of times that a line item will occur in a sample of size **n** is derived in theorem 4.13. Theorem 4.14 derives the mean and variance of the incidence of repeated occurrences of a particular line item in a sample of size n. The proofs assume that nNB_{max}/B is an integer. When nNB_{max}/B is not an integer, the results are approximate.

Theorem 4.13

When a sample of size **n** is selected using Lahiri selection, the probability that line item i is selected at least once in the sample is

$$\pi_i = 1 - [1 - (\frac{1}{N}\frac{B_i}{B_{max}})]^{\frac{nNB_{max}}{B}}$$

on average.

Proof:

In theorem 4.12 it was shown that the average number of trials necessary for **n** selections in Lahiri sampling is

$$\frac{nNB_{max}}{B}$$

The probability that the ith line item is included at least once in nNB_{max}/B trials is;

1 - Prob(ith item is not selected in any of the selections)

The probability that the ith line item is not included in the sample is in $[nNB_{max}]/B$ selections is

$$[1 - \frac{B_i}{N\,B_{max}}]^{\frac{nNB_{max}}{B}}$$

Therefore, the probability that the ith line item is included at least once in a sample of size n is

$$1 - [1 - \frac{B_i}{NB_{max}}]^{\frac{nNB_{max}}{B}}$$

This completes the proof.

Theorem 4.14

In sampling with replacement, with probability B_i/NB_{max} that line item i is chosen in any 1 selection, the mean and variance of the number of times line item i is selected in the sample in $[nNB_{max}]/B$ selections are given by

$$\mu_i = \frac{nB_i}{B}$$

$$\sigma_i^2 = \frac{nB_i}{B}(1 - \frac{B_i}{NB_{max}})$$

Proof:

The probability that line item i is selected in any one Lahiri selection is B_i/NB_{max}. Therefore, the number of times line item i is selected in (nNB_{max}/B) selections is a binomial distribution with parameters;

$$\mu_i = \frac{nNB_{max}}{B} \cdot \frac{B_i}{NB_{max}} = \frac{nB_i}{B}$$

$$\sigma_i^2 = \left(\frac{nNB_{max}}{B}\right)\left(\frac{B_i}{NB_{max}}\right)\left(1 - \frac{B_i}{NB_{max}}\right) = \frac{nB_i}{B}\left(1 - \frac{B_i}{NB_{max}}\right)$$

Clearly, the mean occurrence rate is less than one when $B_i < B/n$ which is the case in this study. So, like SRS, each line item does not occur more than once, on average, in samples of size **n**. However, the variance of the rate of occurrence of line items in Lahiri selection is greater than SRS since

$$V_{srs} = \frac{nB_i}{B}\left(1 - \frac{B_i}{B}\right) \leq \frac{nB_i}{B}\left(1 - \frac{B_i}{NB_{max}}\right) = V_{lah}$$

The extent of the differences in the variances depends on how much NB_{max} differs from B.

4.6.2 Summary

The fundamental characteristics of Lahiri sampling are derived in theorems 4.10 to 4.14. The selection method chooses an epsem sample of monetary units. The line items are selected with probability proportional to their recorded book amounts in any one selection. The average number of selections to obtain a sample of size **n** is $(nNB_{max})/B$. The amount of sampling required to obtain a sample of size **n** varies depending on the line item structure of the population. If the largest line item in the population is very much larger than the smaller ones, Lahiri sampling may involve many rejections and hence a lot of sampling before the sample of size **n** is selected. Lahiri sampling involves fewest rejections when the B_i do not differ too much in size. The sampling method is with replacement of line items and the average rate of occurrence of a particular line item is nB_i/B, similar to SRS. The variance of the number of occurrences of each line item using Lahiri sampling, is greater than or equal to the variance using SRS. Unlike SRS, Lahiri sampling does not require the accumulation of book value totals and consequently, Lahiri sampling may be implemented before the total book value amount is known accurately. Leslie, Teitlebaum and Anderson (1979, p101) point out that the total book value amount may not always be known accurately during the planning stage and it may not be known for transaction streams prior to the end of year.

4.7 Sieve Sampling

A sieve sample of **n** monetary units is selected as follows:

(i) The population total B is divided by the nominal sample size **n**. Let $k = B/n$. It is assumed that each line item is less than or equal to k.

(ii) For each line item, a random integer is selected uniformly from [1,k], r_i say.

(iii) If $r_i \leq B_i$ then the i^{th} line item is chosen.

(iv) The monetary unit selected into the sample from line item i is that which corresponds to r_i.

4.7.1 The Properties of Sieve Sampling

The main characteristics of sieve sampling are derived in theorems 4.15 to 4.18.

Theorem 4.15

The probability that line i is selected in a sample of size n chosen using sieve sampling is nB_i/B.

i.e. sieve sampling selects line items with probability proportional to their recorded book amounts.

Proof:

For each line item, a random integer r_i is selected uniformly from the interval (0,k]. The probability that line item i will be selected is therefore

$$Prob\ (1 \leq r_i \leq B_i) = \frac{B_i}{k} = \frac{nB_i}{B}$$

which is proportional to the size of the line item.

Another way of looking at sieve sampling is to imagine that every item is spread on a sieve with random mesh size, uniformly distributed between 0 and B/n. For each item, a mesh size r_i is generated. Item i remains on the sieve with probability $B_i/(B/n)$. The random number r_i is called the item sieve for item i.

Theorem 4.16

Sieve sampling is epsem for monetary units

Proof:

For each line item, a random number is selected uniformly from (0,k]. A particular monetary unit in line item i say will be selected with probability

$$\frac{1}{k} = \frac{n}{B}$$

Therefore sieve sampling is epsem for monetary units.

Theorem 4.17

Sieve sampling is sampling without replacement of line items provided $B_i < k$ for all i.

Proof:

This proof follows immediately from the selection procedure which makes it clear that a line item can be selected at most once.

Sieve sampling is equivalent to cell sampling when all the items are of equal size. A practical disadvantage of sieve sampling is that the achieved sample size n_o is not always equal to the nominal sample size n. The achieved sample size is variable and depends on the random numbers selected. However, when all the line items are less than or equal to B/n the achieved sample size n_0 has the properties derived in theorem 4.18

Theorem 4.18

(1) The achieved sample size n_0 is an unbiased estimator of the nominal value n.

i.e.

$$E(n_0) = n$$

(2) The variance of the achieved sample size n_0 is

$$Var(n_0) = n - \frac{n^2}{B^2}\sum_{i=1}^{N} B_i^2$$

Proof:

For each line item, define a variable x_i so that

$x_i = 1$ when line item i is selected into the sample

$= 0$ when line item i is not selected into the sample.

Then x_i is a Bernoulli variable with the following probabilities;

$$x_i = 1 \quad \textit{with probability} \quad \frac{n B_i}{B}$$
$$= 0 \quad \textit{with probability} \quad 1 - \frac{n B_i}{B}$$

The achieved sample size n_0 is the sum of N Bernoulli variables. i.e.

$$n_0 = \sum_{i=1}^{N} x_i$$

and

$$E(n_0) = \sum_{i=1}^{N} E(x_i)$$
$$= \sum_{i=1}^{N} \frac{n B_i}{B}$$
$$= \frac{n}{B} \sum_{i=1}^{N} B_i$$
$$= n$$

i.e. n_0 is unbiased for n

(2) the variance of n_0

$$V(n_0) = V\left(\sum_{i=1}^{N} x_i\right)$$

$$= \sum_{i=1}^{N} V(x_i)$$

$$= \sum_{i=1}^{N} \frac{n B_i}{B}\left(1 - \frac{n B_i}{B}\right)$$

$$= \frac{n}{B}\sum_{i=1}^{N} B_i - \frac{n^2}{B^2}\sum_{i=1}^{N} B_i^2$$

$$= n - \frac{n^2}{B^2}\sum_{i=1}^{N} B_i^2$$

This completes the proof.

Corollary 1: It follows from theorem 4.18, part 2 that the standard deviation of the sample size is

$$\sigma_{n_0} = \sqrt{n\left(1 - \frac{n}{B^2}\sum_{i=1}^{N} B_i^2\right)} \leq \sqrt{n}$$

Corollary 2: The coefficient of variation of the sample size

$$\frac{\sigma_{n_0}}{E(n_0)} = \frac{\sqrt{(n(1 - \frac{n}{B^2}\sum_{i=1}^{N} B_i^2)}}{n} \leq \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$$

The variation in the achieved sample size may be of concern to the auditor when the costs of carrying out the audit are being estimated prior to the audit. Clearly, the relative variation compared to the nominal sample size decreases as the sample size increases but it may not be negligible for any sample size.

4.7.2 Summary

The characteristics of sieve sampling are derived in theorems 4.15 to 4.18. Sieve sampling selects line items without replacement and with probability proportional to their recorded book amounts. Selection of monetary units is epsem under sieve sampling. The sample size is variable but the achieved sample size n_0 is an unbiased estimator of the nominal sample size **n** and the standard deviation is less than or equal to $\sqrt{n}$.

4.8 Stabilised Sieve Sampling

Stabilised sieve sampling is a sampling method based on sieve sampling which returns a constant sample of **n** monetary units. Stabilised sieve sampling chooses the sample in two stages as follows:

Stage 1

An initial sample is selected by means of sieve sampling. Suppose the achieved sample size is n_0.

Stage 2

The second stage reduces or increases the initial sample size if n_0 is not equal to **n**.

Case 1 Sample Reduction

When n_0 is greater than **n**, a simple random sample of **n** monetary units is selected from the initial sample.

Case 2 Sample Augmentation

If n_0 is less then **n**, then the sample is augmented by selecting a further $(n - n_0)$ monetary units from the original population using the Lahiri selection method (see 4.6).

When the sample is augmented, the extra line items selected at the second stage are chosen without excluding the line items originally selected. This is necessary to maintain PPS. Wright (1991) documents the problem of augmenting a sample selected with probabilities proportional to size. Using systematic sampling, he derives the inclusion probabilities for an augmented sample obtained by resampling the population with probabilities proportional to size of line items, after excluding the line items selected at the first stage. He proves that the inclusion probabilities of the augmented sample are not, in general, proportional to the size of the line items. The results show that an augmented sample obtained in this manner possesses biased inclusion probabilities. Wright (1991) concludes that expanding a PPS sample is not a trivial exercise. Resampling from the original population, after excluding the line items already selected in the original sample, does not preserve the appropriate PPS inclusion probabilities. In general, the augmented sample will tend to be comprised of items which over-represent the smaller population items and under-represent the larger items. This is because the inclusion probabilities for the augmented sample are higher (lower) for the smaller (higher) population items than those for a PPS sample of the same size selected in a single stage. If the errors are concentrated in the smaller (larger) items, then an over (under) estimate of the total error amount will result. Therefore, in order to maintain PPS and to avoid biased inclusion probabilities, stabilized sieve sampling does not exclude the items selected at the first stage, when augmenting a sample at the second stage.

4.8.1 Properties of Stabilised Sieve Sampling

The main properties of stabilized sieve sampling are derived in theorems 4.19 to 4.21.

Theorem 4.19

Stabilised sieve sampling selects line items with probabilities proportional to their recorded book amounts.

Proof:

(1) Sample Reduction ($n_0 > n$)

In sample reduction, the probability of inclusion of any line item, line item i say, in the final sample is the probability of inclusion of the item at the first stage and the probability of inclusion of the item at the second stage.

i.e.

$$prob(stage1) = \frac{nB_i}{B} \quad (\text{ sieve sampling })$$

$$prob(stage2) = \frac{n}{n_0} \quad (\text{ SRS })$$

$$prob(item\ \ i\ is\ selected) = \frac{nB_i}{B}\ \frac{n}{n_0}$$

which is proportional to B_i.

(2) Sample Augmentation ($n_0 < n$)

In sample augmentation, the probability of inclusion of any item in the final sample is the probability of inclusion in the first stage or the probability of inclusion in the second stage.

The probabilities of selecting item i at each stage are

$$prob(stage1) = \frac{nB_i}{B} \qquad (sieve\ sampling)$$

$$prob(stage2) = \frac{n'}{N}\frac{B_i}{B_{max}} \qquad (Lahiri\ sampling)$$

where n′ is the number of trials necessary to obtain ($n - n_0$) line items. It can be deduced from theorem 4.12 that the average number of trials necessary to obtain a sample of size ($n-n_0$) is $(n-n_o)NB_{max}/B$, i.e. $n' = (n-n_o)NB_{max}/B$.

Since each line item is considered for inclusion at both stages, the probability of selection of each line item is proportional to the line item size at each selection stage.

This completes the proof.

It should be pointed out that SRS, systematic or cell sampling could also be used to augment the sample and the properties derived in theorem 4.19 would prevail under these selection methods. In this study, Lahiri sampling is chosen to augment the sample, because it uses the line item structure of the population when selecting the items.

Theorem 4.20

Stabilised sieve sampling is epsem for monetary units.

Proof:

Case 1: Sample Reduction ($n_0 > n$)

With sample reduction, the probability that a monetary unit is selected into the sample is the probability that it is selected at stage one and stage two, i.e.,

$$\frac{n}{B} * \frac{n}{n_0} = \frac{n^2}{Bn_0}$$

Case 2: Sample Augmentation ($n_0 < n$)

With sample augmentation, monetary units may be selected more than once into a sample.

The probability that a particular monetary unit is selected at the first stage is n/B.

At stage two, a monetary unit is selected with probability

$$\frac{n'}{NB_{max}}$$

where n′ is the number of trials necessary to obtain $(n - n_0)$ line items using Lahiri sampling.

Therefore, the probability that a particular monetary unit is selected in the sample is

$$\frac{n}{B} + \frac{n'}{NB_{max}} - \frac{nn'}{BNB_{max}}$$

Clearly, the monetary units have the same selection probability irrespective of the line items to which they belong.

The selection process is epsem for monetary units.

Theorem 4.21

The expected number of trials to get a sample of size n is

(i) N + n when the initial sample is reduced

and

(ii) $N + (n - n_0)N\ B_{max}/B$ when the initial sample is augmented

Proof:

(i) In sample reduction, the initial sieve sample yields a sample of size $n_0 > n$ in N trials and a simple random sample of n is chosen from the n_0 monetary units originally selected. Therefore, the total number of trials necessary to obtain a sample of size n is (N+n).

(ii) In sample augmentation, the initial sieve sample yields a sample of size $n_0 < n$. The expected number of trials necessary to augment the sample by $(n-n_0)$ is

$$\frac{(n - n_0)\,NB_{max}}{B}$$

This follows from theorem 4.12.

Therefore the expected number of trials necessary to obtain a sample is size n is

$$N + \frac{(n - n_0)\,NB_{max}}{B}$$

Obviously, the number of repeated selections needed to augment a sample of size n_0 to n depends on (i) the difference between n and n_0 and (ii) the size of B_{max} relative to B.

Stabilized sieve sampling does not, in general, select line items without replacement. At the first stage, the monetary units are chosen without replacement provided all the line items have book value amounts less than B/n.

When a reduction is necessary at the second stage, the **n** selected line items are distinct. Therefore stabilised sieve sampling with reduction is sampling of line items without replacement.

However, when it is necessary to augment the sample, each line item, including those already selected, is considered for selection at the second stage. This is necessary in order to preserve PPS. Therefore the sample selection method at the second stage is sampling with replacement of line items.

Theorem 4.22

When an initial sieve sample of size $n_0 < n$ is augmented to n by Lahiri sampling, the probability that line item i is selected at least once in the sample is

$$1 - \left(1 - \frac{nB_i}{B}\right)\left(1 - \frac{B_i}{NB_{max}}\right)^{(n-n_0)\frac{NB_{max}}{B}}$$

Proof:

The probability that a particular line item, line item i say, is included in the sample at least once is

1 - Prob(ith item is not included in either stage one or two)

It follows from theorem 4.15, that the probability that line item i is not included in a sieve sample of nominal size **n** is

$$1 - \frac{nB_i}{B}$$

Also, when a sample of size $(n - n_0)$ is chosen by means of Lahiri sampling, the probability that line item i is not included is

$$\left(1 - \frac{B_i}{NB_{max}}\right)^{(n - n_0)\frac{NB_{max}}{B}}$$

This follows from theorem 4.13.

The probability that line item i is selected at least once in a sample of size **n** is

1 - Prob(ith item is not included in either stage one or two) which is

$$1 - \left(1 - \frac{nB_i}{B}\right)\left(1 - \frac{B_i}{NB_{max}}\right)^{(n - n_0)\frac{NB_{max}}{B}}$$

This completes the proof.

Since sieve sampling is sampling without replacement of line items, then the maximum number of times a line item can occur in an augmented stabilised sieve sample is $(n - n_0)$.

4.8.2 Summary

Theorems 4.19 to 4.22 derive the fundamental properties of stabilised sieve sampling. When the initial sample size is reduced, the sampling method is sampling without replacement of line item. When the initial sample is increased, the sampling method is not sampling without replacement of the line items but the maximum number of times any particular line item can be selected into the final sample is $(n - n_0)$. At each selection, a line item is selected with probability proportional to its recorded book amount. Stabilised sieve sampling is epsem for monetary units.

Chapter 5

The Point Estimator

5.1 Introduction

This chapter analyses the properties of a point estimator of the total error amount using different monetary-unit sampling methods. The purpose of the analysis is, as stated in 1.8.2, to examine the precision of the point estimator for the six sampling methods under investigation. The mean and variance of the estimator with the taint and AON error assignment are derived theoretically for simple random, systematic, cell, Lahiri, sieve and stabilised sieve sampling of monetary units. The design effect (Kish, 1965, p258) is used to compare the precision of systematic, cell, Lahiri, sieve and stabilised sieve sampling relative to simple random sampling of monetary units. The implications of the sample designs for the real accounting populations are also discussed.

The theoretical properties of the point estimator are investigated in order to gain some insight into the behaviour of the heuristic upper bound estimates of the total error amount for the different sampling methods. Previous research by Wurst, Neter and Godfrey (1989a and 1989b) shows that the effects of simple random, cell and sieve sampling on the precision of the point estimator is similar to the effects of the sampling methods on the precision of the Stringer and Cell bounds. This study extends the work of these authors by investigating the theoretical properties of a point estimator using six monetary-unit sampling methods, including a new sampling method 'Stabilised Sieve Sampling' and one not used previously in auditing 'Lahiri Sampling'.

Some of the findings obtained are new to this study. Others, included for completeness, are in agreement with results obtained by other authors.

5.2 Notation

In addition to the notation given in Table 4.1, further notation used in this chapter is outlined in Table 5.1.

Table 5.1 Point Estimator Notation

$T = \sum_{i=1}^{N} E_i$	The total error amount
$\hat{T}$	Sample estimate of the total error amount
$T_{r.} = \sum_{j=1}^{n} u_{r+(j-1)k}$	The total error amount in the r^{th} cluster (systematic sampling)
$T_{r.}(taint) = \sum_{j=1}^{n} t_{r+(j-1)k}$	The total taint amount in the r^{th} cluster (systematic sampling)
$T_i = \sum_{r=1}^{k} u_{r+(i-1)k}$	The total error amount in the i^{th} stratum (cell sampling)
$T_i(taint) = \sum_{r=1}^{k} t_{r+(i-1)k}$	The total taint amount in the i^{th} stratum (cell sampling)

5.3 The Point Estimator

A point estimator is a statistic calculated from the sample data to estimate a parameter in the population. In this study, a point estimator is used to estimate the total error amount in the population. The point estimator is defined as

$$\hat{T}_{taint} = \frac{B}{n} \sum_{i=1}^{n} t_i \qquad \textit{with taint error assignment}$$

$$\hat{T}_{aon} = \frac{B}{n} \sum_{i=1}^{n} u_i \qquad \textit{with AON error assignment}$$

and is used to estimate the total error amount T in the population, where

$$T = \sum_{i=1}^{N} (B_i - A_i) = \sum_{i=1}^{N} E_i$$

In sieve sampling, the summation is over the achieved sample size n_0,

i.e.

$$\hat{T} = \frac{B}{n} \sum_{i=1}^{n_0} t_i \qquad \textit{with taint error assignment}$$

$$= \frac{B}{n} \sum_{i=1}^{n_0} u_i \qquad \textit{with AON error assignment}$$

For stabilised sieve sampling, it is necessary to adjust the estimator to eliminate bias. The estimator used in stabilised sieve sampling is

$$\hat{T}_{taint} = \frac{B}{2n-n_0}\sum_{i=1}^{n} t_i \qquad \textit{with taint error assignment}$$

$$\hat{T}_{aon} = \frac{B}{2n-n_0}\sum_{i=1}^{n} u_i \qquad \textit{with AON error assignment}$$

In the remainder of the chapter, the mean and variance of these estimators are obtained for each sampling method with each error assignment method. The precision of the point estimator based on systematic, cell, Lahiri, sieve and stabilised sieve sampling is compared to the precision of the point estimator based on simple random sampling of monetary units.

5.4 Simple Random Sampling

In simple random sampling of monetary units, the population is looked upon as consisting of **B** monetary units from which a sample of size **n** monetary units is drawn. Each monetary unit has a probability of 1/B of being chosen in any selection. The mean and variance of the point estimator using simple random sampling of monetary units are derived in 5.4.1 and 5.4.2. These derivations are similar to the general results for simple random sampling available in standard texts, (see, for example Cochran, 1977).

5.4.1. Simple Random Sampling with the Taint Error Assignment

The mean and variance of the point estimator with the taint error assignment are derived in theorems 5.1 and 5.2 for simple random sampling of monetary units.

Theorem 5.1

With simple random sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} t_i$$

is an unbiased estimate of the total error amount T.

i.e

$$E\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \sum_{i=1}^{N} (B_i - A_i) = \sum_{i=1}^{N} E_i = T$$

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\, n\, E\,(t_i)$$

$$= B\sum_{i=1}^{B} t_i \frac{1}{B}$$

$$= B\sum_{i=1}^{N} t_i \frac{B_i}{B}$$

$$= \sum_{i=1}^{N} E_i = T$$

Theorem 5.2

With simple random sampling of monetary units,

$$V_{srs}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

With simple random sampling of monetary units,

$$V\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n^2}\sum_{i=1}^{n} V(t_i)$$

$$= \frac{B^2}{n} V(t_i)$$

Now

$$V(t_i) = E(t_i^2) - (E(t_i))^2$$

$$= \sum_{i=1}^{B} t_i^2 \frac{1}{B} - \left(\sum_{i=1}^{B} \frac{t_i}{B}\right)^2$$

$$= \sum_{i=1}^{N} \frac{t_i^2 B_i}{B} - \left(\sum_{i=1}^{N} \frac{t_i B_i}{B}\right)^2$$

$$= \frac{1}{B}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Therefore

$$V\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n}\left[\frac{1}{B}\left\{\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right\}\right]$$

$$= \frac{B}{n}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.4.2 Simple Random Sampling with the AON Error Assignment

With the AON error assignment, the value of each monetary unit, u_i, $1<=i<=N$, is either 1 or 0. In simple random sampling of monetary units

$$u_i = 1 \textit{ with probability } \frac{T}{B}$$

$$= 0 \textit{ with probability } (1 - \frac{T}{B})$$

Therefore, Σu_i is a binomial variable.

Hence

$$E(\sum_{i=1}^{n} u_i) = n\frac{T}{B}$$

$$V(\sum_{i=1}^{n} u_i) = n\frac{T}{B}(1 - \frac{T}{B})$$

The mean and variance of the point estimator with the AON error assignment are derived in theorems 5.3 and 5.4 for simple random sampling of monetary units.

Theorem 5.3

With simple random sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} u_i$$

is an unbiased estimate of the total error amount T.

i.e

$$E(\frac{B}{n}\sum_{i=1}^{n} u_i) = \sum_{i=1}^{N} E_i = T$$

Proof:

$$E(\frac{B}{n}\sum_{i=1}^{n} u_i) = \frac{n}{B}E(\sum_{i=1}^{n} u_i)$$

$$= \frac{B}{n}(\frac{nT}{B})$$

$$= T$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.4

With simple random sampling of monetary units

$$V_{srs}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

$$V_{srs}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B^2}{n^2} V_{srs}\left(\sum_{i=1}^{n} u_i\right)$$

$$= \frac{B^2}{n^2} n\left(\frac{T}{B}\right)\left(1 - \frac{T}{B}\right)$$

$$= \frac{B}{n}\left[T - \frac{T^2}{B}\right]$$

$$= \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.5 Systematic Sampling

To select a systematic sample of **n** monetary units, one monetary unit is selected at random from the first **k** monetary units and every **k**th unit is selected thereafter. The selection of the first monetary unit determines the whole sample. Systematic sampling may be looked upon as a special case of cluster sampling (Cochran, 1977, p 207). The population is divided into **k** clusters each containing **n** monetary units. Table 5.2 illustrates the structure of the population with the taint error assignment.

Table 5.2 A Population Divided into k Clusters

Cluster 1	Cluster 2	. . . . .	Cluster r	. . . . .	Cluster k
t_1	t_2	. . . . .	t_r	. . . . .	t_k
t_{1+k}	t_{2+k}	. . . . .	t_{r+k}	. . . . .	t_{2k}
t_{1+2k}	t_{2+2k}	. . . . .	t_{r+2k}	. . . . .	t_{3k}
.					
$t_{1+(j-1)k}$			$t_{r+(j-1)k}$	. . . . .	t_{jk}
.					
$t_{1+(n-1)k}$	$t_{2+(n-1)k}$	. . . . .	$t_{r+(n-1)k}$	. . . . .	t_{nk}

One cluster is chosen at random from the **k** clusters and the whole cluster is selected into the sample. The probability of selecting any one of the **k** clusters is 1/k.

In the following analysis, $t_{r+(j-1)k}$ and $u_{r+(j-1)k}$ represent respectively the jth taint and the jth AON error amount of the rth cluster, r = 1,2, ...k and j = 1,2,....n.

The mean and variance of the point estimator with each error assignment are derived in 5.5.1 and 5.5.2 for systematic sampling of monetary units. Madow and Madow (1944) derived the theoretical properties of the mean of a sample drawn from a finite population using systematic sampling. The results obtained in this study are similar to those obtained by Madow and Madow (1944) except that they are adapted to the auditing situation.

5.5.1 Systematic Sampling with the Taint Error Assignment.

The mean and variance of the point estimator with the taint error assignment are derived in theorems 5.5 and 5.6 for systematic sampling.

Theorem 5.5

With systematic sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} t_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = E\left(\frac{B}{n}\sum_{j=1}^{n} t_{r+(j-1)k}\right)$$

$$= \frac{B}{n}\sum_{r=1}^{k}\left[\left(\sum_{j=1}^{n} t_{r+(j-1)k}\right)\right]\frac{1}{k}$$

$$= \sum_{i=1}^{B} t_i \qquad \textit{since } n.k = B$$

$$= \sum_{i=1}^{N} t_i B_i \quad \textit{since each } t_i \textit{ occurs } B_i \textit{ times}$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.6

With systematic sampling of monetary units and the taint error assignment,

$$V_{sys}\left(\frac{B}{n}\sum_{j=1}^{n} t_{r+(j-1)k}\right) = \frac{B}{n}\left[T_{r.}^2(taint) - \frac{n}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

$$V_{sys}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = V_{sys}\left(\frac{B}{n}\sum_{j=1}^{n} t_{r+(j-1)k}\right)$$

$$= E\left(\frac{B}{n}\sum_{j=1}^{n} t_{r+(j-1)k} - T\right)^2$$

$$= \sum_{r=1}^{k}\left[\frac{B}{n}\sum_{j=1}^{n} t_{r+(j-1)k} - T\right]^2\frac{1}{k}$$

$$= \frac{n}{B}\sum_{r=1}^{k}\left[\frac{B}{n}T_{r.}(taint) - T\right]^2$$

$$= \frac{n}{B}\left[\sum_{r=1}^{k}\left(\frac{B}{n}\right)^2 T_{r.}^2(taint) - 2T\frac{B}{n}\sum_{r=1}^{k} T_{r.}(taint) + \frac{B}{n}T^2\right]$$

$$= \frac{n}{B}\left[\sum_{r=1}^{k}\left(\frac{B}{n}\right)^2 T_{r.}^2(taint) - 2T\frac{B}{n}\sum_{r=1}^{k}\sum_{j=1}^{n}\frac{E_{r+(j-1)k}}{B_{r+(j-1)k}} + \frac{B}{n}T^2\right]$$

$$= \sum_{r=1}^{k}\frac{B}{n}T_{r.}^2 taint) - 2T\sum_{i=1}^{N}\frac{E_i}{B_i}B_i + T^2$$

$$= \sum_{r=1}^{k}\frac{B}{n}T_{r.}^2(taint) - 2T^2 + T^2$$

$$= \sum_{r=1}^{k}\frac{B}{n}T_{r.}^2(taint) - T^2$$

$$= \frac{B}{n}\left[\sum_{r=1}^{k} T_{r.}^2(taint) - \frac{n}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.5.2 Systematic Sampling with AON Error Assignment

The mean and variance of the point estimator under systematic sampling with the AON error assignment method are derived in theorems 5.7 and 5.8.

Theorem 5.7

With systematic sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} u_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = E\left(\frac{B}{n}\sum_{j=1}^{n} u_{r+(j-1)k}\right)$$

$$= \frac{B}{n}\sum_{r=1}^{k} \left[\sum_{j=1}^{n} u_{r+(j-1)k}\right] \frac{1}{k}$$

$$= \sum_{r=1}^{k}\sum_{j=1}^{n} u_{r+(j-1)k}$$

$$= \sum_{i=1}^{B} u_i$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.8

With systematic sampling of monetary units

$$V_{sys}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\left[\sum_{r=1}^{k} T_{r.}^2 - \frac{n}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

$$V_{sys}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = V_{sys}\left(\frac{B}{n}\sum_{j=1}^{n} u_{r+(j-1)k}\right)$$

$$= E\left(\frac{B}{n}\sum_{j=1}^{n} u_{r+(j-1)k} - T\right)^2$$

$$= \sum_{r=1}^{k} \left[\frac{B}{n}\sum_{j=1}^{n} u_{r+(j-1)k} - T\right]^2 \frac{1}{k}$$

$$= \frac{n}{B}\sum_{r=1}^{k} \left[\frac{B}{n}T_{r.} - T\right]^2$$

$$= \frac{n}{B}\left[\sum_{r=1}^{k} \left(\frac{B}{n}\right)^2 T_{r.}^2 - 2T\frac{B}{n}\sum_{r=1}^{k} T_{r.} + \frac{B}{n}T^2\right]$$

$$= \frac{n}{B}\sum_{r=1}^{k} \left(\frac{B}{n}\right)^2 T_{r.}^2 - 2T\sum_{r=1}^{k}\sum_{j=1}^{n} u_{r+(j-1)k} + T^2$$

$$= \sum_{r=1}^{k} \frac{B}{n}T_{r.}^2 - 2T\sum_{i=1}^{N} E_i + T^2$$

$$= \sum_{r=1}^{k} \frac{B}{n}T_{r.}^2 - 2T^2 + T^2$$

$$= \sum_{r=1}^{k} \frac{B}{n}T_{r.}^2 - T^2$$

$$= \frac{B}{n}\left[\sum_{r=1}^{k} T_{r.}^2 - \frac{n}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.6 Cell Sampling

To select a cell sample of **n** monetary units, the population of **B** monetary units is divided into **n** subsets or cells and an independent selection is made in each cell. Cell sampling may be looked upon as a special case of stratified sampling where the population of **B** monetary units is divided into **n** strata each of size **k** and one monetary unit is chosen independently from each stratum. The mean and variance of the point estimator under cell sampling with each error assignment method are derived in 5.6.1 and 5.6.2. The results are similar to those obtained by Wurst, Neter and Godfrey (1989b).

5.6.1 Cell Sampling with the Taint Error Assignment.

With the taint error assignment, each t_i is selected independently from the i^{th} cell, i<=i<=n. i.e. $t_i = t_{r+(i-1)k}$ where **r** is chosen randomly from [1, k]. The mean and variance of the point estimator under cell sampling with the taint error assignment are derived in theorems 5.9 and 5.10.

Theorem 5.9

With cell sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} t_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\sum_{i=1}^{n} E(t_i)$$

Now $$E(t_i) = \sum_{r=1}^{k} t_{r+(i-1)k}\frac{1}{k}$$

$$\therefore \quad E\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\sum_{i=1}^{n}\sum_{r=1}^{k} t_{r+(i-1)k}\frac{1}{k} \quad \text{where } k = \frac{B}{n}$$

$$= \frac{B}{n}\sum_{i=1}^{N} t_i \frac{B_i}{k}$$

$$= \sum_{i=1}^{N} \frac{E_i}{B_i} B_i$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.10

With cell sampling, the variance of the point estimator is

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} \frac{E_i^{\,2}}{B_i} - \frac{n}{B}\sum_{i=1}^{n} T_i^2(taint)\right]$$

Proof:

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n^2}\sum_{i=1}^{n} V(t_{r+(i-1)k}) \qquad r \in [1,k]$$

$$V(t_{r+(i-1)k}) = E(t_{r+(i-1)k})^2 - [E(t_{r+(i-1)k})]^2$$

$$E(t^2_{r+(i-1)k}) = \sum_{r=1}^{k} t^2_{r+(i-1)k}\frac{1}{k}$$

$$[E(t_{r+(i-1)k})]^2 = \left[\sum_{r=1}^{k} t_{r+(i-1)k}\frac{1}{k}\right]^2$$

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n^2}\left[\sum_{i=1}^{n}\sum_{r=1}^{k} t^2_{r+(i-1)k}\frac{1}{K} - \sum_{i=1}^{n}\left(\sum_{r=1}^{k} t_{r+(i-1)k}\frac{1}{K}\right)^2\right]$$

$$= \frac{B^2}{n^2}\left\{\sum_{i=1}^{N} \frac{E_i^2}{B_i^{\,2}}B_i\cdot\frac{n}{B} - \frac{n^2}{B^2}\sum_{i=1}^{n}\left[\sum_{r=1}^{k} t_{r+(i-1)k}\right]^2\right\}$$

$$= \frac{B}{n}\left\{\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{n} T_i^{\,2}(taint)\right\}$$

5.6.2 Cell sampling with the AON Error Assignment

With the AON error assignment, each u_i is selected independently from the ith cell, 1<=i<=n. i.e. $u_i = u_{r+(i-1)k}$ where r is chosen randomly from [1,k]. The mean and variance of the point estimator under cell sampling with the AON error assignment are derived in theorems 5.11 and 5.12.

Theorem 5.11

With cell sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} u_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\sum_{i=1}^{n} E(u_i)$$

$$E(u_i) = \sum_{r=1}^{k} u_{r+(i-1)k}\frac{1}{k}$$

$$\therefore \quad E\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\sum_{i=1}^{n}\sum_{r=1}^{k} u_{r+(i-1)k}\frac{1}{k}$$

$$= \frac{B}{n}\sum_{i=1}^{N} E_i\frac{1}{k}$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.12

With cell sampling of monetary units

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{n} T_i^2\right]$$

Proof:

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B^2}{n^2}\sum_{i=1}^{n} V(u_{r+(i-1)k}) \qquad r\ e\ [1,k]$$

$$V(u_{r+(i-1)k}) = E(u_{r+(i-1)k})^2 - [E(u_{r+(i-1)k})]^2$$

$$E(u^2_{r+(i-1)k}) = \sum_{r=1}^{k} u^2_{r+(i-1)k}\frac{1}{k} = \sum_{r=1}^{k} u_{r+(i-1)k}\frac{1}{k}$$

$$[E(u_{r+(i-1)k})]^2 = \left[\sum_{r=1}^{k} u_{r+(i-1)}\frac{1}{k}\right]^2$$

$$V_{cell}\left(\frac{B}{n}\sum_{i=1}^{u} u_i\right) = \frac{B^2}{n^2}\left[\sum_{i=1}^{n}\sum_{r=1}^{k} u_{r+(i-1)k}\frac{1}{k} - \sum_{i=1}^{n}\left(\sum_{r=1}^{k} u_{r+(i-1)k}\frac{1}{k}\right)^2\right]$$

$$= \frac{B^2}{n^2}\left\{\sum_{i=1}^{N} E_i\frac{n}{B} - \frac{n^2}{B^2}\sum_{i=1}^{n}\left[\sum_{r=1}^{k} u_{r+(i-1)k}\right]^2\right\}$$

$$= \frac{B}{n}\left\{\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{n} T_i^2\right\}$$

5.7 Lahiri Sampling

The mean and variance of the point estimator with each error assignment method are derived for Lahiri sampling in 5.7.1 and 5.7.2. As far as can be determined, these results are presented here for the first time.

It has been shown in theorem 4.11 that the probability of selecting a monetary unit in any one trial with Lahiri sampling is

$$\frac{1}{NB_{max}}$$

In theorem 4.12, it was shown that the expected number of trials necessary to obtain a sample of size **n** is

$$\frac{nNB_{max}}{B}$$

Therefore, the expected number of trials necessary to obtain one selection is

$$\frac{NB_{max}}{B}$$

These properties will be used when deriving the mean and variance of the point estimator.

5.7.1 Lahiri Sampling with the Taint Error Assignment

The mean and variance of the point estimator with Lahiri sampling using the taint error assignments are derived in theorems 5.13 and 5.14.

Theorem 5.13

With Lahiri sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n} t_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E(\frac{B}{n}\sum_{i=1}^{n} t_i) = \frac{B}{n}\sum_{i=1}^{n} E(t_i)$$

$$= B\ E\ (t_i)$$

$$= B\sum_{i=1}^{B} t_i \frac{N.B_{\max}}{B} \cdot \frac{1}{N.B_{\max}}$$

$$= \sum_{i=1}^{N} t_i \frac{B_i}{B}$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.14

With Lahiri sampling,

$$V_{lah}\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

With Lahiri sampling,

$$V\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n^2}\sum_{i=1}^{n} V(t_i)$$

$$= \frac{B^2}{n} V(t_i)$$

Now

$$V(t_i) = E(t_i^2) - (E(t_i))^2$$

$$= \left[\sum_{i=1}^{N} t_i^2 \frac{N.B_{max}}{B} \frac{B_i}{N.B_{max}} - \left(\sum_{i=1}^{N} t_i \frac{N.B_{max}}{B} . \frac{B_i}{N.B_{max}}\right)^2\right]$$

$$= \frac{1}{B}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Therefore

$$V\left(\frac{B}{n}\sum_{i=1}^{n} t_i\right) = \frac{B^2}{n}\left[\frac{1}{B}\left\{\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right\}\right]$$

$$= \frac{B}{n}\left[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.7.2 Lahiri sampling with the AON error assignment

The mean and variance of the point estimator with Lahiri sampling and the AON error assignment are derived in theorems 5.15 and 5.16. In any one trial, with the AON error assignment, u_i is a Bernoulli variable with the following probabilities

$$u_i = 1 \text{ with probability } \frac{E_i}{NB_{max}}$$

$$= 0 \text{ with probability } 1 - \frac{E_i}{NB_{max}}$$

Theorem 5.15

With Lahiri sampling of monetary units

$$\frac{B}{n}\sum_{i=1}^{n} u_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E(\frac{B}{n}\sum_{i=1}^{n} u_i) = B.E(u_i)$$

$$= B\sum_{i=1}^{N} \frac{E_i}{NB_{max}} \cdot \frac{NB_{max}}{B}$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.16

With Lahiri sampling of monetary units

$$V_{lah}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Proof:

With Lahiri sampling of monetary units,

$$V\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B^2}{n^2}\sum_{i=1}^{n} V(u_i)$$

$$= \frac{B^2}{n} V(u_i)$$

Now

$$V(u_i) = E(u_i^2) - (E(u_i))^2$$

$$= \left[\sum_{i=1}^{N} \frac{E_i}{N.B_{max}} \frac{N.B_{max}}{B} - \left(\sum_{i=1}^{N} \frac{E_i}{N.B_{max}} \frac{N.B_{max}}{B}\right)^2\right]$$

$$= \frac{1}{B}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Therefore

$$V_{lah}\left(\frac{B}{n}\sum_{i=1}^{n} u_i\right) = \frac{B^2}{n}\left[\frac{1}{B}\left\{\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right\}\right]$$

$$= \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

5.8 Sieve Sampling

With sieve sampling, and the taint error assignment the estimator used for the total error amount is

$$\frac{B}{n}\sum_{i=1}^{n_0} t_i$$

and with sieve sampling and the AON error assignment, the estimator used for the total error amount is

$$\frac{B}{n}\sum_{i=1}^{n_0} u_i$$

The mean and variance of the point estimator under sieve sampling for each assignment method are derived in 5.8.1 and 5.8.2. Similar results have been obtained previously by Wurst, Neter and Godfrey (1989b).

5.8.1 Sieve Sampling with the Taint Error Assignment

In sieve sampling with the taint error assignment, each line item is considered for selection and is selected if a random number generated from [1, k] is less that or equal to the value of the line item. The selection process may be modelled by a Bernoulli variable k_i, where $k_i = 1$ if the line item i is selected and k_i =0 if the line item i is not selected, $1<=i<=N$.

Then

$$k_i = 1 \text{ with probability } \frac{nB_i}{B}$$

$$= 0 \text{ with probability } 1 - \frac{nB_i}{B}$$

Therefore

$$E(k_i) = \frac{nB_i}{B}$$

$$V(k_i) = \frac{nB_i}{B}\left(1 - \frac{nB_i}{B}\right)$$

and

$$\sum_{i=1}^{n_0} t_i = \sum_{i=1}^{N} t_i k_i$$

where n_0 is the achieved sample size.

These properties are used in theorems 5.17 and 5.18 to derive the mean and variance of the point estimator under sieve sampling with the taint error assignment.

Theorem 5.17

With sieve sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n_0} t_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$E\left(\frac{B}{n}\sum_{i=1}^{n_0} t_i\right) = E\left(\frac{B}{n}\sum_{i=1}^{N} t_i\, k_i\right)$$

$$= \frac{B}{n}\sum_{i=1}^{N} t_i . E\,(k_i)$$

$$= \frac{B}{n}\sum_{i=1}^{N} t_i \frac{nB_i}{B}$$

$$= \sum_{i=1}^{N} E_i$$

Theorem 5.18

With sieve sampling, the variance of the point estimator is

$$V_{sie}\left(\frac{B}{n}\sum_{i=1}^{n_0} t_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N}\frac{E_i^{\,2}}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^{\,2}\right]$$

Proof:

$$V_{sie}\left(\frac{B}{n_0}\sum_{i=1}^{n} t_i\right) = V\left(\frac{B}{n}\sum_{i=1}^{N} t_i k_i\right)$$

$$= \frac{B^2}{n^2}\sum_{i=1}^{N} t_i^2\ V(k_i)$$

$$= \frac{B^2}{n^2}\left[\sum_{i=1}^{N} t_i^2\left(\frac{nB_i}{B}\right)\left(1 - \frac{nB_i}{B}\right)\right]$$

$$= \frac{B^2}{n^2}\left[\sum_{i=1}^{N} t_i^2\frac{nB_i}{B} - \sum_{i=1}^{N} t_i^2\left(\frac{nB_i}{B}\right)^2\right]$$

$$= \frac{B}{n}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right]$$

5.8.2 Sieve Sampling with the AON Error Assignment

In sieve sampling, with the AON error assignment, a random number from [1,k] is generated for each line item. A monetary unit is selected from line item i provided the random number is less than or equal to B_i. With the AON assignment method, the selected monetary unit is given the value one if the random number is less than or equal to E_i. Otherwise it is given the value 0. This may be represented by a Bernoulli variable x_i as follows

$$x_i = 1 \text{ with probability } \frac{nE_i}{B}$$

$$= 0 \text{ with probability } 1 - \frac{nE_i}{B}$$

Then

$$\sum_{i=1}^{n_0} u_i = \sum_{i=1}^{N} x_i$$

This is used in theorems 5.19 and 5.20 to derive the mean and variance of the point estimator under sieve sampling with the AON assignment method.

Theorem 5.19

With sieve sampling of monetary units,

$$\frac{B}{n}\sum_{i=1}^{n_0} u_i$$

is an unbiased estimate of the total error amount T.

Proof:

$$\begin{aligned}(\frac{B}{n}\sum_{i=1}^{n_0} u_i) &= E(\frac{B}{n}\sum_{i=1}^{N} x_i)\\ &= \frac{B}{n}\sum_{i=1}^{N} E(x_i)\\ &= \frac{B}{n}\sum_{i=1}^{N} \frac{nE_i}{B}\\ &= \sum_{i=1}^{N} E_i\end{aligned}$$

Theorem 5.20

With sieve sampling, the variance of the point estimator is

$$V(\frac{B}{n}\sum_{i=1}^{n_0} u_i) = \frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

Proof:

$$V_{sie}(\frac{B}{n}\sum_{i=1}^{n_0} u_i) = V_{sie}(\frac{B}{n}\sum_{i=1}^{N} x_i)$$

$$= \frac{B^2}{n^2}\sum_{i=1}^{N} V(x_i)$$

$$= \frac{B^2}{n^2}\sum_{i=1}^{N} \frac{nE_i}{B}(1 - \frac{nE_i}{B})$$

$$= \frac{B^2}{n^2}\sum_{i=1}^{N} \frac{nE_i}{B} - \frac{n^2}{B^2}\sum_{i=1}^{N} E_i^2]$$

$$= \frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

5.9 Stabilised Sieve Sampling

In stabilised sieve sampling, the sample is selected in two stages as follows:

Stage 1: An initial sieve sample of size n_0 is selected.

Stage 2: The initial sample is increased or decreased to ensure that a the final sample size is equal to the nominal sample size n. The initial sample is reduced or augmented when necessary as follows;

(i) Sample Reduction: When $n_0 > n$, a simple random sample of size n is selected from the initial sample.

(ii) Sample Augmentation: When $n_0 < n$, a further $(n-n_0)$ monetary units are selected by means of Lahiri sampling.

With stabilised sieve sampling and the taint error assignment the estimator used for the total error amount is

$$\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i$$

And with stabilised sieve sampling and the AON error assignment the estimator used for the total error amount is

$$\frac{1}{2n-n_0}\sum_{i=1}^{n} u_i$$

The mean and variance of the point estimator of total error amount for each error assignment method are derived in 5.9.1 and 5.9.2. As far as can be ascertained, these results are derived here for the first time.

5.9.1 Stabilised Sieve Sampling with the Taint Error Assignment

The mean and variance of the point estimator with stabilised sieve sampling and the taint error assignment are derived in theorems 5.21 and 5.22.

Theorem 5.21

With stabilised sieve sampling

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = \sum_{i=1}^{N} E_i$$

Proof:

Case 1. Sample reduction ($n_0 > n$)

When the initial sample n_0 is greater than the nominal sample size n, the sample is reduced from n_0 to n. Therefore

$$\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i = \frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i - \frac{B}{2n-n_0}\sum_{i=(n+1)}^{n_0} t_i$$

assuming for convenience that the $(n+1)^{th}$, $(n+2)^{th}$, $\ldots n_0^{th}$ units are randomly rejected.

Hence

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i) = E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i - \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i)$$

$$= E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i) - E(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i)$$

Now, the n_0 are selected using sieve sampling with a sieve size of B/n. Therefore, from theorem 5.17

$$E(\frac{B}{n}\sum_{i=1}^{n_0} t_i) = \sum_{i=1}^{N} E_i$$

Hence

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i) = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i$$

and the set of $(n_0 - n)$ monetary units is a simple random sample of monetary units. Thus, it follows from theorem 5.1

$$E(\frac{B}{n_0-n}\sum_{i=n+1}^{n_0} t_i) = \sum_{i=1}^{N} E_i$$

Therefore

$$E\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i\right) = \frac{n_0-n}{2n-n_0}\sum_{i=1}^{N} E_i$$

So

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i\right) - E\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i\right) = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i - \frac{n_0-n}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \frac{2n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \sum_{i=1}^{N} E_i$$

Case 2. Sample augmentation ($n_0 < n$)

When the initial sample n_0 is less than the nominal sample size n, the sample is augmented from n_0 to n by selecting ($n - n_0$) units by means of Lahiri sampling.

Therefore

$$\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i = \frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i + \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i$$

And

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i) = E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i + \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i)$$

$$= E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i) + E(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i)$$

Now, the n_0 are selected using sieve sampling, with a sieve size of B/n. Therefore, from theorem 5.17

$$E(\frac{B}{n}\sum_{i=1}^{n_0} t_i) = \sum_{i=1}^{N} E_i$$

So

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i$$

and the last $(n - n_0)$ units are selected using Lahiri sampling. Therefore, it follows from theorem 5.13

$$E(\frac{B}{n-n_0}\sum_{i=n_0+1}^{n} t_i) = \sum_{i=1}^{N} E_i$$

So

$$E\left(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i\right) = \frac{n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

Therefore

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i\right) + E\left(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i\right)$$

$$= \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i + \frac{n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \frac{2n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \sum_{i=1}^{N} E_i$$

This completes the proof.

Theorem 5.22

With stabilised sieve sampling of monetary units,

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right] + \frac{B|n-n_0|}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right.$$

Case 1. Sample reduction ($n_0 > n$)

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i\right) - \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i)$$

$$= V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i\right) + V\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i\right)$$

$$-2Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i, \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i\right)$$

Now, it can be shown that

$$Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i, \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i\right) = 0$$

(see Cochran, 1977, p48).

Since the first n_0 units are selected using sieve sampling, then from theorem 5.18

$$V(\frac{B}{n}\sum_{i=1}^{n_0} t_i) = \frac{B}{n}[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

Therefore

$$V(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i) = \frac{n^2}{(2n-n_0)^2} V(\frac{B}{n}\sum_{i=1}^{n_0} t_i)$$

$$= \frac{nB}{(2n-n_0)^2}[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

And since the set of $(n_0 - n)$ units is a simple random sample of monetary units, it follows from theorem 5.2

$$V(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} t_i) = \frac{(n_0-n)^2}{(2n-n_0)^2} V(\frac{B}{n_0-n}\sum_{i=n+1}^{n_0} t_i)$$

$$= \frac{(n_0-n)^2}{(2n-n_0)^2}\frac{B}{n_0-n}[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

$$= \frac{B(n_0-n)}{(2n-n_0)^2}[\sum_{i=1}^{N} \frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

Therefore

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N}E_i^2\right] + \frac{B(n_0-n)}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N}E_i\right)^2\right]$$

$$= \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N}E_i^2\right] + \frac{B|n-n_0|}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}\left(\sum_{i=1}^{N}E_i\right)^2\right]$$

Case 2. Sample augmentation ($n_0 < n$)

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i + \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i\right)$$

$$= V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i\right) + V\left(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i\right)$$

$$+ 2Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i, \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i\right)$$

Now, the covariance term is zero because independent selections are made to augment the sample and sampling is done with replacement.

Also, since the first n_0 monetary units are selected using sieve sampling, it follows from theorem 5.18 that

$$V(\frac{B}{n}\sum_{i=1}^{n_0} t_i) = \frac{B}{n}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

Therefore

$$V(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} t_i) = \frac{n^2}{(2n-n_0)^2} V(\frac{B}{n}\sum_{i=1}^{n_0} t_i)$$

$$= \frac{nB}{(2n-n_0)^2}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

And since the next $(n- n_0)$ units are selected using Lahiri sampling, then, from theorem 5.16

$$V(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} t_i) = \frac{(n-n_0)^2}{2n-n_0)^2} V(\frac{B}{n-n_0}\sum_{i=n_0+1}^{n} t_i)$$

$$= \frac{(n-n_0)^2}{(2n-n_0)^2}\frac{B}{n-n_0}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

$$= \frac{B(n-n_0)}{(2n-n_0)^2}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

Therefore

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i\right) = \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N}E_i^2\right] + \frac{B(n-n_0)}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} + \frac{1}{B}\left(\sum_{i=1}^{N}E_i\right)^2\right]$$

$$= \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N}E_i^2\right] + \frac{B|n-n_0|}{(2n-n_0)^2}\left[\sum_{i=1}^{N}\frac{E_i^2}{B_i} + \frac{1}{B}\left(\sum_{i=1}^{N}E_i\right)^2\right.$$

This completes the proof.

5.9.1 Stabilised sieve sampling with the AON error assignment

The mean and variance of the point estimator with sieve sampling and the AON error assignment are derived 5.23 and 5.24.

Theorem 5.23

With stabilised sieve sampling

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = \sum_{i=1}^{N} E_i$$

Proof:

Case 1. Sample reduction ($n_0 > n$)

When the initial sample n_0 is greater than the nominal sample size n, the sample is reduced from n_0 to n. Therefore

$$\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i = \frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i - \frac{B}{2n-n_0}\sum_{i=(n+1)}^{n_0} u_i$$

Hence

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i) = E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i - \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i)$$

$$= E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i) - E(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i)$$

Now, the n_0 are selected using sieve sampling with a sieve size of B/n. Therefore, from theorem 5.19

$$E(\frac{B}{n}\sum_{i=1}^{n_0} u_i) = \sum_{i=1}^{N} E_i$$

Hence

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i$$

and the set of (n_0-n) monetary units is a simple random sample of monetary units. Thus it follows from theorem 5.3

$$E\left(\frac{B}{n_0-n}\sum_{i=n+1}^{n_0} u_i\right) = \sum_{i=1}^{N} E_i$$

Therefore

$$E\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right) = \frac{n_0-n}{2n-n_0}\sum_{i=1}^{N} E_i$$

So

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) - E\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right) = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i - \frac{n_0-n}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \frac{2n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

$$= \sum_{i=1}^{N} E_i$$

Case 2. Sample augmentation ($n_0 < n$)

When the initial sample n_0 is less than the nominal sample size n, the sample is augmented from n_0 to n by selecting $(n - n_0)$ units by means of Lahiri sampling. Therefore

$$\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i = \frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i + \frac{B}{2n-n_0}\sum_{i=(n_0+1)}^{n} u_i$$

And

$$E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i + \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i\right)$$

$$= E\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) + E\left(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i\right)$$

Now, the n_0 are selected using sieve sampling, with a sieve size of B/n. Therefore, from theorem 5.19

$$E\left(\frac{B}{n}\sum_{i=1}^{n_0} u_i\right) = \sum_{i=1}^{N} E_i$$

So

$$E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i = \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i$$

and the last $(n - n_0)$ units are selected using Lahiri sampling. Therefore, it follows from theorem 5.15

$$E(\frac{B}{n-n_0}\sum_{i=n_0+1}^{n} u_i) = \sum_{i=1}^{N} E_i$$

So

$$E(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i) = \frac{n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i$$

Therefore

$$\begin{aligned} E(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i) &= E(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i) + E(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i) \\ &= \frac{n}{2n-n_0}\sum_{i=1}^{N} E_i + \frac{n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i \\ &= \frac{2n-n_0}{2n-n_0}\sum_{i=1}^{N} E_i \\ &= \sum_{i=1}^{N} E_i \end{aligned}$$

This completes the proof.

Theorem 5.24

With stabilised sieve sampling of monetary units,

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right] + \frac{B|n-n_0|}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Case 1. Sample reduction ($n_0 > n$)

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) - \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\Big)$$

$$= V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) + V\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right)$$

$$-2Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i, \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right)$$

Now, it can be shown that

$$Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i, \frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right) = 0$$

(see Cochran, 1977, p48).

Since the first n_0 units are selected using sieve sampling, then from theorem 5.20

$$V\left(\frac{B}{n}\sum_{i=1}^{n_0} u_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right]$$

Therefore

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) = \frac{n^2}{(2n-n_0)^2} V\left(\frac{B}{n}\sum_{i=1}^{n_0} u_i\right)$$

$$= \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right]$$

And since the set of $(n_0 - n)$ units is a simple random sample of monetary units, it follows from theorem 5.4

$$V\left(\frac{B}{2n-n_0}\sum_{i=n+1}^{n_0} u_i\right) = \frac{(n_0-n)^2}{(2n-n_0)^2} V\left(\frac{B}{n_0-n}\sum_{i=n+1}^{n_0} u_i\right)$$

$$= \frac{(n_0-n)^2}{(2n-n_0)^2}\frac{B}{n_0-n}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

$$= \frac{B(n_0-n)}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Therefore

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = \frac{nB}{2n-n_0}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right] + \frac{B(n_0-n)}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

$$= \frac{nB}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right] + \frac{B|n-n_0|}{(2n-n_0)^2}\left[\sum_{i=1}^{N} E_i - \frac{1}{B}\left(\sum_{i=1}^{N} E_i\right)^2\right]$$

Case 2. Sample augmentation ($n_0 < n$)

$$V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i\right) = V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i + \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i\right)$$

$$= V\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i\right) + V\left(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i\right)$$

$$+ 2Cov\left(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i, \frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i\right)$$

Now, the covariance term is zero because independent selections are made to augment the sample and sampling is done with replacement.

Also, since the first n_0 monetary units are selected using sieve sampling, it follows from theorem 5.20 that

$$V\left(\frac{B}{n}\sum_{i=1}^{n_0} u_i\right) = \frac{B}{n}\left[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2\right]$$

Therefore

$$V(\frac{B}{2n-n_0}\sum_{i=1}^{n_0} u_i) = \frac{n^2}{(2n-n_0)^2} V(\frac{B}{n}\sum_{i=1}^{n_0} u_i)$$

$$= \frac{nB}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$$

And since the next $(n- n_0)$ units are selected using Lahiri sampling, then, from theorem 5.18

$$V(\frac{B}{2n-n_0}\sum_{i=n_0+1}^{n} u_i) = \frac{(n-n_0)^2}{2n-n_0)^2} V(\frac{B}{n-n_0}\sum_{i=n_0+1}^{n} u_i)$$

$$= \frac{(n-n_0)^2}{(2n-n_0)^2}\frac{B}{n-n_0} [\sum_{i=1}^{N} E_i - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

$$= \frac{B(n-n_0)}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

Therefore

$$V(\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i) = \frac{nB}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2] + \frac{B(n-n_0)}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i + \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

$$= \frac{nB}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2] + \frac{B|n-n_0|}{(2n-n_0)^2} [\sum_{i=1}^{N} E_i + \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$$

This completes the proof.

5.10 Comparison of the Sampling Methods

'The study of any sampling technique is incomplete unless some comparisons are made with other sampling methods' (Madow and Madow, 1944). In this section, systematic, cell, Lahiri, sieve and stabilised sieve sampling of monetary units are compared with simple random sampling of monetary units. The design effect (deff), a measure for comparing a sample design with simple random sampling, is used as a basis of comparison. The deff of a particular sample design is defined as the ratio of the variance of an estimate with the sample design to the variance of the estimate with simple random sample of the same number of elements (Kish, 1965, p258). If deff is less than one, the sampling method is more precise than simple random sampling. If deff is greater than one, the precision of the sampling method under consideration is less than simple random sampling and if deff is equal to one, the precision of the two sampling methods are the same. A sampling method with a deff of less than one is said to be more efficient than simple random sampling and this implies that the same level of confidence is obtained with a smaller sample size using this method of selection than with simple random sampling.

The unbiased point estimator for each assignment method and for each sample design is given in Table 5.3.

Table 5.3 Unbiased Point Estimator of the Total Error Amount T

Sampling Method	Error Assignment	
	Taint	AON
SRS, Systematic, Cell, Lahiri	$\frac{B}{n}\sum_{i=1}^{n} t_i$	$\frac{B}{n}\sum_{i=1}^{n} u_i$
Sieve Sampling	$\frac{B}{n}\sum_{i=1}^{n_0} t_i$	$\frac{B}{n}\sum_{i=1}^{n_0} u_i$
Stabilised Sieve Sampling	$\frac{B}{2n-n_0}\sum_{i=1}^{n} t_i$	$\frac{B}{2n-n_0}\sum_{i=1}^{n} u_i$

From Table 5.3, the following observations are made

(i) The point estimator is unbiased for simple random, systematic, cell, and Lahiri sampling of monetary units.

(ii) With sieve sampling, the point estimator is obtained by summing over the n_0 sample values and dividing by the nominal sample size **n**.

(iii) In modified sieve sampling, the estimator is adjusted to eliminate bias. As can be seen from table 5.3, it is adjusted downwards when the initial sample size is increased and adjusted upwards when the initial sample size is reduced.

The variances of the unbiased point estimator with the two error assignments for each sample design are summarised in Tables 5.4 and 5.5. These variances are used to investigate the precision of systematic, cell, Lahiri, sieve and stabilised sieve sampling compared to the precision of simple random sampling of monetary units.

Table 5.4 Variances of the Point Estimator with the Taint Error Assignment

Sampling method	Variance
Simple Random	$\frac{B}{n}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N}E_i)^2]$
Systematic	$\frac{B}{n}[\sum_{r=1}^{k}T^2_{r.}(taint) - \frac{n}{B}(\sum_{i=1}^{N}E_i)^2]$
Cell	$\frac{B}{n}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{n}T_i^2(taint)]$
Lahiri	$\frac{B}{n}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{(\sum_{i=1}^{N}E_i)^2}{B}]$
Sieve	$\frac{B}{n}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}(\sum_{i=1}^{N}E_i^2)]$
Stabilised Sieve	$\frac{nB}{(2n-n_0)^2}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{n}{B}\sum_{i=1}^{N}E_i^2] + \frac{B\lvert n-n_0\rvert}{(2n-n_0)^2}[\sum_{i=1}^{N}\frac{E_i^2}{B_i} - \frac{1}{B}(\sum_{i=1}^{N}E_i)^2]$

Table 5.5 Variances of the Point Estimator with the AON Error Assignment

Sampling Method	Variance
Simple Random	$\frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$
Systematic	$\frac{B}{n}[\sum_{r=1}^{k} T_{r.}^2 - \frac{n}{B}(\sum_{i=1}^{N} E_i)^2$
Cell	$\frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{n} T_i^2]$
Lahiri	$\frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$
Sieve	$\frac{B}{n}[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2]$
Stabilised Sieve	$\frac{nB}{(2n-n_0)^2}[\sum_{i=1}^{N} E_i - \frac{n}{B}\sum_{i=1}^{N} E_i^2] + \frac{B\lvert n-n_0\rvert}{(2n-n_0)^2}[\sum_{i=1}^{N} E_i - \frac{1}{B}(\sum_{i=1}^{N} E_i)^2]$

5.10.1 The Design Effect of Systematic Sampling.

From Table 5.4, it can be seen that the magnitude of the variance of the point estimator, with the taint error assignment, under systematic sampling depends on the magnitude of

$$\sum_{r=1}^{k} T_{r.}^{2}\,(taint)$$

and Table 5.5 shows that the magnitude of the variance of the point estimator, with the AON error assignment, depends on the magnitude of

$$\sum_{r=1}^{k} T_{r.}^{2}$$

In both cases, for a fixed error amount in the population, these summations are likely to be small if the errors are distributed uniformly among the **k** systematic samples, and they are likely to be large if the total error amount differs greatly in each systematic sample. This concept applies to cluster sampling in general, where the ideal situation is that each cluster is a mirror image of the population. Equivalently, systematic sampling is precise when monetary units within the same systematic sample are heterogeneous and imprecise when they are homogeneous. This is intuitive because one sample contains all the information about the population parameter when the systematic samples are similar to each other.

Cochran (1977, 207-208) derived the following results for the variance of systematic samples compared to simple random sampling:

(i) Variance estimates based on systematic samples are less than variance estimates based on simple random samples provided the variance within the systematic sample is larger than the population variance as a whole, i.e. systematic sampling is more precise when units within the same systematic sample are heterogeneous and imprecise when they are homogeneous.

(ii) Variance estimates based on systematic sampling are greater than variance estimates based on simple random sampling if the correlation between elements in the same systematic sample is positive.

In accounting populations, identical taints and AONs are clustered in the associated line items and hence in adjacent monetary units. This leads to systematic samples being similar in terms of the total error amount. Also, because similar monetary units are clustered, systematic sampling excludes some extreme sample combinations from the sampling distribution of the estimator. This is the case even if the line items are randomly ordered since identical taints and AON values belonging to the same line item are still adjacent.

The extent of the decrease in variation for systematic sampling will depend on the population characteristics. It is likely to be substantial when the line items and the error amounts in the line items are large because a large number of identical taints and AONs will then be adjacent. Therefore, substantial gains in precision of systematic sampling over simple random sampling would be expected when sampling from the audit populations generated from Population 2 which consists of relatively large line items and relatively large error amounts in the line items in error.

In summary, for accounting populations in general, systematic sampling should have a design effect of less than or equal to one. The design effect will be substantially less than one if the population contains relatively large line items and relatively large error amounts in the line items in error.

5.10.2 The Design Effect of Cell Sampling

Cell sampling can be looked upon as a special case of stratified sampling where the population is divided into **n** strata and one unit is selected from each stratum. As stratification nearly always results in a smaller variance for the estimated mean or total than a simple random sample of the same size (Cochran, 1977, p99), the point estimator under cell sampling should have a variance less than or equal to the variance with simple random sampling, i.e. a deff less than or equal to one. Cochran (1977, p100) showed that the improvement in precision depends on the differences between the stratum (cell) means and the greatest gains are achieved when the sum of the squares of the deviations of the stratum means from the overall means is maximised. This is a general desirable property for stratification, where the ideal situation is that strata (cells) are internally homogeneous and externally heterogeneous (Cochran, 1977, p101). In accounting populations, taints associated with the same line item are equal and adjacent and hence are likely to be in the same stratum (cell). Therefore strata or cells will tend to be internally homogeneous and externally heterogeneous. Thus, it is to be expected that cell sampling will result in a more precise point estimator than simple random sampling, i.e. the deff of cell sampling is likely to be less than one. The greatest gains in precision should occur in accounting populations, like Population 2, which have relatively large line items.

From Table 5.4 it follows that, with the taint error assignment, cell sampling will lead to equal or greater precision than simple random sampling, provided

$$n\sum_{i=1}^{n} T_i^2(taint) \geq (\sum_{i=1}^{N} E_i)^2$$

Now

$$\sum_{i=1}^{N} E_i = \sum_{i=1}^{n}\sum_{r=1}^{k} \frac{E_{r+(i-1)k}}{B_{r+(i-1)k}}$$

Therefore, the point estimator with cell sampling and the taint error assignment will have equal or better precision than simple random sampling provided

$$n\sum_{i=1}^{n}\left(\sum_{r=1}^{k}\frac{E_{r+(i-1)k}}{B_{r+(i-1)k}}\right)^2 \geq \left[\sum_{i=1}^{n}\left(\sum_{r=1}^{k}\frac{E_{r+(i-1)k}}{B_{r+(i-1)k}}\right)\right]^2$$

i.e. provided

$$n\sum_{i=1}^{n}\left(\sum_{r=1}^{k} t_{r+(i-1)k}\right)^2 \geq \left[\sum_{i=1}^{n}\sum_{r=1}^{k} t_{r+(i-1)k}\right]^2$$

The Cauchy-Schwarz inequality guarantees that

$$n\sum_{i=1}^{n}\left(\sum_{r=1}^{k} t_{r+(i-1)k}\right)^2 \geq \left[\sum_{i=1}^{n}\sum_{r=1}^{k} t_{r+(i-1)k}\right]^2$$

Therefore, the point estimator with cell sampling and the taint error assignment has equal or greater precision than the equivalent point estimator with simple random sampling. The gains in precision are obviously greater for larger sample sizes.

Also, the gains in precision of cell sampling over simple random sampling will depend on the pattern of errors in the population. If the errors are distributed uniformly over the cells, i.e. if the summations

$$\sum_{r=1}^{k} t_{r+(i-1)k}$$

are similar for each i, then

$$\sum_{i=1}^{n} (\sum_{r=1}^{k} t_{r+(i-1)k})^2$$

will be small.

Conversely, if the errors are not distributed uniformly over the cells, then

$$\sum_{i=1}^{n} (\sum_{r=1}^{k} t_{r+(i-1)k})^2$$

will tend to be large, thus leading to substantial gains in precision over simple random sampling.

Similarly, with the AON error assignment (Table 5.5), cell sampling will lead to a point estimator with equal or greater precision than simple random sampling, provided

$$n\sum_{i=1}^{n} T_i^2 \geq (\sum_{i=1}^{N} E_i)^2$$

Now,

$$\sum_{i=1}^{N} E_i = \sum_{i=1}^{n} \sum_{r=1}^{k} u_{r+(i-1)k}$$

Therefore, cell sampling leads to a point estimator with equal or greater precision provided

$$n\sum_{i=1}^{n} (\sum_{r=1}^{k} (u_{r+(i-1)k})^2 \geq [\sum_{i=1}^{n} \sum_{r=1}^{k} u_{r+(i-1)k}]^2$$

This condition follows immediately from the Cauchy-Schwarz inequality. The greatest gains in precision occur when the errors are not distributed uniformly among the cells.
i.e. when the

$$\sum_{r=1}^{k} u_{r+(i-1)k}$$

are dissimilar over the i.

Equivalently, for both error assignments, the gains in precision of cell sampling over simple random sampling are greatest when the cells are internally homogeneous and externally heterogeneous. This is a desirable condition for stratified sampling in general (Cochran, 1977, pp 89-90).

In accounting populations, where the taints and AONs associated with the same line item are equal and are likely to be in the same stratum, cell sampling would be expected to be more precise than simple random sampling of monetary units. The greatest gains in precision should occur in accounting populations, like Population 2, which have relatively large line items.

In summary, the design effect of cell sampling is less than or equal to one. The value of deff will be small in populations with large line items and when the sample size is large.

5.10.3 The Design Effect of Lahiri Sampling.

Clearly, from Tables 5.4 and 5.5, the variances of the point estimator with Lahiri sampling is identical to the variance of the point estimator with simple random sampling for both error assignments. Hence, Lahiri sampling has a design effect of one or equivalently Lahiri sampling has the same precision as simple random sampling. In this study, Lahiri sampling is proposed as an alternative to simple random sampling and therefore from the point of view of the point estimator at least, the decision to use Lahiri sampling instead of simple random sampling will be made on non-statistical grounds.

5.10.4 The Design Effect of Sieve Sampling

From Tables 5.4 and 5.5, it can be seen that, with either error assignment method, sieve sampling will have a deff of less than or equal to one provided

$$n\sum_{i=1}^{N} E_i^2 \geq \left(\sum_{i=1}^{N} E_i\right)^2$$

i.e. provided

$$n \geq \frac{\left(\sum_{i=1}^{N} E_i\right)^2}{\sum_{i=1}^{N} E_i^2}$$

The value of **n** for which sieve sampling is more precise than simple random sampling, depends on the pattern of errors in the population. Obviously, if the errors are distributed uniformly throughout the line items, the denominator will tend to be small and hence, the sample size at which the design effect is less than or equal to one, will be large. Conversely, if the errors are not distributed uniformly throughout the line items, the denominator will tend to be large and hence, the sample size at which the design effect becomes less than one will be small.

The minimum sample sizes for which the design effect for sieve sampling is less than one in each set of populations are given in Tables 5.6 and 5.7 respectively.

Table 5.6 Minimum Sample Size for which the Design Effect of Sieve Sampling is less than 1 for Audit Populations generated from Population 1

Error Rate	Taint 1	Taint 2	Taint 3
Error Rate 1	16	17	22
Error Rate 2	33	35	47
Error Rate 3	47	50	67
Error Rate 4	93	99	132
Error Rate 5	140	150	201

Table 5.7 Minimum Sample Size for which the Design Effect for Sieve Sampling is less than 1 for Audit Populations generated from Population 2

Error Rate	Taint 1	Taint 2	Taint 3
Error Rate 1	4	5	5
Error Rate 2	7	7	7
Error Rate 3	11	12	12
Error Rate 4	18	19	19
Error Rate 5	31	33	34

From these tables, the following observations are made:

(i) In both sets of populations, the sample size for which the design effect of sieve sampling is less than one is small for populations with small taints and low error rates.

(ii) In audit populations generated from Population 2 where the line items are large, the design effect of sieve sampling is less than one for relatively small samples for all error rates and for all taint sizes. The maximum sample size to ensure a deff of less than one is 34 in audit populations generated from Population 2. This contrasts with the audit populations generated from Population 1, where the sample sizes for which the design effect of sieve sampling becomes less than one is quite large in populations with large line item error rates.

These results concur with the results obtained by Wurst, Neter and Godfrey (1989a) who found that

(i) For accounting populations where the total error amount is small, the sample size at which sieve sampling becomes more efficient than simple random sampling tends to be small.

(ii) Errors concentrated in large line items and large proportional errors per line item tend to favour sieve sampling in terms of the precision of the point estimator.

Further insights into the gains in precision due to sieve sampling can be obtained by calculating the design effects for each sample size drawn from each audit population. These are given in Tables 5.8 and 5.9 for each error assignment method respectively.

The following observations are made from these tables.

(i) The design effect of sieve sampling is similar for both error assignments in all cases.

(ii) Although there is a tendency in Population 1 for the audit populations with low line item error rates and large sample sizes to have a deff of less than one and the populations with high line item error rates have a deff greater than one, the deff is not substantially different from one in any of the audit populations generated from Population 1 for any sample size.

(iii) The design effect of sieve sampling is substantially less than one for all sample sizes in the audit populations created from Population 2. The greatest gains in precision for sieve sampling over simple random sampling occur in the large sample sizes. Populations with low line item error rates have more precise point estimators under sieve sampling than populations with high line item error rates.

Table 5.8 Design Effect of Sieve Sampling with the Taint (AON) Error Assignment for Audit Populations generated from Population 1

Sample Size	n = 30	n = 60	n = 100
Error Rate 1			
Taint 1	0.99 (0.99)	0.97 (0.97)	0.95 (0.95)
Taint 2	0.99 (0.99)	0.97 (0.98)	0.95 (0.95)
Taint 3	1.00 (1.00)	0.98 (0.98)	0.95 (0.96)
Error Rate 2			
Taint 1	1.00 (1.00)	0.98 (0.98)	0.96 (0.96)
Taint 2	1.00 (1.00)	0.98 (0.99)	0.96 (0.96)
Taint 3	1.01 (1.01)	0.99 (0.99)	0.96 (0.97)
Error Rate 3			
Taint 1	1.01 (1.01)	0.99 (0.99)	0.96 (0.97)
Taint 2	1.01 (1.01)	0.99 (0.99)	0.97 (0.97)
Taint 3	1.03 (1.02)	1.01 (1.00)	0.98 (0.98)
Error Rate 4			
Taint 1	1.04 (1.04)	1.02 (1.02)	1.00 (1.00)
Taint 2	1.05 (1.04)	1.03 (1.02)	1.00 (1.00)
Taint 3	1.07 (1.06)	1.05 (1.04)	1.02 (1.02)
Error Rate 5			
Taint 1	1.06 (1.07)	1.05 (1.05)	1.03 (1.03)
Taint 2	1.08 (1.08)	1.06 (1.06)	1.03 (1.03)
Taint 3	1.13 (1.10)	1.11 (1.08)	1.08 (1.06)

Table 5.9 Design Effect of Sieve Sampling with the Taint (AON) Assignment for Audit Populations generated from Population 2

Sample Size	n = 30	n = 60	n = 100
Error Rate 1			
Taint 1	0.85 (0.85)	0.69 (0.69)	0.46 (0.46)
Taint 2	0.86 (0.86)	0.69 (0.70)	0.47 (0.48)
Taint 3	0.86 (0.86)	0.70 (0.70)	0.48 (0.49)
Error Rate 2			
Taint 1	0.89 (0.89)	0.75 (0.75)	0.56 (0.57)
Taint 2	0.89 (0.89)	0.75 (0.76)	0.57 (0.58)
Taint 3	0.89 (0.90)	0.76 (0.76)	0.58 (0.58)
Error Rate 3			
Taint 1	0.93 (0.93)	0.81 (0.81)	0.66 (0.67)
Taint 2	0.93 (0.93)	0.82 (0.82)	0.66 (0.67)
Taint 3	0.93 (0.93)	0.82 (0.82)	0.67 (0.67)
Error Rate 4			
Taint 1	0.94 (0.94)	0.81 (0.80)	0.62 (0.62)
Taint 2	0.95 (0.95)	0.81 (0.81)	0.62 (0.62)
Taint 3	0.95 (0.95)	0.81 (0.81)	0.63 (0.63)
Error Rate 5			
Taint 1	1.00 (1.00)	0.88 (0.88)	0.71 (0.71)
Taint 2	1.01 (1.01)	0.88 (0.89)	0.72 (0.71)
Taint 3	1.02 (1.02)	0.89 (0.89)	0.72 (0.71)

5.10.5 The Design Effect of Stabilised Sieve Sampling.

No general statement can be made on the design effect of stabilised sieve sampling. The variance of the point estimator with stabilised sieve sampling consists of two variance components, one due to the initial sieve sample selection and the other due to the reduction or the augmentation process. The component due to sieve sampling has the greatest weight so one would expect that the design effect of stabilised sieve sampling would be somewhat similar to the design effect of sieve sampling.

Obviously, the point estimator with stabilised sieve sampling will have a precision less than that with simple random sampling if the initial gains in precision due to the sieve sample component are larger than the extra variance component caused by the reduction or the augmentation process. If the improvement in the precision of sieve sampling relative to simple random sampling is substantial, then the variance of stabilised sieve sampling will be less than the variance of simple random sampling. For example, from Table 5.9, it can be seen that the gains in precision due to sieve sampling are, in some cases, as high as 50% in audit populations generated from Population 2. In these cases also, the variance of stabilised sieve sampling will be substantially less than the variance of simple random sampling, i.e. stabilised sieve sampling will have a **deff** less than one. On the other hand, in the audit populations generated from Population 1, the design effect of sieve sampling is near one in all cases (Table 5.8), therefore stabilised sieve sampling may not lead to gains in precision over simple random sampling.

5.10.5.1 Stabilised Sieve Sampling compared to Sieve Sampling.

Since the variance component due to the initial sieve sample in stabilised sieve sampling is the main part of the variance of stabilised sieve sampling, it is to be expected that the variance of the two sampling methods will be similar, i.e. the efficiency of stabilised sieve sampling relative to sieve sampling should be near one.

5.11 Summary

In this chapter, the mean and variance of an estimator of the total error amount is derived for the six monetary-unit sampling methods. The precision of a point estimator of the total error amount with systematic, cell, Lahiri, sieve and stabilised sieve sampling relative to simple random sampling of monetary units is compared.

It was found that

(i) The design effect of systematic sampling depends on the population characteristics. The design effect is likely to be less than or equal to one in all audit populations. The greatest gains in precision occur in populations with large line items.

(ii) Cell sampling has a design effect of less than or equal to one in all the audit populations. The greatest gains in precision occur in populations with large line items and with large sample sizes.

(iii) Lahiri sampling has a design effect of one in all audit populations.

(iv) Sieve sampling has a design effect of less than or equal to one for most sample sizes and for most of the audit populations. Substantial gains in precision occur when the line items are large and the error rates are low.

(v) No general statement can be made on the design effect of stabilised sieve sampling. It is likely to be less than one when the gains in precision due to sieve sampling are substantial and near one otherwise.

In the following chapters, an empirical investigation is carried out to establish whether these sampling methods have the same effects for upper bounds of the total error amount.

Chapter 6

Upper Bound Comparisons of Simple Random, Systematic, Cell and Sieve Sampling of Monetary Units.

6.1 Introduction

This chapter investigates how different monetary-unit sampling methods perform when obtaining estimates of the total error amount in substantive auditing. Four methods, currently used in practice are examined, namely, simple random, systematic, cell and sieve sampling of monetary units. Simple random sampling is widely used in the United States (Wurst,1990). Systematic and cell sampling are used in audit practice in Canada (CICA, 1990), and sieve sampling has been proposed as an alternative to simple random sampling of monetary units (Gill, 1983) and has gained some acceptance in the Netherlands (Wurst, Neter and Godfrey, 1989a). Studies done on the use of monetary-unit sampling in the UK (for example, McRae, 1982 and Abdul-Hamid, 1993) have not specified the type of monetary-unit sampling plans used in substantive testing.

A comparative investigation of the sampling methods is carried out by means of a large scale simulation study using the thirty audit populations created from the two actual accounting populations as described in Chapter 3. Samples of sizes 30, 60 and 100 are drawn from each audit population. Upper bounds are calculated using the Stringer, Cell and Moment bounds, with both error assignments at three nominal confidence levels, .70, .85 and .95. One thousand replications are performed for each sample size and for each sample design. Analysis of variance models are constructed to compare the performance of the sampling methods.

A number of previous investigations have been carried out to compare the performances of monetary-unit sampling methods. Jenne (1982) compared simple random, cell and systematic sampling but concentrated on the number of observed errors found in the sample with each selection method. Plante, Neter and Leitch (1985) compared simple random, cell and systematic sampling using the Stringer, Cell and Multinominal bounds. Wurst, Neter and Godfrey (1989b) compared sieve sampling with simple random and cell sampling of monetary units using the Stringer and Cell bounds. Wurst (1990) called for further research on the comparative performance of sieve sampling and systematic sampling of monetary units. Wurst (1990) also called for research on the performance of the Moment bound using sieve sampling.

This study addresses the issues raised by Wurst. Extending previous research, it provides the first comparative investigation of some of the sampling methods and the first investigation of the sampling methods using the Moment bound.

In the remainder of the chapter, the analysis of variance models used are detailed, the comparative performance of the sampling methods for each error rate, taint size, sample size and bound is analysed and a comparison is made between the performance of the point estimator and the performance of the bounds under each sampling method.

6.2 The Analysis of Variance Models

To compare the performance of the sampling methods, five-factor Analysis of Variance (ANOVA) models are constructed using the performance measures outlined in 2.11 as the dependent variables. The purpose of the models is to determine to what extent the performance measures are influenced by the different sampling methods and to what extent the sampling methods interact with other factors to affect the performance measures.

6.2.1 The Independent Factors.

Five independent factors are included in the ANOVA models. They are

(i) line item error rate (α)

(ii) taint size (β)

(iii) sample size (γ)

(iv) bound (δ)

(v) sampling method (ε)

The independent factors are detailed in table 6.1.

Table 6.1 Description of the Independent Factors for the ANOVA Models

Factor	Description	Levels
α	Line Item Error Rate	i=1 Lowest i=2 i=3 i=4 i=5 Highest
β	Taint Size	j=1 Low j=2 Medium j=3 High
γ	Sample Size	k=1 n = 30 k=2 n = 60 k=3 n = 100
δ	Bounds	l=1 Stringer l=2 Cell l=3 Moment
η	Sampling Method	m=1 SRS m=2 Systematic m=3 Cell m=4 Sieve

The factors outlined in Table 6.1 have been identified in previous studies as likely to influence the performance measures. Neter and Loebbecke (1975), Leitch, Neter, Plante and Sinha (1982) and Atkinson (1990) demonstrated that the estimates of the total error amount are influenced by population characteristics such as the line item error rate and the magnitude of the errors (taints) in the population. The five line item error rates and the three taint levels for each population have been chosen to provide a range of population characteristics which may occur in real audit situations, as detailed in Chapter 3.

Neter and Loebbecke (1975), Reneau (1978), Wurst, Neter and Godfrey (1989b) and Atkinson (1990) demonstrated that the sample size influences the accuracy of the estimates of the total error amount. McRae (1982, p191) estimated that the modal sample size used in substantive testing by professional accounting firms in the UK fell in the range 40 to 80 and in a recent study carried out by Abdul-Hamid (1993), it was found that the size of the audit sample used by medium sized firms in England ranged from 25 to 100. In keeping with these studies, it was decided that samples of sizes 30, 60 and 100 should represent the range of sample sizes currently used in audit practice.

Three bounds commonly used in auditing i.e. the Stringer, Cell and Moment bounds are used to estimate the total error amount. The Stringer bound is widely used in audit practice in the United States (Wurst, 1990). The Cell bound is used in audit practice in Canada (CICA, 1990) and the Moment bound has recently been adopted by Arthur Anderson as a replacement for the Stringer bound (Felix, Grimlund, Koster and Roussey, 1990). Each of these bounds can be used with each of the monetary-unit sampling methods and the two error assignment methods (Wurst, Neter and Godfrey (1989b). The analysis of the Moment bound for the different sampling methods is given here for the first time.

Sampling method is included as an independent factor so that its influence on the performance criteria could be assessed and the extent to which it interacts with other factors in affecting the performance criteria may be investigated. The four sampling methods considered are those which are currently used in audit practice. The comparative performance of the estimates with sieve sampling and systematic sampling is provided here for the first time.

6.2.2 The ANOVA Models

Separate ANOVA models are examined using coverage, tightness, and standard deviation as the dependent variables for each set of audit populations (2), for each nominal confidence level (3) and for each error assignment method (2). Therefore, twelve (2x3x2) different ANOVA models are examined for each of the three dependent variables. Each model includes all first-order interactions of the independent factors and can be expressed as

$$y_{ijklm} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \eta_m + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\delta)_{il} + (\alpha\eta)_{im}$$
$$+(\beta\gamma)_{jk} + (\beta\delta)_{jl} + (\beta\eta)_{jm} + (\gamma\delta)_{kl} + (\gamma\eta)_{km} + (\delta\eta)_{lm} + \epsilon_{ijklm}$$

subject to

$$\sum \alpha_i = \sum \beta_j = \sum \gamma_k = \sum \delta_l = \sum \eta_m = 0$$

$$\epsilon_{ijklm} \text{ is } N(0, \sigma)$$

where, for i=1,.5, j=1,.3, k=1..3, l=1,..3, m=1,..4,

μ is the overall mean of the dependent variable.

α_i is the main effect of the line item error rate at the ith level.

β_j is the main effect of the taint size at the jth level.

γ_k is the main effect of the sample size at the kth level

δ_l is the main effect of the bound at the lth level.

η_m is the main effect of sampling method at the mth level

$(\alpha\beta)_{ij}$ is the interaction of the line item error rate and the taint size at the ith and jth level respectively.

$(\alpha\gamma)_{ik}$ is the interaction of the line item error rate and the sample size at the ith and kth level respectively.

$(\alpha\delta)_{il}$ is the interaction of the line item error rate and the bounds at the ith and lth level respectively.

$(\alpha\eta)_{im}$ is the interaction of the line item error rate and the sample size at the ith and mth level respectively.

$(\beta\gamma)_{jk}$ is the interaction of the taint size and the sample size at the jth and kth level respectively.

$(\beta\delta)_{jl}$ is the interaction of the taint size and the bound at the jth and lth level respectively.

$(\beta\eta)_{jm}$ is the interaction of the taint size and the sampling method at the jth and mth level respectively.

$(\gamma\delta)_{kl}$ is the interaction of the sample size and the bound at the kth and lth level respectively.

$(\gamma\eta)_{km}$ is the interaction of the sample size and the sampling method at the kth and mth level respectively.

$(\delta\eta)_{lm}$ is the interaction of the bound and the sampling method at the lth and mth level respectively.

In each ANOVA model, there is a total of 540 cells (5 line item error rates x 3 taint sizes x 3 sample sizes x 3 bounds x 4 sampling methods).

6.2.3 Transformations on the Dependent Variables.

To correct for lack of normality and to eliminate heteroscedasticity, a transformation was carried out on each of the dependent variables. The reliability variable is a proportion near one and the arc-sine transformation is used to stabilise the variance of this variable. Draper and Smith (1981, p223) suggested that this transformation is an effective variance stabiliser if the proportion is greater than or equal to 0.70 which is the lowest nominal confidence level in this study. The logarithmic transformation to the base 10 was used on the tightness and precision variables. The effectiveness of the transformations were assessed by examining the fit of the model and the patterns of the residuals. Normal probability plots indicated that the residuals were close to normal for each of the dependent variables. Also, examination of residual plots against the predicted values confirmed the adequacy of the model and the lack of heteroscedasticity in each case. A high r^2 occurred in each of the models. The ANOVA tables are given in Appendix D.

6.3 Comparisons of the Sampling Methods

The main effects of the sampling-method factor and the first-order interactions of the sampling-method factor with each of the other factors for each dependent variable are analysed.

The main effects represent general tendencies in the data and the main effect for a particular factor represents the average of the dependent variable across all levels of the other factors. A detailed analysis of the main effects for all the models is given.

The first-order interactions measure the relationship between the effects of the factors. When comparing two factors, factor 1 and factor 2 say, the interaction will measure to what extent the dependent variable varies for each level of factor 1 with different levels of factor 2. Only the interaction terms including the sampling method factor are investigated here as the study is concerned with the differential performance of the sampling methods in each population for each error rate, taint size, sample size and bound. Interactions terms which do not include sampling method are not important in the context of this study and are not analysed. When a first-order interaction of the sampling method with a particular factor is found to be significant, Tukey's pairwise comparison of means test is used to test the significance of the difference of the means of the sampling methods for each level of that factor.

A detailed breakdown of the first-order interactions of the sampling method with each of the other factors for the taint error assignment at the 95% nominal confidence level is given for each of the performance measures. The comparative results for the other models (i.e. the taint models at the 85% and 70% nominal confidence level and the AON models) are similar and are given in appendix E.

In the ANOVA models, a p-value for the test statistic of less than 0.05 is said to be statistically significant, i.e. when p <.05, for a particular factor or interaction, then that factor or interaction is said to have a significant effect on the dependent variable. Equivalently, the null hypothesis that the factor or interaction has no effect on the dependent variable is rejected. Conversely, when p >= .05 for a particular factor or interaction, the hypothesis that there is no effect is not rejected.

6.3.1 Comparisons of the Reliability of the Sampling Methods

This section analyses the main effects of the sampling method and the first-order interactions of the sampling method with each of the other factors for the reliability dependent variable.

6.3.1.1 The Main Effects of the Reliability Dependent Variable

The reliability main effects of the sampling methods for each model are given in tables 6.2 and 6.3.

Table 6.2 Mean Coverage for Each Sampling Method for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.72	99.01	94.67	93.88	84.94	83.86
Sys	99.43	99.45	96.40	95.16	85.61	85.00
Cell	98.88	99.14	94.92	94.29	84.74	84.36
Sieve	98.63	98.84	94.38	93.70	84.74	83.88

Table 6.3 Mean Coverage for Each Sampling Method for Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.55	98.71	92.87	93.06	80.66	79.96
Sys	99.51	99.60	96.17	96.40	85.61	85.38
Cell	98.18	99.23	94.17	94.09	82.96	83.14
Sieve	99.18	99.28	95.13	95.18	83.52	82.84

The following general observations are made from these tables.

(i) The overall mean coverage does not differ substantially between sampling methods in any of the models.

(ii) The mean coverage is above the nominal in all cases.

6.4.1.2 The First-Order Interactions of the Reliability Dependent Variable

The results of the analysis of the interactions between each of the independent factors and the sampling method for the reliability dependent variable are given. Tables 6.4 and 6.5 give the significance of the first-order interactions which include the sampling method factor for the reliability dependent variable for audit populations generated from Population 1 and Population 2 respectively.

Table 6.4 Significance of the First-order Interactions of Each Factor with the Sampling Method for Coverage for Audit Populations generated from Population 1 with the Taint Error Assignment

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Rate	NS	p <.0001	NS	NS	NS	NS
Taint	NS	NS	NS	NS	NS	NS
Samsize	p <.041	p <.0001	p <.01	NS	NS	NS
Bound	NS	NS	NS	NS	NS	NS

Table 6.5 Significance of the First-order Interactions of Each Factor with the Sampling Method for Coverage for Audit Populations generated from Population 2 with the Taint Error Assignment.

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
First-order	Interactions		with		Sampling Method	
Rate	p <.0001	p <.0001	p <.0001	p <.0001	p < .018	NS
Taint	NS	NS	NS	NS	NS	NS
Samsize	p <.0001	p <.0001	p <.0001	p <.0001	p <.0001	p<.0001
Bound	NS	NS	NS	NS	NS	NS

The following observations are made from tables 6.4 and 6.5.

(i) The first-order interaction of the sampling method and the line item error rate is insignificant for all except the model with the AON error assignment at the 95% nominal level in Population 1. In Population 2, the first-order interaction of the sampling method and the line item error rate is significant for all models except the model with the AON error assignment at the 70% nominal confidence level.

(ii) The first-order interaction of the sampling method and the taint size is insignificant in all the models.

(iii) The first-order interaction of the sampling method and the sample size is significant at the 95% nominal confidence level with both error assignment methods and at the 85% confidence level with the taint error assignment in Population 1. It is significant in all models in Population 2.

(iv) The first-order interaction of the sampling method and the bound is insignificant in all the models.

It would appear from this that audit populations generated from Population 2 are more sensitive. (i.e. more significant first-order interactions) to the sampling method used than the audit populations generated from Population 1, particularly with respect to the interactions of the sample size and the error rate with the sampling method.

6.3.1.2.1 The First-Order Interaction of the Sampling Method and Error Rate for the Reliability Dependent Variable.

The first-order interactions of the sampling method and the error rate for the taint error assignment at the 95% nominal confidence level are given in tables 6.6 and 6.7 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.6 Mean Coverage of the First-Order Interaction of Sampling Method by Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	100.00	99.72	99.28	97.44	97.13
Systematic	100.00	99.91	99.76	98.90	98.60
Cell	100.00	99.70	99.36	97.64	96.70
Sieve	100.00	99.73	99.36	96.77	97.27

Table 6.7 Mean Coverage of the First-Order Interaction of Sampling Method by Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	99.80	99.80	98.84	97.21	97.09
Systematic	100.00	100.00	100.00	98.47	99.07
Cell	99.89	99.83	98.23	98.57	98.35
Sieve	99.91	99.91	99.44	98.70	98.12

From tables 6.6-6.7, the following observations are made

(i) The mean coverage is similar for all the sampling methods in each error rate with the exception of error rate 3 in Population 2, where systematic sampling has a significantly higher coverage than simple random and cell sampling.

(ii) Systematic sampling has the highest mean coverage for all the error rates.

6.3.1.2.2 The First-Order Interaction of the Sampling Method and the Taint Size for the Reliability Dependent Variable.

The first-order interactions of the sampling method and the taint size for the taint error assignment at the 95% nominal confidence level are given in tables 6.8 and 6.9 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.8 Mean Coverage of the First-Order Interaction of the Sampling Method by the Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Taint	1	2	3
SRS	98.99	98.91	98.25
Systematic	99.63	99.57	99.10
Cell	99.15	99.09	98.40
Sieve	98.94	98.85	98.08

Table 6.9 Mean Coverage of the First-Order Interaction of the Sampling Method by the Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	1	2	3
SRS	98.62	98.51	98.52
Systematic	99.66	99.48	99.38
Cell	98.22	98.15	99.15
Sieve	99.22	98.18	99.18

The following observations are made from tables 6.8 and 6.9

(i) There are no significant differences between the mean coverage of the sampling methods for any of the taint sizes.

(ii) Systematic sampling has the highest mean coverage for all the taint sizes.

6.3.1.2.3 The First-Order Interaction of the Sampling Method and the Sample Size for the Reliability Dependent Variable.

The first-order interactions of the sampling method and the sample size for the taint error assignment at the 95% nominal confidence level are given in tables 6.10 and 6.11 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.10 Mean Coverage of the First-Order Interaction of the Sampling Method by the Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n= 30	n= 60	n = 100
SRS	99.66	98.34	98.15
Systematic	99.69	98.27	99.33
Cell	99.60	98.53	98.51
Sieve	99.53	98.28	98.07

Table 6.11 Mean Coverage of the First-Order Interaction of the Sampling Method by the Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n= 30	n= 60	n = 100
SRS	99.26	98.73	97.65
Systematic	99.47	99.32	99.74
Cell	99.44	99.04	99.05
Sieve	99.24	99.03	99.31

The following observations are made from tables 6.10 and 6.11

(i) Systematic sampling has the highest mean coverage for all the sample sizes for audit populations generated from Population 1 and Population 2

(ii) In Population 2, systematic sampling has a significantly higher mean coverage than simple random sampling when the sample size is 100.

6.3.1.2.4 The First-Order Interaction of the Sampling Method and the Bound for the Reliability Dependent Variable.

The first-order interactions of the sampling method and the sample size for the taint error assignment at the 95% nominal confidence level are given in tables 6.12 and 6.13 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.12 Mean Coverage of the First-Order Interaction of the Sampling Method by the Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	99.22	99.02	97.91
Systematic	99.78	99.70	98.81
Cell	99.42	99.18	98.04
Sieve	99.22	98.99	97.67

Table 6.13 Mean Coverage of the First-Order Interaction of the Sampling Method by the Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	98.87	98.82	97.98
Systematic	99.64	99.64	99.64
Cell	99.42	99.40	98.72
Sieve	99.40	99.39	98.78

The following observations are made from these tables

(i) There are no significant differences in the mean coverage between sampling methods for any of the three bounds.

(ii) Systematic sampling has the highest mean coverage of the three bounds.

(iii) The Stringer bound has the highest mean coverage in all cases.

6.3.1.3 Summary and Comparisons with other Studies

The mean coverage is above the nominal for the main effects and for the first-order interactions in all cases.

The sample selection method appears to have little effect on the reliability of the estimates. The four sample selection methods have similar reliabilities for each error rate, taint size, sample size and bound. Systematic sampling has a higher reliability for each error rate, taint size, sample size and bound than the other sampling methods but the difference is not substantial. Significant first-order interactions exist but these are not of practical importance since the differences are not substantial in any case.

The Stringer bound has a higher reliability than the Moment bound but again, the difference is not substantial.

The reliability findings are similar to those of other studies. Plante, Neter and Leitch (1985) found that the method of sample selection (simple random, cell and systematic sampling) appeared to have little effect on the coverage of the Stringer and Cell bounds. Wurst, Neter and Godfrey (1989b) found that, in almost all cases, the coverage for simple random, cell and sieve were similar for a given sample size, nominal confidence level and bound. They also found that the Stringer bound had the highest coverage in all cases.

6.3.2 Comparisons of the Tightness of the Sampling Methods

The analysis of the main effects of the sampling methods and the first-order interactions of the sampling methods with each of the other independent factors for the tightness dependent variable is reported below.

6.3.2.1 The Main Effects of the Tightness Dependent Variable

The tightness main effects for each sampling method for each model are given in tables 6.14 and 6.15

Table 6.14 Mean Tightness for Each Sampling Method for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	276.12	287.27	175.26	178.22	105.13	103.33
Sys	274.21	286.31	173.34	177.23	103.29	102.32
Cell	272.47	288.16	172.14	178.93	102.46	103.86
Sieve	273.86	290.19	173.52	180.80	103.78	105.57

Table 6.15 Mean Tightness for Each Sampling Method in Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	191.05	191.42	118.83	118.58	67.64	67.02
Sys	191.00	193.63	118.45	118.27	67.00	68.27
Cell	191.69	191.51	119.20	118.56	67.86	66.87
Sieve	191.15	190.19	118.70	117.35	62.29	65.77

The following observations are made from these tables 6.14-6.15

(i) The mean tightness measure does not differ substantially between sampling methods.

(ii) The lower nominal confidence levels have tighter estimates than the higher nominal confidence levels.

6.3.2.2 The First-Order Interactions of the Tightness Dependent Variable

The results of the analysis of the interactions between each of the independent factors and the sampling method for the tightness dependent variable are given. Tables 6.16 and 6.17 give the significance of first-order interactions which include the sampling method factor for the tightness dependent variable.

Table 6.16 Significance of the First-order Interactions of Each Factor with the Sampling Method for Tightness for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Rate	NS	NS	NS	p<.024	p<.026	p<.019
Taint	NS	NS	NS	NS	NS	NS
Samsize	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001
Bound	NS	NS	NS	NS	NS	NS

Table 6.17 Significance of the First-order Interactions of Each Factor with the Sampling Method for Tightness for Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Rate	NS	p <.016	NS	p<.014	NS	p <.035
Taint	NS	NS	NS	NS	NS	NS
Samsize	p <.0001	p <.0001	p <.0001	p <.0001	p <.0001	p<.0001
Bound	NS	NS	NS	NS	NS	NS

The following observations are made from tables 6.16 and 6.17

(i) The first-order interaction of the sampling method and the line item error rate is significant for the models at the 95% nominal level in Population 1 and at the 85% nominal confidence with the taint error assignment in Population 1. In Population 2, the first-order interactions of the sampling method and the line item error rate are insignificant for all models with the taint error assignment and significant for all models with the AON error assignment.

(ii) The first-order interactions of the sampling method with the taint size and with the bound are insignificant in all cases.

(iii) The first-order interaction of the sampling method and the sample size is significant for all models.

6.3.2.2.1 The First-Order Interactions of the Sampling Method and the Error Rate for the Tightness Dependent Variable.

The first-order interactions of the sampling method and the error rate for the tightness dependent variable for the taint error assignment at the 95% nominal confidence level are given in tables 6.18 and 6.19 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.18 Mean Tightness of the First-Order Interaction of the Sampling Method by the Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	622.29	312.14	219.80	129.73	96.64
Systematic	617.18	312.14	217.73	128.61	95.41
Cell	612.71	306.65	216.11	130.62	96.24
Sieve	617.10	310.32	216.57	129.34	95.97

Table 6.19 Mean Tightness of the First-Order Interaction of the Sampling Method by the Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	316.49	252.27	187.26	114.26	84.38
Systematic	315.35	253.26	186.31	114.58	85.48
Cell	316.08	253.27	186.57	116.53	86.01
Sieve	318.61	252.30	186.84	114.27	84.38

The following observations are made from these tables

(i) The sampling method does not appear to have any effect on the tightness of the error estimate for any given error rate.

(ii) The estimates of the total error amount are extremely conservative with low error rates especially in Population 1. For example, the average error estimate is over six times the total error amount for all sampling methods in the audit populations generated from Population 1 with error rate 1.

(iii) The estimates of the total error amount become less conservative as the error rate increases.

6.3.2.2.2 The First-Order Interactions of the Sampling Method and the Taint Size for the Tightness Dependent Variable.

The first-order interactions of the sampling method and the taint size for the tightness dependent variable with the taint error assignment at the 95% nominal confidence level are given in tables 6.20 and 6.21 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.20 Mean Tightness of the First-Order Interaction of the Sampling Method by the Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error

Taint	Taint 1	Taint 2	Taint 3
SRS	301.73	289.47	235.15
Systematic	299.44	287.38	235.83
Cell	297.94	285.72	233.74
Sieve	299.18	287.03	235.38

Table 6.21 Mean Tightness of the First-Order Interaction of the Sampling Method by the Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	Taint 1	Taint 2	Taint 3
SRS	195.41	190.79	170.94
Systematic	195.73	190.75	186.53
Cell	198.22	191.66	187.39
Sieve	195.78	190.95	186.73

From these tables, the following observations are made

(i) The sampling method does not have any effect on the tightness of the bound for any given taint size.

(ii) The estimates of the total error amount are tighter in populations with higher taint sizes.

6.3.2.2.3 The First-Order Interactions of the Sampling Method and the Sample Size for the Tightness Dependent Variable.

The first-order interactions of the sampling method and the sample size for the tightness dependent variable with the taint error assignment at the 95% nominal confidence level are given in tables 6.20 and 6.21 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.22 Mean Tightness of the First-Order Interaction of the Sampling Method by the Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sampling Size	n = 30	n = 60	n = 100
SRS	430.22	238.22	159.92
Systematic	423.93	235.49	163.22
Cell	419.48	237.93	159.99
Sieve	424.60	237.08	159.90

Table 6.23 Mean Tightness of the First-Order Interaction of the Sampling Method by the Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sampling Size	n = 30	n = 60	n = 100
SRS	278.64	173.63	120.87
Systematic	279.50	172.88	120.65
Cell	283.78	169.80	121.50
Sieve	282.45	170.64	120.37

The following observations are made from these tables

(i) The sampling method does not have any effect on the tightness of the estimates for any given sample size.

(ii) The estimates of the total error amount are extremely conservative for small sample sizes but become less conservative as the sample size increases.

6.3.2.2.4 The First-Order Interactions of the Sampling Method and the Bound for the Tightness Dependent Variable.

The first-order interactions of the sampling method and the bound for the tightness dependent variable with the taint error assignment at the 95% nominal confidence level are given in tables 6.24 and 6.25 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.24 Mean Tightness of the First-Order Interaction of the Sampling Method by the Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	293.26	292.22	239.86
Systematic	293.90	289.85	238.88
Cell	292.78	288.68	235.95
Sieve	295.53	245.15	238.73

Table 6.25 Mean Tightness of the First-Order Interaction of the Sampling Method by the Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	195.95	195.14	182.04
Systematic	195.33	194.61	183.07
Cell	196.83	196.06	182.18
Sieve	196.42	195.67	181.17

The following observations are made from these tables.

(i) The sampling method does not have any effect on the tightness of the estimates for any given bound.

(ii) The Moment bound is the tightest bound in all cases.

(iii) The Stringer bound is the most conservative of the bounds for all the sampling methods in audit populations generated from Populations 1 and 2.

6.3.2.3 Summary and Comparisons with other Studies

The sample selection method appears to have little effect on the tightness of the estimates. The average tightness is similar for each sample selection method. The four sample selection methods have similar tightness for each error rate, taint size, sample size, and bound.

Differences in tightness that exist, are caused by factors other than the sample selection method. For example, low error rate populations give extremely conservative estimates of the total error amount. Estimates are also very conservative for samples of size thirty and less so for higher sample sizes. The Stringer bound is the most conservative of all the bounds. The Moment bound is the tightest in all cases.

The tightness findings are similar to those obtained by other authors. Wurst, Neter and Godfrey (1989b) found that simple random, cell and sieve sampling have no effect on the tightness of the bound as measured by the mean bound. Wurst, Neter and Godfrey (1989b) showed that the mean Stringer bound with the taint error assignment is consistently smaller than the mean Stringer bound with the AON assignment. Plante, Neter and Leitch (1985) and Wurst, Neter and Godfrey (1989b) demonstrated that the mean Cell bound is consistently smaller that the mean Stringer bound. Grimlund and Dworkin (1984) found that the Moment bound is tighter than the Stringer bound for most error distributions.

6.3.3 Comparisons of the Precision of the Sampling Methods

The analysis of the main effects of the sampling methods and the first-order interactions of the sampling methods with each of the other factors for the precision dependent variable is reported below.

6.4.3.1 The Main Effects of the Precision Dependent Variable

The precision main effects for each sampling method for each model are given in terms of the standard deviation of the estimate in tables 6.26-6.27.

Table 6.26 Mean Standard Deviation (000s) for Each Sampling Method for Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.74	102.46	87.40	91.46	77.41	81.25
Sys	90.03	94.27	79.48	83.97	70.22	74.41
Cell	94.88	100.43	83.88	89.60	74.21	79.53
Sieve	99.43	104.95	88.12	93.76	78.14	83.36

Table 6.27 Mean Standard Deviation(000s) for Each Sampling Method for Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	145.86	146.50	130.78	131.73	116.99	117.96
Sys	126.19	128.13	112.99	115.01	100.79	102.73
Cell	134.69	134.57	120.68	120.90	107.83	108.16
Sieve	133.93	133.77	120.15	120.26	107.46	107.66

As can be seen from tables 6.26 and 6.27, substantial differences in precision exist between the sampling methods. In particular, the following observations are made

(i) Systematic sampling produces the most precise estimates of the total error amount in all cases.

(ii) Cell sampling is more precise than simple random sampling in all cases.

(iii) In Population 1, sieve sampling is less precise than simple random sampling but this is not substantial in any case. In Population 2, sieve and is more precise that simple random sampling.

(iv) The models with the AON error assignment are less precise than the models with the taint error assignment.

(v) The models with the lower nominal confidence levels are more precise than the higher nominal confidence levels.

6.3.3.2 The First-Order Interactions of the Precision Dependent Variable

This section reports the results of the analysis of the interactions between each of the independent factors and the sampling method for the precision dependent variable. Tables 6.28 and 6.29 give the significance of the first-order interactions which include the sampling method factor for the precision dependent variable.

Table 6.28 Significance of the First-order Interactions for Precision in Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Rate	p <.0001	p<.0001	p <.0001	p <.0001	p <.0001	<.0001
Taint	NS	NS	NS	NS	NS	NS
Sample Size	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001
Bound	NS	NS	NS	NS	NS	NS

Table 6.29 Significance of the First-order interactions for Precision in Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Rate	p <.0001	p <.0001	p <.0001	p <.0001	p <.0001	p<.0001
Taint	NS	NS	NS	NS	NS	NS
Samsize	p <.0001	p <.0001	p <.0001	p <.0001	p <.0001	p<.0001
Bound	NS	NS	NS	NS	NS	NS

The following observations are made from tables 6.28 and 6.29 with respect to the precision dependent variables.

(i) The first-order interaction of the sampling method and the line item error rate is significant in all the models.

(ii) The first-order interaction of the sampling method and the taint is not significant in any of the models.

(iii) The first-order interaction of the sampling method and the sample size is significant in all models.

(iv) The first-order interaction of the sampling method and the bound is insignificant in all models.

6.3.3.2.1 The First-Order Interactions of the Sampling Method and the Error Rate for the Precision Dependent Variable.

The first-order interactions of the sampling method and the error rate for the precision dependent variable with the taint error assignment at the 95% nominal confidence level are given in tables 6.30 and 6.31 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.30 Mean Standard Deviation (000s) of Sampling Method by Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	56.52	81.95	96.44	121.70	137.08
Systematic	53.80	79.34	90.95	107.08	118.98
Cell	54.20	77.20	91.56	118.54	132.56
Sieve	56.73	80.17	96.91	123.91	139.93

Table 6.31 Mean Standard Deviation (000s) of Sampling Method by Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	107.43	118.66	137.04	173.22	192.93
Systematic	85.91	102.20	120.36	167.99	154.47
Cell	95.99	110.67	128.45	157.58	180.57
Sieve	90.08	104.63	126.59	161.11	187.30

Although no significant differences were found between sampling methods for each error rate, the following observations are made from these tables.

(i) Systematic sampling has the highest precision across all error rates. The greatest gains in precision due to systematic sampling over the other sampling methods occur in the high error rate populations.

(ii) In Population 1, cell and sieve sampling have approximately the same precision as simple random sampling in most cases. In Population 2, cell and sieve sampling are consistently more precise than simple random sampling. With sieve sampling, the greatest gains in precision occur in the low error rate populations generated from Population 2.

(iii) The estimates of the total error amount are more precise in low error rate populations than in high error rate populations.

6.3.3.2.2 The First-Order Interactions of the Sampling Method and the Taint Size for the Precision Dependent Variable.

The first-order interactions of the sampling method and the taint size for the precision dependent variable with the taint error assignment at the 95% nominal confidence level are given in Tables 6.30 and 6.31 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.32 Mean Standard Deviation (000s) of Sampling Method by Taint Size at 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Taint	1	2	3
SRS	97.13	97.87	101.21
Systematic	88.28	89.02	92.79
Cell	93.04	93.88	97.71
Sieve	97.39	98.29	102.62

Table 6.33 Mean Standard Deviation (000s) of Sampling Method by Taint Size at 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	1	2	3
SRS	144.33	145.46	147.77
Systematic	125.31	125.81	127.45
Cell	133.63	134.46	135.91
Sieve	132.72	133.60	135.47

The following observations are made from these tables

(i) Systematic sampling has the highest precision across all taint sizes.

(ii) Sieve and cell sampling are consistently more precise than simple random sampling for all taint sizes in audit populations generated from Population 2.

(iii) The precision decreases as the taint size increases for all sampling methods.

6.3.3.2.3 The First-Order Interactions of the Sampling Method and the Sample Size for the Precision Dependent Variable.

The first-order interactions of the sampling method and the sample size for the precision dependent variable with the taint error assignment at the 95% nominal confidence level are given in Tables 6.34 and 6.35 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.34 Mean Standard Deviation (000s) of Sampling Method by Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30	n = 60	n = 100
SRS	135.47	92.60	68.14
Systematic	125.74	85.54	61.80
Cell	128.84	90.44	65.37
Sieve	136.48	92.84	68.97

Table 6.35 Mean Standard Deviation (000s) of Sampling Method by Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30	n = 60	n = 100
SRS	195.34	136.57	105.67
Systematic	199.89	106.88	71.81
Cell	187.87	127.84	88.35
Sieve	194.98	124.78	82.04

From these tables the following observations are made

(i) Systematic sampling has the highest precision for all sample sizes with the exception of sample size 30 in the audit populations generated from Population 2. Systematic sampling is significantly more precise than simple random sampling for samples of sizes sixty and one hundred in audit populations generated from Population 2.

(ii) Cell and sieve sampling are more precise than simple random sampling for all sample sizes in audit populations generated from Population 2 and significantly more precise than simple random sampling for samples of size 100.

(iii) Cell sampling is more precise than sieve sampling for all sample sizes in audit populations generated from Population 1.

(iv) The precision of the estimate of the total error amount increases as the sample size increases.

6.3.3.2.4 The First-Order Interactions of the Sampling Method and the Bound for the Precision Dependent Variable.

The first-order interactions of the sampling method and the bound for the precision dependent variable with the taint error assignment at the 95% nominal confidence level are given in tables 6.34 and 6.35 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.36 Mean Standard Deviation (000s) of Sampling Method by Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	96.63	97.41	102.173
Systematic	88.21	88.90	92.98
Cell	92.96	93.75	97.93
Sieve	97.52	98.25	102.51

Table 6.37 Mean Standard Deviation (000s) of Sampling Method by Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	141.13	141.23	155.20
Systematic	122.73	122.69	133.15
Cell	130.26	130.39	143.40
Sieve	129.70	129.81	142.28

From tables 6.30-6.37 the following observations are made

(i) Systematic sampling gives the most precise estimates in all cases.

(ii) In Population 2, the estimates are more precise with cell and sieve sampling than with simple random sampling.

(iii) The Moment bound is the least precise of the bounds in all cases.

6.3.3.3 Summary and Comparisons with other Studies

The sample selection method appears to have substantial effects on the precision of the estimates. Systematic sampling is consistently more precise than the other sampling methods with the greatest gains in precision occurring in the high error rate populations. Cell and sieve sampling give more precise estimates than simple random sampling in audit populations generated from Population 2, i.e. audit populations with large line items.

The precision findings are similar to the findings of other authors. Jenne (1982) compared simple random, cell and systematic and found that systematic sampling is the most precise in all cases. Plante, Neter and Leitch (1985) found that systematic and cell selection reduce the variability of the distribution of the bounds compared to simple random sampling of monetary units. They also found that populations with large line items have a greater reduction in variability for cell and systematic sampling. Wurst, Neter and Godfrey (1989b) found that while simple random, cell and sieve sampling do not have any profound effects on the coverage and tightness, for any given bound, confidence level and sample size, they appear to have definite effects on the variability of the bounds. They found that cell sampling consistently leads to a smaller variability of the bounds from sample to sample, for any given bound, confidence level, and sample size, than do sieve and simple random sampling. Wurst, Neter and Godfrey (1989b) also noted that the Stringer and Cell bounds with the AON error assignments are less precise than with the taint error assignment.

6.4 The Design Effect of Systematic, Cell and Sieve Sampling

Another way of comparing the precision of the sampling methods is by means of the design effect (Kish, 1965). The design effect measures the amount of increase or decrease in the variability of a particular sampling method compared to simple random sampling. It is defined in Chapter 2. It was used in Chapter 5 to compare the variability of systematic, cell and sieve sampling with simple random sampling of monetary units using the point estimator. The design effect is calculated for systematic, cell and sieve sampling, for each audit population, error assignment, nominal confidence level, sample size, using the Stringer, Cell and Moment bounds at the three nominal confidence level with each taint error assignment. The results for the bounds at the 95% nominal confidence level with the taint error assignment are reported in 6.4.1 and 6.4.2. The design effect for the bounds with the taint error assignment at the 85% and 70% nominal confidence levels and the models with the AON error assignment are given in Appendix H. These results are similar to the results discussed below.

6.4.1 The Design Effect of Systematic Sampling.

The design effect of systematic sampling is calculated for each error rate, taint size, sample size and bound at the 95% nominal confidence level are given in Tables 6.38 and 6.39 for audit populations generated from Population 1 and 2 respectively.

Table 6.38 Design Effects of Systematic Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.97	0.97	0.97	0.92	0.92	0.93	0.81	0.81	0.76
Taint 2	0.97	0.97	0.97	0.92	0.92	0.93	0.81	0.81	0.77
Taint 3	0.97	0.97	0.97	0.92	0.92	0.90	0.82	0.82	0.81
Error Rate 2									
Taint 1	1.03	1.03	1.03	0.95	0.95	0.92	0.82	0.82	0.83
Taint 2	1.02	1.02	1.02	0.94	0.94	0.92	0.83	0.83	0.83
Taint 3	1.00	1.00	1.00	0.92	0.92	0.97	0.84	0.85	0.83
Error Rate 3									
Taint 1	0.98	0.99	0.92	0.87	0.87	0.85	0.82	0.82	0.90
Taint 2	0.96	0.99	0.92	0.86	0.87	0.86	0.83	0.83	0.90
Taint 3	0.98	0.98	0.95	0.85	0.85	0.88	0.82	0.84	0.85
Error Rate 4									
Taint 1	0.66	0.66	0.66	0.87	0.88	0.88	0.69	0.69	0.69
Taint 2	0.66	0.66	0.66	0.88	0.88	0.88	0.69	0.69	0.69
Taint 3	0.69	0.69	0.69	0.81	0.89	0.85	0.74	0.74	0.71
Error Rate 5									
Taint 1	0.75	0.75	0.73	0.63	0.63	0.63	0.90	0.90	0.90
Taint 2	0.75	0.75	0.74	0.62	0.62	0.64	0.90	0.90	0.90
Taint 3	0.79	0.79	0.77	0.65	0.65	0.63	0.96	0.96	0.94

Table 6.39 Design Effects of Systematic Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	1.02	1.02	0.97	0.54	0.59	0.47	0.26	0.26	0.29
Taint 2	1.02	1.02	1.01	0.56	0.55	0.48	0.27	0.27	0.32
Taint 3	1.03	1.03	0.99	0.59	0.58	0.53	0.27	0.27	0.51
Error Rate 2									
Taint 1	1.27	1.27	1.27	0.61	0.61	0.57	0.27	0.27	0.31
Taint 2	1.26	1.26	1.26	0.61	0.60	0.56	0.27	0.27	0.30
Taint 3	1.25	1.25	1.25	0.62	0.61	0.57	0.28	0.29	0.32
Error Rate 3									
Taint 1	1.09	1.09	1.09	0.65	0.65	0.59	0.44	0.45	0.46
Taint 2	1.10	1.10	1.05	0.71	0.69	0.65	0.41	0.41	0.41
Taint 3	1.09	1.09	1.07	0.78	0.77	0.86	0.44	0.44	0.53
Error Rate 4									
Taint 1	1.28	1.29	1.25	0.84	0.85	0.85	0.53	0.53	0.52
Taint 2	1.27	1.28	1.26	0.85	0.85	0.85	0.52	0.52	0.50
Taint 3	1.25	1.25	1.25	0.84	0.84	0.86	0.53	0.57	0.50
Error Rate 5									
Taint 1	0.76	0.76	0.71	0.92	0.92	0.92	0.78	0.79	0.80
Taint 2	0.77	0.76	0.71	0.91	0.91	0.91	0.74	0.75	0.75
Taint 3	0.77	0.76	0.76	0.91	0.91	0.91	0.62	0.62	0.61

From these tables the following observation are made

(i) Systematic sampling has a design effect of less than one in most cases in audit populations generated from Population 1

(ii) Systematic sampling has a design effect of less than one in most cases in audit populations generated from Population 2. Exceptions occur for samples of size thirty in Population 2 for error rates 1, 2, 3 and 4 where the design effect is greater than one.

(iii) The greatest gains in precision occur for samples of sizes sixty and one hundred from the audit populations generated from Population 2. This is consistent with the point estimator analysis where it was deduced that the greatest gains in precision occur if the population contain relatively large line items.

(iv) The design effect decreases as the error rate increases.

6.4.2 The Design Effect of Cell Sampling

The design effects of cell sampling for each error rate, taint size, sample size and bound at the 95% nominal confidence level are given in 6.40 and 6.41 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.40 Design Effects of Cell Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.84	0.84	0.78	1.03	1.03	1.03	0.92	0.92	0.88
Taint 2	0.84	0.84	0.89	1.03	1.03	1.09	0.92	0.92	0.88
Taint 3	0.81	0.84	0.89	1.01	1.02	1.06	0.93	0.93	0.97
Error Rate 2									
Taint 1	0.86	0.86	0.79	0.94	0.95	0.95	0.87	0.87	0.87
Taint 2	0.86	0.86	0.81	0.95	0.95	0.95	0.87	0.87	0.89
Taint 3	0.86	0.86	0.86	0.95	0.95	0.96	0.89	0.99	0.89
Error Rate 3									
Taint 1	0.91	0.91	0.86	0.94	0.95	0.93	0.87	0.87	0.91
Taint 2	0.91	0.91	0.87	0.94	0.94	0.96	0.87	0.88	0.90
Taint 3	0.91	0.91	0.92	0.94	0.94	0.95	0.88	0.88	0.88
Error Rate 4									
Taint 1	0.93	0.93	0.93	0.95	0.95	0.96	0.96	0.96	0.96
Taint 2	0.93	0.93	0.92	0.95	0.95	0.95	0.96	0.96	0.96
Taint 3	0.94	10.94	0.95	0.97	0.97	0.97	0.97	0.97	0.96
Error Rate 5									
Taint 1	0.94	0.94	0.91	0.92	0.92	0.92	0.93	0.93	0.93
Taint 2	0.94	0.94	0.94	0.92	0.92	0.92	0.93	0.93	0.93
Taint 3	0.96	0.96	0.95	0.92	0.92	0.92	0.94	0.94	0.94

Table 6.41 Design Effects of Cell Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.89	0.89	0.82	0.83	0.83	0.82	0.63	0.63	0.66
Taint 2	0.89	0.89	0.82	0.84	0.84	0.84	0.63	0.63	0.66
Taint 3	0.89	0.89	0.92	0.84	0.84	0.84	0.63	0.63	0.66
Error Rate 2									
Taint 1	0.97	0.97	0.96	0.87	0.87	0.89	0.68	0.68	0.71
Taint 2	0.97	0.96	0.96	0.87	0.87	0.89	0.67	0.67	0.71
Taint 3	0.97	0.96	0.96	0.87	0.87	0.90	0.67	0.67	0.71
Error Rate 3									
Taint 1	0.92	0.91	0.92	0.88	0.88	0.88	0.74	0.64	0.76
Taint 2	0.93	0.92	0.93	0.88	0.85	0.86	0.72	0.72	0.73
Taint 3	0.94	0.94	0.95	0.87	0.86	0.89	0.71	0.71	0.73
Error Rate 4									
Taint 1	0.90	0.90	0.95	0.88	0.86	0.88	0.71	0.71	0.71
Taint 2	0.89	0.90	0.92	0.88	0.89	0.88	0.70	0.70	0.69
Taint 3	0.88	0.89	0.92	0.88	0.89	0.88	0.70	0.70	0.69
Error Rate 5									
Taint 1	0.91	0.92	0.93	0.90	0.90	0.90	0.62	0.62	0.65
Taint 2	0.91	0.92	0.91	0.89	0.89	0.90	0.62	0.62	0.65
Taint 3	0.90	0.91	0.91	0.87	0.87	0.87	0.63	0.62	0.66

The following observations are made from these tables.

(i) The design effect is less than one in most cases for the audit populations generated from Population 1. In the few cases where it is greater than one, the difference between the variance of cell sampling and simple random sampling is not substantial.

(ii) In the populations generated from Population 2, the design effect is less than one in all cases. The greatest reduction in the variability occur for the large sample sizes.

(iii) The greatest gains in precision occur for samples of sizes sixty and one hundred from the audit populations generated from Population 2. This is consistent with the point estimator analysis where it was deduced that the design effect will be small in populations with large line items.

6.4.3 The Design Effect of Sieve Sampling

The design effect of sieve sampling is calculated for each error rate, taint size, sample size and bound at the 95% nominal confidence level are given in 6.42 and 6.43 for audit populations generated from Population 1 and Population 2 respectively.

Table 6.42 Design Effects of Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	1.00	1.00	1.06	1.05	1.05	1.14	0.92	0.92	0.91
Taint 2	1.06	1.07	1.06	1.04	1.02	1.14	0.92	0.92	0.92
Taint 3	1.02	1.01	1.04	1.03	1.04	1.07	0.92	0.92	0.92
Error Rate 2									
Taint 1	0.97	0.97	0.92	1.00	1.00	1.01	0.88	0.88	0.90
Taint 2	0.97	0.97	0.93	1.00	1.00	1.01	0.89	0.88	0.97
Taint 3	0.98	0.98	0.96	1.00	1.00	1.01	0.89	0.89	0.89
Error Rate 3									
Taint 1	0.96	0.96	0.91	1.09	1.09	1.09	0.97	0.99	0.99
Taint 2	0.95	0.96	0.92	1.09	1.09	01.09	0.97	0.97	0.97
Taint 3	0.97	0.97	0.95	1.09	1.09	1.09	0.96	0.97	0.98
Error Rate 4									
Taint 1	1.02	1.01	1.01	0.97	0.97	0.96	1.15	1.14	1.14
Taint 2	1.02	1.02	1.01	0.97	0.97	0.96	1.23	1.15	1.14
Taint 3	1.05	1.05	1.02	0.98	0.98	0.99	1.18	1.18	1.16
Error Rate 5									
Taint 1	1.07	1.07	1.01	0.95	0.95	0.94	1.07	1.06	1.06
Taint 2	1.08	1.08	1.03	0.95	0.95	0.95	1.08	1.08	1.08
Taint 3	1.14	1.13	1.11	0.97	0.97	0.96	1.12	1.11	1.10

Table 6.43 Design Effects of Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.87	0.87	0.83	0.70	0.71	0.71	0.43	0.43	0.49
Taint 2	0.88	0.88	0.88	0.72	0.72	0.71	0.43	0.43	0.49
Taint 3	0.88	0.88	0.81	0.72	0.72	0.72	0.44	0.43	0.51
Error Rate 2									
Taint 1	0.96	0.96	0.95	0.77	0.77	0.78	0.51	0.50	0.55
Taint 2	0.96	0.95	0.93	0.77	0.78	0.80	0.51	0.51	0.55
Taint 3	0.95	0.95	0.95	0.77	0.79	0.81	0.52	0.52	0.55
Error Rate 3									
Taint 1	1.00	1.00	0.99	0.85	0.85	0.84	0.61	0.61	0.63
Taint 2	1.00	1.00	0.99	0.85	0.85	0.86	0.61	0.61	0.63
Taint 3	1.00	1.00	0.97	0.85	0.85	0.85	0.62	0.62	0.64
Error Rate 4									
Taint 1	1.02	1.02	1.01	0.85	0.85	0.96	0.63	0.62	0.62
Taint 2	1.01	1.02	1.01	0.85	0.85	0.86	0.62	0.62	0.62
Taint 3	1.00	1.01	1.00	0.85	0.85	0.86	0.60	0.60	0.60
Error Rate 5									
Taint 1	1.10	1.10	1.10	0.91	0.92	0.91	0.74	0.74	0.74
Taint 2	1.08	1.08	1.09	0.91	0.91	0.91	0.73	0.73	0.79
Taint 3	1.06	1.08	1.05	0.92	0.92	0.91	0.73	0.73	0.72

The following observations are made from these tables.

(i) The design effect is near one in most cases in audit populations generated from Population 1.

(ii) In audit populations generated from Population 2, the design effect is substantially less than one in most cases. Exceptions occur in the high error rate populations for samples of size thirty where the design effect is somewhat greater than one.

(iii) The greatest gains in precision of sieve sampling over simple random sampling occur for the large sample sizes and the low error rates. This is consistent with the point estimator analysis where it was shown that substantial gains in precision of sieve sampling over sampling random sampling occur when the line items are large and the error rates are low.

6.4.4 Summary

The empirical investigation for the upper bounds give similar results with respect to precision as the theoretical results derived in Chapter 5 for the precision of the point estimator. The sampling methods have similar effects on the upper bound estimates of the total error amount as on the point estimator. These results are consistent with the findings of Wurst, Neter and Godfrey (1989a and 1989b) who studied simple random, cell and sieve sampling using the Stringer and Cell bounds.

6.5 Conclusions

In this chapter, the performance of simple random sampling, systematic, cell and sieve sampling were compared in terms of the reliability, tightness and precision of upper bound estimates of the total error amount. The design effects of systematic,cell and sieve were calculated for each error rate, taint size, sample size and bound and compared to the design effect of the sampling methods for the point estimator. It was found that:

(i) The sample selection method appears to have little effect on the reliability of the upper bound estimates of the total error amount for any given error rate, taint size, sample size or bound. Some significant differences exist but these are not substantial in any case.

(ii) The tightness of the estimates are similar for each sample selection method for any given error rate, taint size, sample size and bound. Differences in tightness that exist are due to factors other than the sampling method. Estimates are more conservative in the AON models than in the taint models. The estimates are extremely conservative for samples of size thirty. The Stringer bound is the most conservative and the Moment bound is the tightest in all cases.

(iii) The sampling methods have substantial effects on the precision of the bounds. Systematic sampling leads to the most precise upper bound estimates for any given error rate, taint size, sample size or bound. Cell and sieve sampling give more precise estimates than simple random sampling in audit populations generated from Population 2. Cell sampling is more precise than sieve sampling in most cases. However, sieve sampling is more precise than cell sampling when the sample size is one hundred in populations with large line items (i.e., audit populations generated from Population 2). The moment bound is somewhat less precise than the Cell and Stinger bounds in most cases.

(iv) The results for the empirical comparisons of the precision of the sampling methods using the bound estimates are similar to the theoretical results obtained using the point estimator analysis.

These findings are consistent with those of other studies (e.g. Wurst, Neter and Godfrey (1989a and 1989b), Plante, Neter and Leitch (1985), Dworkin and Grimlund, 1984). New findings relate to (i) the comparative performance of sieve sampling and systematic sampling and (ii) the performance of the Moment bound with the different sampling methods.

(i) The Comparative Performance of Sieve Sampling and Systematic Sampling

Systematic sampling has a higher mean coverage than sieve sampling for each error rate, taint size, sample size and bound in audit populations generated from Populations 1 and 2. While some of these differences are significant, they are not of practical importance since the differences are not substantial in any case.

Systematic sampling is similar to sieve sampling with respect to the tightness for each error rate, taint size, sample size and bound.

Systematic sampling is more precise than sieve sampling for most error rates, taint sizes, sample sizes and bounds in audit populations generated from Populations 1 and 2. The greatest reductions in the variation of systematic sampling over sieve sampling occur in the high error rate populations and for large sample sizes where reductions in variability of over 10% occur.

(ii) The Performance of the Moment Bound with the Different Sampling Methods.

The differential effects of the sampling methods on the Moment bound estimates of the total error amount are similar to the differential effects of the sampling methods on the Stringer and Cell bound estimates of the total error amount. The coverage and the tightness of the Moment bound are not significantly affected by the sampling method for any error rate, taint size, sample size and bound. Systematic sampling gives the most precise estimates and simple random sampling gives the least precise estimates of the total error amount using the Moment bound.

6.6 Final Comments

In choosing between the sampling methods, the auditor needs to consider a number of practical issues in addition to the performance measures. For example, simple random sampling requires that the book value total be known accurately in advance of sampling and this requirement may impede the planning and implementation of the auditing process. In addition, simple random, systematic and cell sampling require that the book value sub totals be cumulated. Also, systematic sampling may lead to biased selection if there are regularities in the error patterns in the population. A practical disadvantage of sieve sampling, which may be an important consideration for the auditor, is that the sample size is not constant.

In Chapters 7 and 8, Lahiri sampling and stabilised sieve sampling of monetary units are proposed as alternatives to the monetary-unit sampling methods currently used in practice. Lahiri sampling, proposed as an alternative to simple random sampling of monetary units, does not require that the book value total is known accurately in advance of sampling. Stabilised sieve, proposed as an alternative to sieve sampling, is a monetary-unit sampling method which attempts to preserve the advantages of sieve sampling while returning a constant sample size. Lahiri sampling and stabilised sieve sampling are discussed in detail in Chapters 7 and 8 respectively.

Chapter 7

Upper Bound Comparisons of Lahiri Sampling and Simple Random Sampling of Monetary Units

7.1 Introduction

Lahiri sampling, defined in Chapter 4, offers potential advantages to the auditor as an alternative to simple random sampling of monetary units. It uses the line item structure of the population when selecting the monetary units and hence avoids the possible implementation problems of simple random sampling referred to by Wurst, Neter and Godfrey (1989a). Also, unlike simple random sampling, Lahiri sampling does not require that the book value total is known accurately in advance of sampling and the selection procedure may begin before the book value total is known accurately.

This chapter compares the performance of Lahiri sampling with that of simple random sampling with respect to upper bound estimates of the total error amount and investigates the Lahiri monetary-unit sampling method as an alternative to simple random sampling of monetary units.

A comparative investigation of Lahiri and simple random sampling of monetary units is carried out by means of a large scale simulation study using the thirty audit populations created from the two actual accounting populations described in Chapter 3. Samples of sizes 30, 60 and 100 are drawn from each audit population using simple random sampling and Lahiri sampling of monetary units. Upper bounds for the total error amount are obtained using Stringer, Cell and Moment bounds with the taint and the AON error assignments at three nominal confidence levels, .70. .85 and .95. One thousand replications are performed for each sample size and for each sample selection method. Analysis of variance models are constructed to assess the differential effects of Lahiri sampling and simple random sampling of monetary units on bound estimates of the total error amount, for different error rates, taint sizes, sample sizes and bounds.

In the remainder of the chapter, the comparative performance of Lahiri sampling and simple random sampling is measured in terms of the reliability, tightness and precision (see 2.11) of the upper bound estimates of the total error amount (7.2). The design effect of is calculated for each error rate, taint size, sample size and bound(7.3). The practical implications of Lahiri sampling are also discussed (7.4)

7.2 Performance of Lahiri and Simple Random Sampling using Upper Bound Estimates of the Total Error Amount

Reliability, tightness and precision, as defined in 2.11 are used as the performance measures for upper bound estimates of the total error amount with Lahiri sampling and simple random sampling of monetary units. ANOVA models, similar to those detailed in 6.3, are constructed for each set of audit populations, for each nominal confidence level and for each error assignment method with the performance measures used as the dependent variables. The independent variables are identical to those described in Table 6.1 except that the sampling-method independent variable is set at two levels, level 1 for simple random sampling and level 2 for Lahiri sampling.

The ANOVA models, detailed in Appendix D, show that the main effect for the sampling method is significant for the tightness dependent variable in audit populations generated from Population 1 with the AON error assignment and for the precision dependent variable in audit populations generated from Population 2 with both error assignments. The main effect for the sampling method is insignificant in all other models. Tables 7.1 and 7.2 give the significance (at the 95% confidence level) of the first order interactions which include the sampling method for each of the dependent variables for audit populations generated from Population 1 and Population 2 respectively with the taint error assignment.

Table 7.1 Significance of the First-order Interactions of Each Factor with the Sampling Method for Audit Populations generated from Population 1 with the Taint (AON) Error Assignment

	Reliability	Tightness	Precision
Rate	NS (NS)	NS (NS)	NS (NS)
Taint	NS (NS)	NS (NS)	NS (NS)
Samsize	NS (NS)	p<.001 (p<.001)	p<.001 (p<.001)
Bound	NS (NS)	NS (NS)	NS (NS)

Table 7.2 Significance of the First-order Interactions of Each Factor with the Sampling Method for Audit Populations generated from Population 2 with the Taint (AON) Error Assignment

	Reliability	Tightness	Precision
Rate	NS (NS)	NS (NS)	NS (NS)
Taint	NS (NS)	NS (NS)	NS (NS)
Samsize	NS (NS)	p<.001 (p<.001)	p<.001 (p<.001)
Bound	NS (NS)	NS (NS)	NS (NS)

The following observations are made from tables 7.1 and 7.2

(i) The first-order interactions of the sampling method with the line item error rate, bound and taint are insignificant for all the dependent variables.

(ii) The first-order interaction of the sampling method and the sample size is significant for the tightness and the precision dependent variables for each error assignment and each nominal confidence level. It is not significant for the reliability dependent variable.

Similar results prevailed for the 85% and 70% nominal confidence levels. These are given in Appendix D.

Further examination of the sampling method and the sample size interaction terms, in the tightness and precision models, using t-tests to compare the two sampling methods (i.e., SRS and Lahiri) shows that, for any given sample size, the mean tightness or precision of Lahiri sampling is not significantly different from the mean tightness or precision of simple random sampling of monetary units.

In summary, Lahiri sampling is not significantly different from simple random sampling with respect to reliability, tightness and precision for any level of the line item error rate, the taint size, the sample size and the bound.

The results for the models with the taint assignment at the 95% nominal confidence level for each of the populations are reported in 7.2.1 and 7.2.2. The comparative results for the other models (i.e. the taint models at the 85% and 70% nominal confidence level and the AON models) are similar and these are given in Appendix F.

7.2.1 Performance of the Upper Bounds using Lahiri and Simple Random Sampling of Monetary Units in Audit Populations Generated from Population 1

The overall averages for the reliability, tightness and precision dependent variables for audit populations generated from Population 1 with the taint error assignment are given in table 7.3. Each entry in the table is the average of a particular performance measure across all levels of the other factors.

Table 7.3 Average Performance Measures (across all levels of the independent factors) for Audit Population Generated from Population 1 with the Taint Error Assignment at the 95% Nominal Confidence Levels

Performance Measures	Coverage	Tightness	Standard Deviation (000s)
Lahiri	98.69	274.50	98.26
SRS	98.72	276.12	98.74

The results in Table 7.3 indicate that the performance measures, averaged over all error rates, taint sizes, sample sizes and bounds are similar for Lahiri and simple random sampling of monetary units.

A detailed breakdown of the first-order interactions of the sampling methods for each error rate, taint size, sample size and bound are given in Tables 7.4-7.7.

Table 7.4 Average Performance Measures for Each Line Item Error Rate for Audit Population Generated from Population 1 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Rate 1	100.00 (100.00)	620.95 (622.23)	57.74 (56.53)
Rate 2	99.7 (99.7)	309.50 (312.14)	79.42 (81.95)
Rate 3	99.4 (99.3)	217.98 (219.98)	93.47 (96.44)
Rate 4	97.3 (967.41	128.56 (129.73)	123.04 (121.70)
Rate 5	97.1 (97.1)	95.53 (96.64)	137.64 (137.08)

Table 7.5 Average Performance Measures for Each Taint Size for Audit Population Generated from Population 1 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Taint 1	98.95 (98.99)	399.00 (301.73)	97.13 (96.50)
Taint 2	98.89 (98.91)	289.47 (287.79)	98.32 (97.87)
Taint 3	98.21 (98.25)	235.71 (237.15)	101.00(101.21)

Table 7.6 Average Performance Measures for Each Sample Size for Audit Population Generated from Population 1 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation
n = 30	99.52 (99.66)	421.22 (430.20)	132.85 (135.47)
n = 60	98.18 (98.34)	239.68 (238.22)	93.61 (92.60)
n = 100	98.35 (98.15)	161.92 (159.92)	68.33 (68.14)

Table 7.7 Average Performance Measures for Each Bound for Audit Population Generated from Population 1 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Stringer	99.22 (99.27)	291.27 (294.48)	96.63 (96.23)
Cell	99.03 (99.02)	290.31 (292.23)	97.01 (97.41)
Moment	97.76 (97.91)	238.71 (239.86)	101.55 (102.17)

The results from tables 7.4-7.7 indicate that the sampling method has no effect on the reliability, tightness and precision of the upper bound estimates for any given error rate, taint size, sample size and bound. The averages are practically identical in each case.

7.2.2 Performance of the Upper Bounds using Lahiri and Simple Random Sampling of Monetary Units in Audit Populations Generated from Population 2

The overall average coverage, tightness and standard deviation for audit populations generated from Population 2 are given in Table 7.8. Each entry in the table is the average of a particular performance measure across all levels of the other factors.

Table 7.8 Average Performance Measures for Audit Population Generated from Population 2 with the Taint Error Assignment at the 95% Nominal Confidence Levels

Performance Measures	Coverage	Tightness	Standard Deviation (000s)
Lahiri	98.34	187.20	148.00
SRS	98.55	191.05	145.86

The results in Table 7.8 indicate that the coverage, tightness and standard deviation averaged over all error rates, taint sizes, sample sizes and bounds are similar for Lahiri sampling and simple random sampling of monetary units. The performance measures of the other models (i.e. the taint models at the 85% and 70% nominal confidence levels and the AON models at the 95%, 85% and 70% nominal confidence levels) are also similar for the two sampling methods. These are given in the Appendix F.

A detailed breakdown of the first-order interactions of the sampling methods for each error rate, taint size, sample size and bound are given in Tables 7.9-7.11 for the audit populations generated from Population 2 with the taint error assignment at the 95% nominal confidence level.

Table 7.9 Average Performance Measures for Each Error Rate for Audit Population Generated from Population 2 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Rate 1	99.8 (99.8)	309.76 (316.49)	103.96 (107.44)
Rate 2	99.8 (99.8)	246.57 (252.27)	119.00 (118.66)
Rate 3	98.4 (98.8)	181.87 (187.84)	139.52 (137.04)
Rate 4	96.9 (97.2)	113.32 (114.26)	177.06 (173.21)
Rate 5	96.8 (97.1)	84.33 (84.38)	200.50 (192.93)

Table 7.10 Average Performance Measures for Each Taint Size for Audit Population Generated from Population 2 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Taint 1	98.34 (98.62)	191.72 (195.41)	146.53 (144.33)
Taint 2	98.27 (97.83)	186.91 (190.79)	147.77 (145.46)
Taint 3	98.41 (98.52)	182.89 (186.94)	149.71 (147.78)

Table 7.11 Average Performance Measures for Each Sample Size for Audit Population Generated from Population 2 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
n = 30	99.16 (99.26)	280.00 (278.64)	197.70 (195.34)
n = 60	98.45 (98.34)	163.84 (173.63)	139.10 (136.57)
n = 100	97.40 (98.15)	117.65 (120.87)	107.14 (105.67)

Table 7.12 Average Performance Measures for Each Bound for Audit Population Generated from Population 2 with the Taint Error Assignment at the 95% Nominal Confidence Level

Lahiri (SRS)	Coverage	Tightness	Standard Deviation (000s)
Stringer	98.68 (98.87)	192.10 (195.96)	143.40 (141.13)
Cell	98.62 (98.82)	191.34 (195.14)	143.53 (141.23)
Moment	97.72 (97.95)	178.06 (182.05)	157.08 (155.20)

The following observations are made from tables 7.10-7.12 for the taint model generated from Population 2 at the 95% nominal confidence level.

(i) The mean coverage of Lahiri and simple random sampling of monetary units is similar for each error rate, taint size, sample size and bound.

(ii) The mean tightness of Lahiri sampling is somewhat less than the mean tightness with simple random sampling in most cases but this is not significant in any case.

(iii) While no significant differences in the standard deviation exist between the two sampling methods for any error rate, taint size, sample size or bound, there is a tendency for the standard deviation to be somewhat higher with Lahiri sampling than with simple random sampling.

7.2.3 Summary

The performance of Lahiri sampling is similar to the performance of simple random sampling in all audit populations generated from Population 1. However, in audit populations generated from Population 2, Lahiri sampling is somewhat less precise and tighter than simple random sampling for each error rate, taint size, sample size and bound. While these differences are not significant, they appear to be consistent across all levels of the independent factors.

7.3 The Design Effect of Lahiri Sampling

A detailed analysis is carried out on the precision of Lahiri sampling compared to simple random sampling of monetary units. The design effect (see 2.7.1) of Lahiri sampling is calculated for each error rate, taint size, sample size and bound. The results are reported in Tables 7.13 and 7.14 for audit populations generated from Population 1 and 2 respectively, with the taint error assignment at the 95% nominal confidence level. The design effect for the bounds with the taint error assignment at the 85% and 70% nominal confidence levels and the design effect for the models with the AON error assignments are similar to the results reported in tables 7.11 and 7.12 and are given in Appendix I.

Table 7.13 Design Effects of Lahiri Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	1.00	1.04	1.00	1.12	1.12	1.21	0.99	0.99	0.97
Taint 2	1.00	1.00	1.02	1.12	1.12	1.20	0.99	0.99	0.97
Taint 3	1.04	1.05	1.01	1.10	1.10	1.14	1.00	1.01	0.97
Error Rate 2									
Taint 1	0.92	0.92	0.88	1.01	1.02	1.04	0.88	0.89	0.89
Taint 2	0.92	0.92	0.91	1.01	1.01	1.03	0.88	0.85	0.89
Taint 3	0.91	0.91	0.93	1.00	1.01	1.02	0.89	0.89	0.89
Error Rate 3									
Taint 1	0.88	0.88	0.82	1.04	1.04	1.05	0.95	0.95	0.97
Taint 2	0.88	0.87	0.83	1.04	1.04	1.05	0.95	0.95	0.97
Taint 3	0.88	0.88	0.90	1.02	1.02	1.03	0.95	0.95	0.98
Error Rate 4									
Taint 1	1.00	1.00	0.96	1.00	1.00	1.00	1.10	1.10	0.89
Taint 2	1.01	1.01	0.97	1.00	1.00	0.98	1.10	1.10	1.10
Taint 3	1.03	1.02	1.03	0.99	0.99	0.98	1.10	1.10	1.10
Error Rate 5									
Taint 1	1.00	1.00	0.95	0.99	0.99	1.00	1.05	1.05	1.07
Taint 2	1.01	1.01	0.97	0.99	0.99	0.99	1.05	1.03	1.05
Taint 3	1.03	1.03	1.02	0.98	0.98	0.99	1.04	1.04	1.05

Table 7.14 Design Effects of Lahiri Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.92	0.92	0.87	0.99	0.99	0.92	0.97	0.96	0.94
Taint 2	0.92	0.92	0.87	1.00	1.00	0.96	0.94	0.94	0.93
Taint 3	0.92	0.92	0.87	1.00	1.00	0.96	0.98	0.98	0.98
Error Rate 2									
Taint 1	1.01	1.01	1.01	1.02	1.02	0.97	0.98	0.98	0.98
Taint 2	1.01	1.01	1.01	1.04	1.04	1.04	0.98	0.98	0.98
Taint 3	1.01	1.01	1.01	1.04	1.03	1.01	0.98	0.98	0.98
Error Rate 3									
Taint 1	1.02	1.02	1.02	1.08	1.08	1.07	1.02	1.02	1.03
Taint 2	1.04	1.04	1.05	1.06	1.07	1.07	1.00	1.01	1.01
Taint 3	1.05	1.05	1.06	1.04	1.04	1.02	0.99	1.00	1.00
Error Rate 4									
Taint 1	1.04	1.04	1.02	1.05	1.05	1.04	1.07	1.07	1.06
Taint 2	1.04	1.04	1.04	1.04	1.04	1.03	1.06	1.06	1.06
Taint 3	1.04	1.04	1.04	1.04	1.04	1.04	1.06	1.06	1.06
Error Rate 5									
Taint 1	1.10	1.10	1.08	1.06	1.06	1.04	1.11	1.11	1.11
Taint 2	1.10	1.10	1.07	1.06	1.07	1.07	1.08	1.08	1.07
Taint 3	1.06	1.06	1.06	1.06	1.06	1.06	1.09	1.09	1.09

The following observations are made from Tables 7.13 and 7.14.

(i) The design effect of Lahiri sampling is near one in most cases for both populations.

(ii) Fluctuations exist in the design effect in both sets of audit populations but the fluctuations show no consistent relationship with respect to sample size, taint size or error rate for any bound estimate. This is consistent with the theoretical results obtained using the point estimator in Chapter 5.

7.4 Practical Comparisons of Lahiri Sampling and Simple Random Sampling

There are practical aspects of Lahiri sampling and simple random sampling of monetary units which should be considered when deciding between the two selection methods. These relate to the preparation necessary before selection can begin, the amount of sampling necessary to achieve the desired sample size and the number of distinct line items selected with each selection method.

7.4.1 Preparation Prior to Sampling

Leslie, Teitlebaum and Anderson (1979) cite some of the practical disadvantages of simple random sampling of monetary units.

> 'The first is the minor nuisance of having to sort the selected random numbers. The second and more important is the need to know the population value accurately *before* selection can begin'.

Regarding the total book value amount, Leslie, Teitlebaum and Anderson (1979) also point out that

> 'often this value is not known accurately during the planning stage; nor is it known for transaction streams prior to the end of the year'.

In addition, simple random sampling ignores the line item structure of the population when selecting the sample and treats the population as a collection of monetary units from which a simple random sample of monetary units is selected. The selected monetary units must be traced back to the associated line items and this may cause implementation problems (Wurst, Neter and Godfrey, 1989a).

Lahiri sampling, on the other hand, uses the line item structure when selecting the monetary units and does not require the accumulation of the book amounts or necessitate that the total book amount be known accurately in advance of sampling. Therefore, Lahiri selection procedure may begin before the total book value amount is known accurately. The use of the line item structure to select the monetary units avoids possible implementation problems associated with tracing monetary units to associated line items.

7.4.2 The Amount of Sampling

One possible drawback of Lahiri sampling is that the number of selections necessary to obtain a sample of size **n** is usually greater than **n**. The amount of sampling required to obtain a sample of size **n** varies depending on the line item structure of the population. If the largest line item in the population is very much larger than the smaller ones, Lahiri sampling may involve many rejections and hence a lot of sampling before the sample of size **n** is selected. In theorem 4.12, the expected number of trials to obtain a sample of size **n** was shown to be

$$\frac{nNB_{max}}{B}$$

Table 7.15 gives the average number of trials to obtain the required samples of size 30, 60 and 100 from each population using Lahiri sampling.

Table 7.15 Amount of Sampling Required in Lahiri Sampling

Sample Size	Population 1	Population 2
n = 30	1101	175
n = 60	2201	352
n = 100	3668	586

It is clear from Table 7.15 that the amount of sampling required to obtain any given sample size is far greater than the actual sample sizes especially when sampling from Population 1 where the largest line item is very much larger than the smaller ones. More than likely, however, this repeated selection process will not pose serious difficulties if the selection procedure is computerised.

7.4.3 The Number of Distinct Line Items

It was shown in Chapter 4 that simple random sampling and Lahiri sampling are selection methods which may select more than one unit within each line item in any one sample. It was proved in theorems 4.4 and 4.14, that the mean number of times the line item i is included in the sample is

$$\frac{nB_i}{B}$$

for simple random and Lahiri selection methods respectively.

The variance of the number of times each line item is selected in any one selection was shown in theorem 4.4 to be

$$\frac{nB_i}{B}(1 - \frac{B_i}{B})$$

for simple random sampling, and in theorem 4.14 it was shown to be

$$\frac{nB_i}{B}(1 - \frac{B_i}{NB_{max}})$$

for Lahiri sampling

Obviously

$$\frac{nB_i}{B}(1 - \frac{B_i}{NB_{max}}) \geq \frac{nB_i}{B}(1 - \frac{B_i}{B})$$

The statistics of the distribution of the number of distinct items obtained for each sample size using Lahiri and simple random sampling in the 1000 replications are given in tables 7.16 and 7.17 for each population respectively.

Table 7.16 Number of Distinct Line Items obtained for Lahiri (Simple Random) Sampling of for each Nominal Sample Size drawn from Population 1

Lahiri (SRS)	Mean	Std Dev	Minimum	Maximum
n = 30	29.23 (29.25)	0.87 (0.87)	26 (26)	30 (30)
n = 60	57.18 (57.23)	1.66 (1.59)	50 (50)	60 (60)
n = 100	92.29 (92.20)	2.63 (2.69)	83 (81)	98 (99)

Table 7.17 Number of Distinct Line Items obtained for Lahiri (Simple Random) Sampling for each Nominal Sample Size drawn from Population 2

Lahiri (SRS)	Mean	Std Dev	Minimum	Maximum
n = 30	28.00 (28.06)	1.30 (1.30)	23 (24)	30 (30)
n = 60	52.62 (52.59)	2.35 (2.39)	45 (45)	58 (59)
n = 100	80.83 (80.84)	3.47 (3.45)	67 (70)	92 (90)

From Tables 7.16 and 7.17, it can be seen that there is very little difference in the number of distinct line items chosen with either selection method.

7.4.4 Summary

In summary, the main practical advantage of Lahiri sampling compared to simple random sampling of monetary units is that Lahiri sampling uses the line item structure of the population when selecting the sample of monetary units. Simple random sampling of monetary units requires that the random numbers identifying the selected monetary units be related to the corresponding line items to which the selected sample monetary units belong because individual monetary units cannot be audited. The need to identify the line items selected for auditing with simple random sampling of monetary units may at times create some practical implementation problems (Wurst, Neter and Godfrey, 1989a)

A disadvantage of Lahiri sampling compared to simple random sampling of monetary units is that a large number of selections may be necessary before the desired sample size is obtained.

7.5 Conclusions

This chapter investigated the comparative performance of Lahiri sampling and simple random sampling for obtaining upper error bounds. No significant differences between the sampling methods were found in the reliability, tightness and precision of the bounds in any of the models in either Population 1 or Population 2.

The main practical advantage of Lahiri sampling compared to simple random sampling of monetary units is that Lahiri sampling relates monetary units to line items in a natural way and therefore avoids the possible implementation problems referred to by Wurst, Neter and Godfrey (1989a). Also, unlike simple random sampling, Lahiri sampling does not require that the book value total be known accurately in advance of sampling and this enables Lahiri selection to begin before an accurate total book value amount is available.

Chapter 8

A Comparison of Stabilised Sieve Sampling with Sieve Sampling and Simple Random Sampling of Monetary Units

8.1 Introduction

This chapter compares the performance of stabilised sieve sampling with that of simple random and sieve sampling of monetary units with respect to upper bound estimates of the total error amount.

Stabilized sieve sampling, defined in Chapter 4, is a new monetary-unit sampling method proposed in this study as an alternative to simple random sampling and sieve sampling. It preserves the main advantages of sieve sampling while overcoming its primary disadvantage of producing a sample size which may not be equal to the nominal.

In Chapter 4, it was shown that stabilised sieve sampling is a monetary-unit sampling method which selects the line items with probabilities proportional to their book value amounts. In Chapter 5, the theoretical properties of a point estimator of the total error amount were derived for stabilised sieve sampling.

To compare the performance of stabilised sieve sampling with simple random and sieve sampling with respect to the upper bound estimates of the total error amount, a large scale simulation study is carried our using the thirty audit populations created from the two actual accounting populations described in Chapter 3. Samples of sizes 30, 60 and 100 are drawn from each audit population using simple random, sieve and stabilised sieve sampling of monetary units. The Stringer, Cell and Moment bounds for the total error amount are calculated with the taint and AON error assignments at three nominal confidence levels, .70. .85 and .95. One thousand replications are performed for each sample size and for each sample selection method. Analysis of variance models are constructed to assess the differential effects of the sampling methods on bound estimates of the total error amount, for different line item error rates, taint sizes, sample sizes and bounds.

In the remainder of the chapter, the comparative performance of bound estimates with stabilised sieve sampling relative to simple random and sieve sampling is measured in terms of the reliability, tightness and precision of the error estimates (8.2). The design effect of stabilised sieve sampling and the efficiency of stabilised sieve sampling relative to sieve sampling are investigated for each error rate, taint size, sample size and bound (8.3). The practical implications of stabilised sieve sampling are also discussed (8.4).

8.2 Performance of Stabilised Sieve Sampling compared to Simple Random and Sieve Sampling using Upper Bound Estimates of the Total Error Amount

Reliability, tightness and precision, as defined in 2.11, are used as the performance measures for upper bound estimates of the total error amount. ANOVA models, similar to those detailed in 6.3, are constructed for each set of audit populations, for each nominal confidence level and for each error assignment method with the coverage, tightness and standard deviation of the upper bound estimates as the dependent variables. The independent variables are identical to those described in table 6.1 except that the sampling method independent variable is set at three levels, level 1 for simple random sampling, level 2 for sieve sampling and level 3 for stabilised sieve sampling of monetary units. The ANOVA tables are given in the Appendix D.

The main effects of the sampling methods are compared. First-order interactions which include the sampling method are investigated and when they are found to be significant, Dunnett's multiple comparisons test of means with a control is used to compare stabilised sieve sampling with simple random and sieve sampling of monetary units. A family significance level of 0.05 is used in the hypothesis tests.

The results of the analysis using the ANOVA models with the taint assignment at the 95% nominal confidence level for each of the populations are reported. The comparative results for the other models (i.e. the taint models at the 85% and 70% nominal confidence level and the AON models) are similar and these are given in Appendix G.

8.2.1 The Reliability of Stabilised Sieve Sampling Compared to Simple Random Sampling and Sieve Sampling of Monetary Units.

The main effects of the sampling method and the first-order interactions of the sampling method with each of the other factors are analysed for the reliability dependent variable.

8.2.1.1 The Main Effects of the Reliability Dependent Variable

The reliability main effects for each sampling method, for each model are given in tables 8.1 and 8.2.

Table 8.1 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.72	99.01	94.67	93.88	84.94	83.86
Sieve	98.63	98.84	95.38	93.70	84 .74	83.88
Stabilised Sieve	98.91	99.09	94.98	94.30	85.42	84.41

Table 8.2 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.55	98.71	92.87	93.06	80.64	79.96
Sieve	99.19	99.28	95.13	95.18	83.52	82.84
Stabilised Sieve	99.41	99.45	95.94	96.03	85.04	85.14

The following general observations are made from these tables.

(i) Like simple random and sieve sampling, stabilised sieve sampling has a mean coverage above the nominal in all cases.

(ii) Stabilised sieve sampling has a higher coverage than simple random and sieve sampling of monetary units in all cases.

8.2.1.2 The First-Order Interactions of the Reliability Dependent Variable.

A detailed breakdown of the first-order interaction of sampling method with each of the other factors for the reliability dependent variable with the taint error assignment at the 95% nominal confidence level is given in tables 8.3 - 8.10.

Table 8.3 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	100.00	99.72	99.29	97.44	97.13
Sieve	100.00	99.73	99.36	96.77	97.27
Stabilised Sieve	100.00	99.70	99.29	97.96	97.61

Table 8.4 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	99.80	99.80	98.84	97.21	97.09
Sieve	99.94	99.91	99.44	98.70	97.97
Stabilised Sieve	99.96	99.94	99.57	98.90	98.67

Table 8.5 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Taint	1	2	3
SRS	98.99	98.91	98.25
Sieve	98.94	98.85	98.08
Stabilised Sieve	98.21	99.13	98.39

Table 8.6 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	1	2	3
SRS	98.62	98.51	98.52
Sieve	99.22	99.18	99.18
Stabilised Sieve	99.44	99.40	99.39

Table 8.7 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n= 30	n= 60	n = 100
SRS	99.66	98.34	98.15
Sieve	99.53	97.28	98.07
Stabilised Sieve	99.61	98.52	98.60

Table 8.8 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n= 30	n= 60	n = 100
SRS	99.26	98.73	97.65
Sieve	99.24	99.03	99.31
Stabilised	99.56	99.08	99.57

Table 8.9 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	99.22	99.02	97.91
Sieve	99.22	98.99	97.67
Stabilised	99.45	99.24	98.08

Table 8.10 Mean Coverage of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	98.87	98.82	97.95
Sieve	99.40	99.39	98.78
Stabilised Sieve	99.58	99.57	99.08

From tables 8.3 to 8.10, the following general observations are made

(i) Stabilised sieve sampling has a mean coverage above the nominal for each error rate, taint size, sample size and bound.

(ii) The mean coverage of stabilised sieve sampling is similar to that of simple random and sieve sampling for each error rate, taint size, sample size and bound. Some significant differences exist notably in audit populations generated from Population 2 where stabilised sieve sampling has a significantly higher mean coverage than simple random sampling for some factor levels. However, these are not of practical importance since the differences are not substantial in any case.

8.2.2 The Tightness of Stabilised Sieve Sampling compared to Simple Random and Sieve Sampling of Monetary Units.

The main effects of the sampling method and the first-order interactions of the sampling method with each of the other factors are analysed for the tightness dependent variable.

8.2.2.1 The Comparative Tightness of Simple Random, Sieve and Stabilised Sieve Sampling

The tightness main effects for each sampling method for each model are given in tables 8.11 and 8.12

Table 8.11 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	276.12	287.27	175.76	178.22	105.13	103.33
Sieve	273.86	290.19	173.52	180.80	103.78	105.57
Stabilised Sieve	271.85	287.43	171.32	178.34	101.75	103.40

Table 8.12 Mean Tightness of Simple Random, Sieve and Stabilised Sieve for Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	191.05	191.42	118.83	118.58	67.65	67.02
Sieve	191.15	190.19	118.70	117.25	67.28	65.77
Stabilised Sieve	193.26	196.99	120.65	123.44	69.01	71.21

The following observations are made from these tables 8.11 and 8.12

(i) The mean tightness of stabilised sieve sampling is less than or similar to that of simple random and sieve sampling in audit populations generated from Population 1.

(ii) In audit populations generated from Population 2, stabilised sieve sampling is more conservative than simple random and sieve sampling but the difference in the mean tightness between stabilised sieve sampling and the other sampling methods is not sufficiently large to be of practical importance in the audit setting.

8.2.2.2 The First-Order Interactions of the Tightness Dependent Variable.

A detailed breakdown of the first-order interaction of sampling method with each of the other factors for the tightness dependent variable with the taint error assignment at the 95% nominal confidence level is given in tables 8.13 - 8.20.

Table 8.13 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	622.29	312.14	219.78	129.73	96.64
Sieve	617.10	310.32	216.57	129.34	95.98
Stabilised	610.21	306.50	214.31	129.81	97.07

Table 8.14 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	316.495	252.27	187.84	114.26	84.38
Sieve	318.61	252.30	186.24	114.78	83.84
Stabilised	314.37	254.65	189.00	118.94	89.35

Table 8.15 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Taint	Taint 1	Taint 2	Taint 3
SRS	301.73	289.47	237.15
Sieve	299.18	287.04	235.38
Stabilised	297.14	284.92	232.68

Table 8.16 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	Taint 1	Taint 2	Taint 3
SRS	195.41	190.79	186.94
Sieve	195.78	190.95	186.73
Stabilised	197.75	193.15	189.02

Table 8.17 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sampling Size	n = 30	n = 60	n = 100
SRS	430.22	238.22	151.91
Sieve	424.61	237.08	159.90
Stabilised Sieve	417.09	235.61	162.04

Table 8.18 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sampling Size	n = 30	n = 60	n = 100
SRS	278.64	173.63	120.87
Sieve	282.45	170.64	120.37
Stabilised Sieve	285.37	172.00	122.41

Table 8.19 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	296.27	292.22	239.86
Sieve	294.36	290.13	237.10
Stabilised Sieve	291.64	287.54	235.56

Table 8.20 Mean Tightness of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	195.96	195.14	182.04
Sieve	196.42	195.67	181.37
Stabilised Sieve	198.30	197.52	183.96

From tables 8.13 -8.20, the following observations are made

(i) The mean tightness of stabilised sieve sampling is similar to that of simple random and sieve sampling for any given error rate, taint size, sample size or bound in audit populations generated from Population1.

(ii) In audit populations generated from Population 2, stabilised sieve sampling is somewhat more conservative than the other two sampling methods in all cases but the mean tightness of stabilised sieve sampling is not significantly different than the mean tightness of simple random or sieve sampling in any case.

8.2.3 Comparison of the Precision of Stabilised Sieve Sampling with Simple Random and Sieve Sampling of Monetary Units.

The main effects of the sampling methods and the first-order interactions of the sampling method with each level of the other factors for the precision dependent variable are analysed.

8.2.3.1 The Main Effects of the Precision Dependent Variable

The precision main effects for each sampling method for each model are given in terms of the standard deviation of the estimate in tables 8.21 and 8.22.

Table 8.21 Mean Standard Deviation (000s) of Simple Random, Sieve and Stabilised Sieve Sampling for Audit Populations generated from Population 1

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	98.74	102.46	87.40	91.46	77.42	81.25
Sieve	99.43	104.95	88.12	93.76	78.14	83.36
Stabilised Sieve	96.59	100.84	85.43	89.94	75.63	79.83

Table 8.22 Mean Standard Deviation(000s) of Simple Random, Sieve and Stabilised Sieve Sampling for Audit Populations generated from Population 2

Confidence Level	95%		85%		70%	
Assignment	Taint	Aon	Taint	Aon	Taint	Aon
Sampling Methods						
SRS	145.86	146.50	130.79	131.73	116.99	117.96
Sieve	133.93	133.77	120.15	120.26	107.46	107.66
Stabilised Sieve	132.86	136.67	119.21	123.16	106.66	110.53

As can be seen from tables 8.21 and 8.22

(i) Stabilised sieve sampling is more precise than simple random sampling in all the models. The greatest improvements in precision of stabilised sieve sampling over simple random sampling occur in audit populations generated from Population 2. This is consistent with the point estimator analysis in Chapter 5.

(ii) The precision of stabilised sieve sampling is similar to the precision of sieve sampling in all the models.

8.2.3.2 The First-Order Interactions of the Precision Dependent Variable

A detailed breakdown of the precision of the sampling methods with each value of the error rate, taint, sample size and bound with the taint error assignment at the 95% nominal confidence level is given in tables 8.23 - 8.30.

Table 8.23 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	56.53	81.95	96.44	121.70	137.08
Sieve	56.73	80.17	96.41	123.91	139.92
Stabilised Sieve	55.15	78.60	93.81	120.22	135.18

Table 8.24 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Line Item Error Rate at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Line Item Error Rate	1	2	3	4	5
SRS	107.44	118.66	137.04	173.21	192.93
Sieve	90.08	104.63	126.59	161.08	187.30
Stabilised Sieve	92.40	107.28	128.91	157.73	177.98

Table 8.25 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Taint	1	2	3
SRS	97.13	97.87	101.21
Sieve	97.39	98.29	102.62
Stabilised Sieve	94.92	95.67	99.10

Table 8.26 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Taint Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Taint	1	2	3
SRS	143.33	135.46	147.78
Sieve	132.72	133.60	135.47
Stabilised Sieve	131.68	132.60	134.29

Table 8.27 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30	n = 60	n = 100
SRS	135.47	92.60	68.14
Sieve	136.48	92.84	68.97
Stabilised Sieve	130.88	91.34	67.56

Table 8.28 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Sample Size at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30	n = 60	n = 100
SRS	195.93	136.56	105.67
Sieve	194.98	124.78	82.04
Stabilised Sieve	192.78	124.40	81.39

Table 8.29 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	96.63	97.41	102.17
Sieve	97.51	98.25	102.51
Stabilised Sieve	94.40	95.12	100.26

Table 8.30 Mean Standard Deviation of Simple Random, Sieve and Stabilised Sieve Sampling for each Bound at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Bound	Stringer	Cell	Moment
SRS	141.13	141.23	155.20
Sieve	129.70	129.82	144.28
Stabilised Sieve	128.52	128.63	141.43

The results of the analysis of the interactions between each of the independent factors and the sampling method for the precision dependent variable confirm the overall trends observed in 8.2.3.1. The following observations are made from tables 8.23 to 8.30

(i) Stabilised sieve sampling is more precise than simple random sampling for most error rates, taints, sample sizes and bounds. In audit populations generated from Population 1, the mean precision of stabilised sieve sampling and simple random sampling are very close in all cases. The greatest gains in the precision of stabilised sieve sampling over simple random sampling occur in audit populations generated from Population 2. This is consistent with the point estimator analysis in Chapter 5.

(ii) The precision of stabilised sieve sampling and the precision of sieve sampling are similar in most cases.

8.3 Efficiency of Stabilised Sieve Sampling relative to Simple Random and Sieve Sampling

A more detailed analysis is carried out on the precision of stabilised sieve sampling compared to simple random and sieve sampling of monetary units. The design effect, defined in Chapter 2, is used to compare stabilised sieve sampling with simple random sampling. The relative efficiency, also defined in Chapter 2, is used to compare stabilised sieve sampling with sieve sampling.

8.3 1 The Design Effect of Stabilised Sieve Sampling

The design effect (see 2.7.1) of stabilised sieve sampling is calculated for each error rate, taint size and sample size and bound. Tables 8.31 and 8.32 give the design effect for each bound for audit populations generated from Populations 1 and 2 respectively with the taint error assignment at the 95% nominal confidence level. The design effect for the bounds with the taint error assignment at the 85% and 70% nominal confidence levels and the design effect for the bounds with the AON error assignments are given in Appendix J.

Table 8.31 Design Effect of Stabilised Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.84	0.84	0.89	1.06	1.06	1.12	0.97	0.97	0.97
Taint 2	0.85	0.84	0.88	1.06	1.06	1.12	0.97	0.97	0.95
Taint 3	0.85	0.84	0.91	1.06	1.06	1.07	0.96	0.96	0.95
Error Rate 2									
Taint 1	0.91	0.91	0.89	0.93	0.94	0.96	0.91	0.90	0.93
Taint 2	0.91	0.91	0.99	0.92	0.93	0.96	0.91	0.91	0.93
Taint 3	0.91	0.91	0.92	0.93	0.93	0.96	0.93	0.93	0.93
Error Rate 3									
Taint 1	0.91	0.91	0.90	0.99	1.00	0.99	0.94	0.94	0.99
Taint 2	0.92	0.92	0.90	0.99	1.00	0.99	0.94	0.94	0.98
Taint 3	0.92	0.91	0.91	0.99	0.99	1.01	0.93	0.91	0.91
Error Rate 4									
Taint 1	0.97	0.97	0.94	0.93	0.93	0.95	1.00	1.00	1.04
Taint 2	0.97	0.97	0.98	0.93	0.93	0.95	1.00	1.00	1.01
Taint 3	0.99	0.99	0.99	0.94	0.94	0.95	1.01	1.01	1.01
Error Rate 5									
Taint 1	0.95	0.95	0.94	0.99	0.97	0.98	1.05	1.06	1.06
Taint 2	0.95	0.95	0.97	0.97	0.97	0.98	1.05	1.06	1.06
Taint 3	0.97	0.98	0.99	0.96	0.96	0.97	1.04	1.05	1.05

Table 8.32 Design Effect of Stabilised Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.93	0.93	0.90	0.67	0.64	0.67	0.47	0.47	0.52
Taint 2	0.93	0.93	0.89	0.73	0.73	0.71	0.48	0.48	0.52
Taint 3	0.93	0.93	0.89	0.73	0.73	0.71	0.49	0.49	0.53
Error Rate 2									
Taint 1	1.00	1.00	1.01	0.78	0.78	0.78	0.55	0.55	0.59
Taint 2	1.00	1.00	1.03	0.79	0.79	0.80	0.56	0.56	0.59
Taint 3	1.00	1.00	1.03	0.80	0.80	0.81	0.58	0.58	0.60
Error Rate 3									
Taint 1	1.09	1.09	1.11	0.83	0.84	0.84	0.62	0.62	0.64
Taint 2	1.09	1.09	1.11	0.84	0.86	0.86	0.62	0.62	0.64
Taint 3	1.08	1.08	1.09	0.82	0.82	0.83	0.62	0.62	0.63
Error Rate 4									
Taint 1	0.94	0.94	0.97	0.85	0.85	0.86	0.60	0.60	0.59
Taint 2	0.94	0.94	0.96	0.89	0.90	0.86	0.59	0.59	0.59
Taint 3	0.92	0.92	0.94	0.86	0.86	0.86	0.59	0.59	0.59
Error Rate 5									
Taint 1	0.95	0.96	0.96	0.89	0.89	0.89	0.64	0.64	0.64
Taint 2	0.95	0.95	0.95	0.89	0.89	0.89	0.64	0.64	0.64
Taint 3	0.91	0.91	0.90	0.89	0.89	0.90	0.64	0.64	0.64

From these tables the following observations are made

(i) The design effect of stabilised sieve sampling is less than one in most cases in audit populations generated from Population 1 but not substantially less than one in any case. This is consistent with the point estimator analysis in Chapter 5.

(ii) In audit populations generated from Population 2, the design effect is substantially less than one in all cases for samples of sizes sixty and one hundred. This is consistent with the point estimator analysis in Chapter 5.

8.3.2 The Efficiency of Stabilised Sieve Sampling Relative to Sieve Sampling

For each bound, the efficiency (see 2.7.1) of stabilised sieve sampling relative to sieve sampling is calculated for each error rate, taint size, sample size. Tables 8.33 and 8.34 give the relative efficiency for each bound in audit populations generated from Populations 1 and 2 respectively with the taint error assignment at the 95% nominal confidence level. The efficiency of stabilised sieve relative to sieve sampling for the models with the taint error assignment at the 85% and 70% nominal confidence levels and for the models with the AON error assignments are given in Appendix J.

Table 8.33 Efficiency of Stabilised Sieve Sampling relative to Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 1 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	0.85	0.85	0.84	1.01	1.01	1.00	1.04	1.04	1.09
Taint 2	0.84	0.85	0.84	1.01	1.01	0.99	1.04	1.05	1.09
Taint 3	0.84	0.84	0.87	1.02	1.02	1.00	1.04	1.04	1.06
Error Rate 2									
Taint 1	0.94	0.94	0.95	0.93	0.93	0.95	1.03	1.03	1.02
Taint 2	0.94	0.94	0.95	0.93	0.93	0.95	1.03	1.03	1.02
Taint 3	0.92	0.93	0.95	0.93	0.93	0.94	1.03	1.05	1.04
Error Rate 3									
Taint 1	0.96	0.96	0.98	0.91	0.91	0.90	0.96	0.96	1.00
Taint 2	0.95	0.95	0.98	0.91	0.91	0.91	0.97	0.97	0.99
Taint 3	0.94	0.94	0.96	0.91	0.91	0.92	0.98	0.98	0.98
Error Rate 4									
Taint 1	0.95	0.95	0.98	0.95	0.95	0.99	0.88	0.87	0.91
Taint 2	0.95	0.95	0.97	0.96	0.96	0.99	0.89	0.89	0.90
Taint 3	0.93	0.93	0.94	0.95	0.95	0.96	0.88	0.88	0.89
Error Rate 5									
Taint 1	0.89	0.89	0.92	1.02	1.02	1.04	0.94	0.94	0.95
Taint 2	0.89	0.88	0.92	1.02	1.02	1.02	0.93	0.93	0.94
Taint 3	0.85	0.85	0.89	0.99	0.99	1.01	0.91	0.91	0.92

Table 8.34 Efficiency of Stabilised Sieve Sampling relative to Sieve Sampling for Bounds at the 95% Nominal Confidence Level for Audit Populations generated from Population 2 with the Taint Error Assignment

Sample Size	n = 30			n = 60			n = 100		
Bound	Str	Cell	Mom	Str	Cell	Mom	Str	Cell	Mom
Error Rate 1									
Taint 1	1.05	1.06	1.08	1.01	1.01	0.98	1.10	1.10	1.07
Taint 2	1.03	1.04	1.10	1.01	1.01	0.97	1.11	1.11	1.07
Taint 3	1.02	1.03	1.10	1.02	1.02	1.00	1.10	1.10	1.05
Error Rate 2									
Taint 1	1.04	1.05	1.07	1.03	1.01	0.99	1.10	1.10	1.07
Taint 2	1.02	1.03	1.08	1.04	1.04	1.00	1.10	1.10	1.09
Taint 3	1.06	1.06	1.09	1.03	1.03	1.01	1.11	1.11	1.08
Error Rate 3									
Taint 1	1.10	1.09	1.13	0.98	0.98	0.98	1.01	1.01	1.02
Taint 2	1.07	1.07	1.11	0.98	0.98	0.98	1.01	1.01	1.01
Taint 3	1.07	1.07	1.10	0.98	0.98	0.98	1.00	1.00	0.99
Error Rate 4									
Taint 1	0.92	0.92	0.96	1.00	1.00	0.99	0.95	0.95	0.96
Taint 2	0.92	0.92	0.96	1.00	1.00	0.99	0.96	0.96	0.96
Taint 3	0.91	0.92	0.95	1.00	1.00	1.00	0.96	0.96	0.97
Error Rate 5									
Taint 1	0.87	0.87	0.88	0.98	0.98	0.98	0.87	0.87	0.88
Taint 2	0.86	0.86	0.88	0.98	0.97	0.98	0.87	0.87	0.88
Taint 3	0.85	0.85	0.86	0.97	0.97	0.98	0.88	0.89	0.88

The following observations are made from these tables.

(i) Stabilised sieve sampling is near one in most cases in both sets of audit populations. It is somewhat more efficient than sieve sampling in most cases in audit populations generated from Population 1 and in some cases in audit populations generated from Population2.

(ii) Fluctuations exist in the efficiency in both sets of audit populations but the fluctuations show no consistent relationship with respect to sample size, taint size or error rate for any bound estimate.

8.4 Practical Aspects of Stabilised Sieve Sampling

Stabilised sieve sampling has some practical advantages over both simple random sampling and sieve sampling of monetary units which should be taken into account when deciding on a sampling method.

Simple random sampling, often referred to as 'unrestricted random sampling' (Leslie, Teitlebaum and Anderson, 1979, p100) considers the population as a collection of monetary units from which a simple random sample of monetary units is chosen. This selection method requires that the book amounts are accumulated and that the total book amount be known in advance of sampling. In addition, with simple random sampling of monetary units it is possible to obtain more than one monetary unit from the same line item. Therefore, while this sampling method returns the desired sample size of **n** monetary units, the number of distinct line items from which these monetary units are chosen may be less than **n**.

Sieve sampling, on the other hand, uses the line item structure of the population when selecting the sample and hence does not require that the book amounts be accumulated or that the total book amount be known accurately in advance of sampling. Sieve sampling has a further advantage over simple random sampling in that it selects monetary units from distinct line items. However, a drawback of sieve sampling, is that the achieved sample size in any selection may not be equal to the nominal sample size. With sieve sampling, the sample size is a variable and this may be of serious concern to the auditor when deciding on a sample selection method.

Stabilised sieve sampling attempts to preserve the advantages of sieve sampling. It uses the line item structure of the population when selecting the sample and it does not require that the book amounts be accumulated or that the total book amount be known accurately in advance of sampling. In addition, stabilised sieve sampling overcomes the primary disadvantage of sieve sampling by returning a constant sample size of monetary units.

However, while sieve sampling selects the monetary units from distinct line items, with stabilised sieve sampling as with simple random sampling, more than one monetary unit may be selected from any specific line item. The selection process for stabilised sieve sampling, detailed in 4.8, selects the sample in two stages. At the first stage, the monetary units are selected from distinct line items. Repeated selections of monetary units from the same line item is restricted to the second stage of selection in stabilised sieve sampling. When the sample needs to be reduced, the final sample is a set of monetary units selected from **n** distinct line items. When the sample is augmented, it is necessary, in order to preserve PPS, to allow the line items selected at the first stage to be considered for selection at the second stage (see Chapter 4). Since the possibility of multiple selections of monetary units from the same line item is restricted to the second stage selection of the stabilised sieve sample, the number of multiple selections in stabilised sieve sampling should be less than that of simple random sampling of monetary units where all the monetary units selected are considered for inclusion in the sample at each selection.

An investigation into the number of distinct line items obtained using simple random, sieve and stabilised sieve sampling is carried out. Samples of sizes 30, 60 and 100 are drawn from Populations 1 and 2 using each selection method. One thousand replications are performed for each sample size and each sampling method. Tables 8.35-8.37 give the number of distinct line items achieved using simple random, sieve and stabilised sieve sampling, with nominal sample sizes of 30, 60 and 100 respectively.

Table 8.35 Number of Distinct Line Items Selected using Stabilised Sieve Sampling, Sieve Sampling and Simple Random Sampling for a Nominal Sample Size of Thirty

n = 30	Mean	Standard Deviation	Minimum	Maximum
Population 1				
SRS	29.25	0.87	26	30
Sieve	30.07	5.53	15	45
Stabilised	29.99	0.08	29	30
Population 2				
SRS	28.06	1.29	24	30
Sieve	30.09	4.74	16	45
Stabilised	29.91	0.31	28	30

Table 8.36 Number of Distinct Line Items Selected using Stabilised Sieve Sampling, Sieve Sampling and Simple Random Sampling for a Nominal Sample Size of Sixty

n = 60	Mean	Standard Deviation	Minimum	Maximum
Population 1				
SRS	57.23	1.59	50	60
Sieve	60.21	7.66	36	85
Stabilised	59.96	0.21	57	60
Population 2				
SRS	52.29	2.39	49	59
Sieve	60.15	6.60	41	85
Stabilised	59.75	0.58	56	60

Table 8.37 Number of Distinct Line Items Selected using Stabilised Sieve Sampling, Sieve Sampling and Simple Random Sampling for a Nominal Sample Size of One Hundred

n = 100	Mean	Standard Deviation	Minimum	Maximum
Population 1				
SRS	92.20	2.69	81	99
Sieve	99.78	9.07	74	125
Stabilised	99.90	0.34	97	100
Population 2				
SRS	80.84	3.45	70	90
Sieve	99.82	7.44	78	129
Stabilised	99.49	0.97	94	100

The following observations are made from these tables

(i) In audit populations generated from Population 1, each selection method has a mean number of distinct line items near to the nominal sample size for all sample sizes. In audit populations generated from Population 2, sieve sampling and stabilised sieve sampling have a mean number of distinct line items near the nominal in all cases. The mean number of distinct line items selected with simple random sampling is below the nominal in all cases in audit populations generated from Population 2 and substantially less than the nominal for samples of size 100.

(ii) The variability in the number of distinct line items is greatest for sieve sampling and least for stabilised sieve sampling for all nominal sample sizes. The variability of the number of distinct line items obtained using sieve sampling is lower in Population 2 than in Population 1. It should be noted that the variability of the achieved sample size in sieve sampling was derived theoretically in theorem 4.18 where it was shown that the achieved sample size in sieve sampling has a standard deviation of

$$\sigma_{n_0} = \sqrt{n(1 - \frac{n}{B^2}\sum_{i=1}^{N} B_i^2)}$$

For samples drawn from Population 1 using sieve sampling, the variability of the achieved sample size is

$$\sigma_{30} = 5.33 \quad \sigma_{60} \ 7.32 \quad \sigma_{100} = 9.07$$

for samples of nominal sizes 30, 60 and 100 respectively.

And for samples drawn from Population 2 using sieve sampling, the variability of the achieved sample size is

$$\sigma_{30} = 5.07 \quad \sigma_{60} = 6.56 \quad \sigma_{100} = 7.27$$

for samples of sizes 30, 60 and 100 respectively.

The simulated results are consistent with the exact results.

(iii) The number of distinct line items obtained with sieve sampling varies greatly from sample to sample. The minimum achieved sample size with sieve sampling is substantially lower than the nominal and the maximum achieved sample size with sieve sampling is substantially higher than the nominal for all sample sizes.

(iv) The minimum and maximum number of distinct line items obtained using stabilised sieve sampling is near the nominal sample size for all sample sizes.

(v) The number of distinct line items obtained with stabilised sieve sampling has the smallest range and lowest variability for each nominal sample size.

8.5 Conclusions

In this chapter, the performance of stabilised sieve sampling was compared to simple random and sieve sampling of monetary units in terms of the reliability, tightness and precision of upper bound estimates of the total error amount. The design effects of stabilised sieve sampling and the efficiency of stabilised sieve sampling relative to sieve sampling were calculated for each error rate, taint size, sample size and bound. Some practical aspects of stabilised sieve sampling were also considered. It was found that:

(i) Stabilised sieve sampling has a mean coverage above the nominal for each error rate, taint size, sample size and bound.

(ii) The tightness of the estimates with stabilised sieve sampling is similar to the tightness with simple random and sieve sampling for any given error rate, taint size, sample size and bound in audit population generated from Population 1. In audit populations generated from Population 2, stabilised sieve sampling is somewhat more conservative than simple random and sieve sampling in all cases but the differences in the mean tightness between the sampling methods are not significant in any case.

(iii) Stabilised sieve sampling is more precise than simple random sampling in all the models. The greatest improvements in precision of stabilised sieve sampling over simple random sampling occur in audit populations generated from Population 2. The precision of stabilised sieve sampling is similar to that of sieve sampling in all the models.

(v) The average number of distinct line items obtained using stabilised sieve sampling is similar to the average obtained using sieve sampling in all cases and greater than simple random sampling is most cases.

(vi) Stabilised sieve sampling overcomes the variable sample size problem of sieve sampling and returns a sample size of monetary units equal to the nominal sample size.

(v) The average number of distinct line items obtained using stabilised sieve sampling has a range smaller than the average using simple random sampling and substantially smaller than the average using sieve sampling for all sample sizes. The variability in the number of distinct line items obtained from sample to sample is lower for stabilised sieve sampling than the other two sampling methods.

In conclusion, stabilised sieve sampling is a reliable monetary-unit sampling method. While some patterns of difference were found in the tightness of the bounds with stabilised sieve, simple random and sieve sampling, the differences in the mean tightness between stabilised sieve sampling and simple random sampling or between stabilised sieve sampling and sieve sampling are not significant for any line item error rate, taint size, sample size or bound. Stabilised sieve sampling was found to be more precise than simple random sampling in most cases and significantly more precise than simple random sampling with samples of size 100 in audit populations generated from Population 2. No significant differences in the precision occurred between stabilised sieve sampling and sieve sampling.

Consequently, the decision to use stabilised sieve sampling as an alternative to simple random or sieve sampling must be made on non-statistical grounds. Wurst, Neter and Godfrey (1989b) made the following recommendation to auditors when deciding to use sieve sampling instead of simple random or cell sampling of monetary units.

> 'Auditors will need to consider whether sample selection in the field is facilitated by sieve sampling by determining whether ordering of random numbers and accumulating of book amounts that is required by random or cell selection but not by sieve sampling is of serious concern, and whether the variability of sample size with sieve sampling is of serious concern, as contrasted with fixed sample sizes for random and cell selection.'

With stabilised sieve sampling, these considerations are no longer necessary. Stabilised sieve sampling does not require that the random numbers be ordered or that the book amounts be accumulated. In addition, the number of monetary units obtained using stabilised sieve sampling is always equal to the nominal sample size. Stabilised sieve sampling has the added advantage that the number of distinct line items obtained in a sample is less variable than the number obtained with simple random sampling and sieve sampling. It is therefore a useful and practical alternative to simple random sampling and sieve sampling of monetary units in substantive auditing.

Chapter 9

Summary, Conclusions and Recommendations for Future Research.

9.1 Introduction

This chapter reviews how the objectives stated in Chapter 1 have been achieved (9.2), presents a summary of the findings and draws conclusions from the results (9.3). Some areas of future research are suggested (9.4).

9.2 Achievement of the Objectives

The purpose of this study was to investigate how different monetary-unit sampling methods perform in obtaining estimates of the total error amount in substantive auditing. Six monetary-unit sampling methods were examined. Four of these are currently used in practice, namely simple random, systematic, cell and sieve sampling. One method 'Lahiri sampling' has not been applied previously in auditing. A new monetary-unit sampling method, 'Stabilised Sieve Sampling' has been developed in this study and its properties analysed. The performance of the sampling methods were compared by studying the behaviour of the estimates of the population total error amount given by each sampling method. The study involved a point estimator analysis and an upper bound analysis. It also investigated practical aspects of the sampling methods which must be considered by the auditor when choosing a selection method. The sampling methods were tested on data obtained from commercial entities in the Public Sector in Ireland.

As previously stated in Chapter 1, the specific objectives of this study were:

(i) To obtain information on the characteristics of book values and patterns of errors in two populations of debtors in the Public Sector;

(ii) To carry out a theoretical analysis of a point estimator of the total error amount for six monetary-unit sampling methods. Four methods are currently used in practice (i.e., simple random, systematic, cell and sieve sampling), one 'Lahiri sampling' has not been previously used in auditing (see also objective (iv) below), and a new monetary-unit sampling method has been developed in this study (see also objective v);

(iii) To compare the performance of the Stringer, Cell and Moment bounds for estimating the total error amount in substantive auditing using monetary-unit sampling methods currently used in practice;

(iv) To investigate Lahiri sampling as an alternative to simple random sampling of monetary units;

(v) To investigate stabilised sieve sampling as an alternative to sieve and simple random sampling of monetary units.

9.2.2 The Methodology Used to Achieve the Objectives

Two accounting populations of debtors from commercial entities in the Irish Public Sector were examined. One contained a relatively large number of small accounts and the other contained a relatively small number of large accounts. An experimental study involving large audits of the two accounting populations was carried out and the relationships between different population parameters were identified for each population. The purpose of the audits was to obtain sufficient information on the error patterns in each population to be able to model the errors. The characteristics of the book values, the sample designs used in the investigative audits and the distribution of errors found in each population are outlined in Chapter 3.

Chapter 4 provides a basis for the subsequent analysis of estimation procedures by deriving the properties of the six monetary-unit sampling methods used in the study.

A theoretical analysis of a point estimator of the total error amount was carried out in Chapter 5. The mean and variance of the estimator with the taint and AON error assignment methods (see 2.6) were derived for each monetary-unit sampling method. The design effect (Kish, 1965) was used to compare the precision of the point estimator for systematic, cell, Lahiri, sieve and stabilised sieve sampling relative to simple random sampling of monetary units. The analysis of the point estimator was undertaken in order to gain some insight into the behaviour of the heuristic upper bound estimates of the total error amount for the different sampling methods. Previous research by Wurst, Neter and Godfrey (1989a and 1989b) showed that the effects of simple random, cell and sieve sampling on the precision of the point estimator was similar to the effects of the sampling methods on the precision of the Stringer and Cell bounds. This study extends the work of these authors by investigating the theoretical properties of six monetary-unit sampling methods, including a new sampling method 'Stabilised Sieve Sampling' and one not used previously in auditing 'Lahiri Sampling'.

To assess the effects of the sampling methods on the upper bound estimates of the total error amount, it was necessary to investigate the results obtained when the methods were applied to a spectrum of accounting populations. Fifteen audit populations were generated from each of the two accounting populations by seeding errors into the populations of book values, with different error rates and error sizes reflecting the patterns found in the data obtained from the investigative audits. Details of the error models used to generate errors into the populations and the resultant audit populations are given in Chapter 3. These audit populations were used to compare the performance of the different sampling methods on the upper bound estimates of the total error amount.

A comparative investigation of the sampling methods was carried out by means of a large scale simulation study using the thirty audit populations created from the two actual accounting populations. Samples of sizes 30, 60 and 100 were drawn from each audit population. Upper bounds were calculated using the Stringer, Cell and Moment bounds, with the taint and AON error assignments at three nominal confidence levels, .70, .85 and .95. One thousand replications were performed for each sample size and for each sample design.

Analysis of variance models were constructed to assess the comparative performance of the monetary-unit sampling methods currently used in practice (i.e. simple random, systematic, cell and sieve sampling), using the Stringer, Cell and Moment bounds to estimate the upper bound for the total error amount. The criteria for assessing the performance of the upper bound estimates were defined in Chapter 2 and include reliability, tightness and precision. Tests of significance were applied to the performance measures to investigate differences between the sampling methods.

The simulation results were also used to compare simple random sampling with its proposed alternative Lahiri sampling and to compare simple random sampling and sieve sampling with the proposed alternative stabilised sieve sampling in terms of the bound performance. The comparative performance of Lahiri and simple random sampling of monetary units is given in Chapter 7. The performance of stabilised sieve sampling compared to sieve sampling and simple random sampling is given in Chapter 8.

9.3 Summary of the Findings

9.3.1 Population Characteristics and Error Patterns

9.3.1.1 Findings

The book value characteristics and error patterns of the two populations of debtors from the Irish Public Sector displayed important contrasting features. Population 1 consists of a relatively large number of small debtors while Population 2 consists of a relatively small number of large debtors. The taints are larger and the proportion of 100% taints is greater in Population 2 than in Population 1. The taint size tends to decrease with book value size in Population 1 but this is not the case in Population 2. The difference in book value and error characteristics provided an excellent opportunity of investigating the sampling methods under different conditions.

9.3.1.2 Comparison with other Databases.

The characteristics of the two populations of book values of debtors of commercial entities in the Public Sector, were similar to those found in other studies (for example, Neter and Loebbecke, 1975), in that they are both highly skewed and in each case a small proportion of the line items account for a large proportion of the book value total.

The error patterns found in the investigative audits suggest that errors in debtors from the Public Sector are similar in some respects to populations that have been studied previously and they are different in other respects. The characteristics of errors in Population 1 are similar to the populations studied by Neter and Loebbecke (1975). In Population 2, the mean taint is larger and the proportion of 100% overstatement errors is greater than those found in the Neter and Loebbecke populations. All the errors found in the investigative audits are overstatements. This contrasts with the UK studies (e.g. Johnson, 1987, McRae, 1982) where only a slight bias towards overstatement errors was found. However, studies carried out in the US (e.g. Neter and Loebbecke, 1975; Johnson, Leitch and Neter, 1981) found that errors in debtors were mostly overstatements. Monetary-unit sampling is an appropriate selection method when overstatements predominate because line items are selected with probabilities proportional to their recorded values but it may not be the most effective method of sampling when understatements play a dominant role (e.g. the UK data).

9.3.2 Theoretical Analysis of the Point Estimator

9.3.2.1 Findings

In the point estimator analysis, systematic and cell sampling were found to have design effects of less than or equal to one for both error assignment methods. The design effect of systematic and cell sampling was substantially less than 1 when the population contained large line items.

The variances of the point estimator with Lahiri sampling was shown to be equal to the variance of the point estimator with simple random sampling for both error assignments. Hence, Lahiri sampling has a design effect of one or equivalently Lahiri sampling has the same precision as simple random sampling.

The design effect of sieve sampling, using the point estimator, was similar for both error assignment methods in all cases. It was smaller with low line item error rates than with high line item error rates. In audit populations generated from Population 1 (small line items), the design effect of sieve sampling was near one in most cases. Also, in audit populations generated from Population 1, there was a tendency for sieve sampling to have design effects of less than one in populations with low line item error rates and to have design effects greater than one in populations with high line item error rates. In audit populations generated from Population 2, where the line items are large and the mean error amount per line item is large, the design effect of sieve sampling was less than one for most sample sizes, error rates and taint sizes. The minimum sample size for which the design effect of sieve sampling was less than one increased as the line item error rate and the mean taint size increased.

The variance of the point estimator with stabilised sieve sampling was shown to consist of two components, one due to the initial sieve sample selection and the other due to the reduction or the augmentation process. The component due to sieve sampling has the greatest weight so the design effect of stabilised sieve sampling is similar to the design effect of sieve sampling. The point estimator with stabilised sieve sampling will have a precision greater than that with simple random sampling if the initial gains in precision due to the sieve sample component are large (for example, in the audit populations generated from Population 2). The efficiency of the point estimator with stabilised sieve sampling relative to ordinary sieve sampling will be near one in all cases.

9.3.2.2 Conclusions

The findings suggest that the point estimator of the total error amount is more precise with systematic and cell sampling than with simple random sampling especially in accounting populations with large line items. The findings on sieve sampling concur with the results obtained by Wurst, Neter and Godfrey (1989a) that sieve sampling favours populations with low line item error rates and large line items. Lahiri sampling has the same precision as simple random sampling. The variance of the point estimator with stabilised sieve sampling is similar to the variance of the point estimator with sieve sampling and hence stabilised sieve sampling favours populations with low line item error rates and large line items.

9.3.3 Upper Bound Comparisons of Simple Random, Systematic, Cell and Sieve Sampling of Monetary Units.

9.3.3.1 Findings

In the simulation study comparing the upper bound performances of simple random, systematic, cell and sieve sampling, it was found that the differential effects of the sampling methods on the reliability, tightness and precision of the bounds were similar for bounds with the taint error assignment and the AON error assignment at the three nominal confidence levels.

The sample selection method did not affect the reliability of the upper bound estimates of the total error amount for any given error rate, taint size, sample size or bound. Some significant differences did exist between the sampling methods with respect to reliability but these were not of practical importance in an audit setting since the coverage was above the nominal in all cases.

The tightness of the estimates was similar for each sample selection method for any given error rate, taint size, sample size and bound. Differences in tightness that did exist were due to factors other than the sampling method. For example, upper bound estimates of the total error amount were more conservative in the AON models than in the taint models. The estimates were extremely conservative in populations with low error rates and for small sample sizes. The Stringer bound was the most conservative and the Moment bound was the tightest bound in all cases.

The sampling methods had substantial effects on the precision of the bounds. The most precise upper bound estimates for any given error rate, taint size, sample size or bound were obtained with systematic sampling. Cell and sieve sampling gave more precise estimates than simple random sampling in audit populations generated from Population 2 (large line items). Cell sampling was more precise than sieve sampling in most cases in audit populations generated from Population 1. However, sieve sampling was more precise than cell sampling for samples of sizes 60 and 100 in populations with large line items (i.e. audit populations generated from Population 2). Upper bound estimates of the total error amount were consistently more variable with the AON error assignment than with the taint error assignment. The results of the empirical comparisons of the precision of the sampling methods using the bound estimates were similar to the theoretical results obtained using the point estimator.

The results of the simulation study comparing the upper bound performances of simple random, systematic, cell and sieve sampling of monetary units were consistent with the findings obtained by Plante, Neter and Leitch (1985), Wurst, Neter and Godfrey (1989b) and Dworkin and Grimlund (1984). New findings in this area relate to the comparative performance of sieve sampling and systematic sampling using the upper bound estimates of the total error amount and the performance of the Moment bound with the different sampling methods.

In the comparison of sieve sampling and systematic sampling it was found that systematic sampling had a higher mean coverage than sieve sampling for each error rate, taint size, sample size and bound in all the audit populations. While some of these differences were significant, the differences were not of practical importance since both sampling methods were found to be reliable at the three nominal confidence levels. No significant differences in tightness between the two sampling methods were found in any case. Systematic sampling was more precise than sieve sampling for most error rates, taint sizes, sample sizes and bounds in all audit populations. The greatest reductions in the variation of systematic sampling over sieve sampling occurred in the high error rate populations and for large sample sizes where reductions in variability of over 10% were common. It can be concluded therefore than systematic sampling favours populations with high line item error rates and large sample sizes.

In the investigation of the performance of the Moment bound with the different sampling methods, it was found that the differential effects of the sampling methods on the Moment bound estimates of the total error amount were similar to the differential effects of the sampling methods on the Stringer and Cell bound estimates of the total error amount. The coverage and the tightness of the Moment bound were not significantly affected by the sampling method for any error rate, taint size, sample size and bound. Systematic sampling gave the most precise estimates and simple random sampling gave the least precise estimates of the total error amount for the Moment bound.

The overall findings of the simulation study indicate that, in general, the differential effects of the sampling methods were independent of the bound, the error assignment method and the nominal confidence level. The reliability and tightness were similar for all sample selection methods. The precision of the bound estimates however, was affected by the sampling method. Cell and systematic sampling favoured populations with large line items and sieve sampling favoured large line item and low error rate populations. The findings on the precision of the sampling methods using the bound estimates are consistent with the results of the point estimator analysis.

9.3.3.3 Implications for Auditors.

In choosing between the simple random, systematic, cell and sieve sampling, the auditor needs to consider a number of practical issues in addition to the performance measures. These include regularities in the patterns of errors in the data, possible implementation difficulties of the sampling methods and a variable sample size.

Systematic sampling was found to be consistently more precise than the other sampling methods in all populations. However, with this selection method the danger of periodic variation in the data should not be ignored. Systematic sampling may lead to biased selection if there are regularities in the error patterns in the population (Leslie, Anderson and Teitlebaum, 1979). Jenne (1982) considers the case where there exists a systematic pattern with respect to the location of errors in the population. If the errors or groups of errors are **k** units, fractions of **k** units or multiples of **k** units apart, where **k** is the sampling interval, then a systematic sample would contain either no errors or an extremely high proportion of errors when compared to the true population error rate. Though this may be unlikely to happen in reality, it does illustrate the potential risk of bias in systematic sample selection.

Cell sampling also appears to be superior to simple random sampling in terms of precision and would not be affected to the same extent as systematic sampling by a periodic error in the monetary units. However, like simple random sampling, both cell and systematic sampling ignore the line item structure of the population when selecting the sample of monetary units. When the monetary units are selected for auditing using simple random, cell or systematic sampling, the units must be traced back to their associated line items and the need to identify the line items may at times create some practical implementation problems (Wurst, Neter and Godfrey, 1989a). The main disadvantage of simple random sampling of monetary units is that the book value total must be known accurately in advance of sampling. Leslie, Teitlebaum and Anderson (1979) point out that the total book value amount may not always be known accurately during the planning stage and it may not be known prior to the end of year. This requirement may impede the planning and implementation of the auditing process.

The results also show that sieve sampling may lead to improvements in precision over simple random sampling especially in populations with large line items. Sieve sampling does not require the accumulation of book values and therefore audit samples may be chosen before an accurate book value total is available. However, a disadvantage of sieve sampling, which may have practical implications in the audit setting, is that the sample size is not constant. It varies depending on the random numbers chosen when selecting the sample and this could be of serious concern to the auditor when attempting to estimate the cost of the audit.

Auditors will need to consider the practical advantages and disadvantages of each sampling method before deciding which one to use. They will need to consider whether the requisites of simple random, cell and systematic sampling, i.e., the need to cumulate the book value sub totals and the need to trace selected monetary-units back to the associated line items, are of serious practical concern when carrying out an audit. They will also have to decide whether the variability of sample size with sieve sampling is of serious concern, as contrasted with fixed sample sizes for simple random, systematic and cell selection.

9.3.4 A Comparison of Lahiri Sampling and Simple Random Sampling of Monetary Units.

9.3.4.1 Findings

In the investigation of the comparative performance of the upper bound estimates of Lahiri sampling and simple random sampling, no significant differences between the sampling methods were found in the coverage, tightness and precision of the bounds with either assignment method, at any nominal confidence level, in any of the audit populations. However, there was a tendency for the standard deviation of the bound estimates to be somewhat higher with Lahiri sampling than with simple random sampling of monetary units in Population 2 (i.e. large line items) but the differences were not significant in any case.

9.3.4.2 Implications for Auditors

The findings outlined above indicate that the choice between Lahiri and simple random sampling of monetary units will depend on criteria other than the performance measures. The main practical advantage of Lahiri sampling compared to simple random sampling of monetary units is that Lahiri sampling relates monetary units to line items in a natural way and therefore avoids the possible implementation problems referred to by Wurst, Neter and Godfrey (1989a). Also, unlike simple random sampling, Lahiri sampling does not require that the book value total be known accurately in advance of sampling and this enables Lahiri selection to begin before an accurate total book value amount is available. Auditors will need to consider whether the requisites of simple random sampling, i.e., the need to know the book value total in advance of sampling and the need to trace selected monetary-units back to the associated line items, are of serious practical concern when carrying out an audit. If they are, then Lahiri sampling should prove to be a useful alternative to simple random sampling of monetary units for the auditor using MUS sampling in substantive testing.

9.3.5 A Comparison of Stabilised Sieve Sampling with Simple Random Sampling and Sieve Sampling of Monetary Units

9.3.5.1 The Findings

In the comparison of stabilised sieve sampling with sieve sampling and simple random sampling of monetary units, it was found that stabilised sieve sampling had a mean coverage above the nominal for each error rate, taint size, sample size and bound with both error assignment methods, at each nominal confidence level.

The mean tightness of the estimates with stabilised sieve sampling was similar to the tightness with simple random and sieve sampling for any given error rate, taint size, sample size and bound in audit population generated from Population 1. In audit populations generated from Population 2, stabilised sieve sampling was found to be somewhat more conservative than simple random and sieve sampling in all cases but the differences in the mean tightness between the sampling methods was not significant in any case.

Stabilised sieve sampling was more precise than simple random sampling for most error rates, taint sizes, sample sizes and bounds. The greatest gains in the precision of stabilised sieve sampling over simple random sampling occurred in audit populations generated from Population 2. The precision of stabilised sieve sampling was similar to the precision of sieve sampling in most cases.

Stabilised sieve sampling overcomes the primary disadvantage of sieve sampling by returning a constant sample size of monetary units. However, while sieve sampling selects the monetary units from distinct line items, with stabilised sieve sampling as with simple random sampling, more than one monetary unit may be selected from any specific line item. The selection process for stabilised sieve sampling (see 4.8) allows for multiple selections when the sample is being augmented at the second stage. In an investigation into the number of distinct line items obtained using simple random, sieve and stabilised sieve sampling, it was found that the average number of distinct line items obtained using stabilised sieve sampling was similar to sieve sampling in all cases and greater than simple random sampling is most cases. The variance of the number of distinct line items obtained using stabilised sieve sampling was less than the variance with simple random sampling and sieve sampling. In addition, the number of distinct line items obtained using stabilised sieve sampling had a range smaller than simple random sampling and substantially smaller than sieve sampling for all sample sizes.

9.3.5.2 Implications for Auditors

The findings outlined above suggest that the decision to use stabilised sieve sampling as an alternative to simple random or sieve sampling must be made on non-statistical and practical grounds. The main practical advantages of stabilised sieve sampling compared to simple random sampling of monetary units is that stabilised sieve sampling, like sieve sampling, does not require that the random numbers be ordered and does not require that the book amounts be accumulated in advance of sampling and therefore the selection process can begin before an accurate total book value amount is available. In addition, stabilised sieve sampling overcomes the primary disadvantage of sieve sampling by returning a fixed sample size, i.e. the number of monetary units obtained using stabilised sieve sampling is always equal to the nominal sample size. Stabilised sieve sampling has the added advantage that the number of distinct line items obtained in a sample is less variable than the number obtained with simple random sampling and sieve sampling and this could be an important consideration of the auditor when planning the cost of an audit. It is therefore concluded that stabilised sieve sampling may be a useful alternative to simple random sampling and sieve sampling of monetary units in real substantive auditing environments.

9.4 Recommendations for Future Research

Since this study defines a new monetary-unit sampling method 'Stabilised Sieve Sampling' and introduces one which has not been applied previously in auditing 'Lahiri Sampling', future research in this area might involve investigating:

(i) The performance of Lahiri and stabilised sieve sampling of monetary units using other accounting populations.

(ii) The performance of other bounds (e.g., the multinominal) with Lahiri and stabilised sieve sampling of monetary units.

(iii) The performance of Lahiri and stabilised sieve sampling using populations with both understatement and overstatement errors.

(iv) The performance of Lahiri and stabilised sieve sampling using larger sample sizes (for example, samples of sizes 150 and 200).

(v) The costs and benefits of the implementation of Lahiri and stabilised sieve sampling of monetary units.

BIBLIOGRAPHY

Abdolmohammadi, M.J. (1986), Efficiency of the Bayesian Approach in Compliance Testing: Some Empirical Evidence, Auditing: A Journal of Practice and Theory, 5, Spring, 1-16.

Abdul-Hamid, M.A. (1993), A Study of Audit Method: Audit Techniques Employed by Medium Sized Accounting Firms in England, Ph.D Dissertation, University of Bradford.

Adams, R. (1989), Risk and the Role of the APC, Accountancy (UK), 104, 1156, December, 130-133.

AICPA (1973), Operational and Management Auditing, New York:AICPA

AICPA (1983), Audit and Accounting Guide: Audit Sampling, New York:AICPA.

AICPA (1981), Statement on Auditing Standards No. 39: Audit Sampling, New York:AICPA.

Anderson, R.J. (1973), Audit Uses of Statistical Sampling, The Internal Auditor, May/June

Anderson, R.J., and Teitlebaum, A.D. (1973), Dollar-Unit Sampling. A Solution to the Audit Sampling Dilemma, CA Magazine, April, 30-39.

APB (1993), Draft Guideline: Audit Sampling, ICAEW, London.

APC (1989), Explanatory Forward of Auditing Standards and Guidelines, ICAEW, London

APC (1980c), Operational Guideline: Internal Control, ICAEW, London.

APC (1980b), The Auditor's Operational Standard, ICAEW, London.

APC (1980a), Auditing Standards and Guidelines, ICAEW, London.

APC (1980d), Operational Guideline: Audit Evidence, ICAEW, London.

APC (1988), Audit Guideline: Analytical Review, ICAEW, London.

APC (1987b), Draft Guideline: Audit Sampling, ICAEW, London.

APC (1987a), Audit Guideline: Applicability to the Public Sector of Auditing Standards and Guidelines, ICAEW, London.

Arens, A.A. and Loebbecke, J.K. (1981), Applications of Statistical Sampling to Auditing, Prentice Hall.

Arkin, H. D. (1961), Discovery Sampling in Auditing, Journal of Accountancy, February.

Arkin, H.D. (1984), Handbook of Sampling for Auditing and Accounting, 3rd Edition, McGraw-Hill.

Atkinson, M. (1990), A Study of Population Characteristics Affecting the Performance of Audit Sampling Techniques in Substantive Tests, Ph.D. Dissertation, Drexel University.

Banker, R.D., Cooper, W.W, and Potter, G. (1992), A Perspective on Research in Government Accounting, The Accounting Review, 67, 3, July, 496-510.

Beck, P.J. (1980), A Critical Analysis of the Regression Estimator in Audit Sampling, Journal of Accounting Research, 18, 1, Spring, 16-37.

Carpenter, B., and Dirsmith, M. (1993), Sampling and the Abstraction of Knowledge in the Auditing Profession: An Extended Institutional Theory Perspective, Accounting, Organisations and Society, 18, 41-63.

Chan, H. and Smieliauskas, W. (1990), Further Tests of the Modified Moment Bound in Audit Sampling of Accounting Populations, Auditing: A Journal of Practice and Theory, 9, 3, 167-182.

CICA (1990), Interactive Data Extraction Analysis, Version 4.0, Users Manual, CICA.

Cochran, W.G. (1977), Sampling Techniques, 3rd Edition, John Wiley and Sons, New York.

Colbert, J. (1991), Statistical or Non-Statistical Sampling: Which Approach Is Best?, The Journal of Applied Business Research, 7, 2, Spring 1990-1991, 117-120.

Committee on Basic Auditing Concepts. (1973), A Statement of Basic Auditing Concepts, American Accounting Association.

Committee of Applied and Theoretical Sciences of the Board on Mathematical Sciences (1989), Panel on Nonstandard Mixtures of Distributions, Statistical Models and Analysis in Auditing, Statistical Science, 4, 1, 2-33.

Coopers and Lybrand (1985), Student Manual of Auditing, 3rd Edition, Coopers and Lybrand International.

Copeland, R.M., and Englebrecht, T.D. (1975), Statistical Sampling: An Uncertain Defense Against Legal Liability, CPA, November, 23-27.

COSO. (1992), Internal Control-Integrated Framework, Executive Summary, 1992, Committee of Sponsoring Organisations of the Treadway Commission.

Cox, D.R. and Snell, E.J. (1979), On Sampling and the Estimation of Rare Errors, Biometrika, 125-132.

Deming, W.E. (1960), Sample Design in Business Research, Wiley and Sons, New York.

DePaula, F. and Attwood, F.A. (1982), Auditing, Principles and Practice, 16th Edition, Pitman.

Draper, N.R. and Smith H. (1981), Applied Regression Analysis, Second Edition, Wiley and Sons, New York.

Dreissen, S. (1988), The Sieve Method, A Sampling Method for Audit Practice, Statistica Neerlandica, 42, 2, 117-135.

Duke, G.L., Neter, J. and Leitch, R.A. (1985), Behaviour of Test Statistics in the Auditing Environment: An Empirical Study, American Accounting Association.

Duke, G. L., Leitch, R. A. and Neter, J., (1982), Power Characteristics of Test Statistics in the Auditing Environment: An Empirical Study, Journal of Accounting Research, 20, 1, Spring, 42 - 68.

Duke, G.L.(1980), An Empirical Investigation of the Power of Statistical Sampling Procedures Used in Auditing Under Different Models of Change of Error Patterns, Ph.D. Dissertation, University of Georgia.

Duncan, A.J. (1986), Quality Control and Industrial Statistics, 5th Edition, Homewood III, Irwin.

Dunmore, P.V. (1986), On the Comparison of Dollar-Unit and Stratified Mean-Per-Unit Estimators, Contemporary Accounting Research, 3, 1, Fall, pp 125-149.

Dworkin, L.D., and Grimlund, R.A. (1986), Dollar-Unit Sampling: A Comparison of the Quasi-Bayesian and Moment Bounds, The Accounting Review, LXI, 1, January, 36-57

Dworkin, L.D., and Grimlund, R.A. (1984), Dollar Unit Sampling for Accounts Receivable and Inventory, The Accounting Review, LIX, 2, April, 218-241.

Efron, B. (1979), Bootstrap Methods: Another Look at the Jackknife, The 1977 Reitz Lecture, The Annals of Statistics, 7,1,1-26.

Felix, W. L., Grimlund, R.A., Koster, F.J., and Roussey, R. S.(1990), Arthur Andersen's New Monetary Unit Sampling Approach, Auditing: A Journal of Practice and Theory, 9, 3, 1-16.

Felix, W.L., Leslie, D. A., and Neter, J. (1982), University of Georgia, Centre for Audit Research Monetary-Unit Sampling Conference, March 24, 1981, Auditing: A Journal of Practice and Theory, Winter, 92-103.

Felix, W.L. and Kinney, W.R. (1982), Research in the Auditor's Opinion Formulation Process-State of the Art, The Accounting Review, April, 245-272.

Fienberg, S.E., Neter, J. and Leitch, R.A. (1977), Estimating the Total Overstatement Amount in Accounting Populations, Journal of the American Statistical Association, June, 295-302.

Frost, P.A., and Tamura, H. (1982), Jackknifed Ratio Estimation in Statistical Auditing, Journal of Accounting Research, Spring, 103-120.

Frost, P.A., and Tamura, H. (1986), Accuracy of Auxiliary Information Estimation in Statistical Auditing, Journal of Accounting Research, 24, 1, Spring, 57-75.

Garstka, S.J., and Ohlson, P.A. (1979), Ratio Estimation in Accounting Populations with Probabilities of Sample Selection Proportional to Size of Book Values, Journal of Accounting Research, 17, Spring, 23-59.

Garstka S.J. (1977), Models for Computing Upper Error Limits in Dollar-Unit Sampling, Journal of Accounting Research, Spring, 179-192.

Gill, R. D., (1983), The Sieve Method as an Alternative to Dollar-Unit Sampling: the Mathematical Background, Report SN 12/83, Unpublished Working Paper, Stichting Mathematisch Centrum, Amsterdam.

Godfrey, J.T., and Andrews, R.W. (1982), A Finite Population Bayesian Model for Compliance Testing, Journal of Accounting Research, Autumn, pp 304-315.

Godfrey J, and Neter J. (1984), Bayesian Bounds for Monetary Unit Sampling in Auditing, Journal of Accounting Research, 22, 2, Autumn, 1984, 497-525.

Goodfellow, J., Loebbecke, J.K., and Neter, J. (1974), Some Perspectives on CAV Sampling Plans, CA Magazine, Part I, October, 23-30, part II, November, 46-53.

Grimlund, R.A. and Felix, W.L. (1987), Simulation Evidence and Alternative Methods of Dollar-Unit Sampling, Accounting Review, July, 213-226

Guy, D.M., Carmichael, D.R. and Whittington (1994), Audit Sampling: An Introduction to Statistical Sampling in Auditing, 3rd Edition, Wiley and Sons, New York.

Gwilliam, D. (1987), A Survey of Auditing Research, ICAEW, London.

Ham, J., Losell, D., and Smieliauskas, W. (1985), An Empirical Study of Error Characteristics in Accounting Populations, The Accounting Review, LX, 3, July, 387-406.

Hamilton, E.R. (1978), The Role of Auditing Theory in Education and Practice, Auditing Symposium IV; Proceedings of the 1976 Touche Ross/University of Kansas Symposium on Auditing Problems, Stettler, H.F. (editor), University of Kansas,

Hansen, M.H.,and Hurwitz, W.N. (1943), On the Theory of Sampling from Finite Populations, Ann. Math. Stat., 14, 333-362.

Hardman, D.J. (1991), Towards a Conceptual Framework for Government Auditing, Accounting and Finance, 22-37.

Haskins and Sells (1970), Audit Sampling: A Programmed Instruction Course, New York: Haskins and Sells.

Higson, A. W. (1987), An Empirical Investigation of the External Audit Process: Impact of Recent Environmental Changes on Audit Technique, Ph.D Dissertation, University of Bradford.

Horvitz, D.G., and Thompson, D.J. (1952), A Generalization of Sampling without Replacement from a Finite Universe, Journal of the American Statistical Association, 47, 663-685.

Hylas, R.E., and Ashton, R.H. (1982), Audit Detection of Financial Statement Errors, The Accounting Review, October, 751-765

IIA (1992), Standards and Guidelines for the Professional Practice of Internal Auditing,The Institute of Internal Auditors U.K.

IMSL (1991), User's Manual, Stat/Library: Fortran Subroutines for Statistical Analysis, Version 2.0, IMSL inc. Houston Texas.

Jenne, S.E. (1982), A Critical Examination of Dollar-Unit Sampling in the Audit of Accounts Receivable, Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

Jenne, S.E. (1986), The Development of Monetary-Unit Sampling in Auditing Literature, Journal of Accounting Literature, 205-220.

Johnson, J.R., and Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions, Vol.1, Houghton Mifflin.

Johnson, J. R., Leitch, R. A., and Neter J. (1981), Characteristics of Errors in Accounts Receivable and Inventory Audits, The Accounting Review, LVI, 2, April, 270-293.

Johnson, R.N. (1987), Auditor Detected Errors and Related Client Traits - A Study of Inherent and Control Risks in a Sample of U.K. Audits, Journal of Business Finance and Accounting, 14, Spring, 39-64.

Kaplan, R.S. (1975), Synthesis, Journal of Accounting Research, Supplement, 134-142.

Kaplan, R.S. (1973), Statistical Sampling in Auditing with Auxiliary Information Estimators, Journal of Accounting Research, 11, 238-258.

Kendall, M.G., Stuart, A., and Ord, J.K. (1987), Kendall's Advanced Theory of Statistics, 5th Edition, London:Griffin.

Kirtland, J.L. and Holstrum, G.H. (1984), Audit Risk Model: A Framework for Current Practice Future Research, Proceeding of Symposium in Auditing Research V, University of Illinois, Urbana Campaign, pp 201-205.

Kish, L. (1965), Survey Sampling, John Wiley and Sons, New York.

Kish, L. and Frankel, M.R., (1974), Inference from Complex Samples, Journal of the Royal Statistical Society, 65, 1071-1094.

Knight, P. (1979), Statistical Sampling in Auditing: an Auditor's Viewpoint, The Statistician, 28, 4, pp 253-266.

Kreutzfeldt, R.W., and Wallace, W.A. (1986), 'Error Characteristics in Audit Populations: Their Profile and Relationships to Environmental Factors' Auditing: A Journal of Practice and Theory, Fall, pp 20-43.

Lahiri, D.B. (1951), A Method of Sample Selection providing Unbiased Ratio Estimators, Bull. Int. Stat. Inst. 33, 2, 133-140.

Lee, T.A. (1982), Company Auditing, 2nd Edition, Gee/ICAS.

Leitch, R.A, Neter, J., Plante, R. and Sinha, P. (1982), Modified Multinominal Bounds for Larger Numbers of Errors in Audits, The Accounting Review, LVII, 2, 384-401.

Leslie, D.A., and Andersley, S.J. (1982), Discussant's Response to 'The Behaviour of Selected Upper Bounds of Monetary Error Using PPS Sampling', Symposium of Auditing Research 1V, University of Illinois, Urbana Champaign, 389-399.

Leslie, D.A., Teitlebaum, A.D. and Anderson, R.J. (1979), Dollar-Unit Sampling: A Practical Guide for Auditors, Toronto: Copp, Clark, Pitman.

Leslie, D.A. (1975), Monetary-Unit Sampling in Auditing, International Journal of Government Auditing, April, 5-20.

MacGuidwin, M.,Roberts, D. and Shedd, M. (1982), The Behaviour of Selected Upper Bounds of Monetary Error Using PPS Sampling, Symposium of Auditing Research 1V, The Department of Accountancy, University of Illinois at Urbana Champaign, 352-585.

Madow, G.M., and Madow, L.H. (1944), On the Theory of Systematic Sampling, Annals of Mathematical Statistics, 2(20), 333-354.

Mautz, R.K. and Sharaf, H.A. (1961), The Philosophy of Auditing, American Accounting Association.

McCray, J. (1984), A Quasi-Bayesian Audit Risk Model For Dollar-Unit Sampling, Accounting Review, January, 35-51.

McRae, T.W. (1974), Statistical Sampling for Audit and Control, Wiley and Sons, London.

McRae, T.W. (1982), A Study of the Application of Statistical Sampling to External Auditing, ICAEW, London.

Meikle, G.R. (1972), Statistical Sampling in an Audit Context: An Audit Technique Study, CICA.

Menzefricke, U., and Smieliauskas, W. (1984), A Simulation Study of the Performance of Parametric Dollar-Unit Sampling Statistical Procedures, Journal of the Accounting Review, 22, 588-603.

Menzefricke, U., and Smieliauskas, W. (1987a), A Survey of Simulation Studies in Statistical Auditing, Journal of Accounting Literature, 6, 26-54

Menzefricke, U., and Smieliauskas, W. (1987b), A Comparison of the Stratified Difference Estimator with some Monetary-Unit Sampling Estimators, Contemporary Accounting Research, Fall, 240-251.

Miller, R.G. (1974), The Jackknife-A Review, Biometrika, 61, 1-15.

Mock, J., and Watkins, P.R. (1982), The Adequacy of Internal Accounting Control Documentation and the Modelling of Auditor Judgement, Symposium of Auditing Research 1V, The Department of Accountancy, University of Illinois at Urbana Champaign, 215-272.

Moser, C.A. and Kalton, G. (1971), Survey Methods in Social Investigation, 2nd Edition, Heinemann Educational Books, London.

Mosteller, F. (1971), The Jackknife, Review of the International Statistical Institute, 39,3, 363-371.

NAG (1988), NAG Fortran Library Manual, Mark 13, NAG Inc, Oxford, UK; Downers Grove III.

Neter, J. and Godfrey, J. (1988), Statistical Sampling in Auditing: A Review, Probability and Statistics; Essays in Honour of Franklin A. Graybill, Srivastava, J.N. (Editor), Elsevier Science Publishers, North-Holland.

Neter, J., Johnson, J.R., and Leitch, R.A. (1985), Characteristics of Dollar-Unit Taints and Error Rates in Accounts Receivable and Inventory, The Accounting Review,LX, 3, July, 448-499.

Neter, J. and Godfrey, J. (1985), Robust Bayesian Bounds for Monetary-Unit Sampling in Auditing, Applied Statistics, 34, 2, 157-168.

Neter, J. (1986), Boundaries of Statistics-Sharp or Fuzzy? Journal of the American Statistical Association, 81, 1-8.

Neter, J., and Loebbecke, J.K. (1975), Behaviour of Major Statistical Estimators in Sampling Accounting Populations - An Empirical Study, New York:AICPA.

Neter, J. and Loebbecke, J.K. (1977), On the Behaviour of Statistical Estimators when Sampling Accounting Populations, Journal of the American Statistical Association, September, 501-507.

Peek, L.E., Neter, J. and Warren, C. (1991), AICPA Nonstatistical Audit Sampling Guidelines: A Simulation, Auditing: A Journal of Practice and Theory, 10, 2, 33-48.

Phillips, J.J (1985), Bayesian Bounds for Monetary-Unit Sampling using Actual Accounting Populations, Ph.D Dissertation, University of Georgia.

Plante, R., Neter, J. and Leitch, R.A. (1985), Comparative Performance of Multinominal, Cell and Stringer Bound, Auditing: A Journal of Practice and Theory, 5, 1, Fall, 40-56.

Ramage, J.G., Krieger, A. M. and Spero, L.L. (1979), An Empirical Study of Error Characteristics in Audit Populations, Journal of Accounting Research, 17, Supplement, 72-101.

Reneau, J.H. (1978), CAV Bounds in Dollar Unit Sampling: Some Simulation Results, Accounting Review, July, 669-680.

Rietveld, C. (1978, 1979), The Sieve method as Selection Method for Statistical Samples in Audit Practice. Parts 1 to 1V (in Dutch), Compact 15, 16 and 17, Klynveld Kraayenhof and Coy., Amsterdam.

Roberts, D.M. (1978), Statistical Auditing, New York:AICPA.

Roberts, D.M. (1986), Stratified Sampling: Using A Stochastic Model, Journal of Accounting Research, 24, Spring, 111-126.

Robertson, J.C. (1984), A Defense of Extant Auditing Theory, A Journal of Practice and Theory, Spring, 3,2, 57-67.

Royall, R.M. (1970), On Finite Population Sampling Theory under Certain Linear Regression Models, Biometrika, 57,2,377-387.

Schlosser, R.E. (1971), The Field of Auditing, Handbook for Auditors, McGraw-Hill.

Shaw, J.C. (1980), The Audit Report, What it Says and What it Means, ICAS.

Sherer, M. and Kent, D. (1988), Auditing and Accountability, Pitman.

Smieliauskas, W. (1986), A Note on Comparison of Bayesian with Non-Bayesian Dollar-Unit Sampling Bounds for Overstatement Errors of Accounting Populations, The Accounting Review, LXI, 1, January, pp 118-128

Smith, T.M.F. (1979), Statistical Sampling in Auditing: A Statisticians Viewpoint, The Statistician, 4, 235-249.

Sneed, C.R. (1979), A Study of the Effects of Conservatism on the Evidential Sample Size Decisions made by Auditors, Ph.D. Dissertation, North Texas State University.

SPSS (1987), SPSSx, User's Guide, 3rd Edition, SPSS Inc.,McGraw Hill, New York.

Stringer, K.W. (1963), Practical Aspects of Statistical Sampling in Auditing, Proc. Bus. Econ.. Statist. Sec., 405-411, American Statistical Association, Washington.

Tamura, H. (1985), Analysis of the Garstka-Ohlson Bounds, Auditing: A Journal of Practice and Theory, 4, 2, 133-142.

Taylor, G (1985), Audit Judgement: Risk and Materiality, Current Issues in Auditing, Kent, D., Sherer, M., and Turley, S. (Eds), 93-106.

Taylor, D.H., and Glezen, G.W. (1994), Auditing: Integrated Concepts and Procedures, 6th Edition, Wiley and Sons, New York.

Teitlebaum, A.D. (1973), Dollar-Unit Sampling in Auditing, Paper presented to the National Meeting of the American Statistical Association, December.

Tsui, K., Matsumura, E.M., and Tsui, K. (1985), Multinomial-Dirichlet Bounds for Dollar-Unit Sampling in Auditing, The Accounting Review, LX, 1, 76-96.

Van Heerden, A. (1961), Steekproeven als Middel van Accountantscontrolx (Statistical Sampling as a Means of Auditing), Maanblad voor Accountancy en Bedrijfshoudkunde, 11, 453.

Vance, L.L. (1947), Statistical Sampling Theory and Auditing Procedure, Proceedings of the Pacific Coast Economic Association.

Vance, L.L. (1950), Auditing Use of Probabilities and Interpreting Test Checks, Journal of Accountancy, 88, 3, September.

Vance, L.L., and Neter, J. (1956), Statistical Sampling for Auditors and Accountants, Wiley and Son.

Warren, W.G. (1971), Correlation and Regression: Bias or Correlation, Applied Statistics, 20, 148-164.

Wilson, E.B. and Hilferty, M.M. (1931), The Distribution of Chi-square, Proceedings of the National Academy of Sciences, Washington, 17, 684-688.

Winkler, R.L., (1972), An Introduction to Bayesian Information and Decisions, New York:Rinehart and Winston.

Wright, D.W. (1991), Augmenting a Sample Selected with Probabilities Proportional to Size, Auditing: A Journal of Practice and Theory, 10, 1, Spring, 145-158.

Wurst, J., Neter, J. and Godfrey, J. (1991), Effectiveness of Rectification in Audit Sampling, The Accounting Review, 66,2, April, 333-346.

Wurst, J. C. (1990), An Investigation of the Sieve Method and Rectification Sampling for Auditing, Ph.D. Dissertation, University of Georgia.

Wurst, J., Neter, J., and Godfrey, J. (1989a), Efficiency of Sieve Sampling in Auditing, Journal of Business and Economic Statistics, The American Statistical Association, April, 7, 2, 199-205.

Wurst, J., Neter, J. and Godfrey, J. (1989b), Comparison of Sieve Sampling with Random and Cell Sampling of Monetary Units, The Statistician, 38, 235-249.

Bibliography classified as follows:

A. Accountancy Bodies Standards and Guidelines
B. Books
C. Computer Manuals
D. Doctoral Dissertations
E. Journal Papers
F. Research Reports
G. Symposia

A. Accountancy Bodies Standard and Guidelines

AICPA (1973), Operational and Management Auditing, New York:AICPA

AICPA (1983), Audit and Accounting Guide: Audit Sampling, New York:AICPA.

AICPA (1981), Statement on Auditing Standards No. 39: Audit Sampling, New York:AICPA.

APB (1993), Draft Guideline: Audit Sampling, ICAEW, London.

APC (1989), Explanatory Forward of Auditing Standards and Guidelines, ICAEW, London

APC (1980c), Operational Guideline: Internal Control, ICAEW, London.

APC (1980b), The Auditor's Operational Standard, ICAEW, London.

APC (1980a), Auditing Standards and Guidelines, ICAEW, London.

APC (1980d), Operational Guideline: Audit Evidence, ICAEW, London.

APC (1988), Audit Guideline: Analytical Review, ICAEW, London.

APC (1987b), Draft Guideline: Audit Sampling, ICAEW, London.

APC (1987a), Audit Guideline: Applicability to the Public Sector of Auditing Standards and Guidelines, ICAEW, London.

Committee on Basic Auditing Concepts. (1973), A Statement of Basic Auditing Concepts, American Accounting Association.

COSO. (1992), Internal Control-Integrated Framework, Executive Summary, 1992, Committee of Sponsoring Organisations of the Treadway Commission.

IIA (1992), Standards and Guidelines for the Professional Practice of Internal Auditing,The Institute of Internal Auditors U.K.

B. Books

Arens, A.A. and Loebbecke, J.K. (1981), Applications of Statistical Sampling to Auditing, Prentice Hall.

Arkin, H.D. (1984), Handbook of Sampling for Auditing and Accounting, 3rd Edition, McGraw-Hill.

Cochran, W.G. (1977), Sampling Techniques, 3rd Edition, John Wiley and Sons, New York.

Coopers and Lybrand (1985), Student Manual of Auditing, 3th Edition, Coopers and Lybrand International.

Deming, W.E. (1960), Sample Design in Business Research, Wiley and Sons, New York.

DePaula, F. and Attwood, F.A. (1982), Auditing, Principles and Practice, 16th Edition, Pitman.

Draper, N.R. and Smith H. (1981), Applied Regression Analysis, Second Edition, Wiley and Sons, New York.

Duncan, A.J. (1986), 'Quality Control and Industrial Statistics' 5th Edition, Homewood III, Irwin.

Guy, D.M., Carmichael, D.R.and Whittington (1994), Audit Sampling: An Introduction to Statistical Sampling in Auditing, 2nd Edition, Wiley and Sons, New York.

Gwilliam, D. (1987), A Survey of Auditing Research, ICAEW, London.

Haskins and Sells (1970), Audit Sampling: A Programmed Instruction Course, New York: Haskins and Sells.

Johnson, J.R., and Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions, Vol.1, Houghton Mifflin.

Kendall, M.G., Stuart, A., and Ord, J.K. (1987), Kendall's Advanced Theory of Statistics, 5th Edition, London:Griffin.

Kish, L. (1965), Survey Sampling, John Wiley and Sons, New York.

Lee, T.A. (1982), Company Auditing, 2nd Edition, Gee/ICAS.

Mautz, R.K. and Sharaf, H.A. (1961), The Philosophy of Auditing, American Accounting Association.

McRae, T.W. (1974), Statistical Sampling for Audit and Control, Wiley and Sons, London.

Leslie, D.A., Teitlebaum, A.D. and Anderson, R.J. (1979), Dollar-Unit Sampling: A Practical Guide for Auditors, Toronto: Copp, Clark, Pitman.

Moser, C.A. and Kalton, G. (1971), Survey Methods in Social Investigation, 2nd Edition, Heinemann Educational Books, London.

Neter, J. and Godfrey, J. (1988), Statistical Sampling in Auditing: A Review, Probability and Statistics; Essays in Honour of Franklin A. Graybill, Srivastava, J.N. (Editor), Elsevier Science Publishers, North-Holland.

Rietveld, C. (1978, 1979), The Sieve method as Selection Method for Statistical Samples in Audit Practice. Parts 1 to 1V (in Dutch), Compact 15, 16 and 17, Klynveld Kraayenhof and Coy., Amsterdam.

Roberts, D.M. (1978), Statistical Auditing, New York:AICPA.

Schlosser, R.E. (1971), The Field of Auditing, Handbook for Auditors, McGraw-Hill.

Shaw, J.C. (1980), The Audit Report, What it Says and What it Means, ICAS.

Sherer, M. and Kent, D. (1988), Auditing and Accountability, Pitman.

Taylor, D.H., and Glezen, G.W. (1994), Auditing: Integrated Concepts and Procedures, 6th Edition,Wiley and Sons, New York.

Vance, L.L., and Neter, J. (1956), Statistical Sampling for Auditors and Accountants, Wiley and Son.

Winkler, R.L., (1972), An Introduction to Bayesian Information and Decisions, New York:Rinehart and Winston.

C. Computer Manuals

CICA (1990), Interactive Data Extraction Analysis, Version 4.0, Users Manual, CICA.

IMSL (1991), User's Manual, Stat/Library: Fortran Subroutines for Statistical Analysis, Version 2.0, IMSL inc. Houston Texas.

NAG (1988), NAG Fortran Library Manual, Mark 13, NAG Inc, Oxford, UK; Downers Grove III.

SPSS (1987), SPSSx, User's Guide, 3rd Edition, SPSS Inc.,McGraw Hill, New York.

D. Doctoral Dissertations

Abdul-Hamid, M.A. (1993), A Study of Audit Method: Audit Techniques Employed by Medium Sized Accounting Firms in England, Ph.D Dissertation, University of Bradford.

Atkinson, M. (1990), A Study of Population Characteristics Affecting the Performance of Audit Sampling Techniques in Substantive Tests, Ph.D. Dissertation, Drexel University.

Duke, G.L.(1980), An Empirical Investigation of the Power of Statistical Sampling Procedures Used in Auditing Under Different Models of Change of Error Patterns, Ph.D. Dissertation, University of Georgia.

Higson, A. W. (1987), 'An Empirical Investigation of the External Audit Process: Impact of Recent Environmental Changes on Audit Technique' Ph.D Dissertation, University of Bradford.

Jenne, S.E. (1982), A Critical Examination of Dollar-Unit Sampling in the Audit of Accounts Receivable, Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

Phillips, J.J (1985), Bayesian Bounds for Monetary-Unit Sampling using Actual Accounting Populations, Ph.D Dissertation, University of Georgia.

Sneed, C.R. (1979), A Study of the Effects of Conservatism on the Evidential Sample Size Decisions made by Auditors, Ph.D. Dissertation, North Texas State University.

Wurst, J. C. (1990), An Investigation of the Sieve Method and Rectification Sampling for Auditing, Ph.D. Dissertation, University of Georgia.

E. Journal Articles

Abdolmohammadi, M.J. (1986), Efficiency of the Bayesian Approach in Compliance Testing: Some Empirical Evidence, Auditing: A Journal of Practice and Theory, 5, Spring, 1-16.

Adams, R. (1989), Risk and the Role of the APC, Accountancy (UK), 104, 1156, December, 130-133.

Anderson, R.J. (1973), Audit Uses of Statistical Sampling, The Internal Auditor, May/June

Anderson, R.J., and Teitlebaum, A.D. (1973), Dollar-Unit Sampling. A Solution to the Audit Sampling Dilemma, CA Magazine, April, 30-39.

Arkin, H. D. (1961), 'Discovery Sampling in Auditing' Journal of Accountancy, February.

Banker, R.D., Cooper, W.W, and Potter, G. (1992), A Perspective on Research in Government Accounting, The Accounting Review, 67, 3, July, 496-510.

Beck, P.J. (1980), A Critical Analysis of the Regression Estimator in Audit Sampling, Journal of Accounting Research, 18, 1, Spring, 16-37.

Carpenter, B., and Dirsmith, M. (1993), Sampling and the Abstraction of Knowledge in the Auditing Profession: An Extended Institutional Theory Perspective, Accounting, Organisations and Society, 18, 41-63.

Chan, H. and Smieliauskas, W. (1990), Further Tests of the Modified Moment Bound in Audit Sampling of Accounting Populations, Auditing: A Journal of Practice and Theory, 9, 3, 167-182.

Colbert, J. (1991), Statistical or Non-Statistical Sampling: Which Approach Is Best?, The Journal of Applied Business Research, 7, 2, Spring 1990-1991, 117-120.

Copeland, R.M., and Englebrecht, T.D. (1975), Statistical Sampling: An Uncertain Defense Against Legal Liability, CPA, November, 23-27.

Cox, D.R. and Snell, E.J. (1979), On Sampling and the Estimation of Rare Errors, Biometrika, 125-132.

Dreissen, S. (1988), The Sieve Method, A Sampling Method for Audit Practice, Statistica Neerlandica, 42, 2, 117-135.

Duke, G. L., Leitch, R. A. and Neter, J., (1982), Power Characteristics of Test Statistics in the Auditing Environment: An Empirical Study, Journal of Accounting Research, 20, 1, Spring, 42 - 68.

Dunmore, P.V. (1986), On the Comparison of Dollar-Unit and Stratified Mean-Per-Unit Estimators, Contemporary Accounting Research, 3, 1, Fall, pp 125-149.

Dworkin, L.D., and Grimlund, R.A. (1986), Dollar-Unit Sampling: A Comparison of the Quasi-Bayesian and Moment Bounds, The Accounting Review, LXI, 1, January, 36-57

Dworkin, L.D., and Grimlund, R.A. (1984), Dollar Unit Sampling for Accounts Receivable and Inventory, The Accounting Review, LIX, 2, April, 218-241.

Efron, B. (1979), Bootstrap Methods: Another Look at the Jackknife, The 1977 Reitz Lecture, The Annals of Statistics, 7,1,1-26.

Felix, W. L., Grimlund, R.A., Koster, F.J., and Roussey, R. S.(1990), Arthur Andersen's New Monetary Unit Sampling Approach, Auditing: A Journal of Practice and Theory, 9, 3, 1-16.

Felix, W.L., Leslie, D. A., and Neter, J. (1982), University of Georgia, Centre for Audit Research Monetary-Unit Sampling Conference, March 24, 1981, Auditing: A Journal of Practice and Theory, Winter, 92-103.

Felix, W.L. and Kinney, W.R. (1982), Research in the Auditor's Opinion Formulation Process-State of the Art, The Accounting Review, April, 245-272.

Fienberg, S.E., Neter, J. and Leitch, R.A. (1977), Estimating the Total Overstatement Amount in Accounting Populations, Journal of the American Statistical Association, June, 295-302.

Frost, P.A., and Tamura, H. (1982), Jackknifed Ratio Estimation in Statistical Auditing, Journal of Accounting Research, Spring, 103-120.

Frost, P.A., and Tamura, H. (1986), Accuracy of Auxiliary Information Estimation in Statistical Auditing, Journal of Accounting Research, 24, 1, Spring, 57-75.

Garstka, S.J., and Ohlson, P.A. (1979), Ratio Estimation in Accounting Populations with Probabilities of Sample Selection Proportional to Size of Book Values, Journal of Accounting Research, 17, Spring, 23-59.

Garstka S.J. (1977), Models for Computing Upper Error Limits in Dollar-Unit Sampling, Journal of Accounting Research, Spring, 179-192.

Godfrey, J.T., and Andrews, R.W. (1982), A Finite Population Bayesian Model for Compliance Testing, Journal of Accounting Research, Autumn, pp 304-315.

Godfrey J, and Neter J. (1984), Bayesian Bounds for Monetary Unit Sampling in Auditing, Journal of Accounting Research, 22, 2, Autumn, 1984, 497-525.

Goodfellow, J., Loebbecke, J.K., and Neter, J. (1974), Some Perspectives on CAV Sampling Plans, CA Magazine, Part I, October, 23-30, part II, November, 46-53.

Grimlund, R.A. and Felix, W.L. (1987), Simulation Evidence and Alternative Methods of Dollar-Unit Sampling, July, Accounting Review, 213-226.

Ham, J., Losell, D., and Smieliauskas, W. (1985), An Empirical Study of Error Characteristics in Accounting Populations, The Accounting Review, LX, 3, July, 387-406.

Hansen, M.H.,and Hurwitz, W.N. (1943), On the Theory of Sampling from Finite Populations, Ann. Math. Stat., 14, 333-362.

Hardman, D.J. (1991), Towards a Conceptual Framework for Government Auditing, Accounting and Finance, 22-37.

Horvitz, D.G., and Thompson, D.J. (1952), A Generalization of Sampling without Replacement from a Finite Universe, Journal of the American Statistical Association, 47, 663-685.

Hylas, R.E., and Ashton, R.H. (1982), Audit Detection of Financial Statement Errors, The Accounting Review, October, 751-765

Jenne, S.E. (1986), The Development of Monetary-Unit Sampling in Auditing Literature, Journal of Accounting Literature, 205-220.

Johnson, J. R., Leitch, R. A., and Neter J. (1981), Characteristics of Errors in Accounts Receivable and Inventory Audits, The Accounting Review, LVI, 2, April, 270-293.

Johnson, R.N. (1987), Auditor Detected Errors and Related Client Traits - A Study of Inherent and Control Risks in a Sample of U.K. Audits, Journal of Business Finance and Accounting, 14, Spring, 39-64.

Kaplan, R.S. (1975), Synthesis, Journal of Accounting Research, Supplement, 134-142.

Kaplan, R.S. (1973), Statistical Sampling in Auditing with Auxiliary Information Estimators, Journal of Accounting Research, 11, 238-258.

Kish, L. and Frankel, M.R., (1974), Inference from Complex Samples, Journal of the Royal Statistical Society, 65, 1071-1094.

Knight, P. (1979), Statistical Sampling in Auditing: an Auditor's Viewpoint, The Statistician, 28, 4, pp 253-266.

Kreutzfeldt, R.W. and Wallace, W.A. (1986), Error Characteristics in Audit Populations: Their Profile and Relationship to Environmental Factors, Auditing: A Journal of Practice and Theory, Fall, 20-43.

Lahiri, D.B. (1951), A Method of Sample Selection providing Unbiased Ratio Estimators, Bull. Int. Stat. Inst. 33, 2, 133-140.

Leitch, R.A, Neter, J., Plante, R. and Sinha, P. (1982), Modified Multinominal Bounds for Larger Numbers of Errors in Audits, The Accounting Review, LVII, 2, 384-401.

Leslie, D.A. (1975), Monetary-Unit Sampling in Auditing, International Journal of Government Auditing, April, 5-20.

Madow, G.M., and Madow, L.H. (1944), On the Theory of Systematic Sampling, Annals of Mathematical Statistics, 2(20), 333-354.

McCray, J. (1984), A Quasi-Bayesian Audit Risk Model For Dollar-Unit Sampling, Accounting Review, January, 35-51.

Menzefricke, U., and Smieliauskas, W. (1984), A Simulation Study of the Performance of Parametric Dollar-Unit Sampling Statistical Procedures, Journal of the Accounting Review, 22, 588-603.

Menzefricke, U., and Smieliauskas, W. (1987a), A Survey of Simulation Studies in Statistical Auditing, Journal of Accounting Literature, 6, 26-54

Menzefricke, U., and Smieliauskas, W. (1987b), A Comparison of the Stratified Difference Estimator with some Monetary-Unit Sampling Estimators, Contemporary Accounting Research, Fall, 240-251.

Miller, R.G. (1974), The Jackknife-A Review, Biometrika, 61, 1-15.

Mosteller, F. (1971), The Jackknife, Review of the International Statistical Institute, 39,3, 363-371.

Neter, J., Johnson, J.R., and Leitch, R.A. (1985), Characteristics of Dollar-Unit Taints and Error Rates in Accounts Receivable and Inventory, The Accounting Review,LX, 3, July, 448-499.

Neter, J. and Godfrey, J. (1985), Robust Bayesian Bounds for Monetary-Unit Sampling in Auditing, Applied Statistics, 34, 2, 157-168.

Neter, J. (1986), Boundaries of Statistics-Sharp or Fuzzy? Journal of the American Statistical Association, 81, 1-8.

Neter, J. and Loebbecke, J.K. (1977), On the Behaviour of Statistical Estimators when Sampling Accounting Populations, Journal of the American Statistical Association, September, 501-507.

Peek, L.E., Neter, J. and Warren, C. (1991), AICPA Nonstatistical Audit Sampling Guidelines: A Simulation, Auditing: A Journal of Practice and Theory, 10, 2, 33-48.

Plante, R., Neter, J. and Leitch, R.A. (1985), Comparative Performance of Multinominal, Cell and Stringer Bound, Auditing: A Journal of Practice and Theory, 5, 1, Fall, 40-56.

Ramage, J.G., Krieger, A. M. and Spero, L.L. (1979), An Empirical Study of Error Characteristics in Audit Populations, Journal of Accounting Research, 17, Supplement, 72-101.

Reneau, J.H. (1978), CAV Bounds in Dollar Unit Sampling: Some Simulation Results, Accounting Review, July, 669-680.

Roberts, D.M. (1986), Stratified Sampling: Using A Stochastic Model, Journal of Accounting Research, 24, Spring, 111-126.

Robertson, J.C. (1984), A Defense of Extant Auditing Theory, A Journal of Practice and Theory, Spring, 3,2, 57-67.

Royall, R.M. (1970), On Finite Population Sampling Theory under Certain Linear Regression Models, Biometrika, 57,2,377-387.

Smieliauskas, W. (1986), A Note on Comparison of Bayesian with Non-Bayesian Dollar-Unit Sampling Bounds for Overstatement Errors of Accounting Populations, The Accounting Review, LXI, 1, January, pp 118-128

Smith, T.M.F. (1979), Statistical Sampling in Auditing: A Statisticians Viewpoint, The Statistician, 4, 235-249.

Stringer, K.W. (1963), Practical Aspects of Statistical Sampling in Auditing, Proc. Bus. Econ.. Statist. Sec., 405-411, American Statistical Association, Washington.

Tamura, H. (1985), Analysis of the Garstka-Ohlson Bounds, Auditing: A Journal of Practice and Theory, 4, 2, 133-142.

Taylor, G (1985), Audit Judgement: Risk and Materiality, Current Issues in Auditing, Kent, D., Sherer, M., and Turley, S. (Eds), 93-106.

Tsui, K., Matsumura, E.M., and Tsui, K. (1985), Multinomial-Dirichlet Bounds for Dollar-Unit Sampling in Auditing, The Accounting Review, LX, 1, 76-96.

Van Heerden, A. (1961), Steekproeven als Middel van Accountantscontrolx (Statistical Sampling as a Means of Auditing), Maanblad voor Accountancy en Bedrijfshoudkunde, 11, 453.

Vance, L.L. (1950), Auditing Use of Probabilities and Interpreting Test Checks, Journal of Accountancy, 88, 3, September.

Warren, W.G. (1971), Correlation and Regression: Bias or Correlation, Applied Statistics, 20, 148-164.

Wilson, E.B. and Hilferty, M.M. (1931), The Distribution of Chi-square, Proceedings of the National Academy of Sciences, Washington, 17, 684-688.

Wright, D.W. (1991), Augmenting a Sample Selected with Probabilities Proportional to Size, Auditing: A Journal of Practice and Theory, 10, 1, Spring, 145-158.

Wurst, J., Neter, J. and Godfrey, J. (1991), Effectiveness of Rectification in Audit Sampling, The Accounting Review, 66,2, April, 333-346.

Wurst, J., Neter, J., and Godfrey, J. (1989a), Efficiency of Sieve Sampling in Auditing, Journal of Business and Economic Statistics, The American Statistical Association, April, 7, 2, 199-205.

Wurst, J., Neter, J. and Godfrey, J. (1989b), Comparison of Sieve Sampling with Random and Cell Sampling of Monetary Units, The Statistician, 38, 235-249.

F. Research Reports

Committee of Applied and Theoretical Sciences of the Board on Mathematical Sciences (1989), Panel on Nonstandard Mixtures of Distributions, Statistical Models and Analysis in Auditing, Statistical Science, 4, 1, 2-33.

Duke, G.L., Neter, J. and Leitch, R.A. (1985), Behaviour of Test Statistics in the Auditing Environment: An Empirical Study, American Accounting Association.

Gill, R. D., (1983), The Sieve Method as an Alternative to Dollar-Unit Sampling: the Mathematical Background, Report SN 12/83, Unpublished Working Paper, Stichting Mathematisch Centrum, Amsterdam.

McRae, T.W. (1982), A Study of the Application of Statistical Sampling to External Auditing, ICAEW, London.

Meikle, G.R. (1972), Statistical Sampling in an Audit Context: An Audit Technique Study, CICA.

Neter, J., and Loebbecke, J.K. (1975), Behaviour of Major Statistical Estimators in Sampling Accounting Populations - An Empirical Study, New York:AICPA.

G. Symposia

Hamilton, E.R. (1978), The Role of Auditing Theory in Education and Practice, Auditing Symposium IV; Proceedings of the 1976 Touche Ross/University of Kansas Symposium on Auditing Problems, Stettler, H.F. (editor), University of Kansas,

Kirtland, J.L. and Holstrum, G.H. (1984), Audit Risk Model: A Framework for Current Practice Future Research, Proceeding of Symposium in Auditing Research V, University of Illinois, Urbana Campaign, pp 201-205.

Leslie, D.A., and Andersley, S.J. (1982), Discussant's Response to 'The Behaviour of Selected Upper Bounds of Monetary Error Using PPS Sampling', Symposium of Auditing Research 1V, University of Illinois, Urbana Champaign, 389-399.

MacGuidwin, M.,Roberts, D. and Shedd, M. (1982), The Behaviour of Selected Upper Bounds of Monetary Error Using PPS Sampling, Symposium of Auditing Research 1V, The Department of Accountancy, University of Illinois at Urbana Champaign, 352-585.

Mock, J., and Watkins, P.R. (1982), The Adequacy of Internal Accounting Control Documentation and the Modelling of Auditor Judgement, Symposium of Auditing Research 1V, The Department of Accountancy, University of Illinois at Urbana Champaign, 215-272.

Teitlebaum, A.D. (1973), Dollar-Unit Sampling in Auditing, Paper presented to the National Meeting of the American Statistical Association, December.

Vance, L.L. (1947), Statistical Sampling Theory and Auditing Procedure, Proceedings of the Pacific Coast Economic Association.