Sampling Policy and Guideline Manual

Rev. April 2022

Multistate Tax Commission
444 North Capitol Street, N.W., Suite 425
Washington, D.C. 20001

Telephone: 202-650-0300
Website: www.mtc.gov

[This Page Intentionally Left Blank]

Preface

This manual is intended to provide guidance in sampling procedures for the Multistate Tax Commission ("Commission"). It is also intended to inform the member states as to the sampling procedures the Commission's audit staff utilizes when conducting audits on their behalf.

It is assumed electronic records are available and statistical methods will be applied. However, under the right circumstances many of the procedures contained within this manual may be utilized when electronic records are not available. The Commission believes statistical sampling should be used in most situations when electronic records are available; therefore, statistical sampling is considered the Commission's default procedure. Statistical formulas referenced in this manual are available upon request.

It is not the intent of this manual to fully explain all sampling procedures, but rather to augment materials and issues covered in the Commission's sampling courses.

The policies in this manual may be applied to any audit done on behalf of a member state when the member state does not have its own sampling policy. If a member state has a sampling policy which deviates from the Commission's policy, the Commission will follow the member state's policy.

The Computer Audit Specialist (CAS) will contact each state to get their sampling policies. It is the responsibility of member states to provide the Commission with any updates to their sampling policies. The Commission encourages its member states to establish written sampling policies.

Acknowledgements – All the content of this manual is the sole responsibility of the Commission. Some contributions were made by individuals and organizations outside the Commission. It is not possible to acknowledge all contributions. However, there are a few individuals we wish to specifically name for providing us with critical background on the subject.

- Dr. Donald Roberts, University of Illinois, for opening our eyes through his valuable instruction and his book, *Statistical Auditing,* of which both proved invaluable.

- Dr. Allan Kvanli, University of North Texas, for his assistance and guidance in developing and conducting our Basic Statistical Sampling manual and class and for his willingness to share his knowledge in answering our many questions.

Table of Contents

1. Introduction

1.01 Goals and Objectives

This Sampling Policy and Guideline Manual ("Manual" or "Sampling Manual") is intended to provide policies and guidelines to aid the Commission auditor in performing sampling in sales and use tax audits. This manual should not be used as a substitute for appropriate sampling training.

The Auditors should have been properly trained in sampling techniques before attempting to implement any of the policies or guidelines contained within this manual. Auditors are ultimately responsible for their work and should consult with either their supervisor and/or the CAS if questions arise.

This manual covers the basic procedures for planning, conducting, and documenting an audit sample, while utilizing the most appropriate sampling techniques, and while considering the audit circumstances. The manual has the following objectives:

- To help the auditor obtain appropriate information about the taxpayer prior to conducting a sample
- To help the auditor collect, organize, and present all information which is pertinent to the performance of the sample
- To inform the taxpayer of the Commission's sampling requirements
- To encourage cooperation between the taxpayer and auditor.

1.02 Goal of the Commission

The goal of the Commission is to perform efficient and accurate audits on behalf of its member states. To achieve efficiency, the Commission utilizes statistical sampling whenever it is possible and reasonable to do so. The proper use of sampling can minimize the time and costs associated with retrieving and examining sampled audit units. These savings benefit both the Commission and taxpayers. The goal of statistical sampling is to minimize sample sizes while maintaining or increasing audit accuracy in order to maximize efficiency.

Probability sampling can be used to provide information about the accuracy of audit results based on a sample. Because accuracy is dependent on sample size, there is a tradeoff between efficiency and precision. Probability sampling can be used to achieve a balance between these objectives.

1.03 Sampling Policy and Guidelines

The Commission recognizes each sampling situation is unique. Guidelines are recommendations which provide a procedural framework; however, deviations can be made based on the experience of the auditor or supervisor. Departures from policies and guidelines should be documented and explained in the sampling plan, which will be contained within the completed work papers.

1.04 When Should Sampling Be Used

In determining whether error exists or whether further audit examination is warranted, auditors may use sampling methods as they deem necessary. When considering sampling to support an audit finding that corrects or adjusts the tax amounts reported by the taxpayer, the auditor should consider the three standards discussed in this section.

(a) Best Evidence Standard

The Commission auditor has the duty to base audit findings which correct or adjust tax amounts reported by taxpayers on the best available evidence.

A Commission auditor can come to an audit finding for the entire period based on an examination of all records, or an examination of part of the records (a sample). While both methods are acceptable, under the best evidence standard it is preferred, wherever practical, that an equally complete coverage of all records be made (a detailed examination). If the auditor believes that a detailed audit is warranted, the auditor is not obligated to sample, irrespective of the taxpayer's preferences. It is not necessary for the auditor to obtain agreement from the taxpayer regarding the performance of a detailed audit.

However, as a matter of fact, in most audits done by the Commission, it may not be feasible to do a detailed examination of all records. Most often, Commission audits involve multistate taxpayers with voluminous and complex records, necessitating sampling procedures.

(b) Standard to Judge an Audit Finding Based on a Sample

When applying sampling, the result from a sampling process is acceptable if, within an acceptable degree of accuracy and using a reasonable degree of confidence, the same result would have been reached had a detailed audit been performed. The auditor can use the results of a sample whenever this standard is met. However, when constructing a sample, it is prudent to consider and respect any reasonable requests made by a taxpayer.

Probability samples and judgmental samples are two acceptable types of samples used in auditing. The preferred sampling method is the use of a probability sample determined by applying statistical sampling procedures to all transactions under the scope of the audit objective for the entire audit period. Use of a probability sample using non-statistical sampling procedures is also acceptable, but not preferred. Any sample other than a probability sample is considered a judgmental sample and may be acceptable under certain circumstances where a probability sample is not possible.

It is possible to compare the audit results from a probability sample to the results of a detailed audit (had one been performed). This is not possible with a judgmental sample unless a detailed audit is later done. Therefore, under the best evidence standard and the standard to judge an audit finding based on a sample, the auditor is obligated to take the probability sample if one is practical under the circumstances. In all other cases, the Commission auditor may take a judgmental sample, subject to the limitations and guidelines expressed in this manual.

(c) Notification and Disclosure Standard

Prior to sampling, the auditor should disclose the method of sampling to be used, regardless of whether a probability or judgmental sample is considered. The Commission is also required to disclose the results of a sample to the taxpayer, regardless of whether a probability or judgmental sample was taken.

The auditor is obligated to record and report any known objections the taxpayer has with sampling procedures.

1.05 Responsibilities

When conducting a Computer Assisted Audit (CAA), the auditor should seek sampling advice from the CAS if constructing the sampling plan warrants assistance of the CAS. Most importantly, the taxpayer must be informed about the sampling process. Care should be taken to involve the taxpayer in the process as soon and as much as possible.

Probability sampling can be performed in a CAA or in any other audit. A CAA is an audit in which the taxpayer can provide the auditor with electronic records. This includes electronic records that may be provided in text form, spreadsheet, and databases or downloads from mainframe computers.

The primary responsibility for any audit lies with the auditor who has the audit assignment. However, this should not preclude the auditor from seeking the advice and assistance from the CAS when conducting a CAA or when using statistical sampling.

(a) Responsibilities of the Auditor

- Discuss the sampling plan with the taxpayer.
- Determine items to be excluded and refine the population. Forward to CAS as warranted.
- Establish dollar cutoffs (low dollar items not audited and high dollar items reviewed in detail).
- Determine items to be reviewed on a detailed basis.
- Secure and review the sampling units.
- Value each sampling unit for error.
- Statistically evaluate audit conclusions. Forward audit conclusions to CAS as warranted.
- Project the sample results in the audit report.
- Discuss the sampling results with the taxpayer.
- Answer any questions the taxpayer may have regarding the results of the audit.

(b) Responsibilities of the CAS

The specialized nature of statistical sampling requires specific guidelines and policies (see section 4). One of the responsibilities of the CAS is to provide guidance to the auditor regarding methods on probability sampling. The CAS is responsible for:

- Assisting the auditor in answering taxpayer questions about statistical sampling.
- Meeting with auditors, taxpayers, and consultants regarding sampling.
- Informing audit management and other interested parties of the consequences of sampling policy.
- Representing the audit section on issues involving statistical sampling in contested audits. When the procedures outlined in this manual are followed, this responsibility applies whether or not the CAS was actually involved in the design of the sample.
- Discussing the sampling procedures with the taxpayer and auditor.
- Assisting the auditor if needed in defining the sampling frame from the data provided.
- Giving advice if needed on the appropriate sample size.
- Assisting the auditor if needed in stratifying the sampling frame.
- Assisting the auditor if needed in selecting a probability sample from the sampling frame.
- Assisting the auditor if needed in evaluating the audited sample results for any statistical projections.
- Providing the auditor with suggested text for the audit narrative concerning the sample, as needed.

1.06 Auditor Training

The audit director and the auditor's supervisor are ultimately responsible for ensuring that the auditor is properly trained and has demonstrated an understanding of the sampling techniques employed by the Commission before allowing the auditor to conduct a sample.

1.07 Deviations from Manual

The auditor-in-charge may deviate from the guidelines with the permission of the audit supervisor to adapt to the particular circumstances of an audit. Before permission is granted on deviations from the policies, the auditor should consult with the CAS and/or audit supervisor. The audit supervisor may need to consult the audit director.

2. Definitions

Attribute – a qualitative characteristic of interest associated with sampling units.

Audited Amount – in financial auditing, the value which should be recorded in the taxpayer's records.

Audit Risk – the risk that there is a misstatement, and the auditor will not detect the error.

Block Sample – a type of judgment sample where the auditor selects a group (or block) of sequential transactions for examination. Although efficient, this approach lacks the accuracy attained in statistical sampling.

Cluster Sample – a sample taken from a population divided into groups, or clusters. If a simple random sample is taken, where the cluster is the sampling unit, this type of sample is a probability sample. If the clusters are chosen using non-random methods, this would not be a probability sample. Regardless of the method of choosing the clusters, clusters chosen in the sample are then audited in their entirety.

Coefficient of Variation (CV) – the ratio of the standard deviation to the corresponding mean. The population CV measures the relative dispersion of the population distribution. The CV of an estimate is the standard error divided by the estimate.

Confidence Interval – the range of values between the lower confidence limit (LCL) and the upper confidence level (UCL) derived from the sample, which contains the true population value with a specified confidence level.

Confidence Level – the probability that the confidence interval will contain the true population value that is of audit interest.

Correlation – a measure of the degree in which two quantities are linearly related. This relationship could be one of dependence or association. Correlation in a sample or a population can be measured by the correlation coefficient, ρ. The extreme values for ρ of -1 or +1 signify an exact linear relationship. If $\rho = 0$, then the two quantities have no linear relationship (a formula for deriving ρ is included in the endnotes).

Correspondence – the method of matching random numbers to sampling units from the sampling frame.

Decision Rules – special valuation rules employed by the auditor in valuing the sample, whereby taxable error has only one chance of being anything other than zero, avoiding selection bias and possible impairment of the sample results.

Detail – a review of all transactions within an audit by the auditor using thorough, complete, and consistent procedures throughout the examination. Statisticians often refer to a detail as a census.

Difference Estimator – an estimator that backs into the total taxable error value of the population by first estimating the total audited value for the population. The total taxable error value equals total invoice value less the total audited value. The total audited value is estimated by taking the average audited value from the sample multiplied by the population count. Each individual audit value for the sample is computed by subtracting the taxable error value for that item with the invoice value for that item.

Distribution – refers to the scattering or diffusion of values in a population.

Error Rate – the occurrence rate of sampling units in the population (or sample) that have a non-zero error value.

Estimator – the methodology used to estimate some unique value from a population.

Estimator Bias – bias that occurs when the average value of the estimator (taken over all possible samples) is not equal to the population mean.

Frequency Distribution – a tabular representation of a population where the population is divided into classes or ranges and the number of population units falling into each class are counted and shown in the table.

Hypergeometric Distribution – The exact sampling distribution for the sample occurrence rate when a simple random sample of a given sample size and population is selected. The distribution will be different for any change in the sample size or population size.

Judgment Sample – any sample other than a probability sample which is picked from the population by the subjective decision of an individual and where the chance of selection is not known.

Lower Confidence Limit (LCL) – the point estimate less the precision amount and is the smallest value in the confidence interval which contains the true population value with a specified confidence level.

Mean-Per-Unit Estimator – an evaluation procedure in which the average taxable error (or other unknown) from the sample is used to estimate the average taxable error for the population. The total taxable error for the population is calculated by multiplying average taxable error from the sample by the population count.

Normal Distribution (Gaussian Distribution) – an important distribution in statistical theory used to estimate probabilities. It is a symmetric and bell-shaped distribution and is the approximate sampling distribution for many statistical estimates.

Occurrence Rate – the rate of occurrence in the population (or sample) of sample units exhibiting an attribute. The error rate represents the occurrence rate or proportion of sampling units in a population having a nonzero error value.

Point Estimate – an estimate of a parameter of a population. Most often in tax auditing, the point estimate of interest is an estimate of total error.

Population – the aggregate of transactions about which information is desired.

Population Mean – or the true mean value of the population.

Population Standard Deviation – a measure of the dispersion about the population mean.

Precision – see precision amount.

Precision Amount (Precision) – a measurement sampling error which is the measure of how close a sample estimate is to the corresponding population characteristic. It can be estimated from the sample results from a probability sample by multiplying the standard error of the estimate by a factor determined by the desired confidence.

Probability Sample – a sample where the chance of selection of every item in the population has a known, but not necessarily equal chance of selection (contrast this definition with that of a judgmental sample).

Range – the difference between the highest and lowest values in a group of items.

Ratio Estimator – the population ratio, R, equals the total taxable error divided by the total invoice amount. The sample ratio, r, is used to estimate the population ratio. The sample ratio is derived by dividing the total taxable error in the sample by the total invoice amount in the sample. The sample ratio is used to estimate total taxable error by taking r and multiplying it by the known total population invoice value.

Regression Estimator – uses the linear relationships of the taxable error values and the invoice values of the sample in addition to the known total population invoice value to estimate total taxable error in the population.

Relative Precision (Precision Percentage) – the precision amount expressed in relative terms to the point estimate.

Sample – a part of the population.

Sample Design (Sampling Plan) – a plan for sampling a population specified before sampling commences. Various sample designs include cluster sampling, multi-stage sampling, simple random sampling, and stratified random sampling. Key elements of the sample design include identification of the sampling frame and sampling unit, sample size, determination of the source of random numbers, definition of the characteristic being estimated, and identification of the estimator used to project the sample results.

Sample Mean – equals the sum of the sample values divided by the sample size. The sample mean is an unbiased estimate of the population mean.

Sample Standard Deviation – the measure of dispersion in the sample and an estimate of the population standard deviation.

Sampling Error – measured by the precision amount and is the difference between a value from a population, usually not known, and an estimate using a sample from that population.

Sampling Frame (Frame) – the list or file sampled after the original population is refined and is the means by which the target population is sampled.

Sampling Distribution – the distribution of all averages, totals, percentages, or other statistics for all possible samples at a given sample size for a certain population. If the sample is "large" then the sampling distribution will approximate the normal distribution.

Sampling Risk – the probability that the confidence interval will not contain the population parameter of interest.

Sampling Unit – each individual element of the sampling frame that can be selected into the sample.

Sampling With Replacement – a sampling procedure where individual sampling units are returned to the population before selecting subsequent units.

Sampling Without Replacement – a sampling procedure where individual sampling units are not returned to the population before selecting subsequent units.

Selection Bias – occurs when a sampling unit has a probability of selection that is different from the planned probability.

Simple Random Sample – a specific form of probability sampling where if taken without replacement from a finite population, each possible sample for sample size n from a population of N is possible and has an equal chance of being selected.

Skew – a description or quality of a distribution that is asymmetrical, where the frequencies of the values are greater on one side or the other of the most frequent value.

Standard Deviation – a measure of variability within a population or a sample. The standard deviation is the square root of the variance.

Standard Error – the standard deviation of a sampling distribution and measures the variability of the estimates.

Statistical sampling - means any approach to sampling that has the following characteristics:

(1) Use of a probability sample; and,
(2) Use of probability theory to evaluate sample results, including measurement of sampling risk.

A sampling approach that does not have characteristics (1) and (2) is considered non-statistical sampling.

Stratified Random Sample – a form of probability sampling where a population is divided up into different groups, or strata. A simple random sample is taken from within each stratum.

Stratifying – dividing a population, or sampling frame, into groups, or strata. In most cases, stratifying is done to take more than one sample – although in some cases the auditor may stratify and take only one sample from one of the groups (which is not considered stratified random sampling).

Systematic Sample – a sample of size n that is taken from a population of size N by selecting one (or more) starting points (where k equals the number of starting points), and then counting through the entire population and taken a specific unit that corresponds to the counting interval to fill the sample of at least size n. The counting interval is the largest integer less than or equal to kN/n. Starting points are generally an integer from 1 to the interval. Starting points may also be established from 1 to N, in which case the sampler must circle back to first population unit after counting through to the end of the population (unit N) to attain at least n units in the sample (circular systematic sampling). Where starting points are established using random selection techniques, that is selecting k random starting points 1 to the interval (or 1 to N for circular systematic sampling), this form of sampling is a probability sample.

Target Population – the population that is of audit interest.

Upper confidence limit (UCL) – the point estimate plus the precision amount and is the largest value in the confidence interval which contains the true population value with a specified confidence level.

Variance – a measure of variability within a population or a sample. The variance is the standard deviation squared.

3. Sampling Characteristics

3.01 What is Sampling?

Sampling is "the application of an audit procedure to less than 100% of the items within an account balance or class of transactions for the purpose of evaluating some characteristic of the balance or class."

A sample can be drawn many different ways, but principally there are two basic types of samples. An objective form of sampling involves selection based on chance, where the probability of selection is known for each item in the population but is not necessarily equal with respect to all sampling units in the population. This type of sampling is called probability sampling. Probability samples include simple random samples and stratified random samples.

All forms of sampling other than a probability sample are considered judgmental samples. Judgmental samples include large block samples that are commonly used in sales and use tax auditing.

3.02 General Reasons for Sampling

In general, an auditor will sample to reach a conclusion about taxable amounts in a population that is too large to examine in detail. Sampling is frequently necessary to realize efficiencies for both the Commission and the taxpayer. Sampling may minimize costs associated with the audit. Other reasons can include:

(a) Estimation

The auditor wishes to estimate the unknown value of tax error. This usually is the main objective for sampling.

(b) Prevention

When a probability sample is properly utilized in an audit, the audited party cannot predict which items will be chosen for review. This may inspire future improved tax compliance.

(c) Detection

Generally, it is not known whether tax errors exist. Sampling can be used to detect the presence of errors.

(d) Cost Benefit – Auditor Efficiency

Sampling is used to achieve greater auditor efficiency. Auditor time is expensive and scarce. A detailed examination is often unreasonable or impractical due to the complexity and volume of taxpayer records.

(e) Taxpayer Efficiency

Sampling can make the audit process less intrusive for the taxpayer. It may result in less audit time in the taxpayer's office and fewer records to provide for review.

3.03 Justification for the Sampling Procedure Used in the Audit

The three standards expressed above (See 1.04) will dictate which method of sampling is appropriate. No one sampling technique is best suited for all audits and a proper determination cannot be made until there is a thorough understanding of the taxpayer's operations. Auditor judgment is relied upon in determining how the three standards apply to any audit and the auditor is allowed discretion in determining sampling procedures. However, the auditor must consider the three standards and be able to justify the procedures accordingly. If judgmental sampling is used, the auditor will include a statement in the work papers explaining the reasons for judgmental sampling. Similarly, the auditor will include a statement in the work papers explaining the reasons for applying non-statistical sampling procedures to probability samples. (See 4.01 Non-statistical Sampling and Judgmental Samples.)

Non-statistical sampling procedures must be applied to judgmental samples. Although not preferred, non-statistical methods also can be applied to probability samples. A "statistical sample" is a probability sample that has been objectively evaluated using statistical methods (probability theory). The International Federation of Accountants (IFAC), an international organization of national accountancy organizations, has developed pronouncements called "International Standards on Auditing" (ISA) on various auditing topics, including sampling. A useful definition separating the two types of sampling can be found in ISA 530:

Statistical sampling means any approach to sampling that has the following characteristics:

(a) Selection of a probability sample; and,

(b) Use of probability theory to evaluate sample results, including measurement of sampling risk.

A sampling approach that does not have characteristics (a) and (b) is considered non-statistical sampling.

Statistical sampling procedures are preferred over non-statistical sampling procedures. Even so, application of non-statistical procedures to probability samples can be justified under certain circumstances. In some cases, a probability sample can be taken because it is expected that a large sample still will not provide audit results that meet the standard to judge an audit finding based on a sample. The probability sample may be evaluated using non-statistical methods.

4. Statistical Sampling Procedures Applied to Probability Samples

4.01 When to Use Statistical Sampling – Large Populations

In Commission audits, according to the standards expressed in section 1.04, unless impractical, a detail is always preferred. Where a detail is not practical, the auditor must either ignore certain audit populations, or consider some form of sample. Generally, if a detail is not possible, statistical sampling is the sampling method of choice when information about the accuracy of the sample results is desired.

When coverage of audit populations is considered through sampling, the standards in 1.04 provide that the auditor should conduct a probability sample where practical. Statistical evaluation of the results is preferred. However, the standards recognize there are times where it is not appropriate or practical to consider statistical evaluation

In determining how to audit, the auditor has five general ways of dealing with a population (in order of preference):

1. Detail examination.
2. Statistical sampling (probability sampling with a statistical evaluation – covered by sections 4.01 - 4.09(a)).
3. Probability sampling without statistical evaluation (covered by sections 5.01 - 5.10).
4. Judgmental sampling (including block sampling).
5. No audit coverage - most often this will be under extenuating facts that require no specific investigation of the detail (for example, the taxpayer makes all sales in a division of one product that is always considered statutorily exempt from tax – a sales tax verification of the detail transactions is unnecessary in that division).

Any of these approaches may be valid under certain circumstances. The standards in section 1.04 determine the approach and the auditor must exercise judgment in determining how these standards apply.

As a practical matter, however, it is presumed that most audit populations in Commission audits will be large, requiring statistical sampling. It is presumed that if electronic records are kept, they can be made available to the Commission in electronic format. Where audit populations are not large and/or not in electronic format, something other than statistical sampling will generally be considered under the standards in section 1.04. These circumstances should be identified in the audit report.

What is considered "large" will differ from audit to audit, but this most often is a population of at least 2,000 – 5,000 sampling units. This number range includes only those items that will be sampled, excluding all items which have been removed from the population during refinement or through the setting of ceilings and floors. Sampling frames under 2,000 units are generally too small for statistical sampling, although there may be exceptions. Any sampling frame over 5,000 sampling units will be considered "large". As a practical matter, certain sampling frames between these amounts need to be analyzed to see if it is practical to apply statistical sampling. To determine if sufficient items exist for statistical sampling, separate criteria should be applied to stratified and un-stratified sampling frames.

Stratification is preferred unless extenuating circumstances justify an un-stratified frame. When statistical sampling is considered, the auditor should generally stratify the frame into at least two strata. For stratified sampling frames, the number of units available for sampling should be of sufficient size where:

- Stratification is possible under the techniques described herein (see sections 4.06 to 4.06(d));
- A sample size of at least the minimums can be taken from each of the strata (see section 4.07(a));
- It is possible that at least two of the strata can be sampled where sample size does not exceed 50% of the total sampling units available in a stratum;
- There is a reasonable chance that a precision goal can be met (see section 4.08(g)).

Populations that do not meet all these criteria are not candidates for statistical sampling, unless it is practical to use an un-stratified frame. In addition, populations under 2,000 sampling units that meet these conditions can be statistically sampled. Also, populations over 5,000 sampling units which do not meet these conditions are not candidates for statistical sampling.

Statistical sampling techniques can be applied to an unstratified frame if the sampling frame is sufficiently large enough to meet both the following conditions:

- Sample size is at least 300; and,
- There is a reasonable chance that a precision goal can be met (see section 4.08(g)).

4.02 Sampling Frame and Stratification Guidelines

Things to consider when developing and stratifying sample frames:

- In statistical sampling, the unbiased selection of the sampling units is important. Judgment is removed in the selection process. On the other hand, experience, intuition, and judgment are effective in refining and stratifying sampling frames.[1]
- One of the best strategies to provide better precision is to remove items that have little or no potential for material error.
- Stratification of the sampling frames often leads to better precision.
- Every time a projection is made, sampling error will result. Optimally, as few projections as possible should be used to minimize overall sampling error. However, there are frequently times where many individual projections are required.[2]

4.03 Handling of Negatives

This section deals with the handling of "negatives" – meaning accounting entries which correct, change, or cancel a business transaction's recorded value in the books and records of the taxpayer. It also includes the accounting entries that reclassify various transactions within the records by transferring recorded amounts in accounts.

Note that the auditor will usually take a sample to determine total taxable error for a target population. For each sample unit, the auditor might have some accounting value, such as an invoice amount, book value, or other recorded amount. The auditor will examine the selected sampling units, and will in most cases, establish a taxable error value for each unit. The meaning of "negative" as used in this section does not refer to tax overpayments that the auditor values as a negative error amount.

In taking a sample, the auditor should sample from only the "positive" amounts. Although negative dollar amounts are not directly sampled, they are not ignored in the sampling process. Any negative relating to a transaction selected into the sample should be reviewed. Also, extraneous positive and negative items that zero out (items in the same transaction offset each other and total zero) should be removed from the sampling frame before sampling. Occasionally, negative items may be found which partially reduce corresponding positive amounts – these can be replaced by a net positive amount in the sample frame. The process of matching and removing related sampling units that total zero should be done prior to making further refinements to the sampling frame. Note that the more extensive and complete the download from the taxpayer, the more likely that matching positives and negatives will be found. Normal procedures for handling negatives in the sampling frame are as follows:

- When the download of transactions contains negative items, every reasonable effort will be made to match them with corresponding positive items. Ultimately the ability to match negatives with corresponding positives is the responsibility of the taxpayer in that the taxpayer is required to maintain adequate records and information on business transactions. The auditor is responsible for performing this procedure once sufficient information is obtained from the taxpayer. The auditor should use all relevant data elements (customer name, invoice number, date, product code, PO number, etc.) in the matching process. The auditor is responsible for identifying and utilizing all relevant elements.
- Remove all zeros and negative items from the file that represent canceled transactions. In some situations, it may be beneficial to consolidate several positive and negative accounting entries into one sampling unit. This is possible if the net amount is a positive value, and the sampling unit does not change. Sampling units need to remain consistent throughout the population, either at an invoice level or an item level.

When the taxpayer is given the sample items, they will have the opportunity to match any remaining negative items with the sample items drawn.

It is often difficult for some to understand why – as a matter of procedure – that no sample is drawn from the negatives. This failure to understand frequently arises for two reasons. First, it may not be immediately apparent that sampling from negatives can cause selection bias. And secondly, taxpayers may feel that part of the accounting records will not be covered adequately for audit purposes if negatives are not directly sampled.

The nature of typical accounting files and inclusion of the negatives as sampling units may bias the sample selection process. Accounting files represent a target population of tax transactions, and their tax error values. (Note that the tax error values can be either positive or negative.) Accounting files often do not provide for “clean” representations of target populations.

A sampling frame that corresponds to the target in such a way that each sampling unit refers once, and only once, to a specific and discrete part of the target population is optimal for sample selection purposes. Here, “discrete” means a separable and unique part of the target population. Therefore, when a sampling unit is randomly selected into a probability sample (without replacement), that portion of the corresponding target has only one chance at being selected into sample along with all other discrete portions of the target. However, if more than one sampling unit relates to a portion of the target population – that part of the target is more likely to be selected and valued for error when compared to the planned probability of selection for that discrete portion.

Accounting adjustments (negatives) will almost always (if not always) represent a transaction in the target that was previously represented by a positive. No negative transaction has ever initiated a sale or purchase. Note that some transactions in the target population may have no corresponding negatives, while others may have one or more corresponding negatives.

All efforts should be made to avoid bias in the sample selection process. Inclusions of the negatives may lead to such a bias where discrete portions of the target have varying probabilities of being selected which are not equal to the planned probabilities of selection. Selection bias cannot be easily cured or handled after the sample is drawn. Random selection from only the positives – after other extraneous positives have also been removed – provides a means of identifying the discrete portions of the target that should be valued for error.

Note that in some cases, it is difficult or impossible to remove extraneous positives (or even negatives when sampling from non-electronic files). If extraneous positives (and/or negatives) still exist in the sampling frame, a decision rule should be implemented as described in section 6.04. An extraneous positive is where more than one sampling unit with positive accounting values exists in the sampling frame and relates to the same discrete part of the target population.

4.04 Defining the Audit Objective and Refining the Population

When performing an audit examination on a target population, the auditor should have a clearly defined audit objective (sometimes the auditor may have multiple objectives). Stating this objective enables the auditor to identify the best sampling frame to cover the target population. Based on the objective, the auditor can also employ further refinement procedures to minimize sampling error.

An attempt should be made to remove items that have little or no potential for error based on the audit objective.[3] This can be done by many different means:

- Removing transactions outside the audit period.
- Identifying transactions in the accounts of interest.
- Identifying transactions with pertinent cost codes, locations or cost centers which are being audited (such as excluding out-of-state locations or cost centers).
- Identification of certain customers or vendors for exclusion.

When refining the sampling frame, the auditor should discuss these procedures with the taxpayer. The auditor should make all efforts to reach an agreement on what portions of the download will be sampled. If agreement is not reached, generally those portions should remain in the sampling frame. However, identified portions should be removed if it is known that inclusion will seriously impact sampling error and these items will likely not be material to overall results if included in the sampling frame.

For example, where it is known errors are very rare and will likely have little impact in the final results ("needles in the hay-stack"), inclusion of those portions of the download could cause sampling error to be much larger than what would have been had these portions been removed. For this reason, auditors are not obligated to construct sampling frames only when complete agreement is reached with the taxpayer.

4.05 Determination of Limits

Consideration should be given to removing very high and low dollar amounts from the sampling frame. The auditor should make all efforts to reach agreement with the taxpayer in setting these limits.

(a) High Dollar Limit – "Ceiling"

Removing high dollar items from the sampling frame and examining them in detail is an effective strategy for improving precision. This procedure is referred to as setting a ceiling. The auditor sets an upper limit and all items above the limit are considered above the ceiling and are removed from the frame.

Removing high dollar items from the sampling frame effectively reduces the range of dollar values in the non-zero errors. This results in a lower standard deviation in the errors; therefore, greater precision of the estimate.

Consideration should be given to how high a dollar amount the taxpayer would be willing to project. Consideration should also be given to how many items can be detailed before compromising efficiency. Setting the upper limit too low creates inefficiencies.

High dollar items examined on a detailed basis are not included in projection calculations.

(b) Low Dollar Limit – "Floor"

Removing low dollar items from the sampling frame by establishing a lower limit is referred to as setting a floor. When establishing a floor, the auditor should consider the impact of the total dollars between $0 and the lower limit, as these dollars will be ignored.

Materiality is an important consideration when setting a lower limit. Although a $2.00 transaction may appear immaterial, that single transaction in a probability sample may represent thousands of transactions and hundreds of thousands or even millions of dollars in some populations. The auditor should analyze the data before establishing a lower limit.

One relevant analytical test is a "what if test". For example, if the error rate was "x%" on all transactions between $2.00 and $3.00 and that error rate was projected against the population of items within those boundaries, what would be the potential tax impact? When performing analytical tests, auditors should use good judgment when selecting criteria such as assumed error rates. Prior audit results may be a useful source in determining such criteria.

If the amount under the proposed lower limit is determined to be immaterial, or if there is not a reasonable cost-benefit ratio, the auditor should consider setting a floor. The Commission does not project audit results to items below the floor.

4.06 Stratification

Stratification of the sampling frame normally leads to better precision when compared to a simple random sample of equal size. However, while initially adding strata typically results in substantial precision gains, at some point adding more strata will decrease efficiency and provide no net benefits. Also, in rare situations, stratification will not provide a material improvement of relative precision.

Stratification in terms of dollars is used to reduce the effect of extreme values (errors). By reducing the variability of the population within each stratum, stratification often reduces the sample size needed to achieve a desired level of precision and reliability. Unless a population is homogeneous, stratified sampling is the Commission's preferred method to be used.

(a) The Number of Strata

The range between the low and high dollar items in the sampling frame is the biggest factor in setting the number of strata. Usually, this will be anywhere from 2 to 6 strata (high dollar items removed from the sampling frame for detailed review via setting a ceiling are not considered a stratum). Typically, setting up more than four to five strata will yield only minimal gains.

(b) Strata Boundaries

There are a few rules which must be applied when employing stratification:

- Every unit of the sampling frame must belong to only one stratum.

- The characteristics which define the strata should be made prior to sampling (usually, the dollar range of the stratum).
- The exact number of units within each stratum must be known prior to sampling.

(c) The Types of Characteristics Used in Stratifying

Strata may be set up in one of two ways:

1. By a quality, such as location or time period (generally referred to as "grouping" rather than stratifying); or,
2. By a quantity such as invoice amounts or GL distribution amounts.

(d) Stratifying On Amount

Generally, stratification is based on invoice amount. This can be done through one of two methods:

1. Cumulative Square Root of the Frequency Method; or,
2. Proportional.

The Commission prefers the cumulative square root of the frequency method. Refer to Appendix D for a detailed explanation. The method is well documented in statistics references and is used by a number of taxing authorities. This method allows for greater efficiency in the sampling process compared with setting strata boundaries using judgment only.

Using judgment to determine the strata break values is perfectly acceptable and will not invalidate the sample or its results but may not be the most efficient method.

4.07 Sample Size

Statistical sampling differs from non-statistical sampling in several ways. One important difference is that sample size can be determined based on prior knowledge. This knowledge can be exploited to set sample size at a point where a desired degree of accuracy is attained.

Formulas may be used to statistically arrive at sample size based on this prior information. This information includes the estimated standard deviation and the average of the taxable error value in the population. Two other factors which influence sample size are the desired confidence level and precision. Unfortunately, though, often information concerning the average and standard deviation of the taxable errors is either not available, too impractical to obtain in advance, or is simply just a guess. However, absence of this information does not preclude statistical sampling or arriving at a statistically based sample size as the auditor can make educated assumptions to replace this information (Note: Assumptions may not hold true in the final sample results).

Both simple random samples and stratified random samples may be statistically evaluated even if sample size was determined using nonstatistical means. The International Federation of Accountants' standard on sampling, ISA 530 discusses sample size:

> When applying statistical sampling, the sample size can be determined using either probability theory or professional judgment. Moreover, sample size is not a valid criterion to distinguish between statistical and nonstatistical approaches.

In deriving sample size, the auditor can use any of the following methods:

- Using classical statistical formulae which estimate sample size based on prior information (typically this information is gained from a probe sample) (section 4.07(b)).
- Using statistical formulae based on the error-rate model that substitutes some of the needed prior information with assumptions or knowledge concerning the rate of occurrence of the non-zeros in the population (section 4.07(c)).
- An attribute method which attempts to observe a minimum number of nonzero errors in the sample given an error rate (section 4.07(d)).
- A negotiated sample size with the taxpayer.
- Consideration of the largest sample which can be practically done with the time and resources available.

Any of the above methods may be used, provided they arrive at a sample size that is equal to or greater than the required minimums discussed in section 4.07(a).

(a) Minimum Sample Size

Sample size should not be less than 300 sample units for unstratified sampling.[4] In stratified samples, no stratum should have a sample size of less than 100. In most stratified sampling frames, the sample size will be 200 or more per stratum depending on the desired relative precision.

(b) Classic Formulary Approach

The classical formulary approach requires the following to arrive at a statistically based sample size:

- An estimate of the standard deviation of the error amounts in the sampling frame;
- An estimate of the average error amount in the sampling frame;
- The desired relative precision; and
- The desired confidence level.

Unfortunately, the first two pieces of information are very difficult to determine without performing a pilot sample. Pilot samples are frequently impractical. It is possible the standard deviation and average taxable error computed in the probe sample may be significantly different than what will be encountered in the sample, meaning material over or under sampling is possible even when a probe sample is used.

Note there is also a rule of thumb concerning sample size and relative precision:

To achieve a two-fold increase in relative precision, such as when 15% is desired and 30% was attained, the sample size needs to be quadrupled. For example, if the sample was initially 800, it likely needs to be set at 3,200 in order to attain desired relative precision.

The confidence and precision levels used to estimate sample size should be identical to those used in the final evaluation. Most often, these will be set at the stated levels shown in sections 4.08(f) and 4.08(g), unless superseded by state policy.

(c) Error Rate Model

Instead of taking a probe sample, which is generally required in the classic formulary approach (section 4.07(b)), an estimation of the error rate of the population can be used to estimate the average and standard deviation of the error amount. The error rate model requires information as follows:

- An estimate of the error rate (rate of occurrence of nonzero taxable errors)
- The standard deviation of the invoice amounts
- The average of the invoice amounts
- The desired relative precision; and,
- A confidence level

An educated guess can be made as to the error rate. In an audit using electronic data, the standard deviation and average are known (in a manual audit, these can be estimated).[5] Commission policy sets the relative precision and confidence level.

The auditor may make an estimate of the error rate through use of a pilot sample. Most of the time this is not practical, and history or auditor judgment must be used to estimate the error rate. If the error rate is set using judgment, a low error rate should be considered.

However, error rates below 2% typically produce impracticably large sample sizes. If there is no history, sample size can be set using a 2% rate of occurrence (error rate) standard.

(d) Sample Allocation in Stratified Sampling

Once the number of strata and strata boundaries have been determined, the overall sample must be allocated to the strata. The primary consideration is that sample size be sufficient within each stratum to find and weight errors if they exist. Two acceptable methods of sample allocation are:

1. Proportional
2. Optimal

The preferred method used by the Commission to allocate samples is the optimal basis. See Appendix E.

4.08 Evaluation and Projection of Sample Results

If the minimum number of errors (3 per stratum) is observed in a probability sample designed by the CAS or auditor, a statistical evaluation should always be done. These results should be disclosed to the taxpayer, whether or not an actual projection is made.

(a) The Four Estimators

In projecting a total error amount, four estimation procedures are available. The first two, mean-per-unit and difference, measure a certain value for each unit in the sample. The other two methods, ratio and regression, measure a relationship between two numbers found for each sample unit. All four estimation methods will be calculated in statistical evaluations.[6]

1. Mean-Per-Unit Estimation - measures the taxable errors in the sample to arrive at the estimated total taxable error. Mean-per-unit estimation ignores the invoice amounts or book values of the sample units. Frequently, this method provides results similar to the ratio and regression estimators, but usually is slightly less precise.
2. Difference Estimation - uses the audited values found in the sample to estimate the total audited value. Using an estimate of the total audited value, an estimate of total taxable error is made. The overall variance of the estimated total audited value is typically large for low error rate populations. Consequently, this method is useful only in high error rate populations.
3. Ratio Estimation - In sales and use tax auditing, the overall relationship between the error values and the invoice values is used to project the total error. There is generally some correlation between the invoice amounts and the taxable error

values in typical populations sampled.[7] When some degree of correlation exists, ratio estimation often provides for better precision than either mean-per-unit estimation or difference estimation.

4. Regression Estimation - like ratio estimation, regression estimation measures the relationship between the invoice amounts and taxable error values found in the sample to arrive at a total estimate of total taxable error. Regression estimation provides a linear measurement of this relationship and is usually the most precise estimator.

(b) Restrictions for Ratio and Regression Estimators

When an estimator is biased, this means that over all possible samples of a given size, the average value of all the estimates does not equal the actual population value. In an unbiased estimator, the two values will be equal.

The mean-per-unit and difference estimators are unbiased. Ratio and regression estimators have some bias. However, because of the expected gain in precision coming from ratio and regression estimators, these two methods are frequently preferred, despite the bias.

There are statistical concerns,[8] including estimator bias, that preclude the use of ratio and regression estimation when one or more of the following occurs in the sample:

- The coefficient of variation (cv) for the estimated invoice amounts is over 10%
- The coefficient of variation for the estimated audited amounts using either the mean-per-unit or difference estimator is over 10%
- The overall sample size is less than 100.

In computing the cv values above, any detailed stratum is ignored. Because the mean-per-unit and difference estimators are inherently unbiased, none of these conditions preclude their use. Finally, the ratio estimator should not be used when the invoice values include both positive and negative values.[9]

(c) Combined Versus Separate Estimation for Ratio and Regression Estimation

In stratified random samples, ratio and regression estimation may be applied using two different approaches: "separate" or "combined". In either the separate or combined approaches, there will be an overall statement with a point estimate and confidence interval for all sampled strata taken as a whole. The separate approach will provide a different overall point estimate and precision computation when compared to the combined approach. In addition, in separate estimation, the point estimates are additive when coming up with an overall point estimate. This is not true of the combined approach.

If requested by the taxpayer, either separate ratio or separate regression estimator can be used as the final estimator to project error if:

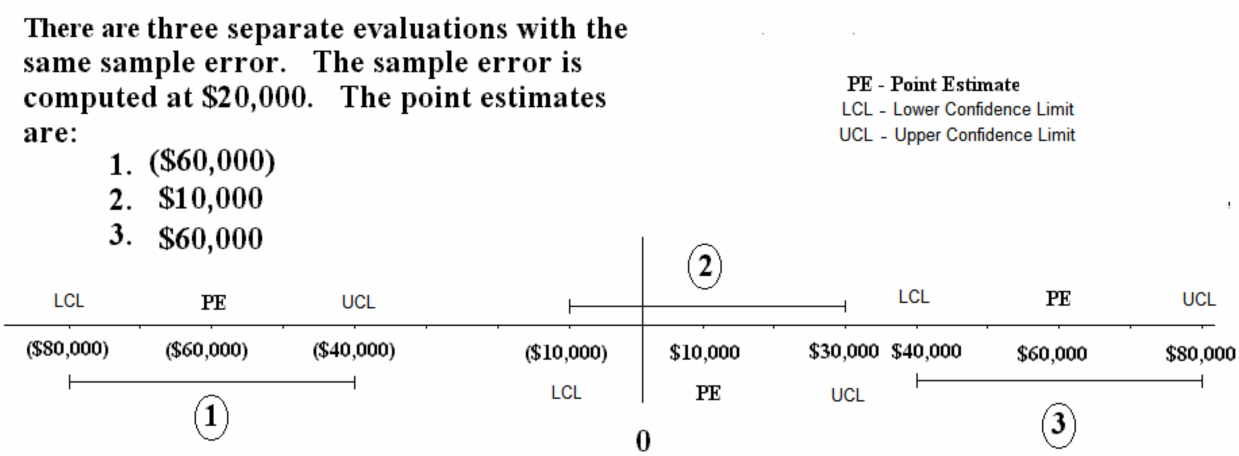
- It is agreed to be part of the sample plan prior to selecting the sample,
- It will provide for the better precision when compared to other estimators (including the combined estimators), and
- The restrictions shown in section 4.08(b), when applied to the strata individually, are satisfied.

(d) Evaluating the Estimators

For each estimator, a precision amount can be calculated using standard statistical formulas. The confidence interval is then computed by adding and subtracting the precision amount to the point estimate (PE). By nature, the width of the confidence interval is always twice the precision amount. The confidence interval is bounded by the upper confidence limit (UCL) and lower confidence limit (LCL).[10]

The procedure for arriving at the UCL and LCL is the same for both positive and negative point estimates, and is shown as follows:

<u>Exhibit 4.1</u>



An auditor can recommend an adjustment at the PE, LCL, or UCL. The PE represents the most likely adjustment (for example #1 of Exhibit 4.1 this is negative $60,000, and for example #3 this is a positive $60,000).

The confidence limit closest to zero represents the least expected amount of the adjustment at a given confidence level (for example #1 the UCL of negative $40,000 and for example #3 the LCL of positive $40,000). The confidence limit furthest from zero represents the greatest expected amount of the adjustment at a given confidence level (for example #1 the LCL of negative $80,000 and for example #3 the UCL of positive $80,000).

Example #2 of Exhibit 4.1 is an example where the possibility of both refund and assessment positions are contained within the confidence interval. Notice that the LCL is a negative $10,000 and the UCL is a positive $30,000. Consequently, the sample results cannot be used to recommend an adjustment based on a projected result. In that event, the only options available are adjusting the sample items without projection or selecting an entirely new and larger sample. Note that a larger initial sample size may have provided a confidence interval where the UCL and LCL are of the same sign.

(e) Determining the Optimal Estimator

Statistically, the narrower the confidence interval, the more reliable the point estimate. Therefore, the Commission's policy is to use the estimator which provides the smallest precision amount as the basis for projection. The relative precision of an estimator will not be used as a basis of deciding amongst the estimators (however, in order to support an adjustment, the estimator must provide sufficient relative precision as set out in section 4.08(g)).

Also, the UCL and LCL of the confidence interval must have the same mathematical sign to support an assessment (examples #1 and #3 in exhibit 4.1). If the signs are different, that is the LCL is negative and the UCL is positive, the sample results indicate both a refund and a tax assessment are reasonably possible. Therefore, no adjustment using the sample results will be recommended. If the UCL and LCL from any of the estimators are of different signs, this does not preclude the use of another estimator where the mathematical signs of the UCL and LCL agree.

(f) Confidence Level

Samples will be evaluated using at least the 90% confidence level (two-sided). If the taxpayer requests a confidence level higher than 90%, this should be addressed before the sample size is determined.

(g) Relative Precision

The goal in setting sample size is to achieve a relative precision of 30% or better. If this is not attained, the following options should be considered:

- The auditor and the taxpayer agree to project the results as is.

- Reevaluate the sample using a 1-sided 95% confidence level. In this case, the projected taxable adjustment that will be recommended is the confidence limit closest to zero, the lower limit, provided both confidence limits (upper and lower) are of the same sign.
- The auditor and the taxpayer agree to increase the sample size over all strata or within a stratum.
- The auditor and the taxpayer agree to re-sample entirely. See Section 4.08(i).

(h) Minimum Number of Errors

Each of the estimation methods described above: mean-per-unit, difference, ratio, and regression - will provide a confidence interval around the point estimate. Creation of a reliable confidence interval assumes the underlying sampling distribution for that estimator is approximately normal. Where the underlying sampling distribution is not normal, the UCL and LCL may not be consistent with the expected probabilities. For example, given a two-sided 90% confidence level, the risk that the true unknown is less than the LCL is about 5% given a normally distributed sampling distribution. Similarly, there is about a 5% risk that the true unknown amount is greater than the UCL. Given a non-normal or materially skewed sampling distribution, the true risks in this example for the LCL and UCL might be materially different than the expected 5%. Further, it is known that as the proportion of sampling units with a zero value increases in the population, the more likely it becomes that the underlying sampling distribution is not normally distributed. To counter this, larger sample sizes are needed for populations with a large proportion of zero values. Instituting a standard whereby a minimum number of non-zero error amounts are required will aid in the effort of achieving a large sample.

To observe the required minimums often calls for very large sample sizes. In Commission audits, most of the concern is over the total amount of the taxable error. Samples are usually taken from populations that contain many non-errors (zeros) and a few errors (non-zeros). The occurrence rate of errors could be smaller than 1%. Many zeros in the sample may cause problems, particularly with the assumption that the underlying sampling distribution is normally distributed as explained above.

Therefore, in Commission audits, at least three errors must be observed in any stratum included in a statistical evaluation. Note that this policy pertains only to probability samples that are statistically evaluated. When one or more strata contain the required minimum of three or more errors, but the remaining strata do not, the strata exceeding the minimum can be evaluated for projection. Actual errors noted in the remaining strata may be assessed or disregarded but may not be projected.

(i) Taxpayer Disagreement with Results

If a taxpayer disagrees with results from a statistical evaluation the auditor should discuss concerns with the taxpayer to determine the areas of disagreement and should review the sampling process to identify and/or correct any deviations from the sampling plan. The auditor should discuss unresolved issues with the CAS, his/her audit supervisor and/or the Audit Director.

If there is concern about precision, Commission and the taxpayer can agree to increase the sample size over all strata or within a stratum. If the taxpayer wishes to expand the sample (rather than re-sampling) when results otherwise satisfied the relative precision requirement, the Commission will not agree to enlarging the existing sample unless the taxpayer agrees to the results of the increased sample before new items are audited. If it is agreed that the results of a prior sample will be increased - rather than taking an entirely new larger sample - no portion of the sample will be ignored or replaced. If the taxpayer requests expansion of the sample, or an entirely new sample, the taxpayer is responsible for retrieving the additional documents for review.

(j) Adjusting the Sample after the Fact

Even if the sample results are distorted, removing any single item from the sample may not be done. However, it is permissible to identify and remove from the target population a class of items. This requires the qualifying class items be removed from both the sampling frame and sample. Therefore, if items identified to be of a certain class are removed from the sample, they should also be removed from the sampling frame along with other sampling units not selected into the sample. Examples of items that can be considered a class include things like shop supplies, assets, divisions, and transactions made in certain locations.[11]

4.09 Documentation Requirements

The following details minimum documentation requirements for statistical samples:

- Maintain a description of the sampling plan (refer to Appendix A)
- Maintain a description of the sampling frame. This description should include details on counts and amounts that have been removed from the data download to arrive at the sampling frame.
- Identification of the source of random numbers. The seed numbers provided by the random number generator should always be documented.
- Any reports made by the program to document the seed should be retained.

- Include instructions to the parties involved in reviewing the sampling units. The auditor should be provided instructions on basic procedures concerning the sample and electronic files (refer to Appendix C).
- Include work papers containing descriptions which support the audit conclusion for each sample item. A zero or non-zero error amount should be established for each sample item. A non-error sample item should not be deleted from the files. Original sample amounts and other data fields should not be edited.
- Identify and explain special valuation items (see Section 6).
- Include work papers which contain the evaluation and projection procedures. This should also include a statistical statement of confidence or precision (refer to Appendix F).
- When appropriate, include a statement as to any unusual factors or complications in the sampling process and their implications.

4.10 Audit Report Requirements

Audit report narratives should describe sampling methods used in conducting the audit. Factors contributing to the auditor's decision to use a particular type of sample should be described. When a statistical sample is used, a separate document should be attached to the narrative that includes the following information:

Objective – a general statement such as "investigate the proper application of retailers' sales tax to sales transactions" or whatever the case may be is sufficient.

Target Population – describe the data in the original download as well as steps the auditor performed when refining the data to arrive at the target population. Examples include removal of appropriately taxed transactions and credit memos. When accounts of interest are used in a purchases sample, the auditor should include a list of the accounts of interest. Include a summary of the sampling frame.

Sampling Procedures – describe the sample (number of strata, floor, and ceiling values) and include a summary of the strata.

Sample Evaluation and Projection – describe the evaluation result and any state specific guidelines that may apply.

5. Procedures Applied to Samples

5.01 Non-Statistical Sampling and Judgmental Samples

If it is not practical to apply probability theory to evaluate the audited results from a probability sample, the auditor can elect to use non-statistical procedures, provided non-statistical methods can be justified under the three standards expressed in section 1.04. Sections 5.01 through 5.10 apply when statistical procedures are not applied to samples.

Although sections 5.01 through 5.10 generally apply to non-statistical sampling procedures applied to probability samples, these concepts can also be applied to judgmental samples.

Non-statistical sampling techniques will normally be used when the taxpayer is unable to provide complete electronic records of the population under examination. Due to the size of most taxpayers the Commission audits, the auditor should initially assume the taxpayer will be able to provide electronic records. Therefore, the use of non-statistical procedures warrants justification within the audit report.

5.02 Defining the Audit Objective

Before conducting any type of sample, the auditor should clearly define the audit objective(s). In some circumstances the auditor may have more than one objective. The auditor's objective(s) should be discussed with the taxpayer and be included in the sampling plan.

Defining the audit objective serves several purposes:

- Provides the taxpayer a clear understanding of the auditor's intentions.
- Aids the auditor in defining the appropriate population(s) to examine.
- Assists the auditor in determining which areas might be sampled, and which should be reviewed in detail.

5.03 Define and Refine the Population

The auditor should give careful thought to the population being examined. Careful and thorough refinement will produce the most accurate audit results. Proper planning at this stage is key to auditor efficiency.

The goal of refinement is to remove known zeros (or non-errors) from the population prior to pulling a sample. When records are not available electronically, performing a thorough refinement is difficult. However, there are still some broad refinement procedures which may be utilized, such as removing all transactions with a specific customer or vendor that we know will likely have zero audit exceptions.

5.04 Sampling Unit, Sampling Frame, and Items of Interest

Sampling units are the individual elements of the population/sampling frame which will be audited if selected into the sample. For example, if the population consists of sales, the sampling unit could be each sales invoice. The sampling frame here will be the list of all sales invoices. Another example might be a sampling unit consisting of transactions grouped together for a time period. Suppose the auditor has a 36-month audit period. All transactions or invoices, either sales or expenses, are grouped (or clustered) together in each month. The auditor will sample the months as sample units. If selected, all invoices for a selected month are audited. How the auditor defines sampling units is generally a function of what is available.

There may be more than one choice in how the auditor could define the sampling unit. The goal is to obtain the sampling unit which provides the least variability between all sampling units in the target population. In cases where more than one sampling unit is possible, a determination will need to be made as to which provides the greatest efficiency. Efficiency in this context is defined as a comparison of the cost associated with locating each sampling unit, the expected return, and the desired accuracy. Even though a sampling unit may yield a higher degree of accuracy (or precision), it may not always be the best choice. Likewise, the most convenient sampling unit may yield results too inaccurate (or imprecise) to be usable.

The following is the Commission's preference of sampling units. These are meant to be general guidelines for helping the auditor decide which sampling unit should be used:

(1) Electronic records…line items or invoice level
(2) Transactions (e.g., invoices, checks, purchase orders)
(3) Clusters
(4) Time periods
    (A) Days/Batches
    (B) Weeks
    (C) Months
    (D) Years

Sampling units can be chosen into the sample using either probability or judgmental sampling methods; however, probability sampling should be used if possible. Note that if the auditor has complete electronic records (1), it is expected statistical methods will be applied as described in sections 4.01 through 4.09. If statistical methods are not applied, this is an exception that the auditor should explain in the audit report.

When using a transactional sampling unit (2), statistical or non-statistical methods may be applied if the sample is a probability sample. The most common reason for applying non-statistical methods over statistical methods is that statistical methods would require an impractically large sample. Whatever the reason, the auditor should explain the choice of using a non-statistical method in the audit report.

The auditor may apply statistical methods to cluster samples (3) that qualify as a probability sample. However, it is more common to apply non-statistical procedures to a cluster sample.

When using time-period samples (4) (actually a form of cluster sampling), it is usually not appropriate to apply statistical evaluations even if these samples qualify as probability samples.

The collection of all sampling units which have a chance of being selected through probability sampling methods, makes up the sampling frame. The sampling frame constructed by the auditor (as described in section 5.03) should include any sampling units which:

(A) Have a likely impact on the amount of reported sales or use tax ; or,
(B) Is included in the scope of the audit objective.

The criteria (A) and (B) describe items of interest. For example, sales of tangible property would likely be of interest, while sales of company stock will likely not be. The auditor is to use judgment when determining items of interest and when removing items not of interest during refinement. The greater the percentage of items not of interest remaining in the frame, the larger the sample must be for both statistical and non-statistical methods. Omitting items of interest from the frame will necessitate the auditor performing a separate examination for these items, leading to inefficiencies.

Pursuant to the disclosure standard in section 1.04(c) the auditor is obligated to describe to the taxpayer what constitutes an item of interest prior to sampling. The auditor should consider any reasonable request made by the taxpayer in modifying or expanding the scope of what is defined as an item of interest and report any objections in the final audit report if no agreement was reached.

Although it is preferred that the auditor remove any items not of interest from the sampling frame, it is not necessary to remove all these items. A sampling frame which contains items not of interest is still valid but will likely require larger samples to attain acceptable accuracy when compared to a frame which has these items removed. Therefore, under the best evidence standard the auditor is obligated to make a reasonable attempt, given the time, information, and resources available, to remove items not of interest from the sampling frame.

5.05 Sampling Method

The primary reason for probability sampling in non-statistical sampling is to provide an unbiased selection from across the sampling frame (note that the sampling frame is created using auditor judgment). Eliminating any bias in the selection process is of utmost concern to the Commission. By using probability sampling, selection bias (known or unknown) - that of the auditor or taxpayer – can be eliminated.

Probability sampling will also lead to more precise results over judgmental sampling where sample size is approximately equal. In many cases probability samples of significantly smaller size will also result in similar or even greater accuracy when compared to judgmental samples. Probability sampling also increases the defensibility of the sample. While the auditor and taxpayer may prefer a monthly sample to be chosen non-randomly (using judgment) for the sake of convenience, random selection of the months should be used.

Probability samples, including simple random samples, should be taken only using software sanctioned by the Commission. If the auditor intends to base audit findings on a probability sample drawn using any other means, including samples drawn by the taxpayer or their representative, this must be approved by the audit supervisor.

If the records are available electronically, the auditor should use stratified random sampling rather than a simple random sample (unless the population is too small to justify a stratified random sample). When the population is not available electronically and stratification is not practical, the auditor should consider a simple random sample, provided that the auditor can establish correspondence which allows the matching of random numbers to sample units. When correspondence is difficult or impractical, the auditor should consider judgmental sampling.

5.06 Sample Sizes

Choosing the appropriate sample size is of critical importance even in a non-statistical sample. One rule to remember about sample size is that more is generally better when considering accuracy. However, this does not mean we want or need larger sample sizes than what is required to obtain the desired results. Choosing the appropriate sample size is always a trade-off between auditor efficiency and precision.

The formula for computing non-statistical sample sizes is:

$$\textit{Minimum Sample Size} = \frac{\textit{Items of Interest}}{\textit{\% of Items of Interest in Population}}$$

The auditor should use the following guidelines when calculating sample sizes for non-statistical samples.

(a) Transaction Sampling – Sample Size

If sampling is conducted on a transactional basis, where a single sampling unit is a complete transaction (typically an invoice), it is common for the transaction to contain both items of interest (potentially taxable) and items which are not of interest (not taxable). The minimum sample size for this population should have a likely probability (at least 50% chance) of uncovering at least 250 items of interest.

Because sampling units contain both items of interest and items not of interest, the minimum sample size will exceed 250 items in most cases. The auditor may not draw a sample of less than 250 items without the approval of the audit supervisor.

Before drawing the sample, the auditor should discuss the nature of the transactional information with the taxpayer to uncover any information which would impact the appropriate sample size.

The Commission may increase the size of the sample if the original sample proves to be insufficient. Before increasing the size of the sample, the insufficiency should be discussed with the taxpayer and the audit supervisor. The taxpayer also has the right to request an increase in sample size. Should the taxpayer request an increase it will be obligated to provide the auditor with the additional sampling units for review.

(b) Cluster Sampling – Sample Size

Cluster sampling is used for efficiency reasons. There is a trade-off between efficiency and precision. Some precision will be sacrificed for efficiency when using cluster sampling. The minimum number of clusters to be selected by the auditor when using cluster sampling is 100. The number of elements in the cluster (cluster width) must be sufficient so that at least 250 items of interest are likely to be uncovered in the sample.

It should be remembered that when using cluster sampling the cluster becomes the sampling unit and not the components which make up the cluster (often an invoice or line amount). The auditor will still audit the same number of transactions computed in the sample size, but it is the cluster which will be evaluated and not the transactions.

(c) Time Period Sampling

Time period sampling is not the preferred method of sampling and should only be used when it is not possible to use the transaction as the sampling unit. The auditor will include in the audit narrative the reasons why transaction sampling could not be used.

Although selection of the time period sampling units should be made using random methods, the samples themselves should never be considered as satisfactory for statistical sampling evaluation.

- **Day as the Sampling Unit – Sample Size**
  When using day sampling as the sampling unit the minimum number of days that will be examined will be 30.
- **Week as the Sampling Unit – Sample Size**
  When using the week as the sampling unit the minimum number of weeks that will be randomly selected and examined will be 20.
- **Month as the Sampling Unit – Sample Size**
  When using the month as the sampling unit the auditor will randomly select 1 month for each 12 months under audit.

In the event the sample size required under this section is not reasonable or achievable, the auditor will discuss the issue with the taxpayer and obtain approval from the audit supervisor before adopting an alternative sample size.

5.07 Systematic Sampling

Under some circumstances, the most convenient method of selecting the sampling units could be the use of systematic selection techniques. If done using random procedures, a systematic sample can be considered a probability sample. Generally systematic sampling will be considered when it is difficult or impossible to match random numbers with sampling units (correspondence).

If conducted properly, a systematic sample can provide sufficient evidence such that a satisfactory, reasonable, and proper audit finding can be made. This is possible since the sample items are spread evenly across the population.

Unlike a simple random sample, a systematic sample has a weakness, in that it is much more susceptible to problems if any continuous and repetitive patterns exist in the population. Should the systematic sample pick up a pattern in the population, it may cause the sample to be deemed insufficient for providing a satisfactory, reasonable, and proper audit finding. To avoid this, the Commission policy is that the auditor will use multiple starting points.

The auditor should use good judgment in determining the number of starting points to use. The minimum number of starting points for any systematic sample is three. Before attempting a systematic sample, the auditor should be familiar with and follow the steps as explained in the MTC sampling courses. If needed, the auditor should seek assistance from a CAS.

5.08 Sufficiency of a Non-Statistical Sample

The aim of the auditor is to substitute partial coverage for total coverage without material dissimilarity between the overall results of either method. Without applying statistical sampling evaluation procedures, it is not possible, using any objective method, to obtain an audit finding which accurately measures, with some reasonable degree of confidence, whether the total error projected from a probability sample estimates the true and unknown total error. However, a sample result may show, without using statistical evaluation procedures that:

- Errors have been made by the taxpayer in the selected sample units
- It can be reasonably be expected, using auditor judgment, that additional sample units in the sampling frame, which were not selected into the sample, will be in error
- It can be expected, again using judgment, that such errors not sampled are of the same or similar nature to those in the sample at approximately the same rate of occurrence and/or proportionate value.

If the nature of the errors found in the sample allows the auditor to conclude other un-audited errors still exist in the population, the auditor should consider coming to an audit finding that recommends an adjustment or change to the taxpayer reporting. This adjustment amount can be made based on further auditing of additional sampling units, or from the projected results from the audited sample using methods described in section 5.10.

Fundamental to determining the sufficiency of the sample, the auditor must determine:

(i) Is the sample itself adequate in size to determine whether a satisfactory, reasonable, and proper audit finding can be made?
(ii) If the audited sample results contain tax reporting errors, does the nature of the errors indicate that more errors exist in the un-audited sampling units? and,
(iii) Is it reasonable to presume that the frequency and amount of the errors found in the sample occur at about the same rate as all the other unaudited sampling units?

If the auditor can answer “yes” to all three of these questions, then the auditor should project (extrapolate) the sample results.

5.09 Judging the Nature of Errors Found in the Sample

Using auditor judgment is really all that can be done using non-statistical methods. The best the auditor can do is decide on the sufficiency of the errors by answering the basic questions (ii) and (iii) stated in Section 5.08.

Further, a reasonableness test can be applied to a non-statistical approach: can it be reasonably expected that an independent and disinterested third party to the audit would agree with the auditor’s conclusions regarding (ii) and (iii) of Section 5.08? If the taxpayer disagrees with the auditor’s conclusions with respect to (ii) and (iii), the auditor should consider expanding the sample. If the expanded sample bears out much the same result, then the auditor should consider projecting (extrapolating) the sample under the provisions of section 5.10.

5.10 Projecting a Non-Statistical Sample

There are many methods of projecting sample results. In non-statistical sampling, the ratio and mean-per-unit estimators are most commonly used. Of the two estimators, the ratio estimator will normally be used. The mean-per-unit estimator is used less frequently because a count of the sample and population items is needed. In cases where the population is not electronically available to the auditor, the counts are usually not available or are difficult to obtain.

In situations where either estimator can be used, the auditor will discuss with the taxpayer and come to an agreement as to which estimator will be used before the sample is drawn. Before using any other projection method, the auditor will discuss with and obtain approval from his/her supervisor.

6. Special Auditing Concerns Relating to Sampling

6.01 Valuing the Sample

If a probability sample is to be projected and included in an audit assessment, the auditor is required to come to an audit conclusion (that is, an error amount), on every sampling unit drawn in the sample. The error amount may be a zero or non-zero amount. The valuation procedures should be, in most instances, the same as if the item was examined in a detailed (actual) audit. However, some sample units drawn in probability samples may require special valuation procedures. This section addresses some of these special circumstances.

6.02 Missing Items

A missing item is a source document representing a sampling unit which has been drawn as a sample item and the supporting documentation cannot be located or was not provided to the auditor. Missing invoices constitute a possible attribute or characteristic of an accounting population. According to Arkin,[12] a probability sample is bound to have missing items if it is drawn from a population that has many missing documents. He further states that:

> It is not sufficient in the case of a document missing from the [population] for an auditor to take another sample item in its place without running down the missing document, unless the auditor is willing to restrict his/her conclusions only to those items remaining in the file to recognize the deficiency.

If the source document cannot be located in the taxpayer's records and is otherwise not available; there are three basic options in valuing the sample unit. They are:

1. The item may be accepted as reported with no adjustment based on auditor's judgment. For example, a missing item represented a common purchase made from a vendor which always charged tax on similar transactions in the audit period.
2. A partial adjustment may be made based on alternative evidence or procedures. For example, a missing invoice from a vendor with a history of errors was included in the sample. The auditor used the proportionate errors found in the other transactions and applied it to the missing invoice.
3. The item may be considered unsubstantiated and totally adjusted. For example, a transaction represented the only purchase from a vendor. The auditor treated the purchase as an error.

Treating a missing item as totally in error is the last resort. This option should be used if the taxpayer is unable to provide sufficient evidence regarding the missing item.

Individual missing items should never be replaced in a sample. Replacing the sample containing the missing items with an entirely new sample is an option. However, if there were a significant number of missing items in the first sample, it is likely that another sample will also contain missing items.

6.03 Extraordinary Items

The auditor should discuss and attempt to reach agreement with the taxpayer on the sample frame and sampling unit. After agreement has been reached (or if all reasonable requests from the taxpayer have been considered), a probability sample is selected from the sample frame.

All items within the probability sample are included in the sample frame and are not 'extraordinary items' separate and apart from the sample frame.

No adjustment should be made to delete items from the probability sample selected. If the taxpayer is not in agreement, the auditor should discuss this issue with the audit supervisor.

6.04 Corrections and Reclassifications

It is acceptable to consider evidence from outside the sample in determining the value of error for the sample unit selected. The transaction, related documents, and accounting entries should be followed to their logical conclusion. Examples include:

- A reclassification of an expense item to another account included within the sampling frame. All entries to both accounts are sampling units within the sampling frame, but only one was selected in the sample.
- A reclassification of an expense item within the sampling frame to an account outside the sampling frame.

In the first example, the transaction had more than one opportunity to be valued for error because the original and correcting entries were both in the sampling frame. If the transaction is in error, this has a potential to bias the sample results, unless a decision rule is followed:

If any entry other than the final entry is selected in the sample, it can never be valued as an error other than $0. Only the final accounting entry, if selected in the sample, should be valued for taxable error.

The following is another example:

> The first invoice is $80, the second invoice is a credit memo for $80, and the third invoice is the corrected invoice in the amount of $75. All invoices are in the sampling frame. The transaction is in error. If the first invoice or second invoice is selected in the sample, the sample item should be valued as a $0 error. The sample item would be valued as a $75 error only if the third invoice is selected in the sample.

The auditor should explain to the taxpayer that it is the taxpayer's responsibility to provide evidence that the transaction has been adjusted.

6.05 Multiple Items

A "multiple item" is one sampling unit that is a collection of more than one transaction, where one transaction is expected. Often, sampling units are summaries of many transactions. Here are some examples of multiple items:

- Credit card entries
- Employee's expense reimbursement
- Petty cash reimbursement
- Checks

All components of the sampling unit should be valued for error.

6.06 Installments

If the sampling unit selected is part of an installment transaction the entire transaction should be examined. If an error is found it should be projected to the sample item proportionate to the entire transaction amount.

The following is an example:

> A transaction is billed on four invoices. Each billing is for $1,250, for a total of $5,000. One of the invoices for $1,250 is selected in the sample. Three more installment payments for $1,250 exist but were not selected. One of the four invoices is not in the sampling frame. Upon review of the entire transaction, the following is determined:
>
> > Of the $5,000 total transaction, $2,000 is in error.

The formula to value the error is:

$$\frac{\text{Total Error}}{\text{Total Transaction Amount}} \times \text{Sample Unit Amount}$$

$$\frac{\$2{,}000}{\$5{,}000} \times \$1{,}250 = \$500$$

The valuation for the sample unit error is $500.

Appendices

Appendix A: Sampling Plan

Sampling Plan and Documentation

Taxpayer Name			**FEIN**	
			State ID	
Audit Period			**Auditor-in-Charge**	
	To			
State				

This form is provided as an aid to the auditor in documenting all sampling procedures and decisions made by the auditor concerning sampling.

Sample Plan Documentation	*Yes*	*No*
Has the sampling plan, along with any modifications, been discussed with the taxpayer?		
Has a sampling plan letter been provided to the taxpayer? *(If yes, please attach)*		
Were there any significant modifications to the sample plan after the sample was drawn? *(If yes, please describe below, or otherwise provide documentation, along with any discussions with the taxpayer regarding changes)*		

Audit Objective – Reason to Sample
Target Population Description *(For example: sales, purchases, fixed assets, or expenses for manufacturing division)*

Sampling Frame Description
Generally, describe the sampling frame below and attach any documentation summarizing the frame. Include the period(s) for which an electronic download was received if it differs from the audit period described above. *(e.g., electronic download of purchase orders, general ledger entries, or list of claimed deductions)*

Describe the sampling unit. *(The individual elements of the sampling frame that will be drawn into the sample, e.g., download line items, clusters of invoices, or all transactions in randomly selected time periods)*
Describe any refinements made to the sampling frame that will not be sampled as they likely will not meet the audit objective, are not under the scope of the audit, or possibly will be handled separately in a detail or another sample. *(Items removed, such as inventory, sales to certain vendors, purchases made in divisions not located in the state or from certain accounts)*

Frame and Sample Summary					
Stratum	**Stratum Description**	**Total # of Sampling Units**	**Total Invoice Value of Sampling Units**	**Sampling Units Selected into Sample or Audited 100%**	**Seed Number**
	Total Sampling Frame				

Sampling Procedures *(Please fill out the table labeled below as "Frame and Sample Summary")*	*Indicate with an "X"*
Indicate the sample design used: *(Please select only one from the following)*	
Stratified random	
Simple random	
Cluster	
Time period (random selection)	
Time period (judgmental selection)	
Judgmental block (describe below)	
Other probability sampling (describe below)	
Other judgmental method (describe below)	

Description of judgmental or “other” sample designs used *(if applicable):*	
	Indicate with an “X”
Indicate the software used to draw the sample or provide the random numbers: *(Please select only one from the following)*	
Multistate Tax Commission Sampling Software (MSS)	
Other software (please describe below)	
Other random method (please describe below)	
Taxpayer drew/provided random numbers (please describe below)	
Description of “other” source of random numbers *(if applicable)*:	
	Indicate with an “X”

Indicate how sample size was determined: *(Please select only one from the following)*	
Software recommendation using expected error rate (describe error rate)	
Using the suggested minimums	
Negotiated with the taxpayer (describe issues that relate to sample size)	
Other method (describe below)	
Comments concerning sample size issues:	

Sample Valuation Issues	*Indicate with an "X*
Indicate the value estimated in the sample: *(Please select only one from the following)*	
Taxable Error	
Tax Error	
Taxable Amount	
Tax Amount	
Other (please explain below)	
Other variables being valued in the sample:	

	Yes	*No*
Were any sample units removed from the sample (not valued)? *(If, yes please explain below)*		

	Indicate with an "X
Indicate if any special valuation issues below were encountered, and explain these issues below: *(Please indicate any that might apply)*	
Missing Items (sample units lacking documentation)	
Sample unit where documentation/records outside the sample frame were considered in valuation	
Installments	
Duplicated items	
Voided transactions	
Other valuation issues (please explain below)	
Describe any special valuation issues encountered in auditing the sample:	

Sample Evaluation and Projection	*Yes*	*No*
Was the sample statistically evaluated? *[If "Yes", please attach work paper showing evaluation using MTC Sampling Software (MSS), Otherwise, if "No", please explain why no evaluation was performed]*		
If the sample was statistically evaluated, was the taxpayer provided the results of the evaluation?		
Will the audit report recommend an adjustment based on a sample projection (whether or not a statistical evaluation was done)?		
If the audit report includes a projection and the sample was statistically evaluated, does the statistically recommended value from MSS agree with the projected value used in the audit report upon which a tax adjustment is made? *(If "No", please explain)*		

If the audit report will include a projection, was the projected value allocated between time periods (tax rate/penalty/interest considerations) and/or locations (local tax rate considerations)? *(If "Yes", please explain and attach work papers showing allocations)*		

Appendix B: Sampling Letter

Date:

Address:

The following explains the proposed sampling procedures to be used by the Multistate Tax Commission in the audit of **[INSERT TAXPAYERS NAME]**. Attached is a document which includes the details to the sampling plan.

The area of the audit to be sampled includes **[INSERT WHETHER SALE/USE OR BOTH]** within the listed states and for the following period(s) (See attached list of participating states and audit periods). The **[INSERT WHETHER SALE/USE OR BOTH]** would be reviewed for possible overpayments and underpayments of tax. The following briefly summarizes the steps in the process.

1. The first step is to identify the necessary data fields required in the electronic data. Typically, such fields include reference date (invoice, purchase order, etc.), reference number, item description, general ledger code, purchase amount, vendor number, vendor name, retail sales tax charged by the vendor, and use tax accrued.
2. Next, the electronic data received will be tested for completeness.
3. Once completeness has been established, these records will be refined and placed into a separate file for the purpose of sampling. Please note the basic 'sampling unit' will also be established at this time. If the sampling unit is the invoice line item, only the selected line items will be valued for possible tax error – not the entire invoice. On the other hand, if the invoice is chosen as the sampling unit, the entire invoice will be valued for error.
4. High dollar transactions may be excluded from the population and examined on an actual basis. A lower-level threshold may also be established. The auditor and taxpayer will agree upon these thresholds.
5. All samples will be selected randomly using a random number generator. The random seed will be documented to allow for expansion or recreation of the sample if necessary. Negative and zero-dollar amounts will not normally be directly sampled.
6. The population (tax relevant items, accounts of interest, other sampling units) will be stratified, by amount, into two to six strata plus a potential detailed stratum for high dollar items. This is to reduce the overall variation in the sample. The procedure for determining the exact number of sampled stratum and stratum breaks will be determined mathematically according to a procedure explained in the

sampling plan details. Generally, sample size will be approximately 200 - 400 items per strata plus the detailed stratum. In rare instances a simple random sample may be warranted.

7. If there are less than three errors in a stratum, that stratum is ignored in the final evaluation. Any error included in a stratum that is not projected may be adjusted on an actual basis.
8. The errors found in the strata sampled will be projected to a statistically conservative estimate of the tax due. One of the following estimators will be used; stratified mean-per-unit, stratified difference, combined ratio, or combined regression. The estimator that provides the best precision will be used to project any taxable errors valued in the sample.
9. Unless the impacted state has its own specific evaluation and projection requirements, the projection method will use the 90% confidence level (two-sided).
10. If the upper and lower limits are of different mathematical signs, no projection will be made.

If you have any further questions, please do not hesitate to contact me at ______________.

Computer Audit Specialist

Appendix C: Sampling Letter Attachment Explaining Sampling Plan Detail

Statistical Sampling Policies

Statistical sampling requires a sampling plan that should be determined before sampling commences. This plan is usually established by referencing relevant state law, through agreements made with the taxpayer, and by agency policy directives. The Commission and some – but not all – of its member states have developed policies concerning statistical sampling. These policy directives may differ in some areas. In the event of conflict between state and Commission procedures and directives, the state's procedures and directives will be followed. In the absence of a state policy, the Commission's procedures and directives will be applied.

This document details our procedures for sampling and issues that may arise as a result of the audit sample. In addition to this document, the Commission may rely on its audit manual and agreement reached with the taxpayer in guiding its audit procedures.

<u>Sample Size</u>

Under no circumstances will a sample of less than 100 items be used. If a simple random sample is to be used, the minimum sample size will be 100. If a stratified random sample is conducted, the minimum sample size is 100 items per stratum.

There are several approaches which can be used to statistically calculate the sample size. In using any of these approaches, determining appropriate sample size for the given precision level and confidence interval is a best guess. The Commission will discuss sample size with the taxpayer and attempt to reach agreement on that issue, Generally, the Commission will use a sample size of 250 for a simple random sample, and 200 – 400 per stratum for stratified random samples.

<u>Number of Strata</u>

In any population in which a statistical sample is considered, the population will almost always be divided into several groups or strata, and independent random samples will be taken from each stratum (a stratified random sample). The primary reason why stratification is preferred is that the accuracy of a stratified random sample is usually better when compared to a sample of the same size taken from an unstratified population (a simple random sample). However, a simple random sample may be taken in cases where the population is too small to justify stratification, or it is known in advance that stratifying will not improve the accuracy of the projected sample results.

The criteria used to stratify are usually invoice value, book value, or some other amount available from the books and records. In some cases, stratification can be done on other characteristics such as business locations or even a combination of characteristics.

Although stratification will place the population into strata for the purpose of sampling, some of the strata may not be sampled.

Sometimes, a low dollar (floor) amount can be established and all records with an invoice value below that amount will be ignored (not audited). Similarly, a high dollar amount (ceiling) is usually established, and all records with an invoice value at or above this amount are not to be sampled but audited 100% (detailed). The Commission will make all efforts to come to some mutually agreeable floor and ceiling amounts with the taxpayer.

In most cases the number of sampled strata (excluding strata detailed or ignored) will be three or four strata. In some cases, only two strata will be sampled, and in cases with very large populations up to six strata may be considered. Constructing more than six strata will rarely provide any benefit. The Commission will negotiate in good faith with the taxpayer regarding sample strata.

Procedure for Determining Strata Breaks:

1. All negative amounts (accounting adjustments) and zero items should be eliminated from the population
2. The total count of the remaining population will be calculated
3. The total dollar value of the population will be calculated
4. The total count for each defined class will be calculated
5. The total dollar value of each defined class will be calculated
6. The Commission will then define the strata breaks using "Cumulative Square Root of the Frequency". After the strata breaks are determined the sample will be drawn from each defined stratum.

Handling of Accounting Adjustments (Negatives):

Most accounting populations examined for error will have accounting adjustments, which include things like credit memos, debit memos, reclassifications, or other similar adjustments. Although the sample will be generally taken from a population excluding the accounting adjustments, this does not mean the accounting adjustments will be ignored in the sample. Any applicable accounting adjustments must be considered by the auditor in determination of any error.

First, before the sample is drawn, an examination will be made which matches up offsetting accounting adjustments These matched accounting entries are excluded from the population sampled. The remaining population often has unmatched accounting adjustments (negatives) which will not be sampled. After drawing a sample, the auditor must consider any unmatched accounting adjustments which correspond to the sampling units drawn.

In some cases, the accounting adjustment which corresponds to the underlying transaction for any sample unit drawn into the sample is outside the time frame covered by the population or audit period. The auditor should also consider these accounting entries in determination of any error.

Missing Items

A "missing item" is any sample unit drawn into the sample where the taxpayer or its representative, for whatever reason, is not able to provide, or has not provided the necessary source documents and other accounting records required to make a satisfactory audit conclusion.

In no case will the Commission consider substituting, replacing, or removing missing items from the sample. Every sample unit drawn into the sample will be valued for error. However, subject to auditor approval, alternative audit evidence may be used in determination of any audit error. In addition, also subject to auditor approval, the audit error may be valued at zero error or only partially in error, based on an examination of similar transactions or other satisfactory evidence. In other cases, the auditor may determine that unless the necessary documentation is provided, the unit will be considered fully in error.

Evaluating the Sample

The final evaluation of the audited sample results should be made using a 2-sided 90% confidence level where possible. If the sample evaluation in fact achieves relative precision of 30% or better, the projected total taxable (or tax) adjustment can be made at the point estimate. However, if relative precision of at least 30% is not achieved, various options exist:

a. The Commission and the Taxpayer agree to project the results as is.
b. Reevaluate the sample using a 1-sided 95% confidence level. In this case, the projected taxable adjustment that will be recommended is the confidence limit closest to zero, the lower limit, provided both confidence limits (upper and lower) are of the same sign.
c. The Commission and the taxpayer agree to increase the sample size over all strata or within a stratum.

d. The Commission and the taxpayer agree to re-sample entirely. See Section 4.08(i).

In the case where the confidence limits are not of the same sign or where relative precision is at 100% or greater, no projection of the sample will be done. Also, to project into a stratum, at least three nonzero errors must be observed in the sample for that particular stratum. Note where one or more strata have less than three nonzero errors, an overall projected taxable amount may be made from the other sampled strata having at least three nonzero errors. In addition to the projected taxable error, the actual error observed in the detail stratum or other strata not projected may also be adjusted.

When evaluating statistical samples, the Commission uses the following statistical estimators. These estimators can be used in simple or stratified random samples. In stratified sampling, the ratio and regression have two approaches in providing an overall estimate: separate or combined estimation.

1. Mean-per-unit
2. Difference
3. Ratio
4. Regression

The estimator that yields the best precision (that is the smallest precision amount) will be used in making any projection at the midpoint (point estimate).

In evaluating the sample results, a statistical sampling statement can be made that computes a confidence interval that is bounded by the lower confidence limit (LCL) and upper confidence limit (UCL). The point estimate (PE) is always between the LCL and UCL. If the LCL, PE, and UCL of the confidence interval have the same mathematical sign, that is either all positive or all negative, the sample is said to "evaluate" and the estimator can be used to project the sample results. If the signs are different, that is the LCL is negative, the PE is either negative or positive, and the UCL is positive, the sample results are said "not to evaluate". If an estimator does not evaluate, the point estimate cannot be used to project the sample results. If an estimator does not evaluate, this does not preclude the use of another estimator which does evaluate. In some cases, all the estimators may not evaluate. In this instance, the Commission will recommend to the states that no projected assessment or refund be made.

Appendix D: Procedures for Using Cumulative Square Root of the Frequency

Cumulative Square Root of the Frequency (CSRF)

Once the sampling frame is established, there are several ways to determine stratum breaks. In Commission audits, the "Cumulative Frequency of the Square Root" is the preferred method to establish strata boundaries.

Background

Stratification boundaries can be determined using several approaches. The most common method is CSRF. The origin of the CSRF method is attributed to an article appearing in the Journal of American Statistical Association in 1959.[13] It is universally accepted in sampling and is referenced in many accounting and survey sampling books.[14] Although it would be optimal to establish strata boundaries based on the error amounts, these amounts are unknown. Therefore, populations are stratified using invoice amounts.[15] CSRF will distribute the variation of invoice amounts in an "optimal" manner to the strata.

CSRF methodology breaks down the population into intervals. These intervals are sometimes referred to as "cells" which can be of equal or unequal width. In most references, the method is described using cells of equal width. When broken into unequal widths, an extra step is required.[16] The Commission generally uses cells of unequal width because the stratum breaks become more exact. Note that CSRF does not dictate the number of strata, only the breaks between the strata.

Mechanics

The following is an explanation of the procedures used to determine stratum breaks for 4 strata assuming cells of unequal width.

Step 1 Evaluate the population and determine if transactions below a certain dollar level should be removed based on the dollar significance of those transactions.

In this example, the taxpayer and the auditor agreed on the accounts of interest and the sample population. No transactions equal to or less than $100.00 will be reviewed.

Step 2 Evaluate the population and determine the dollar amounts which will be reviewed on an actual basis.

In this example, all items greater than $25,000.00 will be reviewed in detail (actual basis).

Step 3 Group the remaining transactions in the population ($100.01 - $25,000.00) into many dollar ranges.

Step 4 Determine the frequency for each dollar range. This is the number of transactions within each range.

Step 5 Calculate the square root of the frequency (number of transactions) for the first range ($100.01 – $150.00). The corresponding frequency was 16,853, therefore the square root of the frequency was 129.82. Next, calculate the square root for the following range ($150.01 – $200.00). This frequency was 10,203 and the square root of the frequency was 101.01. Continue this process for each dollar range.

Step 6 Calculate the square root of the range width for the first range ($150.01 – $200.00). This value is 7.07. Then calculate the square root of the next range ($150.01 – $200.00). Again, this value equals 7.07. Continue this process for each of the ranges.

Step 7 For each range, multiply the square root of the frequency (step 5) times the square root of the range (step 6).

Step 8 Accumulate the values calculated in step 7 for each range. For the first range, the cumulative square root value is 917.83. The cumulative value for the second range is 1,632.97 (714.14 + 917.83 - accumulated value for the preceding range).

For each range, add the square root value (step 7) to the preceding cumulative square root value (step 8) to arrive at the next cumulative square root value. Continue this process for each of the ranges.

Step 9 Divide the total cumulative square root value (say 18,422.66) by the number of strata (4 in this audit) to arrive at cumulative square root value for each stratum. In this example the value is 4,605.67.

Step 10 Multiply the result of step 9 by 1 to calculate the first stratum range, or 4,605.67. Locate the cumulative square root value closest to this amount. In this case, it is 4,623.62. Locate the ending range associated with this value ($500.01 - $550.00). The first stratum range (boundary) is $100.01 – $550.00.

Step 11 Multiply the result of step 9 (4,605.67) by 2 to calculate the second stratum range, or 9,211.34. Locate the cumulative square root value closest to this amount. In this case, it is 9,213.44. Locate the ending range associated with this value ($2,100.01 – $2,200.00). The second stratum range (boundary) is $550.01 – $2,200.00.

Step 12 Multiply the result of step 9 (4,605.67) by 3 to calculate the third stratum range, or 13,817.07. Locate the cumulative square root value closest to this amount. In this case, it is 13,871.18. Locate the ending range associated with this value ($8,250.01 – $8,500.00). The third stratum range (boundary) is $2,200.01 – $8,500.00.

Step 13 Stratum 4 is whatever is left after calculating the stratum breaks for stratum1, 2, and 3.

Step 14 Calculate the coefficient of variation (standard deviation divided by the mean) for each of the strata (4 strata in this audit). Compute the average coefficient of variation for all the strata. An average coefficient of variation of less than 50% is the target range. In this example, the combined coefficient of variation is 40.20%.

Step 15 If the average coefficient of variation is greater than 50%; consideration should be given to adding another stratum. On the other hand, if the percentage is significantly less than 50%, then the total number of strata may be reduced.

Cumulative Square Root of the Frequency - Strata Boundary Determination
CSRF-Unequal

Range Beginning Amount	Range Ending Amount	Range Width Value	Range Count (#)	Square Root of the Count (#)	Square Root of the Width (Value)	Sum of the Square Root Count (#) * Width (Value)	Cumulative Square Root Values Count (#) * Width (Value)	Stratum
100.01	150.00	49.99	16853	129.82	7.07	917.87	917.87	1
150.01	200.00	50.00	10204	101.01	7.07	714.28	1,632.15	1
200.01	250.00	50.00	7046	83.94	7.07	593.55	2,225.70	1
250.01	300.00	50.00	5476	74.00	7.07	523.26	2,748.96	1
300.01	350.00	50.00	4339	65.87	7.07	465.78	3,214.74	1
350.01	400.00	50.00	4016	63.37	7.07	448.11	3,662.84	1
400.01	450.00	50.00	2352	48.50	7.07	342.93	4,005.77	1
450.01	500.00	50.00	2010	44.83	7.07	317.02	4,322.79	1
500.01	550.00	50.00	1810	42.54	7.07	300.83	4,623.62	1

Cumulative Square Root of the Frequency - Strata Boundary Determination
CSRF-Unequal

Range Beginning Amount	Range Ending Amount	Range Width Value	Range Count (#)	Square Root of the Count (#)	Square Root of the Width (Value)	Sum of the Square Root Count (#) * Width (Value)	Cumulative Square Root Values Count (#) * Width (Value)	Stratum
550.01	600.00	50.00	1639	40.48	7.07	286.27	4,909.89	2
600.01	650.00	50.00	1198	34.61	7.07	244.74	5,154.64	2
650.01	700.00	50.00	1174	34.26	7.07	242.28	5,396.92	2
700.01	750.00	50.00	1008	31.75	7.07	224.50	5,621.42	2
750.01	800.00	50.00	1255	35.43	7.07	250.50	5,871.92	2
800.01	850.00	50.00	717	26.78	7.07	189.34	6,061.26	2
850.01	900.00	50.00	873	29.55	7.07	208.93	6,270.18	2
900.01	950.00	50.00	615	24.80	7.07	175.36	6,445.54	2
950.01	1,000.00	50.00	611	24.72	7.07	174.79	6,620.33	2
1,000.01	1,050.00	50.00	574	23.96	7.07	169.41	6,789.74	2
1,050.01	1,100.00	50.00	424	20.59	7.07	145.60	6,935.34	2
1,100.01	1,150.00	50.00	388	19.70	7.07	139.28	7,074.62	2
1,150.01	1,200.00	50.00	347	18.63	7.07	131.72	7,206.34	2
1,200.01	1,250.00	50.00	417	20.42	7.07	144.40	7,350.74	2
1,250.01	1,300.00	50.00	493	22.20	7.07	157.00	7,507.74	2
1,300.01	1,350.00	50.00	312	17.66	7.07	124.90	7,632.64	2
1,350.01	1,400.00	50.00	271	16.46	7.07	116.40	7,749.04	2
1,400.01	1,450.00	50.00	230	15.17	7.07	107.24	7,856.28	2
1,450.01	1,500.00	50.00	316	17.78	7.07	125.70	7,981.98	2
1,500.01	1,600.00	100.00	312	17.66	10.00	176.64	8,158.62	2
1,600.01	1,700.00	100.00	421	20.52	10.00	205.18	8,363.80	2
1,700.01	1,800.00	100.00	457	21.38	10.00	213.78	8,577.57	2
1,800.01	1,900.00	100.00	288	16.97	10.00	169.71	8,747.28	2
1,900.01	2,000.00	100.00	350	18.71	10.00	187.08	8,934.36	2
2,000.01	2,100.00	100.00	210	14.49	10.00	144.91	9,079.28	2
2,100.01	2,200.00	100.00	180	13.42	10.00	134.16	9,213.44	2
2,200.01	2,300.00	100.00	193	13.89	10.00	138.92	9,352.37	3
2,300.01	2,400.00	100.00	117	10.82	10.00	108.17	9,460.53	3
2,400.01	2,500.00	100.00	199	14.11	10.00	141.07	9,601.60	3
2,500.01	2,600.00	100.00	98	9.90	10.00	98.99	9,700.59	3
2,600.01	2,700.00	100.00	202	14.21	10.00	142.13	9,842.72	3
2,700.01	2,800.00	100.00	133	11.53	10.00	115.33	9,958.05	3
2,800.01	2,900.00	100.00	169	13.00	10.00	130.00	10,088.05	3
2,900.01	3,000.00	100.00	203	14.25	10.00	142.48	10,230.52	3
3,000.01	3,100.00	100.00	114	10.68	10.00	106.77	10,337.30	3
3,100.01	3,200.00	100.00	94	9.70	10.00	96.95	10,434.25	3
3,200.01	3,300.00	100.00	99	9.95	10.00	99.50	10,533.75	3
3,300.01	3,400.00	100.00	71	8.43	10.00	84.26	10,618.01	3
3,400.01	3,500.00	100.00	82	9.06	10.00	90.55	10,708.56	3
3,500.01	3,600.00	100.00	91	9.54	10.00	95.39	10,803.96	3
3,600.01	3,700.00	100.00	63	7.94	10.00	79.37	10,883.33	3
3,700.01	3,800.00	100.00	142	11.92	10.00	119.16	11,002.49	3
3,800.01	3,900.00	100.00	80	8.94	10.00	89.44	11,091.94	3
3,900.01	4,000.00	100.00	93	9.64	10.00	96.44	11,188.37	3
4,000.01	4,100.00	100.00	120	10.95	10.00	109.54	11,297.92	3
4,100.01	4,200.00	100.00	117	10.82	10.00	108.17	11,406.08	3
4,200.01	4,300.00	100.00	96	9.80	10.00	97.98	11,504.06	3
4,300.01	4,400.00	100.00	45	6.71	10.00	67.08	11,571.14	3
4,400.01	4,500.00	100.00	59	7.68	10.00	76.81	11,647.96	3
4,500.01	4,600.00	100.00	30	5.48	10.00	54.77	11,702.73	3
4,600.01	4,700.00	100.00	49	7.00	10.00	70.00	11,772.73	3
4,700.01	4,800.00	100.00	38	6.16	10.00	61.64	11,834.37	3
4,800.01	4,900.00	100.00	23	4.80	10.00	47.96	11,882.33	3
4,900.01	5,000.00	100.00	93	9.64	10.00	96.44	11,978.77	3
5,000.01	5,250.00	250.00	66	8.12	15.81	128.45	12,107.22	3
5,250.01	5,500.00	250.00	64	8.00	15.81	126.49	12,233.71	3
5,500.01	5,750.00	250.00	99	9.95	15.81	157.32	12,391.03	3
5,750.01	6,000.00	250.00	121	11.00	15.81	173.93	12,564.96	3
6,000.01	6,250.00	250.00	63	7.94	15.81	125.50	12,690.46	3
6,250.01	6,500.00	250.00	91	9.54	15.81	150.83	12,841.29	3
6,500.01	6,750.00	250.00	168	12.96	15.81	204.94	13,046.23	3
6,750.01	7,000.00	250.00	64	8.00	15.81	126.49	13,172.72	3
7,000.01	7,250.00	250.00	63	7.94	15.81	125.50	13,298.22	3
7,250.01	7,500.00	250.00	47	6.86	15.81	108.40	13,406.61	3
7,500.01	7,750.00	250.00	48	6.93	15.81	109.54	13,516.16	3
7,750.01	8,000.00	250.00	82	9.06	15.81	143.18	13,659.34	3
8,000.01	8,250.00	250.00	23	4.80	15.81	75.83	13,735.17	3
8,250.01	8,500.00	250.00	74	8.60	15.81	136.01	13,871.18	3

Cumulative Square Root of the Frequency - Strata Boundary Determination
CSRF-Unequal

Range Beginning Amount	Range Ending Amount	Range Width Value	Range Count (#)	Square Root of the Count (#)	Square Root of the Width (Value)	Sum of the Square Root Count (#) * Width (Value)	Cumulative Square Root Values Count (#) * Width (Value)	Stratum
8,500.01	8,750.00	250.00	64	8.00	15.81	126.49	13,997.67	4
8,750.01	9,000.00	250.00	34	5.83	15.81	92.20	14,089.87	4
9,000.01	9,250.00	250.00	24	4.90	15.81	77.46	14,167.33	4
9,250.01	9,500.00	250.00	27	5.20	15.81	82.16	14,249.48	4
9,500.01	9,750.00	250.00	29	5.39	15.81	85.15	14,334.63	4
9,750.01	10,000.00	250.00	45	6.71	15.81	106.07	14,440.70	4
10,000.01	10,500.00	500.00	42	6.48	22.36	144.91	14,585.61	4
10,500.01	11,000.00	500.00	42	6.48	22.36	144.91	14,730.53	4
11,000.01	11,500.00	500.00	62	7.87	22.36	176.07	14,906.59	4
11,500.01	12,000.00	500.00	77	8.77	22.36	196.21	15,102.81	4
12,000.01	12,500.00	500.00	72	8.49	22.36	189.74	15,292.54	4
12,500.01	13,000.00	500.00	51	7.14	22.36	159.69	15,452.23	4
13,000.01	13,500.00	500.00	48	6.93	22.36	154.92	15,607.15	4
13,500.01	14,000.00	500.00	41	6.40	22.36	143.18	15,750.33	4
14,000.01	14,500.00	500.00	67	8.19	22.36	183.03	15,933.36	4
14,500.01	15,000.00	500.00	51	7.14	22.36	159.69	16,093.05	4
15,000.01	15,500.00	500.00	38	6.16	22.36	137.84	16,230.89	4
15,500.01	16,000.00	500.00	53	7.28	22.36	162.79	16,393.68	4
16,000.01	16,500.00	500.00	23	4.80	22.36	107.24	16,500.91	4
16,500.01	17,000.00	500.00	58	7.62	22.36	170.29	16,671.21	4
17,000.01	17,500.00	500.00	29	5.39	22.36	120.42	16,791.62	4
17,500.01	18,000.00	500.00	21	4.58	22.36	102.47	16,894.09	4
18,000.01	18,500.00	500.00	26	5.10	22.36	114.02	17,008.11	4
18,500.01	19,000.00	500.00	20	4.47	22.36	100.00	17,108.11	4
19,000.01	19,500.00	500.00	24	4.90	22.36	109.54	17,217.65	4
19,500.01	20,000.00	500.00	39	6.24	22.36	139.64	17,357.30	4
20,000.01	25,000.00	5,000.00	227	15.07	70.71	1065.36	18,422.66	4
			74506					

Cumulative Value	18,422.66
# of Strata	4
Amt per Stratum	4,605.67

STRATUM	AMT PER STRATUM	APPROX BREAK	COUNT	STD DEV	AVE	TOTAL VALUE	CV (SEE NOTE)	CV * TOTAL
1	4,605.67	4,605.67	54,106	116.58	236.44	12,792,823	49%	6,307,677
2	4,605.67	9,211.33	15,080	412.88	1,020.02	15,381,902	40%	6,226,230
3	4,605.67	13,817.00	3,986	1,679.82	4,203.43	16,754,872	40%	6,695,763
4	4,605.67	18,422.66	1,334	4,608.09	14,884.44	19,855,843	31%	6,147,192
			74,506			64,785,439		

Coefficient of Variation

AVE =>	40%

STRATUM	Range Beginning Amount	Range Ending Amount
1	100.01	550.00
2	550.01	2,200.00
3	2,200.01	8,500.00
4	8,500.01	25,000.00

Appendix E: Sample Allocation

Sample Allocation

Once the sampling frame has been stratified, the sample size within each stratum is determined. This can be done proportionately or "optimally".

Proportional sampling allocates the sample based solely on the strata sizes. Optimal allocation considers the size of the strata and the variability within each stratum.

In the following example of a stratified sampling frame, the initial population was refined to exclude items that were not of interest or detailed:

	COUNT	AMOUNT	AVERAGE	STD DEV	CV
TOTAL DOWNLOAD	27,478	64,852,762.11	2,360.17	33,964.35	1439.06%
SMALL INVOICES & NEGATIVES	8,957	(3,433,107.08)	(383.29)	20,572.87	-5367.48%
LARGE INVOICES	802	52,496,892.54	65,457.47	175,018.56	267.38%
EXCLUDED_EMP	1,036	223,236.18	215.48	252.32	117.10%
EXCLUDED_VEND	249	242,266.22	972.96	1,484.38	152.56%
DELETED VENDORS	1,857	1,368,113.13	736.73	1,336.31	181.38%
OTHER REFUND ITEMS	1,448	305,021.71	210.65	413.50	196.30%
OTHER INERCOMPANY	28	22,370.10	798.93	1,127.79	141.16%
SAMPLING FRAME	13,101	13,627,969.31	1,040.22	1,662.41	159.81%
	27,478	64,852,762.11			
STRATUM 1 (50 TO 874.99)	9,162	2,485,764.06	271.31	210.47	77.57%
STRATUM 2 (875 TO 3449.99)	2,877	4,961,067.32	1,724.39	670.92	38.91%
STRATUM 3 (3450 TO 10000)	1,062	6,181,137.93	5,820.28	1,809.21	31.08%
	13,101	13,627,969.31		Average ->	49.19%

The overall sample size is 900.

1. Proportional Allocation

The allocation of the sample to each stratum could be as follows:

DESCRIPTION	COUNT	%	SAMPLE DISTRIBUTION
STRATUM 1 (50 TO 874.99)	9,162	70%	629
STRATUM 2 (875 TO 3449.99)	2,877	22%	198
STRATUM 3 (3450 TO 10000)	1,062	8%	73
	13,101	100%	900

2. Optimal Allocation - Neyman

If optimal allocation is used, the sample size for each stratum could be as follows:

DESCRIPTION -A-	COUNT -B-	STANDARD DEVIATION (see note) -C-	COUNT X STD DEV (B * C) -D-	% of D -E-	SAMPLE DISTRIBUTION (E * 900) -F-
STRATUM 1 (50 TO 874.99)	9,162	210.47	1,928,326.14	33.36%	300
STRATUM 2 (875 TO 3449.99)	2,877	670.92	1,930,236.84	33.40%	301
STRATUM 3 (3450 TO 10000)	1,062	1,809.21	1,921,381.02	33.24%	299
	13,101		5,779,944.00	100.00%	900

note: This is the standard deviation of the book or invoice amounts. To really be "optimal", the distribution should be based on the standard deviation of the error amounts, which are unknown prior to sampling - so the best information available is used.

The Optimal allocation is the preferred method by the Commission. However, if strata breaks are determined using the Cumulative Square Root of the Frequency method, the sample is “approximately optimal” when the sample size within each stratum is equal. The allocation step can be ignored if strata breaks are determined using the Cumulative Square Root of the Frequency method and sample size is equal for each stratum. In this instance, sample size would be set at 300 items per stratum (900 divided by 3 strata). The minor variances from 300 (301 for stratum 2 and 299 for stratum 3) as noted in column F can be ignored.

End Notes

1 Note Arkin's discussion of "extraneous units" in Sampling Methods for the Auditor, page 145.

2 If we have a sampling frame that is broken up into, say four strata, with a single overall projection across the four strata, this will count as one projection – not four.

3 An error can be a debit (deficiency) or a credit (overpayment).

4 In non-statistical sampling, sample size will never go below 250, but will probably be much greater than 300 according to the formula shown in section 5.06(b) (transaction sampling). The minimum of 300 for statistical sampling is needed to ensure a reasonable chance of finding errors in the each of the strata, if significant error exists in the population.

5 For audits where the electronic file is unavailable, the standard deviation of the invoice amounts can be estimated by the following formula:

$$\frac{\text{(High Invoice Amount in the Frame – Low Invoice Amount in the Frame)}}{4}$$

The average can be estimated by multiplying the estimated error rate by the total invoice amount divided by the total number of invoices.

6 Because difference estimation performs poorly in low error rate populations, this estimator will be effectively ignored for populations where the error-rate is less than 20%. This means that for samples where the error rate is small, this estimator will never be used.

7 A correlation coefficient, ρ, can easily be computed in Excel with the function "=correl([range of x_i values],[range of y_i values]", where the x_i values represent the invoice amounts in the population (or sample), and the y_i values represent the taxable error values in the population (or sample). The formula for correlation is:

$$\rho = \frac{\sum_{i=1}^{n}\left((x_i - \bar{x}) * (y_i - \bar{y})\right)}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 * \sum_{i=1}^{n}(x_i - \bar{x})^2}}.$$

8 In effort to control bias in ratio and regression estimation, and the fact that the estimated standard error calculation for ratio estimation is approximate, the

conditions noted here should be followed. Refer to Cochran's Sampling Techniques at pages 153 and 166.

9 The basis for the concern is found in Cochran's book, Sampling Techniques, at pages 153 and166. He recommends that the coefficient of variation of the sample mean for the invoice amounts should be less than 10%. Application of Cochran's concern over bias has been extended to the coefficients of variation of the sample mean for the audited amounts using both mean-per-unit and difference estimation.

10 The Commission does not use other approaches, such as the bootstrap and empirical likelihood to calculate a confidence interval.

11 In random sampling, every sample of the size of n should be equally probable. Indiscriminately removing items may violate this basic rule unless the item is removed from both the sample and the population.

12 Refer to the Handbook of Sampling for Auditing and Accounting, page 21.

13 The following are some of the references:

1. Statistical Auditing, Donald Roberts, AICPA, 1978, pages 97 to 98.
2. Sampling Techniques, William Cochran, John Wiley and Sons, 1977, pages 127-131.
3. Applications of Statistical Sampling to Auditing, Alvin Arens & James K Loebbecke, Prentice Hall, Englewood CA 1981, pages 256 – 258.
4. IRS training publication: Advanced Statistical Sampling, Training 3174-002 Rev (05-92) TPDS 87030A in Lesson 10, pages 6 – 10.
5. Survey Sampling, Leslie Kish, John Wiley & Sons, 1995, pages 104 – 106.
6. Sampling of Populations: Methods and Applications, 3rd Edition, Paul Levy and Stanley Lemeshow, John Wiley & Sons, New York, 1999, pages 179 - 183.
7. Elementary Survey Sampling, 5th edition, Richard Scheaffer, William Mendenhall, R. Lyman Ott, Duxbury Press, New York, 1996, page 165.

14 According to Elementary Survey Sampling, 5th edition, Richard Scheaffer et al. footnote13, page 189, CSRF was developed by T. Dalenius and J.L. Hodges Jr.

15 Generally, there is a relationship between error and invoice amounts. We know that the variation in the errors has something to do with the range of values in the non-zero error amounts. However, the biggest contributor to the variation of errors is usually the number of zero errors as percentage of the total (the more zero errors, the more variable the population). The number of zero errors in the population will not have anything to do with variation in values of invoice amounts. CSRF will not consider the effect of the zeros.

16 In Statistical Auditing, Roberts describes the methodology of unequal lengths in a footnote on page 98. In Sampling Techniques, Cochran describes this modification on page 130.

[This Page Intentionally Left Blank]