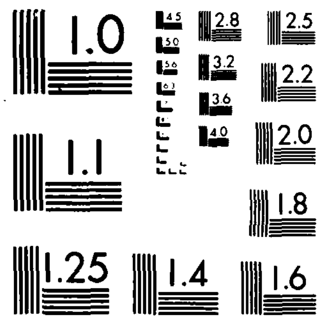
AD-A174 444 SELECTED UPPER ERROR LIMIT METHODS APPLIED TO AN ACCOUNTING POPULATION(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST.. A J CUKR 1/1

UNCLASSIFIED SEP 86 AFIT/GSM/LSY/86S-5 F/G 5/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

SELECTED UPPER ERROR LIMIT METHODS APPLIED TO AN ACCOUNTING POPULATION

THESIS

Anita J. R. Cukr
Capt, USAF

AFIT/GSM/LSY/86S-5

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

2

AFIT/GSM/LSY/86

SELECTED UPPER ERROR LIMIT
METHODS APPLIED TO AN
ACCOUNTING POPULATION

THESIS

Anita J. R. Cukr
Capt, USAF

AFIT/GSM/LSY/86S-5

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GSM/LSY/86S-5

SELECTED UPPER ERROR LIMIT METHODS APPLIED TO AN ACCOUNTING POPULATION

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Anita J. R. Cukr, B.A.

Captain, USAF

September 1986

Acknowledgements

I wish to thank my faculty advisor, Lt. Col. Jeffrey J. Phillips for his assistance throughout this research effort. I give special thanks to my husband, Jeffrey, for his understanding, patience, and calming effect on me in times of crisis. Special thanks also to my mother, Helen Rowley, for her understanding and assistance throughout the months of this research effort. Without the help of these people, I could not have completed this research.

Anita J. R. Cukr

Table of Contents

List of Tables

Abstract

The purpose of this research was to study the performance of the Cox and Snell Upper Error Limit Model. Twelve Cox and Snell bounds were constructed by changing the prior error probabilities incorporated in the model. The performance of the twelve bounds was compared to the performance of the Stringer bound.

The study constructed twenty-four study populations from an actual accounting population, sampled from these study populations, and compared the error estimates produced by the Cox and Snell model with the true dollar error in the study populations. The objective was to find the best upper error limit estimator from the Cox and Snell bounds tested. The bounds were examined in terms of robustness, mean relative tightness, effect of nominal confidence level, and statistical power.

In order to analyze the results of the simulation, the study populations were divided into four categories by error amount intensity. Robust bounds were found for the low error amount intensity study populations only. For the medium, high, and highest error amount intensity populations, bound coverage was significantly less than the nominal confidence levels of the bounds. The Cox and Snell

bounds were significantly tighter than the Stringer bound. The effect of lowering the bound confidence level on the relative tightness of the bounds was negligible. Some of the bounds which provided adequate coverage were statistically powerful.

SELECTED UPPER ERROR LIMIT METHODS APPLIED TO AN ACCOUNTING POPULATION

I. Introduction

General Issue

Periodic audit programs allow organizations to find instances of inefficiency and to draw conclusions on the accuracy of their financial statements and their compliance with prescribed management rules. They also allow management to make accounting estimates and to uncover instances of inefficiency. These are important objectives both in industry and government. Auditors have an obvious interest in expressing as accurate an opinion as possible on the error amount in the audit population because of the serious implications of a material error going undetected. For instance, an undetected material error in an account balance can have serious earnings implications for a company, and an undetected large error in government accounts could have serious adverse effects on public opinion about the competence of government in managing public funds (11:no page). A number of statistical models for predicting errors in accounting populations are available, but there is a need for research to validate and refine these models and to find new models (4:39).

Ideally, auditors would audit all accounting records in order to render an opinion on the error in those records with as high a level of confidence as possible. Cost considerations make this impractical, so auditors have turned to statistical sampling procedures to evaluate the correctness of audit populations. Using samples and sample statistics, they can construct acceptable confidence intervals around their estimates of the total error in the population. They have an obvious interest in finding the best statistical tools to help them (11:7).

Various researchers have noted problems in the application of classical statistical sampling and estimating techniques to accounting populations. Neter and Loebbecke provided empirical evidence that error bounds based on classical confidence interval estimators do not perform well for accounting populations with low error rates. Since many accounting populations do have low error rates, they do not meet the assumption of normality of the distribution of sample statistics which is the basis for the use of classical statistical estimators. As a result, classical estimators tend to produce poor error prediction bounds for accounting populations (20:76-77).

In response to the difficulties associated with the use of classical statistical estimators, researchers developed new techniques which can be classified as Bayesian and non-Bayesian. Godfrey and Andrews refer to the non-Bayesian approaches as the classical procedure and divide the Bayesian approaches into the infinite Bayesian procedure and the finite Bayesian procedure. The classical procedure uses the hypergeometric, binomial, or Poisson distributions to estimate the number of errors in a population. The Bayesian procedures differ from the classical procedure in that they incorporate auditor prior expectations about the distribution of errors in the population. This feature may give the Bayesian procedures an advantage over non-Bayesian approaches (7:304). However, the incorporation of incorrect auditor prior expectations about accounting populations could adversely affect the accuracy of Bayesian error prediction models. The difficulty of quantifying prior expectations is probably the main reason that Bayesian methods have not received more widespread acceptance (7:313).

There is still room for research into improved error prediction models for auditing. Empirical research to find more accurate and concise bounds for error prediction than currently exist will improve and facilitate the audit task. This is the general problem that this thesis addresses.

Specific Problem

This study compares the performance of the Cox and Snell upper error limit model and the Stringer bound. The Cox and Snell model is a Bayesian upper error limit bound that has been used in a number of research studies. The Stringer bound is a widely used non-Bayesian upper error limit bound. Both bounds were tested using 24 study populations constructed by varying the error characteristics of an actual accounting population of U.S. Air Force aircraft parts inventory accounts.

A major objective of this study was to find a robust Cox and Snell bound among the bounds tested. A $(1-\alpha)$ Bayesian bound is considered robust if it is correct in repeated sampling $(1-\alpha)*100$ percent of the time for a wide variety of populations. A robust bound is accurate at a certain level of confidence for populations with widely varying error characteristics and therefore auditors can use it without being concerned about their ability to assess prior probabilities. Godfrey and Neter studied the characteristics of the Cox and Snell bound and searched for robust Cox and Snell and modified Cox and Snell bounds. They found that the Cox and Snell and modified bounds are very sensitive to changes in the prior error parameters incorporated into them. They also showed that the Cox and Snell bound performs better than the Stringer bound for low total error accounting populations. Specifically, they found the Cox and Snell bound to be tighter than the Stringer bound and to provide high coverage for low error rate populations (8:520). Coverage refers to the number of times that a bound equals or exceeds the true population dollar error in repeated sampling. In comparing bounds, a bound which exceeds other bounds by a lesser amount while providing at least the same coverage, is the tighter bound (18:9-10). The Stringer bound is considered a conservative, or loose bound. The looser a bound is, the less desirable it is for the auditor because to obtain a given level of confidence, the auditor must accept a larger interval estimate than he would with a tighter bound. Godfrey and Neter found that by choosing conservative prior parameters, auditors could find a Cox and Snell bound that is robust, much tighter than the Stringer bound for low error amount populations, and moderately tighter than the Stringer bound for populations whose total error amount is large. In the conclusion to their 1984 study, Godfrey and Neter addressed the need for further research to examine the robustness of the Cox and Snell bound with conservative prior distributions for other error characteristics and sample sizes than they used (8:521). In line with this, then, this study tested the Cox and Snell bound with different error characteristics than previous researchers used. The next section addresses the research objectives in more detail.

Research Questions

The first research question addressed the robustness of selected Cox and Snell error limit bounds using study populations with known error characteristics. The Cox and Snell bounds used had different prior error rate and error taint probabilities than those used in earlier research. Error rate is defined as the probability that an individual population item is in error. Error taint is defined as the error amount divided by the book value of the line item in error (10:281). The objective was to find new robust bounds, if possible. Godfrey and Neter found some robust bounds in their 1984 and 1985 studies of the Cox and Snell bound. As already noted, the importance of finding such a robust bound is that auditors could use such a bound for a wide variety of populations without having to quantify their prior judgement on the error characteristics of the population they are auditing. The prior error characteristics incorporated in a robust bound would be conservative enough to ensure that the bound accurately predicts errors $(1-\alpha)$ $*100$ percent of the time. See Chapter III for a description of the error characteristics of the study populations.

The second research question addressed bound tightness. The objective was to find the tightest robust bound from the bounds tested. Godfrey and Neter measured relative tightness by comparing the median Cox and Snell bounds with the median Stringer bounds. Another possible measure of relative tightness is the mean ratio of a given Cox and Snell bound and a given Stringer bound. Phillips used this measure to reinforce the Godfrey and Neter measure (18:10). The present study measured only the mean relative tightness of the bounds.

The third research question examined the effect on bound robustness of changing the nominal confidence level of the bounds tested. The nominal confidence level is the expected frequency with which the bound yields correct error estimates in repeated sampling (9:5). Godfrey and Neter examined Cox and Snell bounds at the 95 percent nominal level and found some robust bounds. Phillips examined Cox and Snell bounds at the 85 percent nominal level (18:11). This study lowered the nominal level even more (to the 75 percent and 80 percent levels) to see if this might yield robust bounds that are tighter than comparable bounds at higher confidence levels. This would give the auditor the choice of balancing confidence level against bound tightness, which are both determinants of the precision of the estimate.

The fourth research question addressed the power of the bounds studied. In order to discuss the power of a statistical test, it is necessary to discuss statistical hypothesis testing. In this kind of testing, an inference is drawn from sample data. The hypothesis to be tested is typically called the null hypothesis and designated H_o. In classical statistical testing, H_o is the hypothesis that is the most costly or damaging if rejected when true. It is also necessary to formulate an alternative hypothesis, typically designated H_1. In hypothesis testing, you can make two kinds of errors, a Type I and Type II error. A Type I error, denoted by α, is the probability of accepting H_1 when in fact H_o is the correct conclusion. A Type II error, denoted by β, on the other hand, refers to the probability of accepting H_o when in fact H_1 is correct. There is a need to control both types of risk, but because of cost considerations which necessitate the use of samples instead of a complete census of a population, some risk is unavoidable. It is important to realize that if α is reduced, β will increase. The tradeoff between α and β risk depends on the seriousness of a Type I or Type II error. Ideally, you would want to minimize both types of risk, but since the more extreme error is typically in rejecting H_o when it is true, it is usual to specify the maximum α tolerable, and then to choose the test that minimizes β.

The power of a statistical test refers to the probability of rejecting the null hypothesis when the null hypothesis is in fact false. In other words, it refers to the probability of making the correct decision. In the audit situation, the null hypothesis typically would be

$$H_o : \mu = X \tag{1}$$

where μ is the population error, and X is the hypothesized amount of μ. For the pupose of this research question, X is a predetermined level of error materiality. Auditors typically set such a level since it is not cost effective to redo records to correct small error amounts.

The alternative hypothesis would be

$$H_1 : \mu \neq X \tag{2}$$

By using a statistical upper error limit method, auditors can estimate with a given level of confidence the maximum amount of error in a population based on a sample. They can then compare that estimate to the predetermined level of materiality to help them in making their recommendation (14:349-356). Auditors thus have a need to know the power characteristics of different accounting populations and test statistics (2:42).

II. Literature Review

This literature review examines previous research into the use of statistical sampling and testing procedures in auditing. Specifically, it addresses the use of classical statistical techniques in auditing, the use of dollar unit sampling, and the use of non-Bayesian and of Bayesian techniques for auditing. It also examines a study by Johnson, Leitch and Neter of the error characteristics of certain accounting populations. To clarify the differences between non-Bayesian and Bayesian models for predicting errors, reference is made to Godfrey and Andrews' classification scheme. They distinguish between non-Bayesian models which they call the classical procedure, and Bayesian models. According to their classification, the non-Bayesian procedures include models based on the hypergeometric, Poisson, or binomial distributions. Thus, the errors in a population are presumed to follow these distributions. Bayesian approaches, on the other hand, incorporate auditor prior knowledge of accounting population error characteristics (7:304).

Classical Statistical Techniques

Initially auditors used classical survey sampling estimators such as ratio and difference estimators to construct confidence limits for the proportion of dollar error amounts in accounting populations. Researchers, however, found that classical techniques did not work well in the typical audit situation. Classical statistical theory presumes that these estimators follow an approximately normal distribution for large samples. With their typically low error rates, accounting populations do not satisfy the normality assumption (20:76-77). Specifically, Stringer pointed out that the use of these estimators will lead to unsatisfactory results when no error rates are found in a sample, a reasonable finding to expect in a low error rate population. Stringer noted that with zero sample errors the estimated standard error would also equal zero which would cause the classical confidence limit to collapse into a point estimate. *Kaplan, and Neter and Loebbecke conducted* simulation studies which showed that the nominal confidence level based on the classical normal distribution assumption often differed significantly from the actual confidence intervals. Given the evidence on the poor performance and lack of suitability of classical estimators in audit sampling, Kaplan concluded that new approaches for audit sampling should be considered (17:77-78).

Dollar Unit Sampling

As a response to the inadequacy of classical statistical techniques in the typical audit setting, researchers developed new sampling methods. Dollar unit sampling (DUS) is one method that has gained acceptance. DUS overcomes the drawbacks of classical estimators because it makes no assumption about the normality of the population and is particularly well suited for low error rate populations. This suitability is due to the fact that the sampling unit in DUS is the individual dollar rather than the more traditional physical sampling unit (for example, accounts receivable or invoice). In DUS each dollar, instead of each physical unit, has an equal chance of selection. As a result, the use of DUS reduces the probability of a material dollar error escaping undetected. This is because each dollar unit of the material dollar error has an equal chance of being selected under DUS whereas in physical unit sampling the likelihood of detection depends on the probability of the physical unit which has the material error being selected. In essence, DUS is a very efficient stratification technique. It breaks the sample unit down to the smallest unit, which has obvious advantages in low error rate populations. Furthermore, each dollar unit sample is the same size so that stratification by book value is not necessary (17:79). Thus, DUS is very well suited for audit sampling. DUS is often used in conjunction with error limit models researchers have developed to deal with the need for accurate error prediction in auditing. The DUS method will be described in more detail in Chapter III of this thesis.

Non-Bayesian Upper Error Limit Models

Stringer developed a non-classical procedure for obtaining bounds for the population error amount. The Stringer bound uses the Poisson probability distribution for the number of errors to obtain upper confidence limits on the error amount in a population. Goodfellow, Loebbecke and Neter reviewed the Stringer bound and found a major drawback of the bound to be that there is no general statistical theory to support it. Empirical research to date has also shown the Stringer bound to be overly conservative (17:78).

Neter, Leitch and Fienburg developed a different error prediction method. They used DUS with a multinomial distribution for error prediction to calculate a bound for total population overstatement and understatement errors. Their bound produced tighter bounds than the Stringer bound. Another advantage of their model compared to the Stringer bound is that the distributional properties of the multinomial bound are fully known (17:77).

Bayesian Upper Error Limit Models

Bayesian upper error limit models enable the auditor to incorporate his suppositions on the error characteristics of the audit population into the upper error limit model. This property of Bayesian models can be an advantage as well as a disadvantage. If the auditor suppositions on the error characteristics are fairly accurate, a Bayesian model incorporating these suppositions would probably produce more accurate estimates than other models (7:304). On the other hand, auditors find it difficult to express their suppositions in quantitative terms. Crosby compared two techniques for eliciting auditor beliefs in the form of a probability distribution. The practical applications of his study could not be ascertained. He pointed out that the results of studies in this area are inconclusive and that further research to find improved ways of eliciting auditor beliefs in terms of probability distributions might increase the usefulness of Bayesian methods to auditors (1:364). Alternatively, auditors can resort to the use of robust Bayesian upper error prediction models. These are models with conservative prior error probabilities built into them that have been shown to perform well over a wide variety of populations. Researchers have found such robust models and one objective of this study is to find a robust model for general auditing applications. Bayesian models have outperformed other models in numerous simulation studies in spite of the difficulty of expressing auditor beliefs in terms of a probability distribution. A review of some of the major studies follows.

Garstka proposed a compound Poisson model for DUS with a discrete prior distribution for the mean error taint only. Sample evidence from his simulations indicated that his models yielded much tighter upper error limits than the Stringer bound used by Anderson and Teitlebaum in 1973.

Vanacek proposed a model using a beta prior distribution for the population error rate only. He calculated bounds based on the Bayesian bounds for the error rate using a Stringer type bound (8:499).

Felix and Grimlund developed a model based on a single beta-normal probability distribution for the total error in a composite account. A composite account consists of a large number of subsidiary accounts. Felix and Grimlund noted that Bayesian approaches in the literature had primarily looked at error rates. They felt that auditors would be interested in error amounts. Their approach incorporates auditor judgement on error rate and size. They assume an auditor can express his prior judgement on error rates using the beta distribution and his prior judgement on error size in terms of a normal-gamma-2 probability distribution. Their goal was simply to analyze the properties of this model; they did not test it against other models or run simulations with it (4:24-27).

Cox and Snell used infinite population sampling theory to develop a Bayesian upper error bound for the total overstatement error in a population (8:499). Godfrey and Neter studied the characteristics of the Cox and Snell bound, the sensitivity of the Cox and Snell bound when its assumptions are modified, and the robustness of the Cox and Snell bound. Their simulation studies showed that the Cox and Snell bound is tighter than the Stringer bound and has a higher coverage when the accounting population has a low total error amount. They also concluded that Cox and Snell unmodified and modified bounds are substantially affected by the prior parameters incorporated in them. Thus, by choosing conservative prior parameters, a Cox and Snell bound could be found that is robust, tighter than the Stringer bound for low total error amount populations, and moderately tighter then the Stringer bound for large total error amount populations. As a result, they felt that auditors need not be overly concerned about their ability to correctly assess the probability distribution of their prior judgement. They could simply rely upon the conservative Cox and Snell bound for a wide range of populations (8:520).

Godfrey and Andrews presented a method for establishing an upper error limit using a Bayesian approach which assumed a finite population instead of the infinite population assumed by Cox and Snell. One of their main conclusions was that the finite Bayesian model is more representative of the actual auditing environment than the infinite Bayesian model or classical procedures. They also concluded that sample sizes required by the finite Bayesian method will always be smaller than or equal to sample sizes using infinite Bayesian approaches (7:304-305).

Dworin and Grimlund published the results of their simulation studies with a new method of analysis for DUS in 1984. Their method, which they called the moment bound, incorporates a beta distribution of possible error rates and a hypothetical error tainting based on the average of the observed taints and the number of these observed taints. Their model assumes the mean error taint is approximated by the gamma distribution. They found that the moment bound was significantly tighter than the Stringer bound and generally tighter than the multinomial bound for accounts receivable populations with low material errors. For inventory accounts with both overstatement and understatement errors, the moment bound produced significantly tighter bounds than the Stringer or multinomial bounds (3:219-220).

Tsui, Matsumura and Tsui published findings from their research using a multinomial-Dirichlet bound with DUS. They combined the numltinomial bound suggested by Fienberg, Neter, and Leitch with an error probability distribution based on auditor expectations. The prior distribution they used to incorporate auditor expectations belongs to the class of Dirichlet distributions. They found that their bound usually was tighter than the multinomial bound (20:79).

Accounting Population Error Characteristics

To facilitate the task of making a prior assessment of population error characteristics, auditors need information about the error characteristics of accounting populations. There has been relatively little research in this area.

The results of a study by Crosby indicated that the beta distribution, which has been widely used in practice because of its analytical simplicity, is in fact an adequate model of auditor prior expectations (1:364).

In response to the need for research on accounting population characteristics, Johnson, Leitch, and Neter studied the error characteristics of 55 accounts receivable and 26 inventory populations. They found:

1. There is a great variability in error rates, with error rates tending to be substantially higher in inventory accounts.

2. There is evidence that error rates in both types of populations may be higher for larger accounts and line items.

3. The distribution of error amounts for both types of populations is far from normal.

4. Receivables have mainly overstatement errors while inventories have more of a balance of overstatement and understatement errors.

5. There are frequent 100 percent overstatement errors in accounts receivables and moderately frequent large negative taintings for inventory population.

6. There is variability in taintings for both types of populations, with surprisingly large mean taintings for accounts receivable. The mean taintings of inventories

studied were smaller because of the offsetting effect of understatement error (10:291).

Conclusion

The foregoing review of some of the research on accounting upper error limit models makes clear that numerous statistical tools have been developed to assist in the audit task and that an auditor's choice of upper error limit model can make a great difference in the accuracy of his estimate. The classical statistical methods have some major flaws when used with accounting populations. DUS, Bayesian and non-Bayesian upper error limit models exist to help the auditor. The Bayesian methods show particular promise because they allow the auditor to incorporate his prior knowledge of the accounting population characteristics into the estimate. This advantage is also a disadvantage of the Bayesian methods because of the difficulty auditors may have expressing their prior expectations in the form of a probability distribution. In order to improve the accuracy of upper error limit models, research is needed to find more useable and accurate models, to find techniques for successfully eliciting auditor judgement in the form of a probability distribution, and to document the error characteristics of real accounting populations.

III. Research Methodology

Introduction

This chapter addresses the sampling approach, the upper error limit models, the design of the simluation study, and the measurement procedures used in this study.

The first step of the research effort was to write a computer program to construct study populations. The program constructed 24 study populations by inserting different error characteristics into an actual population of U.S. Air Force aircraft parts inventory accounts. The second step was to use a computer program to draw 500 random dollar unit samples of size 200 from each study population. For each sample, the computer program computed upper error limit bounds using the Stringer and 12 Cox and Snell bounds. Each Cox and Snell bound differed in the prior error characteristics incorporated into it. The fourth step was to measure the performance of each bound in terms of coverage, relative tightness, and power. As defined in Chapter I and as used in this study, coverage refers to the number of times in 500 replications that an upper error limit bound includes the true study population total error amount. Relative tightness refers to the size of one bound compared to another, more conservative or looser bound. In this study, the Stringer bound is the bound used for comparison purposes. A tighter bound is preferable to a looser bound with comparable coverage (18:10). Power refers to the number of times a bound projected a material error when there was in fact a material error. The fifth step was to analyze the results of the study in terms of the research objectives.

Accounting Population

This population was obtained from the United States Air Force (USAF) Logistics Data Supply Center at Gunter Air Force Station, Alabama. It consists of 5412 USAF aircraft parts inventory accounts. Table 3.1 gives some descriptive statistics on this population.

Sampling Approach

In this study, DUS was used. The rationale for using DUS in auditing was already discussed in Chapter II. To summarize, in DUS the sampling unit is the individual dollar unit instead of a physical unit such as an invoice or account receivable. The advantage of using dollar unit samples is that, in the typical low error rate accounting population, DUS reduces the probability of a large material error going undetected. Since DUS is the method employed in this study, it is necessary to understand precisely how the technique works. Neter, Leitch, and Fienberg summed the procedure up succinctly as follows:

TABLE 3.1

Accounting Population Descriptive Statistics

Mean Book Value	$ 197.24
Standard Deviation	$ 947.30
Minimum Book Value	$ .01
Maximum Book Value	$ 42,248.24
Total Book Value	$1,067,465.00
Variance	$ 897,383.67
Skewness	24.08
Kurtosis	849.85
Coefficient of Variation	4.08

1. Determine the total cumulative book value for the line items in the population. The total cumulative book value, X, is the total number of dollar units in the population.

2. If the desired sample size is 200, the auditor selects 200 random numbers from 1 to X. These are the dollar unit samples.

3. Each dollar unit sample is associated with a physical unit. The auditor examines the physical unit to which a dollar unit sample belongs. If the audit value for the physical unit differs from the book value of that unit, there is an error. Thus, the error in the dollar unit is defined as

$$t_e = PU_e/Y_i \qquad (3)$$

where

Y_i = the book value of the physical unit in error

X_i = the audit value of the physical unit in error

$PU_e = Y_i - X_i$ (18:79).

Upper Error Limit Models

Upper error limit models calculate an upper limit or bound on the estimated amount of error in a population based on a sample of that population. The models used in this research are the Stringer and Cox and Snell bounds. A discussion of each follows.

The Stringer Bound. In the following analysis of the Stringer bound, the assumption is made that all errors in the population are overstatement errors. Since inventory and accounts receivable populations both have overstatement errors, this assumption is not unreasonable. Most research with the Stringer bound and the Cox and Snell bounds has dealt with only overstatement errors. According to Godfrey and Neter, substantial modifications to the Cox and Snell bound would be necessary to make it applicable to understatement errors also (8:499). For these reasons, this research will be restricted to overstatement errors. The reader should be aware, however, that this model is not as applicable to populations which show large understatement errors, which is more often the case with inventory accounts than with accounts receivable.

The Stringer bound almost always uses the Poisson distribution to obtain the upper confidence limits for the error in a population. It has been found to be a conservative error estimator (18:5). However, Neter, Leitch and Fienberg point out that use of the binomial distribution with the Stringer bound would yield tighter bounds (17:5).

The Stringer bound for the total overstatement error in the population is

$$D \leq XP_u(0;1-\alpha)+X\sum_{m=1}^{K}[P_u(m;1-\alpha)-P_u(m-1;1-\alpha)]d_m \qquad (4)$$

where

Y = total population audit value

X = total population book value

D = X - Y = total population overstatement error amount

d_1, d_2, d_3 represent the observed sample errors

where $dl \geq d_2 \ . \ . \ . \geq d_k$

n = number of dollar units in random sample

m = number of errors observed

P $(m;1-\alpha)$ stand for the upper error limit for population proportion P based on a confidence coefficient of $1 - \alpha$ when n dollar units are sampled and m errors observed. P $(0;1-\alpha)$ stands for the upper error limit for zero errors, and P $(m-1;1-\alpha)$ has been called the "incremental factor" by Anderson and Teitlebaum (17:79-80). The Stringer bound starts with an assumption that all error taints are 100 percent taints. We know from the assumptions of the model that the error amounts $d_1 \geq d_2 \ldots \geq d_k$. The maximum error incorporated in the Stringer bound is then incrementally reduced by the excess of the maximum error over the actual sample error by the inclusion of the incremental factor (18:17).

The Cox and Snell Bound. Cox and Snell based their Bayesian upper error limit model for the total overstatement error in a population on infinite population sampling theory. Thus, they treated a finite population of N items with total book value, Y, as a continuous density function d(y). They let P(y) be the probability that a line item is in error. Since the total book value of line items containing errors is a function of P(y), Cox and Snell defined π, the probability that a dollar unit is in error, as

$$\pi = TY_{(e)}/Y \qquad (5)$$

where $TY_{(e)}$ is the total book value of line items in error.

Letting $k(e_i | Y_i)$ be the conditional density function of the error amount e_i in a line item, Cox and Snell defined t_i, the taint for the ith dollar unit sample, as

$$t_i = e_i/Y_i \quad (0 < t_i \leq 1) \qquad (6)$$

where Y_i is the book value of line item i.

Note that equation 6 constrains errors to overstatement errors. Then, the mean taint for dollar units in error, μ, is

$$\mu = T_e/TY_{(e)} \qquad (7)$$

where T_e is the total population error amount. Then

$$T_e = Y(\pi)(\mu) \qquad (8)$$

Since Y is known, we must estimate $[(\pi)(\mu)]$, the population mean error amount per dollar unit, which Cox and Snell define as

$$\Upsilon = (\pi)(\mu) \qquad (9)$$

Multiplying Y by Υ then yields the total error amount in the population, T_e.

The assumptions of the Cox and Snell model are:

1. The unconditional error probability, π, follows a gamma prior distribution:

$$\phi_1(\pi) = [(a/\pi_o)^a \pi^{(a-1)} \exp(-a\pi/\pi_o)]/\Gamma(a) \qquad (10)$$

where π_o is the expected value of π and a is a gamma distribution parameter controlling the prior distribution variance.

2. The reciprocal of the mean taint, μ^{-1}, follows a gamma prior distribution:

$$\phi_2(\mu) = [\mu_o^{(b-1)}]^b \mu^{-(b+1)} \exp[-(b-1)\mu_o/\mu]/\Gamma(b) \qquad (11)$$

where μ_o is the expected value of μ and b is the gamma distribution parameter controlling the prior distribution variance.

3. π and μ are independent random variables. As a result, the prior density function of Υ, $f(\Upsilon)$, is a multiple of the F distribution.

4. The observed number of dollar units in error, m, in the sample of n dollar units follows the Poisson distribution with parameter $(n\pi)$. Its probability is

$$L_1 \ (m|\pi) = \exp[(-n\pi)](n\pi)m/m! \tag{12}$$

5. The sample error taints, $t_1, \ldots, t_m$, are m random observations that follow an exponential distribution with parameter $(1/\mu)$. Their probability is

$$L_2(\bar{t}|\mu) = \mu^{-m} \exp(-m\bar{t}/\mu) \tag{13}$$

where

$$\bar{t} = (\Sigma t_i)/m. \tag{14}$$

6. Since the gamma prior distributions for π and $1/\mu$ are conjugate prior probabilities to the likelihoods $L_1(m|\pi)$ and $L_2(\bar{t}|\mu)$, it follows that the posterior density functions of π and $1/\mu$ are also gamma densities. Therefore, the posterior density function of Υ is also a multiple of the F distribution.

Given the prior values of π_o, a, μ_o, b, and the observed sample values of m and $\bar{t}$, the $(1-\alpha)*100$ percentile of the F distribution with 2(m+a) and 2(m+b) degrees of freedom can be calculated. Based upon the above assumptions, the Cox and Snell upper error limit model can be calculated:

$$UEL_{cs} = [[m\bar{t}+(b-1)\mu_o]/[a/\pi_o+n]] \\ [(m+a)/(m+b)]F[(1-\alpha);2(m+a),2(m+b)] \qquad (15)$$

where UEL_{cs} is the Cox and Snell upper error limit (18:19-22).

Simulation Study Design

Simulation Procedures. The following is a summary of the steps used to design the study populations:

1. The accounting population was sorted by line item from lowest to highest book value and split into two strata with equal total book values. The rationale for splitting the population into two strata (low book value/high book value) originates with findings by Johnson, Leitch, and Neter that error rates in accounts receivable and inventory populations are higher for higher value accounts and line items (10:291).

2. A computer program randomly selected line items to be in error. The number of line items selected was in accordance with the mean error rates decided upon for use in this study.

3. For stratum 1 (low book value line items), 100 percent error taints were randomly assigned to 10 percent of the line items receiving errors, in accordance with findings by Johnson, Leitch, and Neter about error characteristics in actual accounting populations. The remaining 90 percent of the line items in error in stratum 1 received random error taints generated from an exponential distribution truncated at 1.0 with the appropriate mean taint. This truncation was done in order to make the study populations more realistic in accordance with the 1984 Godfrey and Neter study (see Appendix A for an explanation of the truncating process). In accordance with the assumptions of the Cox and Snell bound, only overstatement errors were seeded into the study populations (18:27-29).

4. For stratum 2 (high book value line items), random taints were assigned to line items selected to receive errors using an exponential distribution truncated at 1.0 (18:28).

5. The above procedure designed the 24 study populations with the error characteristics shown in Appendix B. Once these were designed, the next step was to take 500 samples of 200 dollar units using systematic random sampling. A computer program using a random starting point and skip intervals accomplished the sampling task. The sampled dollar units and the line items to which they belonged were placed in arrays for analysis.

6. The next step was to compute the Stringer and Cox and Snell bounds. For each dollar unit sample, the Stringer bound and 12 Cox and Snell bounds used in this research were calculated at the 0.95, 0.80 and 0.75 nominal confidence levels.

7. The last step was to analyze the results of the simulation in terms of coverage, relative tightness, and power.

Prior Probability Parameters. Godfrey and Neter, Neter and Godfrey, and Phillips conducted simulation studies to test the performance of Cox and Snell bounds with varying error characteristics. Table 3.2 shows the range of selected prior parameter values for the Cox and Snell bounds used in these studies. It also shows the prior parameters selected for use in this research.

As is evident from the above table, this research focused on a more limited range of prior parameter settings than previous research. This was done in accordance with the findings by Neter and Godfrey that the most promising region in which to find conservative Cox and Snell bound prior parameter values is the range from (μ_o = .40, π_o = .10) to (μ_o = .05, π_o = .20). Furthermore, this research focused on the region in which Neter and Godfrey found three robust bounds (μ_o = .40, π_o = .10, μ_o = .30, π_o = .15, μ_o

TABLE 3.2

Range of Prior Parameter Settings
of Cox and Snell Bounds

	Godfrey and Neter Study	Phillips Study	Cukr Study
Error Rate, π_o	.01-.20	.01-.20	.12-.25
Standard Deviation, $\sigma\pi_o$	.005-.20	.005-.20	.10-.30
Mean Taint, μ_o	.05-.40	.05-.40	.20-.35
Standard Deviation, $\sigma\mu_o$	.025-.40	.025-.40	.10-.30

= .20, π_o = .20) (16:10-11). The specific bounds this research tested are identified in Table 3.3.

Each bound was tested for $\sigma\mu_o$, $\sigma\pi_o$ = .10, .15, .30. This resulted in twelve Cox and Snell bounds tested for each sample. See Appendix C for a complete listing of the prior expectation parameters used for the bounds.

Study Population Characteristics. Twenty-four study populations were constructed with different error rates and mean error taints. Table 3.4 shows the planned error characteristics of the study populations. Actual error characteristics are outlined in Appendix B. They differ from the planned figures because the exponential distribution was

TABLE 3.3

Cox and Snell Bound Prior Parameter Settings

μ_o	π_o
.35	.12
.30	.17
.25	.22
.20	.25

truncated at 1.0 when seeding the error taints (see Appendix A for an explanation).

Table 3.4 shows that the Cukr population tests the Cox and Snell bounds under even more adverse conditions than the previous research in this area. In accordance with the assumptions of the Cox and Snell bound, only overstatement errors were seeded into the study populations.

Table 3.4 also shows that this study uses a smaller error rate (.005) than the Godfrey and Neter studies and the Phillips study. Also, this study uses the highest error rate (.80) that Godfrey and Neter used. The error rates chosen for this thesis were intended to cover a large spectrum of possible error rates and to differ from the error rates used in previous research. The error rates selected are also in accordance with research into actual

TABLE 3.4

Study Populations Error Characteristics

Cukr Study	Godfrey and Neter Studies	Phillips Study
	Error Rates, π	
.80	.80	.50
.25	.50	.30
.10	.30	.15
.05	.20	.15
.01	.10	
.005	.01	
	Approximate Mean Error Taints, μ	
.10/.05*	.40	.40/.20
.20/.10*	.20	.20/.10
.40/.20*	.20	
.80/.40*		

* (low book value/high book value)

accounting population error characteristics by Johnson, Leitch, and Neter. These researchers found error rates in the inventory audits to be substantially higher then error rates in the accounts receivable audits they examined (10:291).

Table 3.4 also shows that this study uses a higher mean taint (.80/.40) than the Godfrey and Neter studies and the Phillips study. The taints chosen for this study cover a broad cross section of possible error taints. They are in accordance with the range of mean taints Johnson, Leitch, and Neter found in their study of actual accounting populations. These researchers found many of their accounts receivable audits to have mostly overstatement errors and high mean taints. The inventory accounts they examined had lower mean taints because of the greater proportion of large understatement errors in these accounts as compared to accounts receivables, which had almost no understatement errors. In accordance with the findings by these researchers that error taintings in the actual accounting populations they studied had frequent 100 percent overstatement errors, 100 percent error taints were randomly assigned to 10 percent of the dollars in error in the low book value strata of the study populations (10:291). The remaining error taints were generated by an exponential distribution to yield the study population mean taints. Additionally, to make the study more realistic in accordance with Phillips, the exponential distribution was truncated at 1.0 (18:28).

Analysis Criteria

The performance of the bounds for each study population was measured in terms of coverage, relative tightness, and power. A description of each measure follows.

Coverage refers to the proportion of times that an upper error limit bound equals or exceeds the true population error amount at the nominal level of confidence. This was measured by summing the bounds over the 500 samples and taking their mean value. To qualify as robust, a bound would have to provide 100 percent coverage $(1-\alpha)*100$ percent of the time.

Relative tightness is measured using the mean ratio of the Stringer bound and each Cox and Snell bound over the 500 samples taken. This measure was used by Phillips (18:35).

Power was measured as the proportion of times a bound correctly predicted a material error. Two levels of materiality, .005 percent and .0025 percent of study population total book value, were used.

IV. Analysis

The Cox and Snell bounds used in the simulation are referred to in the analysis as shown in Table 4.1.

TABLE 4.1

Simulation Study - Cox and Snell Bounds

Cox and Snell Bound	Error Rate π	Error Standard Deviation $\sigma\pi$	Error Taint μ	Taint Standard Deviation $\sigma\mu$
CS1	.12	.10	.35	.10
CS2	.12	.15	.35	.15
CS3	.12	.30	.35	.30
CS4	.17	.10	.30	.10
CS5	.17	.15	.30	.15
CS6	.17	.30	.30	.30
CS7	.22	.10	.25	.10
CS8	.22	.15	.25	.15
CS9	.22	.30	.25	.30
CS10	.25	.10	.20	.10
CS11	.25	.15	.20	.15
CS12	.25	.30	.20	.30

In order to analyze the results of the simulation, it was necessary to first sort the 24 study populations by error amount intensity (EAI), which is calculated by dividing total dollars in error by the total population book value. Algebraically,

$$EAI = T_{LI} \times T_t$$

where T_{LI} is the dollar-unit error rate, and t_t is the mean taint included in the Cox and Snell bound (18:36).

The sorted study populations were logically sorted into four categories: (1) low, (2) medium, (3) high, and (4) highest EAI. Table 4.2 shows the number of study populations and the EAI range in each category.

TABLE 4.2

EAI Categories

Category	EAI Range	Number of Study Populations
Low	.000344 - .008763	11
Medium	.014495 - .060927	8
High	.072104 - .118089	3
Highest	.247347 - .349158	2
Total		24

Table 4.3 shows the distribution of study populations by EAI category.

TABLE 4.3

Study Populations (SP) by EAI Category

Low		Medium		High		Highest	
SP	EAI	SP	EAI	SP	EAI	SP	EAI
21	.000344	15	.014495	7	.072104	3	.247347
17	.000374	10	.017958	8	.112713	4	.349158
22	.000702	5	.018480	2	.118089		
18	.000757	16	.022046				
24	.001067	11	.031940				
23	.001493	6	.036288				
19	.002379	12	.060417				
20	.002806	1	.060927				
13	.004045						
9	.007882						
14	.008763						

Bound Robustness

The first research question addressed bound robustness. To assess upper error limit bound robustness two approaches, point estimate and statistically achieved, were used. A robust upper error limit bound is one that correctly estimates the population error from repetitive samples (1-α) x 100 percent of the time for a wide variety of accounting populations. Using the point estimate approach, a bound is robust as long it meets or exceeds its nominal confidence level in repetitive sampling for all study populations. Using the statistically achieved approach, sampling variation is taken into account. Thus, if the bound has a 95 percent nominal level, an observed coverage of .9309 or greater for all study populations would be sufficient for that bound to be considered robust at the 95 percent nominal confidence level. For the 80 percent and 75 percent nominal confidence levels, the observed coverages would have to be 76.49 percent and 71.20 percent, respectively, for all study populations in order for the bounds to pass the robustness test. The algebraic equation which determines the statistically achieved robustness level for the 95 percent confidence level is

$$.95-1.96[(.95)(.05)/500]\exp\tfrac{1}{2} = .9309$$

Thus, if a bound achieved coverage of at least .9309 in repetitive sampling most of the time (97.5%) at a certain nominal confidence level over all study populations, it could be considered robust (18:39).

Using the point estimate approach, none of the bounds qualified as robust. Bound robustness did not improve using the statistically achieved approach. Tables 4.4

TABLE 4.4

Number of Populations for Which Coverage by Cox and Snell Bounds Met or Exceeded the Nominal Level (Point Estimate Approach)

95 Percent Nominal Level			
	EAI Category		
Cox and Snell Bound	Low	Medium	High + Highest
CS1	11/11	4/8	0/5
CS2	11/11	1/8	0/5
CS3	6/11	0/8	0/5
CS4	11/11	1/8	0/5
CS5	11/11	0/8	0/5
CS6	7/11	0/8	0/5
CS7	11/11	0/8	0/5
CS8	9/11	0/8	0/5
CS9	6/11	0/8	0/5
CS10	9/11	0/8	0/5
CS11	9/11	0/8	0/5
CS12	6/11	0/8	0/5

TABLE 4.5

Number of Populations for Which Coverage by Cox and Snell Bounds Met or Exceeded the Nominal Level (Point Estimate Approach)

80 Percent Nominal Level

	EAI Category		
Cox and Snell Bound	Low	Medium	High + Highest
CS1	11/11	3/8	0/5
CS2	7/11	0/8	0/5
CS3	0/11	0/8	0/5
CS4	11/11	0/8	0/5
CS5	8/11	0/8	0/5
CS6	2/11	0/8	0/5
CS7	11/11	0/8	0/5
CS8	9/11	0/8	0/5
CS9	4/11	0/8	0/5
CS10	9/11	0/8	0/5
CS11	8/11	0/8	0/5
CS12	4/11	0/8	0/5

TABLE 4.6

Number of Populations for Which Coverage by Cox and Snell Bounds Met or Exceeded the Nominal Level (Point Estimate Approach)

	75 Percent Nominal Level		
	EAI Category		
Cox and Snell Bound	Low	Medium	High + Highest
CS1	11/11	3/8	0/5
CS2	7/11	0/8	0/5
CS3	0/11	0/8	0/5
CS4	11/11	0/8	0/5
CS5	8/11	0/8	0/5
CS6	2/11	0/8	0/5
CS7	11/11	0/8	0/5
CS8	9/11	0/8	0/5
CS9	4/11	0/8	0/5
CS10	9/11	0/8	0/5
CS11	8/11	0/8	0/5
CS12	4/11	0/8	0/5

through 4.6 show how many Cox and Snell bounds provided coverage at or above the nominal confidence level. From these tables it is apparent that there were some robust bounds for the low EAI study populations. Specifically, CS4, CS1, and CS7 were robust at all confidence levels for the low EAI study populations. CS2 and CS5 were robust at the 95 percent confidence level only. None of the bounds were robust for the medium, high, and highest EAI study populations. These results are consistent with findings by Godfrey and Neter, Neter and Godfrey, and Phillips. These studies showed the Cox and Snell bounds in general to be robust for the lowest EAI study populations. They also showed that the Cox and Snell bound is very sensitive to changes in its prior parameter settings (8:520).

Since the only robust bounds were for low EAI study populations, the bounds were analyzed further by EAI category, with particular emphasis on the low EAI category. Of the Cox and Snell bounds that were robust for the low EAI study populations, CS4 had the highest mean coverage for all three confidence levels, followed closely by CS1 and CS7, respectively. CS2 and CS5, although only robust at the 95 percent nominal level, had mean coverages above the nominal levels for the other nominal levels. Additionally, CS8 and CS10 had mean coverages above the nominal levels for the 80 percent and 75 percent nominal confidence levels. CS10 also had a fairly high mean coverage (.9164) for the 95 percent nominal confidence level. Tables 4.7 to 4.10 show the mean coverages of the Cox and Snell bounds by EAI category and nominal confidence level.

The bounds that were robust for all nominal levels for the low EAI study populations, CS4, CS1, and CS7 all had moderate (.10) prior parameter settings for $\sigma\mu$ and $\sigma\pi$. Their prior parameter settings for error rate, π, ranged from .12 to .22 and their prior parameter settings for mean taint, μ, ranged from .25 to .35. The next best performing bounds were CS2 and CS5. These bounds differed from CS1 and CS4 only in their higher prior parameter settings for $\sigma\mu$ and $\sigma\pi$. CS10 came next in terms of mean coverage. It had moderate (.10) prior parameter settings for $\sigma\mu$ and $\sigma\pi$, and a higher error rate and lower error taint than any of the other bounds that performed well. CS8, which followed CS10 in terms of mean coverage, had an expected error rate and mean taint between that of CS5 and CS10, but a higher prior parameter setting for $\sigma\pi$ and $\sigma\mu$.

From the above results, it appears that the bounds, when applied to low EAI study populations, are particularly sensitive to changes in $\sigma\pi$ and $\sigma\mu$. A lower prior setting for these parameters seems to result in increased coverage. Also, high prior error taint settings (.25 to .35) combined with moderate (.12 to .22) prior error rate settings seem to increase coverage.

TABLE 4.7

Mean Coverages of Cox and Snell Bounds
for Low EAI Study Populations

Cox and Snell Bound	Mean Coverage Nominal Confidence Level 95%	80%	75%
CS1	1.0000	1.0000	0.9669
CS2	1.0000	0.8725	0.8362
CS3	0.7247	0.4209	0.3558
CS4	1.0000	1.0000	0.9998
CS5	1.0000	0.8507	0.7922
CS6	0.7544	0.4073	0.3596
CS7	1.0000	0.9902	0.9515
CS8	0.8451	0.8184	0.8158
CS9	0.6642	0.4625	0.3547
CS10	0.9164	0.8200	0.8184
CS11	0.8182	0.7273	0.6944
CS12	0.6124	0.4151	0.3038

TABLE 4.8

Mean Coverages of Cox and Snell Bounds
for Medium EAI Study Populations

	Mean Coverage		
	Nominal Confidence Level		
Cox and Snell Bound	95%	80%	75%
CS1	0.5403	0.4005	0.3795
CS2	0.2508	0.0830	0.0665
CS3	0.0840	0.0377	0.0308
CS4	0.4265	0.1303	0.0910
CS5	0.1000	0.0455	0.0355
CS6	0.0730	0.0308	0.0240
CS7	0.0863	0.0268	0.0175
CS8	0.0650	0.0258	0.0175
CS9	0.0620	0.0255	0.0172
CS10	0.0400	0.0063	0.0027
CS11	0.0505	0.0122	0.0068
CS12	0.0570	0.0215	0.0122

TABLE 4.9

Mean Coverages of Cox and Snell Bounds
for High EAI Study Populations

Cox and Snell Bound	Mean Coverage Nominal Confidence Level 95%	80%	75%
CS1	0.2507	0.1187	0.0893
CS2	0.2073	0.0840	0.0620
CS3	0.1740	0.0653	0.0507
CS4	0.1573	0.0440	0.0267
CS5	0.1633	0.0527	0.0340
CS6	0.1660	0.0567	0.0393
CS7	0.0940	0.0133	0.0053
CS8	0.1380	0.0320	0.0213
CS9	0.1573	0.0513	0.0327
CS10	0.0633	0.0033	0.0000
CS11	0.1167	0.0233	0.0140
CS12	0.1527	0.0480	0.0313

TABLE 4.10

Mean Coverages of Cox and Snell Bounds
for Highest EAI Study Populations

	Mean Coverage		
	Nominal Confidence Level		
Cox and Snell Bound	95%	80%	75%
CS1	0.1680	0.0060	0.0000
CS2	0.3490	0.0230	0.0070
CS3	0.4460	0.0480	0.0210
CS4	0.0800	0.0000	0.0000
CS5	0.2660	0.0100	0.0000
CS6	0.4200	0.0380	0.0190
CS7	0.0460	0.0000	0.0000
CS8	0.2120	0.0050	0.0000
CS9	0.4120	0.0350	0.0160
CS10	0.0350	0.0000	0.0000
CS11	0.1830	0.0020	0.0000
CS12	0.4040	0.0350	0.0100

In his 1985 study, Phillips found one robust bound. In Phillips' study, the highest EAI for a study population was .1259, compared to .3492 for this study. This must be borne in mind when comparing the results of the two studies. Phillips' robust bound had high prior parameter settings (.40 and .20) for μ and $\sigma\mu$, and moderate settings (.10 and .10) for π and $\sigma\pi$. CS1 and CS2 are the bounds with the closest prior parameter settings to the settings of Phillips' robust bound. Interestingly, CS1 is also the bound with the highest mean coverage for the medium EAI study populations. CS4 has the second highest mean coverage for the medium EAI study populations. CS2 is in third place. CS4 and CS1 have their low prior parameter settings for $\sigma\pi$ and $\sigma\mu$ in common. CS2, on the other hand, has a higher prior parameter setting for $\sigma\pi$ and $\sigma\mu$. It also has higher mean coverage for the medium EAI study populations then CS5, which has the same prior parameter setting for $\sigma\pi$ and $\sigma\mu$ but a lower prior parameter setting for μ and a higher setting for π. Thus, it seems that for the medium EAI study populations, low settings for $\sigma\mu$ and $\sigma\pi$ increase coverage. At the same time, a higher mean taint may be necessary to achieve higher coverage for these study populations. This supports the results of Phillips, Godfrey and Neter, and Neter and Godfrey (18:105).

For the highest EAI study populations the mean coverages for CS3, CS6, CS9, and CS12, were significantly higher than the mean coverages of the bounds that performed well for the low and medium EAI study populations. CS3, CS6, CS9, and CS12 have in common high prior parameter settings (.30) for $\sigma\pi$ and $\sigma\mu$.

To summarize, moderate to high prior settings for mean taint (.25 to .35), moderate prior error rates (.12 to .22) and moderate prior parameter settings (.10 to .15) for $\sigma\mu$ and $\sigma\pi$ resulted in high coverage for low EAI study populations. For medium EAI study populations, a high prior mean taint (.35), moderate error rate prior parameter setting (.12), and moderate prior parameter setting (.10) for $\sigma\mu$ and $\sigma\pi$ increased the mean coverage. Phillips found a robust bound for all EAI categories. It had a high prior parameter setting (.40) for error taint, a moderate prior parameter setting (.20) for $\sigma\mu$ and moderate settings (.10) for error rate and $\sigma\pi$ (18:60). Thus, the difference between high coverages for low and medium EAI study populations appears to lie in the prior parameter setting for mean taint. A higher mean taint increased the coverage for the medium EAI populations. The results of this study also show that for the highest EAI study populations, a high (.30) prior parameter setting for $\sigma\mu$ and $\sigma\pi$ resulted in increased coverage. Thus, in selecting prior parameter settings, the probable EAI of the study population should be considered. Research into the error characteristics of accounting populations will help in this task. Also, researchers can narrow their search for robust bounds to the prior parameter settings identified by Godfrey and Neter and verified by Phillips.

Bound Mean Relative Tightness

The second research question addressed bound relative tightness. Bounds were evaluated in terms of mean relative tightness. Mean relative tightness is measured by comparing the dollar value of one bound to another, more conservative bound (eg., the Stringer bound). Given acceptable coverage, the bound with the greater relative tightness is the one auditors would prefer. A tighter bound reduces the risk of incorrectly stating that an account contains a material error because it gives upper error limits closer to the true error amount than a looser bound. Thus, a tighter bound with acceptable coverage will make any required adjustment in financial data smaller. As a result, of two bounds with the same coverage, the tighter is preferred (18:62).

In this study, the measure of relative tightness is the arithmetic mean of the Stringer bound divided by each Cox and Snell bound for the 500 samples from each study population. The Stringer bound was presumed to be the more conservative bound and was thus placed in the numerator of the equation to compute mean relative tightness

$$MRT = \left(\sum_{i=1}^{500} STR_i / CS_i\right)/500$$

where

MRT = mean relative tightness

STR_i = the Stringer bound value for iteration i

CS_i = the Cox and Snell bound value for iteration i (18:63)

Tables 4.11 to 4.14 show the mean relative tightness of each Cox and Snell bound. With regard to relative tightness, the robust bounds for the low EAI study populations are of most interest. These and the other Cox and Snell bounds that provided adequate coverage are examined in the next paragraphs.

Of the three bound that were robust for the three confidence levels, CS1 had the highest mean relative tightness. CS4 and CS7 also showed significant increases in relative tightness over the Stringer bound. CS2 and CS5 provided even more significant increases in relative tightness. Thus, an auditor could choose CS2 or CS5 over CS1, CS4, or CS7 if a tighter bound and lower but adequate coverage is acceptable to that auditor. On the other hand, the more risk averse auditor would choose CS1, CS4, or CS7. CS8 and CS11 also had high mean relative tightnesses compared to the Stringer bound. CS11 has the highest mean relative tightness of the bounds with high mean coverage.

For the highest EAI category, where a number of bounds achieved mean coverages of over 40 percent for the 95 percent nominal level, the increase in relative tightness

TABLE 4.11

Mean Relative Tightness of the Cox and Snell Bounds Compared to the Stringer Bound Low EAI Study Populations

Cox and Snell Bound	Nominal Confidence Level 95%	80%	75%
CS1	1.8511	1.7156	1.6797
CS2	2.6958	2.9305	2.9929
CS3	5.7023	11.7460	15.2107
CS4	1.6409	1.4386	1.3873
CS5	2.5060	2.4767	2.4544
CS6	5.0995	7.0387	7.7213
CS7	1.6065	1.3900	1.3351
CS8	2.5721	2.4620	2.4157
CS9	5.2742	6.3660	6.6379
CS10	1.8809	1.6503	1.5885
CS11	3.0746	2.9494	2.8916
CS12	6.2194	7.2580	7.4702

TABLE 4.12

Mean Relative Tightness of the Cox and Snell Bounds Compared to the Stringer Bound Medium EAI Study Populations

Cox and Snell Bound	Nominal Confidence Level 95%	80%	75%
CS1	1.7628	1.6413	1.6151
CS2	2.2752	2.1191	2.0845
CS3	3.2041	2.9527	2.8962
CS4	2.0840	1.9296	1.8962
CS5	2.6619	2.4624	2.4182
CS6	3.5451	3.2443	3.1771
CS7	2.4733	2.2750	2.2321
CS8	3.0969	2.8417	2.7854
CS9	3.8909	3.5341	3.4551
CS10	2.9865	2.7262	2.6698
CS11	3.5960	3.2693	3.1977
CS12	4.2423	3.8227	3.7306

TABLE 4.13

Mean Relative Tightness of the Cox and Snell Bounds Compared to the Stringer Bound High EAI Study Populations

Cox and Snell Bound	Nominal Confidence Level		
	95%	80%	75%
CS1	1.9551	1.8975	1.8851
CS2	2.1431	2.0800	2.0663
CS3	2.3255	2.2545	2.2388
CS4	2.1425	2.0747	2.0600
CS5	2.2677	2.1972	2.1818
CS6	2.3768	2.3023	2.2858
CS7	2.2989	2.2214	2.2046
CS8	2.3649	2.2880	2.2711
CS9	2.4169	2.3395	2.3224
CS10	2.4184	2.3333	2.3147
CS11	2.4353	2.3536	2.3857
CS12	2.4478	2.3682	2.3506

TABLE 4.14

Mean Relative Tightness of the Cox and Snell Bounds Compared to the Stringer Bound Highest EAI Study Populations

Cox and Snell Bound	Nominal Confidence Level		
	95%	80%	75%
CS1	1.2551	1.2465	1.2456
CS2	1.2265	1.2196	1.2190
CS3	1.2090	1.2033	1.2028
CS4	1.2847	1.2758	1.2747
CS5	1.2406	1.2336	1.2330
CS6	1.2135	1.2076	1.2072
CS7	1.3051	1.2958	1.2948
CS8	1.2506	1.2435	1.2428
CS9	1.2169	1.2110	1.2106
CS10	1.3135	1.3042	1.3031
CS11	1.2552	1.2481	1.2474
CS12	1.2189	1.2129	1.2126

over the Stringer bound was not significant, particularly since the Stringer bound was robust for these study populations. As a result, the Stringer bound may well be a better choice for the highest EAI study populations.

Nominal Confidence Level Effects

The third research objective was to examine the effects of lowering the nominal confidence levels of the bounds on bound coverage and mean relative tightness. The assumption was that by lowering the nominal confidence level, it might be possible to find tighter bounds with acceptable coverages. For bounds with comparable coverages but different nominal confidence levels, the auditor might choose the lower confidence level if it yielded sufficient increases in mean relative tightness (18:70).

The effect of nominal confidence levels was only examined for bounds that performed well for low EAI study populations. Three bounds were robust for all confidence levels, CS1, CS4, and CS7. All these bounds had actual mean coverages over 95 percent even at the 75 percent nominal confidence level. CS2 and CS5, on the other hand, experienced larger losses in mean coverage as the nominal confidence level was lowered from 95 percent to 80 percent. These bounds still had mean coverages that exceeded their nominal confidence levels at all confidence levels. CS8 and CS10 were not robust for any of the low EAI study populations, but had mean coverages exceeding the nominal confidence level for the 80 percent and 75 percent nominal levels. These two bounds only showed small losses in mean coverage as the nominal levels were lowered.

A lower nominal confidence level will generally result in a tighter bound. Thus, if high coverage is an acceptable criterion for selecting a bound, an auditor might prefer a bound that provides high mean coverage and increases in tightness to a looser bound. Table 4.15 shows the improvements in relative tightness for bounds as the nominal confidence level is lowered. The percentages in this table were derived by dividing the mean coverage for each bound at the 95 percent nominal level by the mean coverage of each bound at the 85 percent nominal level and multiplying the result by 100 percent. The same was done for the 85 percent and 75 percent nominal confidence levels (18:73).

Table 4.15 shows that increases in tightness were only significant for CS2, CS5, and CS10. These are the bounds with lower mean coverages at the lower confidence levels. In order to achieve the gain in tightness that comes from lowering the confidence level, it would be necessary to sacrifice a significant amount of coverage.

Power of the Bounds

The fourth research objective was to examine the power of the bounds tested. The more power a bound has, the more likely the auditor is to correctly predict a material error. Two levels of materiality were tested for: ½ percent and

TABLE 4.15

Effects of Lowering the Nominal Confidence
Level on Relative Mean Tightness

Mean 95% Bound Coverage Relative to Mean 85% Bound Coverage for low EAI Study Populations

Cox and Snell Bound	Percentage Comparison
CS1	1.0000
CS2	1.1461
CS4	1.0000
CS5	1.1755
CS7	1.0099
CS8	1.0326
CS10	1.1176

Mean 85% Bound Coverage Relative to Mean 75% Bound Coverage for low EAI Study Populations

Cox and Snell Bound	Percentage Comparison
CS1	1.0340
CS2	1.0434
CS4	1.0002
CS5	1.0738
CS7	1.0407
CS8	1.0032
CS10	1.0020

½ percent of total book value. Thus, if a bound had a power of 1 at the 95 percent nominal confidence level, the auditor could say that 95 percent of the time that bound would correctly predict a material error.

For the high and highest EAI study populations, all bounds correctly predicted a material error 100 percent of the time at all nominal confidence levels tested. For the medium EAI study populations, most of bounds correctly predicted a material error 100 percent of the time at all confidence levels. Since none of the bounds in these EAI categories provided adequate coverage, they are not discussed further.

Tables 4.16 through 4.17 contain power statistics for the low EAI populations. Note that, of the bounds that were robust for the low EAI populations at all confidence levels, only CS7 had a power of 1 for both levels of materiality at all nominal confidence levels. The power of CS1 and CS4 dropped when the nominal confidence level was lowered for the ½ percent level materiality. CS2 and CS5, which had mean coverages above the nominal confidence levels at all nominal levels, had fairly low power. CS10 had a power of 1 for all nominal confidence levels and for both levels of materiality. Thus, in terms of power only, CS7 and CS10 performed best of the bounds with adequate coverage, followed by CS4 and CS2. CS4 and CS2 both had a power of 1 at the 95 percent nominal confidence level, however. Since

TABLE 4.16

Statistical Power of Selected Cox and Snell Bounds
Low EAI Study Populations

Level of Materiality: .005 x Total Book Value

Cox and Snell Bound	Nominal Confidence Level 95%	80%	75%
CS1	1.0000	0.4655	0.5655
CS2	0.7053	0.3424	0.2971
CS3	0.3424	0.1664	0.1375
CS4	1.0000	0.7053	0.7053
CS5	0.7053	0.2971	0.2713
CS6	0.1825	0.0318	0.0205
CS7	1.0000	1.0000	1.0000
CS8	1.0000	0.1884	0.1825
CS9	0.0547	0.0082	0.0040
CS10	1.0000	1.0000	1.0000
CS11	0.8920	0.0256	0.0178
CS12	0.0105	0.0007	0.0000

TABLE 4.17

Statistical Power of Selected Cox and Snell Bounds
Low EAI Study Populations

Level of Materiality: .0025 x Total Book Value

Cox and Snell Bound	Nominal Confidence Level 95%	80%	75%
CS1	1.0000	1.0000	1.0000
CS2	1.0000	0.7053	0.7053
CS3	0.7053	0.4655	0.3424
CS4	1.0000	1.0000	1.0000
CS5	1.0000	0.7053	0.7053
CS6	0.7053	0.3165	0.2844
CS7	1.0000	1.0000	1.0000
CS8	1.0000	1.0000	1.0000
CS9	0.7053	0.2120	0.1909
CS10	1.0000	1.0000	1.0000
CS11	1.0000	1.0000	1.0000
CS12	0.6582	0.1825	0.1802

lowering the nominal confidence level did not result in significantly tighter bounds for CS1 and CS4, then, these bounds could be used at the 95 percent confidence nominal level where they correctly predicted a material error 100 percent of the time.

V. Summary and Recommendations for Future Research

This study builds upon research into the properties of the Cox and Snell Upper Level Limit Model by Godfrey and Neter and Phillips. The prior probability settings for error rate, mean taint, and the standard deviations of these parameters used in this study differed from the settings used in earlier research. Twenty-four study populations were constructed out of an actual accounting population of USAF aircraft inventory parts. The Cox and Snell bounds were then tested to see how well they captured the true error in various study populations. Specifically, this study focused on the coverage and mean relative tightness of the bounds. It also looked at the effects of lowering the nominal confidence level on coverage and mean relative tightness and at the statistical power of the bounds.

The bounds only provided coverage at or above the nominal confidence levels for the low EAI study populations. Therefore, none of the bounds was robust as defined in Chapter I. For the low EAI study populations, however, there were some robust bounds. CS1, CS4, and CS7 were robust for all nominal confidence levels. CS4 had the highest mean coverage of all the bounds. CS2 and CS5 were robust at the 95 percent nominal confidence level only but had mean coverages above the nominal confidence level for the other nominal levels. CS8 and CS10 were not robust, but had high mean coverages. Their mean coverages were lower than the nominal confidence level at the 95 percent level, but exceeded the nominal levels for the other confidence levels.

Of the bounds that were robust or had high mean coverages for the low EAI study populations, all provided significant increases in relative tightness over the Stringer bound. Of the robust bounds, CS1 had the highest mean relative tightness. CS2 and CS5 had even more significant increases in mean relative tightness. CS10 had the highest mean relative tightness of the bounds which provided high mean coverages. Thus, if an auditor could accept a mean coverage of 0.916 at the 95 percent nominal confidence level, but wanted a tight bound, CS10 would be a good choice. On the other hand, the more risk averse auditor would choose CS1, CS4, or CS7, which were robust and provided significant increases in relative tightness over the Stringer bound. Lowering the nominal confidence level only resulted in very insignificant increases in tightness for CS1, CS4, and CS7. CS2, CS5, and CS10 had fairly significant (above 10 percent) increases in tightness as the nominal level was lowered from 95 percent to 80 percent.

The most powerful of the bounds were CS7 and CS10, which provided $(1-\alpha) \times 100$ percent accuracy in predicting a material error correctly for all confidence levels and both levels of materiality. CS1, CS4, and CS8 provided (1-α) x 100 percent accuracy in predicting a material error correctly at the 95 percent nominal confidence level for both levels of materiality. Lowering the confidence levels or raising the level of materiality lowered the power of some of the bounds significantly.

Recommendations for Future Research

Further research is called for in the following areas:

1. The error characteristics of actual accounting populations should be researched. The robust Cox and Snell bounds found in this research could then be applied to the kinds of populations more likely to have low error amount intensities.

2. Other upper error limit models should be tested against the populations used in this study to see how their performance compares to that of the Cox and Snell bound.

Appendix A: Relationship Between Untruncated and Truncated Exponential Distribution

When the exponential distribution is truncated at 1.0, its mean value relationship to the untruncated distribution can be expressed as

$$\mu^* = \{1-1/(1/\mu)[e^{-1/\mu}/(1-e^{-1/\mu})]\}$$

where

μ = the untruncated mean

μ^* = the truncated mean (18:306)

Appendix B: Study Population Error Characteristics

Study Population Number	Total Error Amount	Error Rate	Error Taint	EAI
1	65037.69	0.80	0.078	0.0609
2	126055.59	0.80	0.151	0.1181
3	264033.97	0.80	0.317	0.2473
4	372714.31	0.80	0.447	0.3492
5	19727.11	0.25	0.074	0.0185
6	38736.07	0.25	0.146	0.0363
7	76968.91	0.25	0.290	0.0721
8	120317.25	0.25	0.454	0.1127
9	8414.06	0.10	0.065	0.0079
10	19169.31	0.10	0.148	0.0180
11	34094.70	0.10	0.264	0.0319
12	64493.30	0.10	0.499	0.0604
13	4318.28	0.05	0.067	0.0040
14	9354.64	0.05	0.145	0.0088
15	15472.69	0.05	0.240	0.0145
16	23533.71	0.05	0.365	0.0220
17	399.57	0.01	0.064	0.0004
18	808.00	0.01	0.130	0.0008
19	2539.08	0.01	0.407	0.0024

Appendix B (continued)

Study Population Number	Total Error Amount	Error Rate	Error Taint	EAI
20	2995.82	0.01	0.481	0.0028
21	366.86	0.005	0.112	0.0003
22	749.29	0.005	0.228	0.0007
23	1594.15	0.005	0.485	0.0015
24	1138.80	0.005	0.347	0.0011

Appendix C: Cox and Snell Bound Prior Expectation Parameters

Bounds	π_o	$\sigma\pi_o$	μ_o	$\sigma\mu_o$	a	b	Υ_o
CS1	0.12	0.10	0.35	0.10	1.44	14.25	0.042
CS2	0.12	0.15	0.35	0.15	0.64	7.44	0.042
CS3	0.12	0.30	0.35	0.30	0.16	3.36	0.042
CS4	0.17	0.10	0.30	0.10	2.89	11.00	0.051
CS5	0.17	0.15	0.30	0.15	1.28	6.00	0.051
CS6	0.17	0.30	0.30	0.30	0.32	3.00	0.051
CS7	0.22	0.10	0.25	0.10	4.84	8.25	0.055
CS8	0.22	0.15	0.25	0.15	2.15	4.78	0.055
CS9	0.22	0.30	0.25	0.30	0.54	2.69	0.055
CS10	0.25	0.10	0.20	0.10	6.25	6.00	0.050
CS11	0.25	0.15	0.20	0.15	2.78	3.78	0.050
CS12	0.25	0.30	0.20	0.30	0.69	2.44	0.050

$\Upsilon_o = \pi_o \times \mu_o$

$a = \pi_o^2/\sigma_{\pi_o}^2$

$b = (\mu_o^2/\sigma_{\mu_o}^2) + 2$

π_o = prior expectation parameter of π

μ_o = prior expectation parameter of μ

Bibliography

1. Crosby, Michael A. "Bayesian Statistics in Auditing: A Comparison of Probability Elicitation Techniques," The Accounting Review, 56: 355-364 (April 1981).

2. Duke, Gordon L., John Neter and Robert A Leitch. "Power Characteristics of Test Statistics in the Auditing Environment: An Empirical Study," Journal of Accounting Research, 20: 42-66 (Spring 1982).

3. Dworin, Lowell and Richard A. Grimlund. "Dollar Unit Sampling for Accounts Receivable and Inventory," The Accounting Review, 59: 218-241 (April 1984).

4. Felix, Wiliam L., Jr. and Richard A. Grimlund. "A Sampling Model for Audit Tests of Composite Accounts," Journal of Accounting Research, 15: 23-40 (Spring 1977).

5. Fienberg, Stephen E., John Neter and R. A. Leitch. "Estimating the Total Overstatement Error in Accounting Populations," Journal of the American Statistical Association, 72: 295-301 (June 1977).

6. Garstka, Stanley J. "Models for Computing Upper Error Limits in Dollar Unit Sampling," Journal of Accounting Research, 15: 179-191 (Autumn 1977).

7. Godfrey, James T. and Richard W. Andrews. "A Finite Population Bayesian Model for Compliance Testing," Journal of Accounting Research, 120: 304-315 (Autumn 1982).

8. Godfrey, James T. and John Neter. "Bayesian Bounds for Monetary Unit Sampling in Accounting and Auditing," Journal of Accounting Research, 22: 497-525 (Autumn 1984).

9. Helton, 1st Lt Michael W. A Validation of an Accounting Upper Error Limit Bound. MS Thesis LSY/85S-17. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1985.

10. Johnson, Johnny R., Robert A. Leitch and John Neter. "Characteristics of Errors in Accounts Receivable and Inventory Audits," The Accounting Review, 56: 270-293 (April 1981).

11. Leslie, Donald A., Albert D. Teitlebaum and Rodney Anderson. Dollar Unit Sample: A Practical Guide for Auditors. Toronto: Copp Clark Pitman Publishers Inc., 1979.

12. McCray, John H. "Quasi-Bayesian Audit Risk Model for Dollar Unit Sampling," The Accounting Review, 59: 35-50 (January 1984).

13. Menzefricke, Ulrich. "Using Decision Theory for Planning Audit Sample Size with Dollar Unit Sampling," Journal of Accounting Research, 22: 570-586 (Autumn 1984).

14. Menzefricke, Ulrich and Wally Smieliauskas. "A Simulation Study of the Performance of Parametric Dollar Unit Sampling Statistical Procedures," Journal of Accounting Research, 22: 588-603 (Autumn 1984).

15. Muskowitz, Herbert and Gordon P. Wright. Statistics for Management and Economics. Columbus: Bell and Howell Company, 1985.

16. Neter, John and James Godfrey. "Robust Bayesian Bounds for Monetary-Unit Sampling." Unpublished article. University of Georgia, College of Business Administration GA, 12 December 1984.

17. Neter, John, Robert A. Leitch and Stephen E. Fienberg. "Dollar Unit Sampling: Multinomial Bounds for Total Overstatement and Understatement Errors," The Accounting Reveiw, 53: 77-92 (January 1978).

18. Phillips, Jeffrey Joseph. Bayesian Bounds for Monetary Unit Sampling Using Accounting Populations. PhD dissertation. University of Georgia, Athens GA, 1985.

19. Reneau, J. Hal. "CAV Bounds in Dollar Unit Sampling: Some Simulation Results," The Accounting Review, 53: 669-680 (July 1978).

20. Tsui, Kam-Wah, Ella Mae Matsumura and Kwok-Leung Tsui. "Multinomial-Dirichlet Bounds for Dollar Unit Sampling in Auditing," The Accounting Review, 60: 76-95 (January 1985).

VITA

Captain Anita J. R. Cukr was born on 9 December 1954 in San Rafael, California. She graduated from high school in Heidelberg, Germany, in 1973 and attended the American College in Paris from which she received the Bachelor of Arts in Business Management in January 1979. In November 1979, she received a commission in the USAF through the Officer Training School program. In December 1979, she was assigned to the Budget Division at Headquarters Air Force Logistics Command, Wright-Patterson AFB, Ohio. In August 1982, she was assigned as a cost analysis officer to the Air Force Plant Representative Office at the Boeing Company in Seattle, Washington, where she worked until entering the School of Systems and Logistics, Air Force Institute of Technology, in May 1985.

Permanent Address: 5675 Sandpiper Lane
Dayton, Ohio 45424

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): AFIT/GSM/LSY/86S-5	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: School of Systems and Logistics; 6b. OFFICE SYMBOL (If applicable): AFIT/LSY	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433-6583	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): See Box 19	
12. PERSONAL AUTHOR(S): Anita J. R. Cukr, B.A., Captain, USAF	
13a. TYPE OF REPORT: MS Thesis; 13b. TIME COVERED: FROM ____ TO ____	14. DATE OF REPORT (Yr., Mo., Day): 1986 September; 15. PAGE COUNT: 85
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES		
FIELD	GROUP	SUB. GR.
05	01	
14	01	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Accounting, Auditing, Sampling, Model Theory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Title: SELECTED UPPER ERROR LIMIT METHODS APPLIED TO AN ACCOUNTING POPULATION

Thesis Chairman: Jeffrey J. Phillips, Lt Col, USAF

Approved for public release: IAW AFR 190-1. 24 Sept 86
LYNN E. WOLAVER
Dean for Research and Professional Development
Air Force Institute of Technology (AFC)
Wright-Patterson AFB OH 45433

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Jeffrey J. Phillips, Lt Col, USAF	22b. TELEPHONE NUMBER (Include Area Code): 513-255-4845; 22c. OFFICE SYMBOL: AFIT/LSY

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

The purpose of this research was to study the performance of the Cox and Snell Upper Error Limit Model. Twelve Cox and Snell bounds were constructed by changing the prior error probabilities incorporated in the model. The performance of the twelve bounds was compared to the performance of the Stringer bound.

The study constructed twenty-four study populations from an actual accounting population, sampled from these study populations, and compared the error estimates produced by the Cox and Snell model with the true dollar error in the study populations. The objective was to find the best upper error limit estimator from the Cox and Snell bounds tested. The bounds were examined in terms of robustness, mean relative tightness, effect of nominal confidence level, and statistical power.

In order to analyze the results of the simulation, the study populations were divided into four categories by error amount intensity. Robust bounds were found for the low error amount intensity study populations only. For the medium, high, and highest error amount intensity populations, bound coverage was significantly less than the nominal confidence levels of the bounds. The Cox and Snell bounds were significantly tighter than the Stringer bound. The effect of lowering the bound confidence level on the relative tightness of the bounds was negligible. Some of the bounds which provided adequate coverage were statistically powerful.

END

12-86

DTIC