This is to certify that the

dissertation entitled

The Use of Empirical

Bayes Estimation In Audit Testing

presented by

David William Wright

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Accounting

Major professor

Date 5/12/86

THE USE OF EMPIRICAL BAYES ESTIMATION IN AUDIT TESTING

By

David William Wright

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Accounting

1986

ABSTRACT

THE USE OF EMPIRICAL
BAYES ESTIMATION IN AUDIT TESTING

By

David William Wright

The objectives of this study are to propose and investigate empirical Bayes estimation as a tool useful in efficiently integrating the results of certain audit procedures. Empirical Bayes integration of audit evidence is analyzed across two domains -- interprocedural integration of evidence gathered from statistical sampling and analytical review procedures and interitem integration of statistical sampling evidence from several similar accounting populations. The use of empirical Bayes procedures is considered for both major phases of the audit -- internal control compliance testing and account balance substantive testing.

The parametric empirical Bayes (PEB) estimator requires the auditor's estimation problem to be routine. Thus, an unknown parameter is to be estimated for each of several similar accounting populations. These parameters are modeled in the Bayesian fashion as realizations from an underlying prior probability distribution. The objective of PEB estimation is to partially exploit the efficiency of pure Bayesian procedures while avoiding the risks from the misspecification of parameters in the prior probability distribution. PEB estimators are essentially Bayesian in style, substituting estimates for the parameters of the prior probability distribution obtained from the sample data itself in lieu of subjective prior specifications.

The PEB estimator for population error rates during internal control compliance testing has a frequentist interpretation as the James-Stein (JS) estimator. The use of the estimator in these circumstances was considered from both the frequentist and Bayesian perspectives.

The efficiency, bias and reliability characteristics of the estimator were examined over numerous realistic audit scenarios using both exact numerical computations and Monte Carlo simulation methods. In general, the results showed the PEB/JS estimator was universally efficient relative to classical sampling procedures currently employed by auditors. Confidence interval procedures for the PEB/JS estimators were shown to produce reliability levels of the same relative magnitude as classical procedures with narrower confidence intervals. Finally, relative to pure Bayesian estimation, PEB procedures were shown to avoid the inefficiency and unreliability induced by subjective misspecification of the prior probability distribution parameters in pure Bayesian estimation through the objective estimation of these prior parameters from the sample data itself.

For April, whose unwavering support and
optimism made it all worthwhile

ACKNOWLEDGMENTS

It is with distinct pleasure that I express my gratitude for the unselfish efforts of the many individuals who played a part in my progress through the Ph.D. program.

The supportive environment maintained by my dissertation committee members, Dr. Ronald Marshall, chairman; Dr. D. Dewey Ward; and Dr. Raoul LePage, was invaluable. In particular, Dr. Ward's uplifting sense of humor, valued counsel and expressions of confidence in my abilities allowed me to persevere over obstacles which, at the time, seemed insurmountable. Dr. Marshall's advice, insight and academic zeal were inspirations to me at all levels of the doctoral program.

In addition, I am appreciative to the entire accounting faculty and my fellow Ph.D. students for maintaining a rich and open academic climate during my years at Michigan State University. The fondest memories of my graduate work will be those of the many hours spent in the interchange of ideas with these talented individuals.

Financial support for this dissertation was provided by a fellowship from the Deloitte, Haskins & Sells Foundation. I also wish to thank Dr. Carl Morris for his helpful comments during the formative stages of this work.

Finally, I take great pride in acknowledging the role of my family in these endeavors. The lifelong efforts of my parents have established a heritage of dedication, self sacrifice, scholarship and integrity for their children to inherit. Undoubtedly, my greatest words of appreciation are reserved for my wife, April. While I may never fully comprehend the depth of her sacrifice and commitment, until such time that I am able to repay with equal measure I simply say, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The objective of an audit is the expression of an opinion by an independent auditor regarding the fair presentation of a set of financial statements. The formulation of this single opinion requires the auditor to integrate the results from a series of subsidiary tests and analyses of the reported account balances and underlying accounting system. This integration is performed across at least two domains -- the various account balances and control procedures which must be tested and the various audit tests employed for each item.

Among the ten generally accepted auditing standards enumerated by the Statements on Auditing Standards (SAS), the fourth standard of reporting specifies that the auditor's "report shall either contain an expression of opinion regarding the financial statements, taken as a whole, or an assertion to the affect that an opinion cannot be expressed." [AICPA, 1985, Section 150.02] This standard requires the auditor to integrate the results of his compliance tests of the client's relevant internal control procedures and his substantive tests of each of the account balances comprising the financial statements into a single opinion regarding their fair presentation when "taken as a whole." This form of integration will be referred to as interitem integration.

The third standard of field work states that "sufficient competent evidential matter is to be obtained through inspection, observation, inquiries, and confirmations to afford a reasonable basis for an opinion regarding the financial statements under examination." [AICPA, 1985, Section 150.02] This wide variety of potential audit procedures (any number of which may apply to even a single account balance) creates a second domain over which results and opinions must be integrated. This form of integration will be referred to as intraitem or interprocedural integration.

The major purpose of this study is to propose the technique of empirical Bayesian estimation as an audit tool which is useful in efficiently integrating the results of certain audit procedures over either of these domains. These techniques are relevant to both major phases of the auditor's examination -- compliance and substantive testing. Specifically, the study will show how empirical Bayes estimation can be used to reduce the auditor's expected error when simultaneously estimating the company's compliance with a set of internal control procedures. Additionally, empirical Bayes estimation is proposed as a method for both interitem and interprocedural integration by efficiently combining the results of analytical review procedures and the direct tests of balances during the substantive testing of several related account balances. The behavior of empirical Bayes estimators in both of these circumstances will be examined by exact numerical computations and Monte Carlo simulations.

The remainder of this introduction discusses the general motivation behind the study of empirical Bayes estimation, and the specific research objectives of this dissertation.

1.1 Motivation

The need for the integration of audit evidence across items and across procedures is widely recognized in both the academic and professional literature.

SAS No. 1 indicates that auditing requires techiniques which can effectively pool the auditor's judgment and experience together with collateral and direct evidence to form a single informed opinion [AICPA, 1985, Section 330.09]:

> The amount and kinds of evidential matter required to support an informed opinion are matters for the auditor to determine in the exercise of his professional judgement after a careful study of the circumstances in the particular case. In making such decisions, he should consider the nature of the item under examination; the materiality of possible errors and irregularities; the degree of risk involved, which is dependent on the adequacy of internal control and susceptibility of the general item to conversion, manipulation, or misstatement; and the kinds and competence of evidential matter available.

As early as 1961 Mautz and Sharaf in their landmark treatise, The Philosophy of Auditing, recognized both forms of evidence integration [Mautz and Sharaf, 1961, pp. 29-30]:

> Having accepted the composite problem of a request for his opinion on financial statements ... he proceeds to divide the composite problem into a host of individual problems, each of which is related to the major issue ... Financial statements consist of a large number of individual assertions each of which becomes a problem or proposition to be tested by the auditor ...
>
> With his 'hypotheses' developed, the auditor sets out to put them to the test. This he does by selecting the audit techniques that apply to the given proposition and then determining the procedures by which the techniques will actually be applied ...
>
> Performance of the audit tests supplies the evidence ... Once the evidence is all in, he then evaluates it in respect to the financial statement propositions. With these judgments in hand, he proceeds to consider them all together and to arrive at a judgment on the composite problem of the reliability of the financial statements themselves. This last step is an important one, of course, and must be understood. In many processes of judgment, the preliminary judgments form something of a chain. A failure of any one of these negates the final conclusion. This is not so with an audit judgment. The final audit judgment is not so much like a chain as like a bundle of sticks. If one of them is weak or broken, it weakens the strength of the entire bundle, but it does not necessarily mean the bundle has no strength. Thus the auditor weighs the negative judgments he has made on individual propositions against the positive judgments, considering the relative importance of each one. This leads him to a final all-inclusive judgment.

While Mautz and Sharaf are lucid in their description of the decomposition of the overall audit objective and the subsequent integration of the results of the subsidiary tests, they fail to recognize the inherent complexity of these tasks and provide little guidance beyond the excercise of professional judgment for accomplishing this feat. The Committee on Basic Auditing Concepts [1973, pp. 38-39] described these operational complexities:

> Auditors have long recognized that independent (or corroborative) sources of evidences gathered on a proposition tend to increase its credibility. What generally has been overlooked is the incredible complexity of the statistical measurements involved ...
>
> Measuring (even roughly) the degree of credibility of the more general auditing assertions presents the auditor with a seemingly insurmountable obstacle ... The mental leap necessary to go from simple evidential propositions to these broad generalities is supported only by the vaguest `system of inference', so vague in fact that our use of the term in this context may be totally unwarranted given the state of auditing art.
>
> We present these problems not to answer them but to illustrate what the Committee considers to be the desirable direction of evolution in auditing and therefore major topics for research.

The design of a comprehensive and operational audit model which recognizes the need for the integration of evidence across each of the two domains is beyond the scope of this study. Two complex and highly conceptual attempts at such a model have been made by Park [1977] and Grimlund [1977]. The intent of this study is to present the technique of empirical Bayes estimation as an audit tool which is both practical given the current state of the auditing art as well as a step in the evolutionary process toward a more integrated audit model.

Theoretically, pure Bayesian estimation procedures could be used as a technique for either interitem or interprocedural integration. Over twenty years ago Birnberg [1964] proposed using Bayesian methods to integrate the results of classical audit procedures with the auditor's ex ante subjective evaluations. However, in reviewing some of the operational difficulties surrounding its use he noted: "The auditor has not yet reached the point of an expressed willingness to perform the full Bayesian analysis. Perhaps this is something for the future." [Birnberg, 1964, p. 114] Apparently, the time has not yet arrived for the practical implementation of Bayesian analysis since Bailey [1981, p. 241] reflects "to the author's knowledge, Bayesian models have yet to be applied in field settings."

The major operational difficulty impeding the adoption of Bayesian methods is the requirement that auditors be able to accurately specify their ex ante beliefs in the form of a subjectively determined prior probability distribution. Section 2.2.4 reviews the literature chronicling the auditor's inability to make these subjective prior probability specifications. However, as shown in Section 2.3, empirical Bayes estimation avoids this major operational drawback by not requiring the auditor to formally specify these prior beliefs. A major purpose of this dissertation is to show that empirical Bayes methods perform nearly as well as Bayesian methods in a wide variety of circumstances and in fact outperform them when the prior beliefs are sufficiently misspecified. Accordingly, empirical Bayes estimation is proposed as a more palatable audit tool than the methods of pure Bayesian estimation which have been considered in the past.

1.2 Specific Research Objectives

There are two major research objectives of this dissertation. The first is to introduce empirical Bayes estimation to the auditing literature as an objective and efficient method of integrating auditor judgments and direct evidence both within and across the items at interest to the auditor. Chapter II reviews classical statistical estimation and Bayesian estimation, contrasting them both with empirical Bayes estimation. The manner in which Bayes or empirical Bayes procedures can assist in the integration of audit tests is explained.

The second major objective is to examine the behavior of various empirical Bayes estimators in the specific contexts of audit sampling and estimation with analytical review procedures. Chapters III and IV present a series of investigations comparing the behavior of empirical Bayes estimators with that of various traditional classical statistical sampling and analytical review estimators, pure Bayesian estimators and other multivariate estimators which have been proposed in the literature. Investigations are made of the mean squared error, bias and reliability of each of the estimators. Chapter III investigates applications of these estimators during the internal control compliance testing phase of the auditor's examination. Chapter IV investigates applications during the auditor's substantive testing of the reported general ledger account balances.

CHAPTER II

CLASSICAL, BAYESIAN AND EMPIRICAL BAYESIAN ESTIMATION

An essential component of auditing is the estimation of unknown quantities. The items at interest range from the rate of compliance (or its complement, the error rate) with the prescribed system of internal accounting controls to the total dollar amount of errors in a reported general ledger balance. The following three sections review and compare three alternative methods for obtaining these estimates. These are classical, Bayesian and empirical Bayesian estimation procedures, respectively.

2.1 Classical Estimation Procedures

Two classical estimation methods currently used by auditors are traditional statistical sampling techniques and certain analytical review procedures.

SAS No. 39, Audit Sampling [AICPA 1985], together with the AICPA audit guide, Audit Sampling [AICPA 1983], govern the use of audit procedures designed to estimate an unknown population parameter from a sample of less than 100% of the population. While both SAS No. 39 and the audit guide discuss both statistical and nonstatistical sampling, only the former is considered in this dissertation. The audit guide [AICPA, 1983, p. 130] defines statistical sampling as "audit sampling that uses the laws of probability for selecting and evaluating a sample from a population for the purpose of reaching a conclusion about the population."

Cochrane [1977] provides a good reference for the general theory of classical statistical sampling techniques. Several texts have been written on the specific techniques of classical statistical sampling as they apply to auditing. Examples include Roberts [1978], Arens and Loebbecke [1981], and Bailey [1981]. In each of these texts as well as in the authoritative literature a distinction is made between attributes and variables sampling.

Attributes sampling is defined as "statistical sampling that reaches a conclusion about a population in terms of a rate of occurrence." [AICPA, 1983, p. 127] The most common example is sampling to determine the rate of noncompliance (error rate) with a prescribed internal accounting control. Typically a random sample is selected from the population of all transactions for which the internal control procedure is applicable. The estimated error rate for the population is simply the proportion of sample items for which noncompliance was established by the auditor. Confidence intervals for the population error rate can be constructed using exact binomial distribution theory or asymptotic normal distribution theory.

Variables sampling is used by the auditor during the direct substantive testing of account balances. It is defined as "statistical sampling that reaches a conclusion on the monetary amounts of a population." [AICPA, 1983, p. 130] Three major classical variables sampling estimators are typically mentioned in the literature. These are the mean-per-unit, difference and ratio estimators.

For notational purposes define:

Y = Recorded book value of the population
X = True but unknown balance of the population
N = Number of items in the population
n = Number of items in the random sample
y = Recorded book value for the sample items
x = True or audited value for the sample items

Using the above notation the three major classical variables sampling estimators can be written as:

$$\text{Mean-per-unit estimator} = N\frac{x}{n}$$

$$\text{Difference estimator} = Y + N\frac{x - y}{n}$$

$$\text{Ratio estimator} = Y\frac{x}{y}$$

Theoretically, all of the above estimators can be used with either simple random sampling or stratified random sampling. Simple random sampling implies that every item in the population has an equal probability, n/N, of being included in the sample. Under stratified random sampling the population is divided into a set of relatively homogenous groups or strata. Within each stratum simple random samples of various sizes are selected and the within strata estimator is calculated. The results are combined across the strata to arrive at the single point estimate for the population.

Confidence intervals around the point estimate for any of the classical variables sampling estimators are constructed using normal distribution theory. An estimation technique which calculates an upper bound on the true account balance without first constructing a point estimate accompanied by a confidence interval based on normal distribution theory is the combined attributes-variables (CAV) or dollar-unit sampling (DUS) method. Under CAV the probability of a population item being included in the sample is directly proportional to its size. The upper bound on the true account balance (often referred to as the Stringer bound) is calculated in a two step procedure. The first step involves determining an upper bound on the number of errors in the account based upon the number of errors observed in the sample using classical attributes sampling techniques. A conservative estimate for the dollar amount associated with this upper bound on the number of errors is then made.

CAV sampling has been shown to be a useful technique in certain limited situations where the expected error rate is low and nearly all of the errors are overstatements. However, it is not a classical sampling technique in the sense of providing a point and interval estimate useful for making audit adjustments [Arens and Loebbecke, 1981, p. 353]. Since CAV sampling does not produce point and interval estimates based on normal probability distribution theory, it is not comparable with the Bayes and empirical Bayes estimation techniques discussed in the subsequent sections. The remainder of this study will apply only to classical statistical sampling techniques.

In addition to classical statistical sampling methods auditors may use a second set of classical estimation techniques known as analytical review procedures. The use of these techniques is governed by SAS No. 23, Analytical Review Procedures. SAS No. 23 identifies three potential uses of analytical review procedures [AICPA, 1985, Section 318.05]:

> a. In the initial planning stages to assist in determining the nature, extent, and timing of other auditing procedures by identifying, among other things, significant matters that require consideration during the exam.
>
> b. During the conduct of the examination in conjunction with other procedures applied by the auditor to individual elements of financial information.
>
> c. At or near the conclusion of the examination as an overall review of the financial information.

It is the second of these three applications which is most relevant to this study. Both regression and ARIMA time series models have been proposed in various levels of detail in the academic literature as techniques for estimating the correct balance of general ledger accounts during substantive audit testing. Examples include Stringer [1975], Albrecht and McKeown [1977], Kaplan [1978], Kinney

[1978], Akresh and Wallace [1980], Neter [1980] and Lev [1980]. These models use past and present observations of variables for the company, its industry or the general economy to estimate the correct balance in an account. Normal distribution theory is used to represent the model's error term(s) so that both confidence intervals and point estimates can be obtained. Bailey [1981, Chapter 10] is a good introduction to these models.

Despite their many differences, each of the classical statistical sampling and analytical review estimation procedures have one characteristic in common. These classical methods result in an estimate which depends only on the sample observations used in the construction of the estimator. In particular, the estimate is not affected by the auditor's prior beliefs about the value for the unknown parameter. These beliefs may naturally arise as a result of the auditor's experience with similar estimation problems on prior engagements or from the results of collateral tests performed during the current engagement.

This is not to say that an application of classical estimation procedures can be done entirely without the use of the auditor's judgment. Subjective evaluations are necessary, for example, to select the estimation procedure which is most likely to provide an efficient and reliable estimate for the parameter at interest. Judgment is required in the selection of the variables to be included in ARIMA and regression models and the selection of the model's functional form. However, once these preliminary judgments have been made the resulting estimate will depend only on the sample observations. Mixing of the auditor's subjectively determined prior beliefs, the results of collateral tests and the results of the various classical estimation procedures into a single estimate of the unknown parameter must be done informally, if at all. One of the objectives of Bayesian estimation procedures is to provide a formal mechanism for the integration of these subjective and objective evaluations into a single estimate.

2.2 Bayesian Estimation Procedures

Tracy [1969, p. 41] stated rather precisely the fundamental difference between Bayesian and classical estimation procedures:

> The classical approach looks at the sample results -- and only the sample results -- to draw an inference about the population test area. Any other audit evidence that may have a bearing on the test area is ignored.
>
> In many cases the auditor may have already gathered evidence by other audit procedures that is relevant to the test area. The Bayesian method incorporates such "collateral" evidence into the statistical evaluation of the sample results. Compared to the classical method, the Bayesian method can yield significantly different interpretations that in many cases would allow optimal allocation of audit effort. The auditor could have the same degree of confidence with a smaller sample size or a greater degree of confidence with the same sample size.

At the core of the Bayesian technique is the assumption that the auditor can assemble his experience together with the results of all relevant collateral tests to form a single subjective evaluation regarding the unknown parameter at interest. This subjective evaluation is in the form of a probability distribution which assigns a prior probability to every possible value for the parameter. As an example, suppose the auditor's experience and the results of his collateral tests are such that he believes an account with a reported balance of $100,000 contains errors which when accumulated may range from a $5,000 net understatement to a $10,000 net overstatement with all possibilities equally likely. In the Bayesian formulation this belief would be stated in the form of a "prior distribution", g, on the unknown correct balance, θ , of the following form:

$$g(\theta) = 1/\ 15000 \quad \text{for } \theta \in (\ 90000,\ 105000)$$
$$= 0 \text{ otherwise}$$

Given the actual value for the parameter, θ , the classical estimator, $\hat{\theta}$, follows some conditional probability distribution, $f(\hat{\theta}|\theta)$. This is, of course, the same distribution which would be used for establishing confidence intervals for the unknown parameter using the classical estimation techniques.

Bayes theorem can be used to construct the conditional posterior distribution, $g(\theta|\hat{\theta})$, for the parameter at interest given the results of the classical estimation procedure. Bayes rule is:

$$g(\theta|\hat{\theta}) = \frac{g(\theta)\ f(\hat{\theta}|\theta)}{\int g(\theta)\ f(\hat{\theta}|\theta)\ d\theta} = \frac{g(\theta)\ f(\hat{\theta}|\theta)}{f(\hat{\theta})} .$$

A second major assumption in Bayesian estimation is that the estimator should be constructed in order to minimize the conditional expected value of the user's loss function, L. The loss function measures the penalty to the user of estimating the true θ by some $\tilde{\theta}$. A standard assumption of a quadratic loss function is often made so that

$$L(\theta,\tilde{\theta}) = c(\theta - \tilde{\theta})^2 \text{ for c a constant.}$$

Scott [1975] provides some evidence that a quadratic loss function is a reasonable choice in the context of the audit attest function. Section 3.3.2 considers a more general asymmetric loss function.

The Bayes estimate, $\hat{\theta}^B$, is that estimate which minimizes the conditional expected loss. Thus, $\hat{\theta}^B$ is chosen in order to minimize $\int L(\theta,\tilde{\theta})\ g(\theta|\hat{\theta})\ d\theta$.

For a quadratic loss function it is easy to show that the Bayes estimate is the mean of the conditional posterior distribution of θ. The first order condition for the minimization of the expected loss is

$$\frac{\partial}{\partial\theta} \int c\ (\theta - \tilde{\theta})^2 g(\theta|\hat{\theta})\ d\theta = 0$$

or

$$2c \int (\theta - \tilde{\theta})\ g(\theta|\hat{\theta})\ d\theta = 0.$$

This condition is satisfied for $\hat{\theta}^B = E\{\theta|\hat{\theta}\}$, the mean of the conditional posterior distribution. Notice that the Bayes estimate is independent of the constant, c, in the loss function.

The following three sections illustrate three potential uses of Bayesian estimation. These are, respectively, the integration of a classical estimate of an unknown parameter with the auditor's subjective judgment about its value, with other classical estimation procedures (interprocedural integration), and with estimates of other related parameters (interitem integration).

2.2.1 Bayesian Techniques for Integrating Classical Estimates with Subjective Evaluations

The traditional use of pure Bayesian estimation assumes the auditor can subjectively consolidate his prior beliefs together with any collateral evidence to form a prior probability distribution on the unknown parameter. As an example, suppose the auditor specifies his prior beliefs as

$$\theta \sim g(\theta) = \text{Normal } (\mu, \tau)$$

Under pure Bayesian estimation values for μ and τ must be specified by the auditor. A classical estimation procedure (such as a regression model or a classical statistical sampling routine) is then employed to yield an estimate, $\hat{\theta}$, whose conditional distribution follows some known form, typically:

$$\hat{\theta}|\theta \sim f(\hat{\theta}|\theta) = \text{Normal}\ (\theta,\sigma^2)$$

Generally, the variance of the classical estimator, σ^2, is not known. However, a consistent estimate is available ex post based upon classical sampling or regression theory.

Under these circumstances the posterior distribution for θ conditional on the classical estimate $\hat{\theta}$ is:

$$\theta|\hat{\theta} \sim g(\theta|\hat{\theta}) = \text{Normal}\ \left(\frac{\mu\sigma^2 + \hat{\theta}\tau}{\tau + \sigma^2}, \frac{\tau\sigma^2}{\tau + \sigma^2} \right)$$

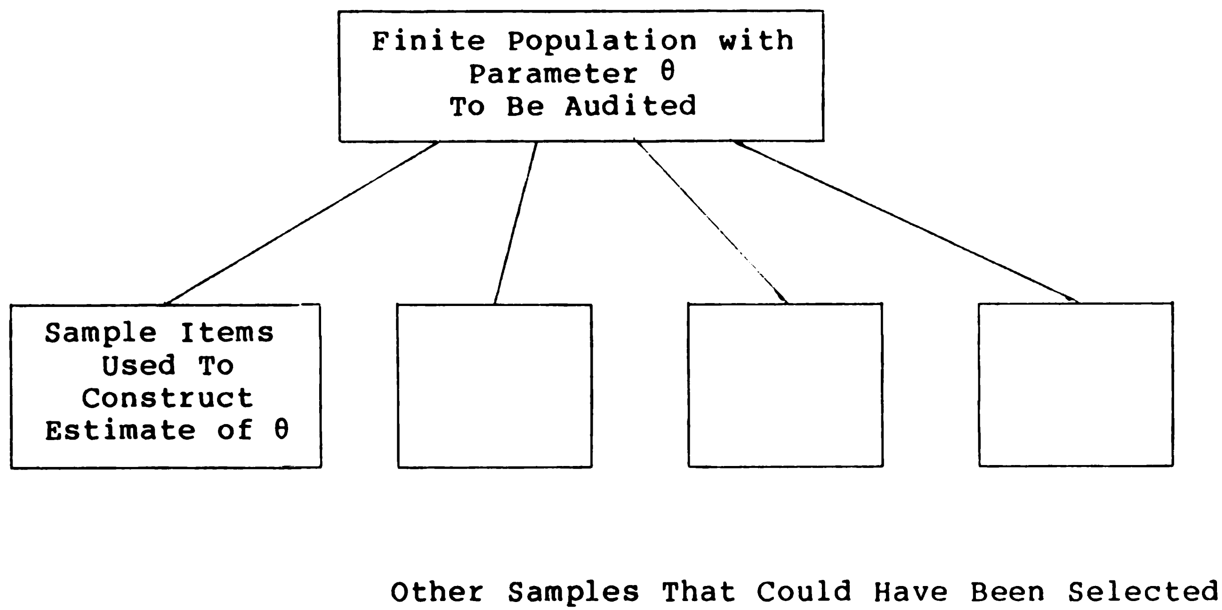
If the auditor's loss function is quadratic, then the Bayes estimate, $\hat{\theta}^B$, is the expectation of the conditional posterior distribution:

$$\hat{\theta}^B = E\ \{\ \theta|\hat{\theta}\ \} = \frac{\sigma^2}{\tau + \sigma^2}\mu + \frac{\tau}{\tau + \sigma^2}\hat{\theta}$$

Since μ and τ are specified in the auditor's subjective prior and since σ^2 is either known or consistently estimated from the data, the realization of the classical estimator results in a specific value for the Bayes estimate.

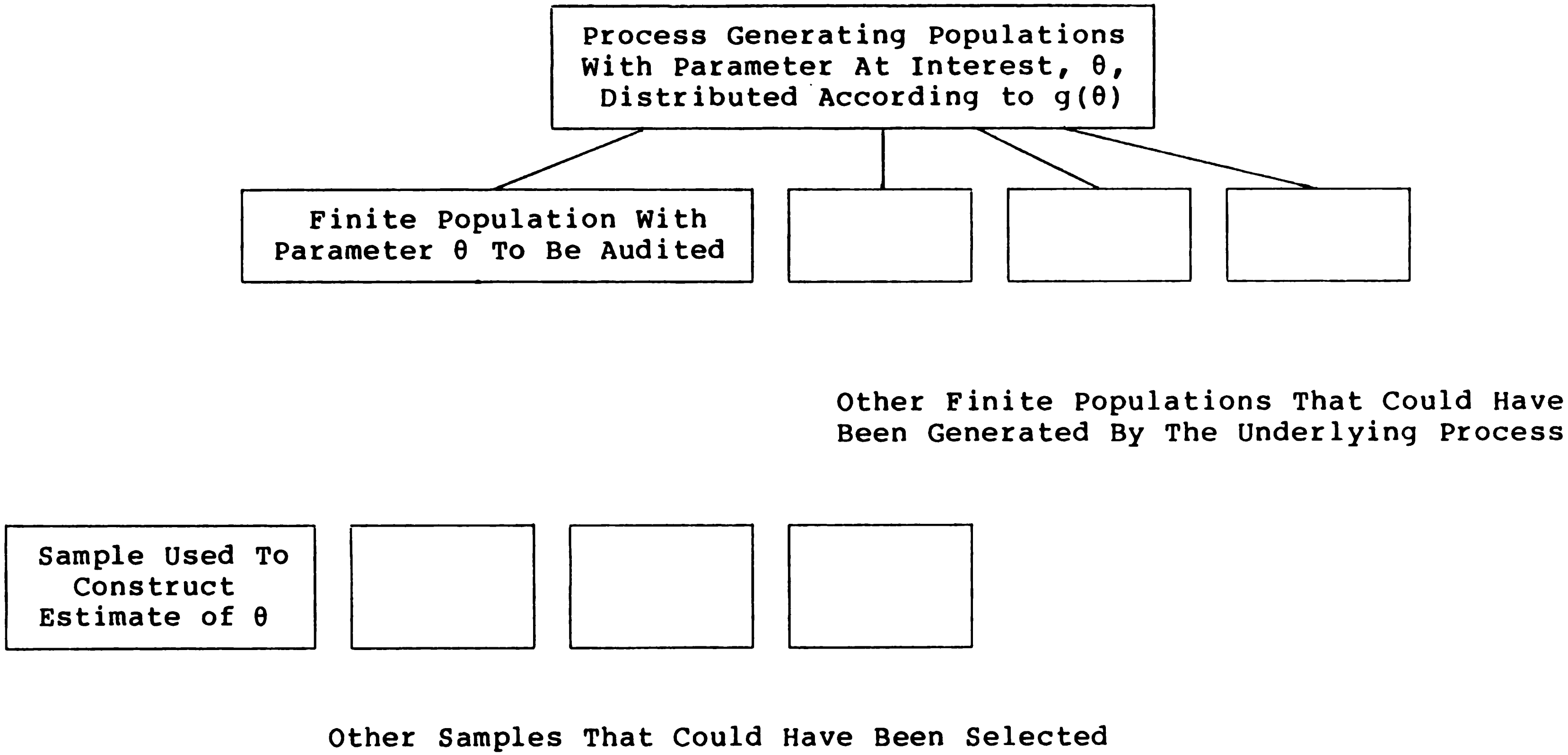
Figures 1 and 2 (modifications of Godfrey and Andrews [1982, p. 307, Figure 1]) give a schematic view of the difference between the Bayesian and classical viewpoints. Figure 1 presents the sampling problem from the classical or frequentist perspective. The population at interest is viewed as fixed. The only source of variation or risk in the estimation procedure arises from variability in the samples which could be selected from the population.

From the Bayesian perspective (Figure 2) two sources of variation are present. The population at interest is generated from some underlying process. The actual



Frequentist View of Estimation

FIGURE 1



Bayesian View of Estimation

FIGURE 2

population which the auditor must test is but one of many possible populations which could have been generated by the underlying process. Thus, Ramage, Krieger and Spero [1979, p. 73] argue that "each audit 'population' is actually a point sample from a stochastic process which produces a set of book values potentially observable at any instant in time." The auditor uses his experience and collateral evidence to specify his prior belief regarding the probability distribution generating the various possible populations. Given the realization of any one population, the second source of variation from the sampling scheme is, of course, still present.

Define the risk, r, of an estimator as the expected value of the loss function. The risk of the classical estimator, $r(\hat{\theta})$, under a quadratic loss function is, of course, simply the variance of the conditional sampling distribution, σ^2, as shown below:

$$r(\hat{\theta}) = \int\int (\theta - \hat{\theta})^2 \; f(\hat{\theta}|\theta) \; g(\theta) \; d\theta \; d\hat{\theta}$$
$$= \int \sigma^2 \; g(\theta) \; d\theta$$
$$= \sigma^2$$

The risk of the Bayes estimator, $r(\hat{\theta}^B)$, is

$$r(\hat{\theta}^B) = \int\int (\hat{\theta}^B - \theta)^2 \; f(\hat{\theta}|\theta) \; g(\theta) \; d\theta \; d\hat{\theta}$$
$$= \int\int (\hat{\theta}^B - \theta)^2 \; g(\theta|\hat{\theta}) \; f(\hat{\theta}) \; d\hat{\theta} \; d\theta$$
$$= \int \frac{\tau\sigma^2}{\tau + \sigma^2} f(\hat{\theta}) \; d\hat{\theta}$$
$$= \frac{\tau\sigma^2}{\tau + \sigma^2}$$

Thus, the Bayes estimator, by exploiting the information contained in the prior distribution, results in smaller risk than the classical estimator since

$$r(\hat{\theta}^B) = \frac{\tau\sigma^2}{\tau + \sigma^2} \leq \sigma^2 = r(\hat{\theta}).$$

It is in this sense of smaller risk or expected loss that Bayesian estimation is efficient relative to classical estimation.

This efficiency, however, depends upon the ability of the auditor to specify a prior belief which is an accurate representation of the underlying process generating the parameter at interest. Inaccuracies in the specification of the parameters of the underlying distribution can lead to Bayes estimates which are inefficient relative to classical estimates. Suppose, for example, the auditor's belief about the underlying process is such that he specifies a normal distribution with mean μ and variance τ. If this subjective specification is accurate with the single exception that the actual mean of the prior distribution is μ' and not μ, then the actual risk of the Bayes estimator is

$$r(\hat{\theta}^B) = \iint (\hat{\theta}^B - \theta)^2 \, g(\theta|\hat{\theta}) \, f(\hat{\theta}) \, d\hat{\theta} \, d\theta$$

$$= \iint \left(\mu \frac{\sigma^2}{\tau + \sigma^2} + \hat{\theta} \frac{\tau}{\tau + \sigma^2} - \theta\right)^2 g(\theta|\hat{\theta}) \, f(\hat{\theta}) \, d\hat{\theta} \, d\theta$$

$$= \iint \left(\mu \frac{\sigma^2}{\tau + \sigma^2} - \mu' \frac{\sigma^2}{\tau + \sigma^2} + \mu' \frac{\sigma^2}{\tau + \sigma^2} + \hat{\theta} \frac{\tau}{\tau + \sigma^2} - \theta\right)^2 \cdot g(\theta|\hat{\theta}) \, f(\hat{\theta}) \, d\hat{\theta} \, d\theta$$

$$= \left[(\mu - \mu') \frac{\sigma^2}{\tau + \sigma^2} \right]^2 + \frac{\tau\sigma^2}{\tau + \sigma^2}$$

In this example, the cost of mispecification is an addition of $[(\mu - \mu') \frac{\sigma^2}{\tau + \sigma^2}]^2$ to the risk of the perfectly specified Bayes estimate. If the prior belief is sufficiently mispecified (i.e., $|\mu - \mu'| > \tau + \sigma^2$), then the benefits from Bayesian estimation are eroded to the extent that it is inefficient relative to classical estimation. For a more complete analysis of the additional risks posed by inaccurate prior probability specifications in audit sampling see Beck, et.al. [1985].

The ability of the auditor to accurately specify the parameters of the underlying distribution is a critical element and a potential weakness in the application of Bayesian methods to audit testing. Empirical Bayes methods avoid this potential difficulty by estimating the parameters of the underlying distribution ex post from the sample data itself instead of requiring their ex ante specification by the auditor. Empirical Bayes estimation is discussed in detail in Section 2.3

2.2.2 Bayesian Techniques for Interprocedural Integration

Bayesian estimation can be used as a technique for formally combining the results of an analytical review procedure and direct classical variables sampling procedures on a single account balance. As an example, suppose a multiple regression model is constructed in order to preliminarily estimate the correct balance, Y_o, of an account. This regression model would take the form:

$$Y_t = \boldsymbol{\beta} \mathbf{X}_t + \varepsilon_t$$

Where : Y_t = Audited balance for accounts in the sample

$\boldsymbol{\beta}$ = Regression coefficients

$\mathbf{X}_t$ = Set of explanatory variables such as other account balances, industry indices, etc.

ε_t = error term distributed i.i.d. Normal $(0,\tau)$

t = $1,\ldots,T$ = index for crossectional or intertemporal observations

The regression coefficients are estimated in the usual manner by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}.$$

The analytical review estimate, $\hat{Y}_o^{AR}$, for the current correct balance would be obtained by inserting the current values for the explanatory variables, $\mathbf{X}_o$, in the regression equation:

$$\hat{Y}_o^{AR} = \hat{\boldsymbol{\beta}} \mathbf{X}_o$$

The regression estimate is distributed as:

$$\hat{Y}_o^{AR} \sim g(\hat{Y}_o^{AR}) = \text{Normal } (Y_o,\ \tau[1 + \mathbf{X}_o(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_o'\,])$$

Let $\hat{Y}_o^S$ be a second independent estimate of the correct account balance obtained from classical statistical sampling procedures. Under classical sampling theory we take the distribution of $\hat{Y}_o^S$ to be:

$$\hat{Y}_o^S \sim f(Y_o^S) = \text{Normal } (Y_o,\ \sigma^2)$$

Interprocedural integration requires the auditor to obtain a single estimate, $\hat{Y}_o^I$, of the account balance based upon the analytical review estimate, $\hat{Y}_o^{AR}$, and the classical statistical sampling estimate, $\hat{Y}_o^{S}$. An obvious choice would be a linear weighting of the two:

$$\hat{Y}_o^{I} = B\,\hat{Y}_o^{AR} + (1 - B)\,\hat{Y}_o^{S}$$

for some $B \in (0,1)$.

Ideally, the linear weighting factor, B, would be chosen in order to minimize the expected loss function for the integrated estimator. Thus, under a quadratic loss function the objective function is:

$$\min_B \iint [\,B\,\hat{Y}_o^{AR} + (1 - B)\,\hat{Y}_o^{S} - Y_o)^2\, g(\hat{Y}_o^{AR})\, f(\hat{Y}_o^{S})\, d\hat{Y}_o^{AR}\, d\hat{Y}_o^{S}$$

$$\min_B \iint [\,B\,(\hat{Y}_o^{AR} - Y_o) + (1 - B)\,(\hat{Y}_o^{S} - Y_o)]^2 \cdot g(\hat{Y}_o^{AR})\, f(\hat{Y}_o^{S})\, d\hat{Y}_o^{AR}\, d\hat{Y}_o^{S}$$

By independence of the two estimators this is equivalent to:

$$\min_B \iint [\,B^2\,(\hat{Y}_o^{AR} - Y_o)^2 + (1 - B)^2\,(\hat{Y}_o^{S} - Y_o)^2\,]\; g(\hat{Y}_o^{AR})\, f(\hat{Y}_o^{S})\, d\hat{Y}_o^{AR}\, d\hat{Y}_o^{S}$$

$$= \min_B\ [\,B^2 \text{ Variance } (\hat{Y}_o^{AR}) + (1 - B)^2 \text{ Variance } (\hat{Y}_o^{S})\,]$$

The first order condition for the minimization of the above expression is:

$$2\,B \text{ Variance } (\hat{Y}_o^{AR}) - 2\,(1 - B) \text{ Variance } (\hat{Y}_o^{S}) = 0$$

or

$$B = \frac{\text{Variance } (\hat{Y}_o^{S})}{\text{Variance } (\hat{Y}_o^{AR}) + \text{Variance } (\hat{Y}_o^{S})}$$

$$= \frac{\sigma^2}{\sigma^2 + \tau\,[1 + \mathbf{x}_o\,(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{x}_o']}$$

While neither σ^2 nor τ are known, consistent estimates of them are available from the results of the regression and sampling procedures. These could be used to estimate the weighting factor, B, to form the single integrated estimator, $\hat{Y}_o^I$, which has smaller expected squared error than either of the initial estimators.

The relationship of this integrated estimator to a Bayes estimator is easily seen. Suppose the auditor uses the results of the analytical review procedure to form his prior belief on the correct level of the account balance. A logical choice for this prior belief would be

$$Y_o \sim \text{Normal} (\hat{Y}_o^{AR}, \ \tau [1 + \mathbf{X_o}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X_o}']).$$

Combining this prior distribution on Y_o with the classical sampling estimator, $\hat{Y}_o^S$, using Bayesian estimation procedures yields a Bayes estimator which is indentical to the integrated estimator, $\hat{Y}_o^I$, as derived above. Thus, by using analytical review procedures to form prior beliefs, Bayesian estimation can be viewed as an efficient mechanism for integrating the results of analytical review and classical statistical sampling procedures.

2.2.3 Bayesian Techniques for Interitem Integration

Suppose the auditor wishes to simultaneously estimate an unknown parameter, θ_i, for each of i = 1,...,k similar populations. As an example θ_i may represent the error rate for each of k attributes pertaining to a particular internal control system. Alternatively, θ_i may represent the error rate for the ith division of a company with k divisions each of which is being tested with respect to the same attribute. As a third example, θ_i may represent the correct balance in each of k similar accounts or subaccounts each of which is to be simultaneously estimated.

In simultaneously estimating k parameters the auditor faces an apparent dilemma. On the one hand he is forced to estimate each of the parameters individually. For example, if the auditor is testing compliance with internal controls, he must establish which controls or which divisions are not operating as prescribed in order to tailor his substantive audit procedures accordingly. On the other hand, classical estimation procedures ignore any commonality which may exist between the populations. Since each of the populations is generated within the same underlying internal control environment, it would appear that some characteristics should be shared by the populations. Furthermore, while the individual estimation of account balances and error rates must be made, the final audit opinion is with respect to the financial statements taken as a whole. This implies that the auditor's estimates require accuracy both individually as well as when consolidated across the various items estimated.

Bayesian estimation is a method by which individual estimates of each of the k parameters can be made while at the same time exploiting any underlying commonality between the populations in order to achieve greater consolidated or ensemble accuracy. Suppose the auditor believes the k parameters at interest are independently generated from the same normal distribution:

$$\theta_i \underset{iid}{\sim} \text{Normal } (\mu, \tau)$$

The classical estimation procedures will yield k independent estimates, $\hat{\theta}_1,\ldots,\hat{\theta}_k$, each with its own conditional distribution:

$$\hat{\theta}_i | \theta_i \sim \text{Normal } (\theta_i, \sigma_i^2)$$

Figure 3 presents a schematic representation of this Bayesian view of the simulataneous estimation problem.

The conditional posterior distributions on the unknown parameters are of the form:

$$\theta_i | \hat{\theta}_i \sim g(\theta_i | \hat{\theta}_i) = \text{Normal } \left(\frac{\mu\sigma_i^2 + \hat{\theta}_i\tau}{\tau + \sigma_i^2} , \frac{\tau\sigma_i^2}{\tau + \sigma_i^2} \right)$$

Let the auditor's composite or ensemble loss function be represented by a general sum of squared errors:

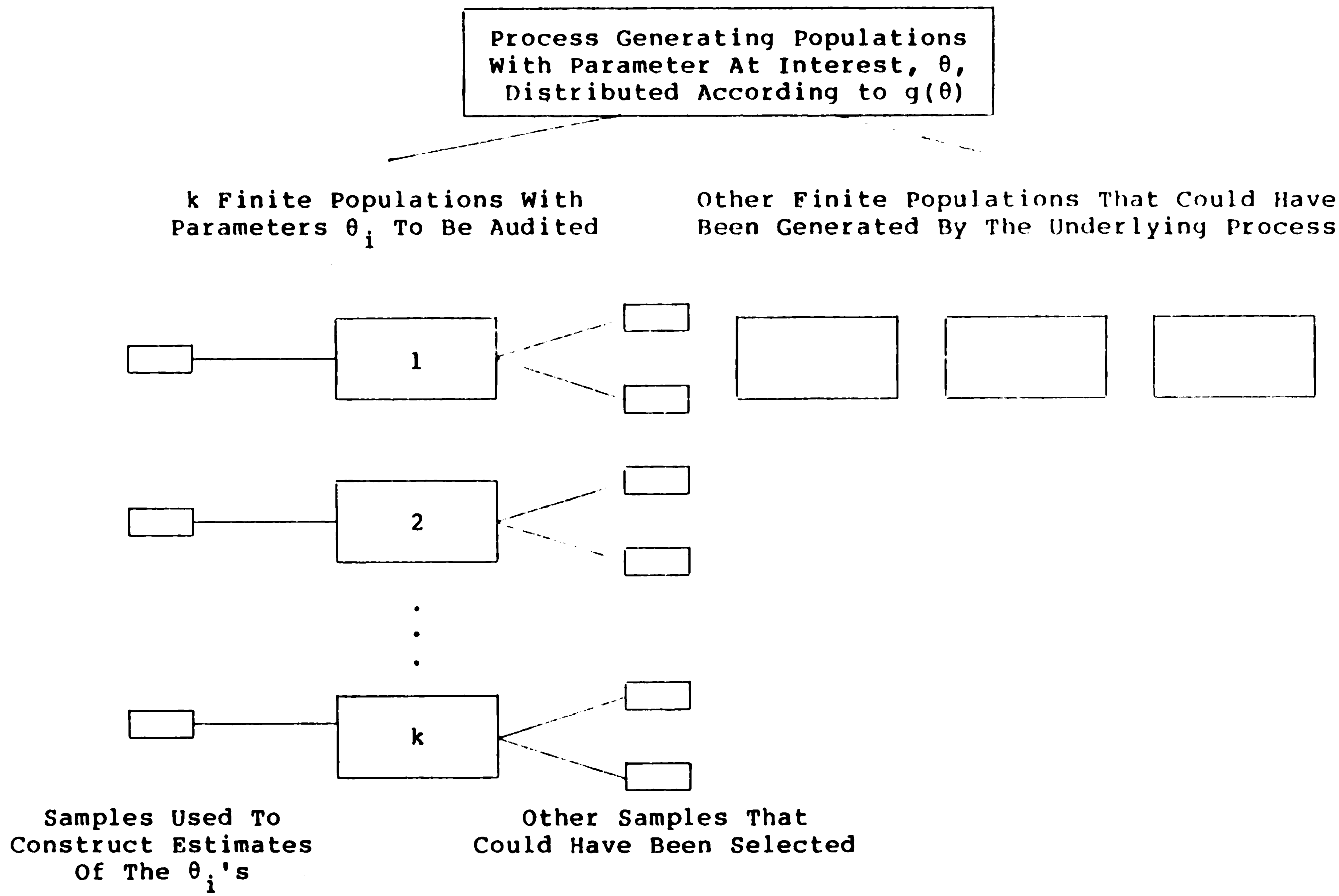
$$L(\boldsymbol{\theta},\tilde{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' \mathbf{C} (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})$$

Where: $\boldsymbol{\theta}$, $\tilde{\boldsymbol{\theta}}$ = k x 1 vectors composed of $\theta_1,\ldots,\theta_k$ and $\tilde{\theta}_1,\ldots,\tilde{\theta}_k$

$\mathbf{C}$ = k x k symmetric matrix

$\mathbf{C}$ represents an arbitrary k x k matrix of weighting factors representing the relative importance to the auditor of estimation errors in each of the k populations both when taken individually and in interaction with each of the other k-1 populations.

The simultaneous Bayes estimates, $\hat{\theta}_1^B,\ldots,\hat{\theta}_k^B$, are those which minimize the expected ensemble loss function



Bayesian View of Simultaneous Estimation

FIGURE 3

conditional on the observed results of the classical estimation procedure, $\hat{\theta}_1, \ldots, \hat{\theta}_k$. Thus $\hat{\boldsymbol{\theta}}^B$ minimizes

$$\int\int \cdots \int (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' \, \mathbf{C} \, (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) \prod_{i=1}^{k} g_i(\theta_i | \hat{\theta}_i) \, d\theta_1 \cdots d\theta_k .$$

The first order conditions for the minimization of the expected ensemble loss are

$$2 \int\int \cdots \int \sum_{j=1}^{k} c_{ij} \; (\theta_j - \tilde{\theta}_j) \quad \sum_{i=1}^{k} g_i(\theta_i | \hat{\theta}_i) \, d\theta_i \, \ldots \, d\theta_k = 0$$

for all i.

These conditions are satisfied by the estimator $\hat{\theta}_i^B = E\{\theta_i | \hat{\theta}_i\}$. Thus, the multivariate Bayes estimate which serves to minimize the expected loss integrated over the k populations is the expectation of each of the conditional posterior distributions, or

$$\hat{\theta}_i^B = \frac{\sigma_i^2}{\tau + \sigma_i^2} \, \mu + \frac{\tau}{\tau + \sigma_i^2} \, \hat{\theta}_i .$$

Notice that the risk minimizing Bayes estimates are the same regardless of the weighting matrix, $\mathbf{C}$, which assigns a measure for each population's relative importance to the auditor. That is to say, the auditor need not make these subjective "relative importance" evaluations since the resulting optimal estimator is independent of the weightings.

2.2.4 Review of the Relevant Literature

The purpose of this section is to review the literature surrounding the use of Bayesian analysis by auditors.

Crosby [1985] also provides a good introduction to this literature.

Birnberg [1964] first proposed the use of Bayesian statistics as an audit tool. Birnberg suggested in a general setting that auditors could use prior probability assessments and Bayesian analysis to reduce required sample sizes and to incorporate prior experience and qualitative judgments into the sampling plan. However, even in this initial work two operational problems in the use of Bayesian analysis were recognized which still remain as significant impediments for its practical application (Birnberg [1964], p.115]):

> The Bayesian approach is not free of operational problems. For the accountant, the following seem paramount:
>
> 1. The difficulty in assessing a prior personal probability distribution.
>
> 2. The measurement of utility in those decisions where the sums involved are large.

It is the first of these two operational difficulties which can be avoided by the use of the parametric empirical Bayes estimation procedures proposed by this dissertation.

Kraft [1968], Tracy [1969a] and Sorenson [1969] considered the use of Bayesian analysis in attributes sampling. They proposed that the auditor subjectively specify prior probabilities to a discrete set of possible error rates. The sample results and this discrete prior probability distribution are combined using Bayes rule to formulate Bayesian hypothesis tests of population error rates. Their results show that Bayesian analysis uses smaller sample sizes than classical sampling to obtain a given confidence level. Tracy [1969b] addresses the same issue from the perspective of constructing upper confidence limits for the unknown error rate rather than the equivalent hypothesis testing view.

Each of the Bayesian models in the aforementioned papers used a discrete subjective prior probability distribution over a finite set of potential population error rates. Felix and Grimlund [1977] generalize these models to a continuous prior distribution. Their model was developed for substantive testing of account balances resulting from the aggregation of a large number of subsidiary components. The model incorporated a beta distribution as the auditor's prior belief about the percentage of subsidiary components in error. Godfrey and Andrews [1982] revise the model to recognize that the auditor is selecting his sample from a finite population. Blocher [1981] investigates the sensitivity of the required sample size in statistical attributes sampling with Bayesian revision to changes in the parameters of the beta prior probability distribution and the auditor's desired reliability and precision levels.

The literature on the use of Bayesian analysis during substantive testing is somewhat more limited than that for compliance tests applications. As noted above, Felix and Grimlund [1977] propose one such model. They use a beta distribution as a prior for the percentage of account components in error and a normal distribution for the size of the error given that one exists. Scott [1973] uses a different sampling model in his application of Bayesian analysis to substantive testing. He models the auditor's sampling scheme as the selection of a set of days from throughout the year. For each sample day one hundred percent of the transactions are audited. He proposes a prior distribution on the true value of the day's change in net assets as:

$$\tilde{a} = \tilde{v}\,\tilde{f}$$

Where $\tilde{a}$ = change in net assets

$\tilde{v}$ = change in true gross book value in net assets

$\tilde{f}$ = adjustment factor or valuation coefficient for such items as the collectibility of accounts receivable

Scott models both $\tilde{v}$ and $\tilde{f}$ as random variables. The true gross book value, $\tilde{v}$, (as oppossed to the recorded book value) is random due the possibility of errors or fraud in the accounts. The valuation coefficient, $\tilde{f}$, is a random realization of the underlying stochastic process generating the various contingencies a business enterprise faces (collectibility of accounts, value of inventory, etc.). Scott assumes the prior underlying distributions for these two random variables each to be normal. This implies that their product, $\tilde{a}$, is _not_ normal. However, Scott assumes that $\tilde{a}$ can be approximated by a normal distribution with mean

$$E(\tilde{a}) = E(\tilde{v})\ E(\tilde{f})$$

and variance

$$Var(\tilde{a}) = Var(\tilde{v})[E(\tilde{f})]^2 + Var(\tilde{f})[E(\tilde{v})]^2 + Var(\tilde{v})\ Var(\tilde{f}).$$

Scott then proceeds to establish Bayesian estimates for the true net book values based upon the auditor's prior beliefs and the sampling results.

In addition to the statistical difficulties involved in Scott's assumption that the product of two normal random variables can be approximated as a normal random variable, there is also a significant operational weakness in his model. His sampling scheme assumes that auditors select a random sample of business days from the year and audit each transaction recorded during the sample days. Generally auditors do not sample in this manner. The typical audit sampling scheme defines the population at interest to be all transactions for the year or all subcomponents of a balance sheet account on a specific date. As Felix and Grimlund [1977] argue:

> For those enterprises with sufficient transaction volume to validate the multivariate central limit theorem assumption, it would probably be economically impossible to analyze more than a few days of accounting work. The sample evidence of such a small sample size usually would not significally alter the auditor's prior judgment.

Despite their various strengths and weaknesses all of the previous studies into the application of Bayesian techniques for audit testing rely on a critical assumption of the auditor's ability to make accurate prior probability judgments about the underlying parameter at interest. A large body of behavioral research has investigated the ability of individuals to make such prior probability assessments. The following reviews this literature within the context of prior probability assessments by auditors.

Winkler [1967] first proposed a set of techniques by which an individual's prior probability assessments can be elicited. These techniques can be categorized as:

Direct Methods

1. Cumulative distribution function (C.D.F.) or fractile method.

2. Probability density function (P.D.F.) method.

Indirect Methods

1. Equivalent prior sample (E.P.S.) information method.

2. Hypothetical future samples (H.F.S.) method.

The ability of auditors to make accurate prior probability assessments using one of these elicitation techniques has been studied using two alternative methodologies. Convergent validity studies investigate either the similarity of assessments among a group of auditors using the same elicitation technique or the consistency of auditor assessments from two or more techniques. A second methodology attempts to assess the accuracy of auditor assessments against an objective criteria. These "calibration" studies compare observed frequencies of an event with auditors subjective prior probability assessments. A review of the relevant research into the auditor's ability to make prior probability assessments follows. For a more complete review of prior probability elicitation techniques outside the auditing context see Chesley [1975].

Corless [1972] first studied the use of these elicitation techniques in applying Bayesian statistical methods in auditing. Corless examined the use of the C.D.F. and P.D.F. methods in eliciting prior probabilities about population error rates. His results showed that while the 88 auditors used as experimental subjects were willing to make prior probability assessments their beliefs were both widely divergent between the subjects and inconsistent across the two techniques. Felix [1976] performed a similar study comparing the E.P.S. and C.D.F. techniques using auditors who had undergone limited training. His results showed a smaller divergence between the E.P.S. and C.D.F. methods then Corless had found between the P.D.F. and C.D.F. techniques. However, it is hazardous to draw many inferences from the Felix study due to the small number of subjects (ten) used in his study.

Crosby [1980] investigated the relationship between two prior probability elicitation techniques (C.D.F. and E.P.S.) and the implied sample sizes required to obtain specified reliability and precision levels. The results showed that the differences between implied sample sizes for the two methods were statistically significant. Additionally, the implied sample sizes under the two elicitation techniques were significantly different from judgmental sample sizes which the subjects indicated they would use to achieve the desired reliability and precision levels. The implications of these results are:

1. The assessed prior probability distributions revealed by the elicitation techniques did not accurately reflect the auditors' true beliefs, or

2. The judgmentally determined sample sizes were inconsistent with the auditors' prior beliefs and the Bayesian probability revision model.

Crosby [1981] performed direct tests on the prior probability distributions elicited under the C.D.F. and E.P.S. techniques. The two distributions were tested for statistically significant differences among various measures of central tendency and dispersion. His results are summarized in the following table.

Distribution Characteristic	Results of Tests for Statistically Significant Difference
Median	No significant difference
Mean	No significant difference
Range of 25th to 75th percentile	Significant difference
Range of 5th to 95th percentile	No significant difference
Variance	No significant difference

Each of the above studies investigates auditors' abilities in making prior probability assessments regarding error rates. Soloman, et.al. [1982] studied the use of the C.D.F. elicitation technique in the auditor's assessment of account balances. Their results showed some evidence of consistency among auditors in the assessments of the inner fractiles of the prior distribution with much less consistency in elicited beliefs about the tails of the distribution. Soloman, et.al. [1984] investigated the use of the E.P.S. elicitation technique in circumstances similar to those in Soloman, et.al. [1982]. The results showed a much higher rate of inconsistency between the auditor subjects. Evidence of inconsistencies between E.P.S. elicitations and pure professional audit judgment was also noted.

The combined results of the previous studies are inconclusive with regard to auditors' abilities in making consistent and accurate prior probability assessments. The convergent validity methodology results in a joint test of auditors' abilities to make consistent subjective prior probability assessments and the ability of the elicitation techniques to capture these subjective beliefs. To the extent that these studies reveal a lack of consistency either across auditors or across elicitation techniques, the results should be troubling for those who would propose the use of pure Bayesian estimation in audit testing. Furthermore, even to the extent that the convergent validity of the methods is supported, without an objective criteria against which the accuracy of the assessments can be evaluated the ability of auditors to make both consistent and accurate assessments for Bayesian analysis remains in doubt.

The purpose of calibration studies is to provide such an objective standard for evaluating prior probability beliefs. Newman and Tomassini [1983] review the use of the calibration methodology in the context of investigating auditors' abilities in forming accurate prior probability distributions. Tomassini, et.al. [1982] examined the accuracy of auditor judgments about distributions of account balances. The results of their study showed "a mixture of well-calibrated and miscalibrated auditor judgments" (p. 398). While in general the study revealed a tendency for auditors to be less overconfident than typically observed in prior studies with nonauditor subjects, evidence was presented of prior probability distributions which were both underconfident (too diffuse) and overconfident (too tight). Beck, et.al. [1985] analyze the impact of miscalibrated prior probability distributions on audit risk and audit effectiveness.

The combined results of these calibration studies reveal that no single elicitation technique appears to produce consistent and accurate auditor prior probability assessments. It is not clear from the results whether auditors are unable to form consistent and accurate prior beliefs or whether the elicitation methods are inadequate techniques for revealing auditor beliefs. Without a technique for eliciting consistent and accurate prior beliefs, the practical implementation of pure Bayesian estimation for audit testing is impaired.

2.3 Empirical Bayes Estimation Procedures

As noted in the prior section, one of the major operational difficulties in the practical application of Bayesian estimation procedures for audit testing is the requirement that auditors subjectively specify ex ante their prior beliefs about the distribution of the parameter at interest. Specifically, auditors must assemble their experience and subjective evaluations together with collateral evidence from alternative tests and observations into a single probability distribution for the parameter to be estimated. Both a particular form for the distribution (e.g., normal) as well as the required identifying parameters of the distribution (e.g., the mean and variance) must be specified. The technique of empirical Bayes estimation provides the auditor relief from these tasks.

Empirical Bayes estimation procedures were first proposed by Robbins [1956, 1964]. A good introduction to this literature is given by Krutchkoff [1969] and Rutherford and Krutchkoff [1969]. A basic outline of the techniques of empirical Bayes estimation follows.

Consider the standard objective of estimating an unknown parameter, θ, with small equared error. Suppose a classical estimate, $\hat{\theta}$, is available which is distributed as

$f(\hat{\theta}|\theta)$. Suppose further that in the Bayesian framework θ itself follows some probability distribution. Unlike pure Bayesian estimation, under empirical Bayes estimation the distribution of θ, call it $g(\theta)$, is assumed to be entirely unknown. Note in particular that a diffuse or so-called "non-informative" prior is not used. Indeed, no distribution for θ is ever specified in the nonparametric empirical Bayes estimation process.

For empirical Bayes techniques to be applicable the estimation problem must be routine. That is, $\hat{\theta}_1$, is observed from $f(\hat{\theta}_1 | \theta_1)$ and θ_1 is to be estimated. Subsequently (or concurrently) a second estimation problem of the same type presents itself so that $\hat{\theta}_2$ is observed from $f(\hat{\theta}_2 | \theta_2)$ and θ_2 is to be estimated. This problem is repeated k times resulting in k classical estimates $\hat{\theta}_1,\ldots,\hat{\theta}_k$ for the k unknown parameters $\theta_1,\ldots,\theta_k$. The classical approach of estimating θ_i by $\hat{\theta}_i$ ignores both the underlying structure on the unknown parameters given by $g(\theta)$ and any information in the other k-1 observations of the classical estimator. The pure Bayesian approach requires the complete specification of the form and parameters of $g(\theta)$ and results in the mean of the posterior distribution of θ as the risk minimizing estimator:

$$\hat{\theta}_i^B = E\{\theta_i | \hat{\theta}_i\} = \int \theta_i \; g(\theta_i | \hat{\theta}_i) \; d\theta_i$$

The objective of empirical Bayes estimation is the generation of an estimate $\hat{\theta}_i^{EB}$ which depends only on the observed values $\hat{\theta}_1,\ldots,\hat{\theta}_k$ and not on any prior specification of $g(\theta)$ and yet nonetheless approaches the Bayes estimate as the number of estimation problems, k, increases. Under empirical Bayes estimation one can do as well (asymptotically) in reducing expected loss as the pure Bayesian, knowing nothing of the underlying marginal distribution of θ.

To illustrate the methods of empirical Bayes estimation it is useful to consider a specific yet widely applicable example. Suppose the classical estimates, $\hat{\theta}_i$, are normally distributed with unknown mean θ_i and variance σ_i^2. For notational convenience the subscript i is dropped. Thus,

$$\hat{\theta} \sim f(\hat{\theta}|\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp \{ -1/2 \frac{(\hat{\theta} - \theta)^2}{\sigma^2} \}.$$

Then,

$$\frac{\frac{\partial f(\hat{\theta}|\theta)}{\partial \theta}}{f(\hat{\theta}|\theta)} = - \hat{\theta}/\sigma^2 + \theta/\sigma^2$$

or, solving for θ,

$$\theta = \hat{\theta} + \sigma^2 \frac{\frac{\partial f(\hat{\theta}|\theta)}{\partial \theta}}{f(\hat{\theta}|\theta)}.$$

Consider the Bayes estimate if $g(\theta)$ and hence $g(\theta|\hat{\theta})$ were known:

$$\hat{\theta}^B = E\{\theta|\hat{\theta}\} = \int \theta \, g(\theta|\hat{\theta}) \, d\theta$$

$$= \int [\hat{\theta} + \sigma^2 \frac{\frac{\partial f(\hat{\theta}|\theta)}{\partial \theta}}{f(\hat{\theta}|\theta)}] \, g(\theta|\hat{\theta}) \, d\theta$$

$$= \theta + \sigma^2 \int \frac{\frac{\partial f(\hat{\theta}|\theta)}{\partial \theta}}{f(\hat{\theta}|\theta)} \, g(\theta|\hat{\theta}) \, d\theta$$

However

$$g(\theta|\hat{\theta}) = \frac{f(\hat{\theta}|\theta)\, g(\theta)}{f(\hat{\theta})}$$

so

$$E\,\{\theta|\hat{\theta}\} = \hat{\theta} + \sigma^2 \int \frac{\frac{\partial f(\hat{\theta}|\theta)}{\partial \hat{\theta}}}{f(\hat{\theta}|\theta)} \; \frac{f(\hat{\theta}|\theta)\; g(\theta)}{f(\hat{\theta})}\, d\theta$$

$$= \hat{\theta} + \frac{\sigma^2}{f(\hat{\theta})} \int \frac{\partial f(\hat{\theta}|\theta)}{\partial \hat{\theta}}\, g(\theta)\, d\theta$$

so that

$$\hat{\theta}^B = E\,\{\theta|\hat{\theta}\} = \hat{\theta} + \sigma^2 \frac{f'(\hat{\theta})}{f\,(\hat{\theta})} .$$

Written in this form the Bayes estimate can be viewed as the classical estimate plus a "correction factor." The correction factor is the variance of the classical estimator times the ratio of the derivative of the classical estimator's marginal density and the marginal density itself. The pure Bayesian specifies a belief for $g(\theta)$ and computes the marginal density accordingly (that is $f(\hat{\theta}) = \int f(\hat{\theta}|\theta)\, g(\theta)\, d\theta$).

The marginal density $f(\hat{\theta})$ remains unknown to the empirical Bayesian since the prior distribution on θ, namely $g(\theta)$, remains completely unspecified. However, the k observations on $\hat{\theta}$ can be used to empirically estimate their own marginal distribution and its derivative. Quite simply, the results of the k simultaneous or sequential estimation problems are used to infer via empirical methods the marginal distribution of $\hat{\theta}$ (and hence the underlying prior distribution on θ). Parzen [1962] gives methods for estimating the density function for a random variable given a finite number of observations on the variable. Clemmer and Krutchkoff [1968] is an early example of the use of these methods to generate empirical Bayes estimates of normally distributed coefficients in a linear regression model.

Theoretically, empirical Bayes estimation could be employed in a variety of auditing problems. Four examples are given in Table 1. The classical estimation procedures currently employed by auditors ignore the information contained in estimates obtained for prior years, other client divisions, or for similar accounts or internal control procedures. Empirical Bayes estimation objectively uses these prior or concurrent estimates to produce an estimator which empirically mimics a pure Bayes estimate with the prior distribution properly specified. However, empirical Bayes techniques do not require the auditor to subjectively specify any prior beliefs in the form of a prior probability distribution.

A major practical difficulty in the application of the empirical Bayes estimation procedures as described above results from the fact that much of their development is based on asymptotic theory. Hence, most of the properties of these "nonparametric" empirical Bayes procedures rely on a large number of prior or concurrent estimation problems of the same type. Such applications in auditing would be rare.

TABLE 1

Examples of Audit Applications of Empirical Bayes Estimation

Interpretation of θ_k	Interpretation of $\theta_1,\ldots,\theta_{k-1}$	Environmental Assumption
Current year's rate of noncompliance with a particular internal control procedure	Rate of noncompliance in the prior k-1 years	General control environment this year is similar to that of the prior k-1 years
Rate of noncompliance with internal control procedure k	Rate of noncompliance with internal control procedures 1,...,k-1	Error rates are similar for the k controls
Rate of noncompliance with a particular internal control procedure for the client's kth division or location	Rate of noncompliance for the other k-1 divisions or locations	Error rates are similar for the k-1 divisions or locations
Ratio of current year's correct balance to reported balance for a particular account	Ratio of correct to reported balance for the prior k-1 years	General control environment this year is similar to that of the prior k-1 years

Indeed it is not surprising that the number of estimation problems, k, must be large in order to infer from their results both the form and the identifying parameters of the prior distribution. However, of more practical interest to auditors is a class of empirical Bayes procedures designed specifically for instances where k is small. These are known as parametric empirical Bayes (PEB) estimation techniques. Morris [1983a] provides a good introduction to these methods.

PEB estimation is similar to pure Bayesian estimation in that a form for the prior distribution of the unknown parameters is specified. In the Bayesian manner the PEB point estimate is the expectation of the posterior distribution conditional on the observed classical estimate. PEB estimation differs from pure Bayesian estimation since only the form and not the identifying parameter values of the prior distribution is specified. Instead of prespecifying the values for the prior distribution parameters, they are estimated from the observed data itself. Table 2 summarizes the relationships between classical, pure Bayesian, nonparametric empirical Bayes and parametric empirical Bayes estimation.

As an example of PEB estimation assume the following prior and conditional sampling distributions:

$$\theta_i \sim \text{Normal } (\mu, \tau)$$

$$\hat{\theta}_i | \theta_i \sim \text{Normal } (\theta_i, \sigma^2) \quad i = 1, \ldots, k \text{ independently}$$

As shown in Section 2.2.1 the Bayes estimate for θ_i is

TABLE 2

Relationships Between Classical, Pure Bayes, Nonparametric Empirical Bayes, and Parametric Empirical Bayes Estimation

Estimation Technique	Form of Prior Distribution	Parameters of Prior Distribution
Classical	Ignored	Ignored
Pure Bayes	Subjectively Specified	Subjectibely Specified
Parametric Empirical	Subjectively Specified	Inferred from the data
Nonparametric Empirical Bayes	Inferred from the data	Inferred from the data

$$\hat{\theta}_i^B = B\,\mu + (1 - B)\;\hat{\theta}_i$$

$$\text{Where } B = \frac{\tau}{\tau + \sigma^2}$$

Under pure Bayesian estimation σ^2, if not known, is easily estimated from the data while the values for μ and τ are prespecified so that the Bayes estimate results. Under PEB estimation values for μ and τ are not prespecified but must be estimated from the sample data as well.

Define

$$\bar{\hat{\theta}} = 1/k \sum_{i=1}^{k} \hat{\theta}_i .$$

$\bar{\hat{\theta}}$ is an unbiased estimate of μ derived strictly from the observed estimates $\hat{\theta}_1, \ldots, \hat{\theta}_k$.

Define

$$S = \sum_{i=1}^{k} (\hat{\theta}_i - \bar{\hat{\theta}})^2 .$$

S is distributed as $(\sigma^2 + \tau)\chi^2_{(k-1)}$, where $\chi^2_{(k-1)}$ is a chi-square random variable with k-1 degrees of freedom. Since $E\{1/\chi^2_{(k-1)}\} = 1/(k-3)$, it follows that an unbiased estimate of $1/(\sigma^2 + \tau)$ is $(k-3)/S$.

Substituting these unbiased estimates for the unspecified parameters in the formula for the Bayes estimate yields the parametric empirical Bayes estimator:

$$\hat{\theta}_i^{PEB} = \hat{B}^{PEB}\,\bar{\hat{\theta}} + (1 - \hat{B}^{PEB})\;\hat{\theta}_i$$

$$\text{Where } \hat{B}^{PEB} = (k-3)\,\sigma^2/\,S$$

If σ^2 is not known, it is easily estimated in the usual manner by within group sums of squares.

Efron and Morris [1972, 1973] show that the risk or expected squared error of the PEB estimator is given by

$$r(\hat{\theta}_i^{PEB}) = \frac{\tau\sigma^2}{\tau + \sigma^2} + 3/k \frac{(\sigma^2)^2}{\tau + \sigma^2} \quad .$$

PEB estimation is efficient relative to classical estimation since

$$r(\hat{\theta}_i^{PEB}) = \frac{\tau\sigma^2}{\tau + \sigma^2} + 3/k \frac{(\sigma^2)^2}{\tau + \sigma^2} \leq \sigma^2 = r(\hat{\theta}_i) .$$

The price for estimating the parameters of the prior distribution by the data is an increase in risk over the Bayes risk when the parameters are specified correctly. The increase is $3/k \dfrac{(\sigma^2)^2}{\sigma^2 + \tau}$ which tends to zero as the number of populations, k, increases.

In Section 2.2.1 it was demonstrated that misspecification of the identifying parameters of the prior distribution (μ and τ) adds risk to the pure Bayes estimator, perhaps to the extent of inefficiency relative to classical estimation. PEB estimates avoid this hazard by objectively estimating these parameters from the data itself. However, as Table 2 shows, PEB estimation still requires the user to subjectively specify the form of the prior distribution. The question arises as to whether particular forms for the prior distribution are more robust than others to protect the user against a cost for functional form misspecification. The answer appears to lie in the use of natural conjugate priors.

The natural conjugate prior is the distribution which, given the conditional sampling distribution, $f(\hat{\theta}|\theta)$, will result in the prior, $g(\theta)$, and posterior, $g(\theta|\hat{\theta})$, distributions having the same functional form. In the example above the use of Normal (μ,τ) for the prior is an illustration of a natural conjugate prior since the posterior distribution (Normal $(\frac{\mu\sigma^2 + \hat{\theta}\tau}{\tau + \sigma^2}, \frac{\tau\sigma^2}{\tau + \sigma^2})$) is of the same functional form. Jackson, et. al. [1970] and Morris [1983b] prove the following important robustness property of natural conjugate priors for sampling distributions which are members of the natural exponential family with quadratic variance function (NEF-QVF). NEF-QVF is a family of probability distributions. There are six basic NEF-QVF distributions. These include the normal, binomial and Poisson distributions as special cases.

Let Π be the class of all possible priors on the unknown parameter θ with mean μ and variance τ. For any estimator t and prior distribution $\pi \varepsilon \Pi$ define the quadratic risk function as $r(\pi,t) = E\{(t-\mu)^2\}$. This is a double expectation over both the sampling distribution given θ and θ distributed according to π. Denote the Bayes estimator for a specified prior $\pi \varepsilon \Pi$ as t_π and denote the natural conjugate prior as π_o. If the sampling distribution is a member of NEF-QVF, then $r(\pi,t_{\pi_o}) = r(\pi_o,t_{\pi_o}) \leq r(\pi_o,t)$ for all possible priors $\pi \varepsilon \Pi$ and any estimator t.

As Morris [1983b, p. 525] concludes, the theorem "justifies using the conjugate prior in Bayes and empirical Bayes practice when one has little knowledge of the distribution of θ beyond its first two moments. In that case choosing $\pi \neq \pi_o$ can be risky because the statistician thinks his risk is $r(\pi, t_\pi) < r(\pi, t_{\pi_o})$ but it may actually be $r(\pi^*, t_\pi) > r(\pi^*, t_{\pi_o}) = r(\pi_o, t_{\pi_o})$ if some other $\pi^* \in \Pi$ obtains. Only the conjugate prior avoids this hazard."

Clearly, if the auditor knew the form of the prior distribution, he would be wise to use that information in constructing (empirical) Bayes estimates. However, in all practical circumstances the proper form will not be known. Guessing at any form other than the conjugate prior is dangerous since the resulting estimator may have unanticipated and unobservable additional risk if the form is misspecified. Only by using the conjugate prior is the auditor assured that his actual risk is as anticipated regardless of the actual form of the prior.

As a final remark it should be noted that the PEB estimator can be improved by placing an upper bound of one on the weighting factor $\hat{B}^{PEB}$. Since τ and σ^2 are variances of probability distributions, they are both nonnegative. This implies $B = \frac{\tau}{\tau + \sigma^2} \leq 1$. Since it is known a priori that $B \leq 1$, it follows that the PEB estimate is improved by estimating B by $\hat{B}^{PEB}$ equal to the minimum of 1 and

$(k-3)\sigma^2/S$. Unless otherwise noted subsequent references to the PEB estimator will refer to this bounded coefficient estimator.

Chapter III considers the specific uses and behavior of empirical Bayes estimates in the context of auditing tests of internal control compliance. Chapter IV considers their use in the context of substantive audit testing of account balances.

2.3.1 A Frequentist Interpretation of Empirical Bayes Estimators

The preceding formulation of the PEB estimator was made using the Bayesian view of the estimation of a set of parameters $\theta_1,\ldots,\theta_k$ which are themselves realizations of some fundamental random process. However, there exists a frequentist interpretation of the PEB estimator when the underlying parameters are taken to be unknown but fixed. The frequentist interpretation arises from the fact that the PEB estimator developed in Section 2.3 is the same as the James-Stein estimator for the means of k normal distributions.

Ijiri and Leitch [1980] introduced James-Stein (JS) estimators to the accounting literature. Efron and Morris [1977] give a good nontechnical introduction to these estimators. Stein [1956] considered the following estimation problem. Suppose one observes k independent normally distributed random variables $\hat{\theta}_1,\ldots,\hat{\theta}_k$. In the auditing context these might represent the results of any of the classical estimation procedures discussed in Section 2.1 for k different populations. Denote the means of these k random variables by $\theta_1,\ldots,\theta_k$ and the variance of each by σ^2. The maximum likelihood estimators (MLE) of these fixed but unknown means are, of course, simply the observed values $\hat{\theta}_1,\ldots,\hat{\theta}_k$ themselves.

However, Stein [1956] proved the rather surprising result that these normally distributed MLE's are inadmissable as estimates of the vector of means under a quadratic loss function when the number of populations exceeds two. Inadmissibility means that there exist some alternative estimators to $\hat{\theta}_1,\ldots,\hat{\theta}_k$, call them $\hat{\theta}^*_1,\ldots,\hat{\theta}^*_k$, which have smaller expected ensemble error. Thus,

$$E\left\{ \sum_{i=1}^{k} (\hat{\theta}^*_i - \theta_i)^2 \right\} \leq E\left\{ \sum_{i=1}^{k} (\hat{\theta}_i - \theta_i)^2 \right\}.$$

The above inequality is a frequentist assertion in the sense that the expectation is taken only over the sampling distributions of the estimators ($\hat{\theta}^*_i$ or $\hat{\theta}_i$) with the unknown parameters $\theta_1,\ldots,\theta_k$ fixed. Furthermore, the inequality holds for <u>any</u> fixed set of parameters $\theta_1,\ldots,\theta_k$.

James and Stein [1961], Lindley [1962], and Efron and Morris [1973] showed that the following James-Stein estimator, $\hat{\theta}^{JS}_i$, although itself inadmissible (i.e., dominated by some other estimator) nonetheless dominates the MLE estimator:

$$\hat{\theta}_i^{JS} = \hat{B}^{JS}\ \bar{\hat{\theta}} + (1 - \hat{B}^{JS})\ \hat{\theta}$$

$$\text{Where:}\quad \hat{B}^{JS} = \min\{1,\ (k-3)\sigma^2/S\}$$

$$\bar{\hat{\theta}} = 1/k \sum_{i=1}^{k} \hat{\theta}_i$$

$$S = \sum_{i=1}^{k} (\hat{\theta}_i - \bar{\hat{\theta}}_i)^2$$

This implies that for <u>any</u> set of fixed underlying parameters to be estimated, the expected ensemble squared error of $\hat{\boldsymbol{\theta}}^{\mathbf{JS}} = (\hat{\theta}_1^{JS},\ldots,\hat{\theta}_k^{JS})'$ is no greater than that for $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1,\ldots,\hat{\theta}_k)'$. Defining the frequentist risk, R, of an estimator to be its expected ensemble squared error we have:

$$R(\hat{\boldsymbol{\theta}}^{\mathbf{JS}},\boldsymbol{\theta}) = E\ \{\ \sum_{i=1}^{k}\ (\hat{\theta}_i^{JS} - \theta_i)^2\ \}$$

$$\leq E\ \{\ \sum_{i=1}^{k}\ (\hat{\theta}_i - \theta_i)^2\ \} = R(\hat{\boldsymbol{\theta}},\boldsymbol{\theta})$$

for <u>any</u> fixed set of unknown parameters $\boldsymbol{\theta}$.

The above expression of the ensemble efficiency of the JS estimator can be viewed as a frequentist interpretation of the PEB estimator since $\hat{\theta}_i^{JS}$ is equivalent to $\hat{\theta}_i^{PEB}$ of the prior section. There is, however, one important difference between the two interpretations. The frequentist property of the JS estimator holds only for ensemble risk and not for each population individually. Thus, among the k populations there may be a set of fixed parameters $\theta_1',\ldots,\theta_k'$ such that for at least one population, j, the James-Stein estimate has higher expected squared error or

$$E\ \{\ (\hat{\theta}_j^{JS} - \theta_j')^2\ \} > E\ \{\ (\hat{\theta}_j - \theta_j')^2\ \}.$$

The empirical Bayes interpretation of smaller expected squared error holds both for the ensemble and for each of the populations individually. Thus,

$$E \{ \sum_{i=1}^{k} (\hat{\theta}_i^{PEB} - \theta_i)^2 \} \leq E \{ \sum_{i=1}^{k} (\hat{\theta}_i - \theta_i)^2 \}$$

and

$$E \{ (\hat{\theta}_i^{PEB} - \theta_i)^2 \} \leq E \{ (\hat{\theta}_i - \theta_i)^2 \} \text{ for all } i.$$

This is since under the PEB view the expected value is taken over both the distribution of the estimator conditional on θ_i and the distribution of θ_i while under the frequentist view the expectation is taken only over the distribution of the estimator with θ_i fixed.

Ijiri and Leitch [1980] consider using JS estimates as a method for an auditor to limit ensemble risk over "(1) a client's set of accounts which comprise a set of financial statements, (2) similar accounts for all clients in his practice, and (3) similar accounts over time" [p. 93]. They consider applications for both attributes and variables sampling problems including "(1) estimates of error rates, (2) population estimates, (3) difference estimates, and (4) regression estimates" [p. 106].

Chapter III considers the use of PEB/JS estimators in the context of compliance testing for internal control error rates. The behavior of the estimator is examined both from the frequentist and the empirical Bayes view. Chapter IV considers the use of PEB estimators during substantive testing.

2.3.2 Review of the Relevant Literature

The prior two sections outlined the general development of the James-Stein or parametric empirical Bayes estimator. Subsequent chapters will examine the estimator's behavior in the specific context of statistical sampling for audit testing. The purpose of this section is to review the literature documenting the practical applications of the estimator in other fields.

Carter and Rolph [1974] used PEB procedures in the estimation of the probability that alarms from New York City fire alarm street boxes signaled serious structural fires. These probability estimates were employed as a component of a logistical model allocating additional equipment in an efficient manner toward those alarms which had the highest likelihood of signaling a serious fire. Fire alarm data from 1967-1969 were used to develop empirical Bayes estimates of underlying signal rates for certain fire alarm street boxes. The performance of the empirical Bayes and maximum likelihood estimators was evaluated by comparing the results of the equipment dispatch policy which would have arisen in 1970 under each of the two estimators. It is interesting to note that the loss function employed is not quadratic on the estimates of the probabilities themselves, but rather a function of the way these probability estimates would be used -- i.e., the results of the dispatch policy. Their results showed that using the empirical Bayes estimates in the dispatch policy model for 1970 could have produced a statistically significant reduction in serious structural fire underresponses while reducing the total number of alarm responses and holding the number of overresponses constant.

Morris and Van Slyke [1978] studied the use of PEB methods for the estimation of automobile insurance claim costs for 27 geographical underwriting territories over the period 1974-1976. Data from the first two years were used to estimate regional differences in claim costs for 1976 using both MLE and PEB methods. The results showed the PEB estimator reduced the 1976 ensemble squared error by 1.5% for property damage claims and 12.5% for bodily injury claims.

Fay and Herriott [1979] considered the estimation of per capita income for small cities. The objective was to provide more accurate estimates since federal revenue sharing allocations under the State and Local Fiscal Assistance Act of 1972 were based, in part, upon estimated per capita income.

Their sampling estimator was based upon the 1970 census sample results for each of the small communities. A second estimate was obtained using Internal Revenue Service and census bureau county data. This estimate was based on a regression of community per capita income on county per capita income, community and county housing values, and IRS adjusted gross income per exemption for the community and the county. PEB methods were used to construct a single estimate comprised of a linear weighting of the census sample and regression equation estimates. The goal was to construct a single PEB estimator which outperformed both of the existing estimators.

A test of their PEB estimator was provided by comparing the 1972 per capita income estimates for 24 communities subjected to a special 100% census in 1973. Their results, summarized below, confirm that the PEB estimator is efficient relative to both of the alternative estimators upon which it is based.

Estimate of Per Capita Income	Mean Squared Error	Mean Absolute Error
Census sampling estimator	$550,055	$589
Regression estimator	567,437	611
PEB estimator	299,654	430

Rubin [1980] employed PEB techniques in his law school admission standards validity study. Two factors which law schools use in admissions decisions are the applicant's score on the Law School Aptitude Test (LSAT) and the applicant's undergraduate grade point average (UGPA). The objective of the study was to develop a linear model predicting first year law school grade point average (FYGPA) based upon LSAT and UGPA in order to assist schools in their admissions decision.

Data from 82 law schools for LSAT, UGPA and FYGPA were obtained for the three years 1974-1976. For each school an ordinary least squares model of FYGPA on LSAT and UGPA was fit for both 1974 and 1975. Predictions for 1975 and 1976 FYGPA were made based upon the prior year's ordinary least squares (OLS) regression estimates. The PEB estimates of FYGPA were linear combinations of the OLS regression estimate and the grand mean of FYGPA for the 82 schools. Rubin concludes that the PEB estimates were superior since 57 of the 1975 and 49 of the 1974 FYGPA predictions had smaller mean squared prediction error than under the pure ordinary least squares model.

Hoadley [1981] was interested in constructing estimates of the percentage of production units which are defective for quality assurance purposes. In his model the production failure rate for the current period is estimated by an empirical Bayes combination of the current period's sample results and the average of the sample results for the past five years. The PEB estimates are incorporated into an overall model for reporting quality assurance and production system control measures to Bell System management by the Bell Laboratories Quality Assurance Center.

Maier, et.al. [1982] construct an empirical Bayes technique for estimating an individual security's market risk. They use the traditional market model:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it}$$

$t = 1,\ldots,T$ indexes intertemporal observations

R_{it} = Rate of return of security i in period t

α_i = Rate of return component on security i which is independent of movements in the market index

R_{mt} = Rate of return on a market index in period t

β_i = Measure of market risk of security i reflecting the responsiveness of the security's return to changes in the market index

ε_{it} = Normal (O, σ_i^2) disturbance term

In numerous capital markets studies in the finance and accounting literature β_i is estimated using ordinary least squares regression over some observation period. A crucial element in most of these studies is an accurate estimate of β_i. Maier, et.al. construct an empirical Bayes model for more efficient estimation of β_i than OLS estimation at the individual security level.

Maier, et.al. model the parameters of the market model for each security as random occurrences from a trivariate normal distribution:

$$(\alpha_i\ ,\ \beta_i\ ,\ \sigma_i^2) \sim \text{Normal}\ (\boldsymbol{\mu},\boldsymbol{\Sigma})$$

Clearly if $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ were known they could be used to provide Bayesian revised estimates of $\hat{\beta}_i^{OLS}$. Maier, et.al. assume that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are neither known nor prespecified but do provide formulas for their unbiased estimation based upon the results of individual OLS market model regressions for each of N securities in a portfolio assumed to be representative of the market. Thus, for example:

$$\hat{\mu}_\beta = 1/N \sum_{i=1}^{N} \hat{\beta}_i^{OLS}$$

These underlying parameter estimates are used to construct empirical Bayes estimates of β_i for each of the securities as a function of both the OLS estimates $\hat{\alpha}_i^{OLS}$, $\hat{\beta}_i^{OLS}$, and $\hat{\sigma}_i^2$ and the estimates of the underlying parameters $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. Presumably these empirical Bayes estimates are efficient relative to the individual market model OLS estimates. However, Maier, et.al. present no empirical evidence to support this claim.

CHAPTER III

THE BEHAVIOR OF PARAMETRIC EMPIRICAL BAYES ESTIMATORS IN COMPLIANCE TESTING

The purpose of Chapters III and IV is to consider specific applications of empirical Bayes estimators within the context of audit testing. Chapter III investigates applications during the internal control compliance testing phase of the auditor's examination. Chapter IV investigates applications during the auditor's substantive testing of the reported general ledger account balances.

3.1 The General Compliance Testing Setting

The purpose of compliance testing is to provide the auditor with evidence that the internal controls upon which he intends to rely are operating as prescribed. Attributes sampling procedures are often employed to estimate the rate of noncompliance or error rate, p, for the population of all transactions during the year for which the internal control procedure was applicable.

Suppose the auditor wishes to simultaneously estimate the unknown error rate, p_i, for each of $i = 1,\ldots,k$ similar attributes populations. For example, i may index k different attributes pertaining to a particular internal control system or k divisions within a company each of which is being tested for the same attribute.

In an error rate attributes application of statistical sampling procedures the auditor might draw a random sample of n transactions from each of the k populations. For example, when testing the cash disbursements system a random sample of n disbursements might be selected by reference to check numbers. Each disbursement would then be examined for compliance with k control points (e.g., vendor's invoice stamped "paid", receiving report and purchase order attached to the voucher, review and approval by a supervisor, etc.). If $\hat{x}_i$ represents the number of observed errors in the sample for the ith attribute population, then $\hat{p}_i = \hat{x}_i/n$ is, of course, the traditional MLE estimate for p_i, and the one which is nearly universally used by auditors. The AICPA audit guide, Audit Sampling [AICPA, 1983] refers to $\hat{p}_i$ as the sample deviation rate.

The sampling distribution for the observed number of errors, $\hat{x}_i$, is binomial if the sample items are selected with replacement and hypergeometric if they are selected without replacement. However, in most realistic audit sampling situations the populations are large enough relative to the sample sizes so that the differences between these two distributions are negligible. For simplicity all results in this study are obtained from a model which assumes a large population relative to the sample sizes so that the use of the binomial distribution is appropriate.

3.2 Comparing Normal and Poisson Based Stein-type Estimators

The MLE sample deviation rate can be viewed as the sample mean of a random variable whose value is one if the item is not in compliance with the internal control and zero otherwise. Thus, the central limit theorem guarantees that the MLE estimates are asymptotically distributed as normal random variables, or

$$\hat{p}_i \sim \text{Normal}\left(p_i, \frac{p_i(1 - p_i)}{n}\right).$$

If each $\hat{p}_i$ is indeed an independent normally distributed random variable, then the risk limiting properties of the PEB/JS estimator apply. For the error rates sampling model the JS estimator may be written as:

(1) $$\hat{p}_i^{JS} = \hat{B}^{JS}\,\bar{\bar{p}} + (1 - \hat{B}^{JS})\,\hat{p}_i$$

Where: $$\bar{\bar{p}} = 1/k \sum_{i=1}^{k} \hat{p}_i$$

$$\hat{B}^{JS} = \min\left\{ \frac{(k-3)\bar{\bar{p}}(1-\bar{\bar{p}})}{nS}, 1 \right\}$$

$$S = \sum_{i=1}^{k} (\hat{p}_i - \bar{\bar{p}})^2$$

The above formulation of the JS estimator for use in attributes sampling was first proposed by Ijiri and Leitch [1980] as their equation (3). The formula assumes that the variances of each of the sampling distributions,

$\sigma_i^2 = \frac{p_i(1-p_i)}{n}$, are close so that $\frac{\bar{\hat{p}}(1-\bar{\hat{p}})}{n} \simeq \frac{p_i(1-p_i)}{n}$ serves as an adequate estimate for each σ_i^2. As an alternative Efron and Morris [1975] suggest a variance stablizing transformation of the form $\hat{y}_i = n^{\frac{1}{2}}$ arcsine $(2\hat{p}_i - 1)$ and apply the James-Stein procedure to the resulting $\hat{y}_i$'s. This study examined the performance characteristics of both the transformed and untransformed estimator. The results were not materially different and further references to the JS estimator for attributes sampling will be to the simpler equation (1) as originally proposed by Ijiri and Leitch [1980].

Matsumura and Tsui [1982] observed, however, that the assumption of normality for the sampling distribution of the $\hat{p}_i$'s is suspect in view of the small sample sizes and low error rates often present in attributes sampling. If the observed sample error rates are not normally distributed, then the risk dominance of the James-Stein estimator is no longer guaranteed. Thus, there may be some set of population error rates, $\mathbf{p}^*$, such that $R(\mathbf{p}^{JS},\mathbf{p}^*) > R(\mathbf{p},\mathbf{p}^*)$.

They suggest modeling the observed number of errors, $\hat{x}_i$, under a Poisson distribution since it is a closer approximation to the binomial distribution when error rates and sample sizes are small. Under the assumption that $\hat{x}_1,\ldots,\hat{x}_k$ are independent observations from Poisson distributions they present three estimators (Peng, Hudson-Tsui, and Tsui) each of which risk dominates MLE in the frequentist sense of smaller expected ensemble error for any given fixed set of underlying population error rates. A more complete discussion of Stein-type Poisson based estimators is given in Ghosh, et. al. [1983].

The forms of these three estimators are:

Peng

$$\hat{p}_i^P = 1/n\,[\hat{x}_i - (k-N_o-2)_+\, h(\hat{x}_i)/(S_h+N_o)] \qquad \text{if } \hat{x}_i > 0$$

$$= \min\{(k-N_o-2)_+/(S_h+N_o)\;,\; [1-(k-N_o-2)_+/(S_h+N_o)]\} \quad \text{if } \hat{x}_i = 0$$

Where: N_j = the number of $\hat{x}_i$ equal to j

$$h(x) = \sum_{j=1}^{x} 1/j \qquad \text{if } x \geq 1$$

$$= 0 \qquad \text{if } x = 0$$

$$S_h = \sum_{j=1}^{k} h^2(\hat{x}_j)$$

$$(k-N_o-2)_+ = \max\{0\;,\;(k-N_o-2)\}$$

Hudson-Tsui

$$\hat{p}_i^{HT} = 1/n\,[\hat{x}_i - r(I)\, H(\hat{x}_i)/S_H]$$

Where: I = any prespecified integer

$$r(I) = \max\{0,\, k - \sum_{j=0}^{I} N_j - 3\}$$

$$H(\hat{x}_i) = h(\hat{x}_i) - h(I)$$

$$S_H = \sum_{j=1}^{k} H^2(\hat{x}_j)$$

h(x) = as defined for the Peng estimator

Tsui

(2) $\hat{p}_i^T = 1/n\ [\hat{x}_i - r_M\ K(\hat{x}_i)/S_M]$

Where: M = median of the $\hat{x}_i$'s

$r_M = \max\ \{(\text{number of } \hat{x}_i\text{'s greater than } M) - 2,\ 0\}$

$$K(x) = 1 + \sum_{j=2}^{x-M} 1/(j+M) \qquad \text{if } x \geq M+2$$

$$= 1 \qquad \text{if } x = M+1$$

$$= 0 \qquad \text{if } x = M$$

$$= -b,\ \text{any positive constant} \quad \text{if } x < M$$

$$S_M = \sum_{j=1}^{k} K^2(\hat{x}_j)$$

These Poisson based estimators represent a potentially valuable tool in the efforts to apply multivariate risk limiting estimators to audit sampling. However, there are several reasons why the auditor might not wish to abandon the JS/PEB normal based estimator in favor of one of the Poisson based procedures. Five such reasons are given here.

First, each of the three Poisson based estimators suffer from operational difficulties in the specific context of audit attributes sampling. The Peng estimator adjusts or "shrinks" all non-zero sample observations down toward zero. This inherent downward bias in error rate estimation is most likely unacceptable to the auditor. The auditor will generally be more concerned with the additional audit risk created by an overreliance on internal accounting controls when the estimated error rates are understated (SAS 39, par. 14, AICPA [1985]) than with the inefficiency generated by underreliance when they are overstated (SAS 39, par. 13, AICPA [1985]). This makes the consistent underestimation of the Peng estimator an unacceptable feature for the auditor. The second Poisson based estimator, Hudson-Tsui, shrinks the MLE point estimates toward the auditor's specified prior belief, I/n, for the error rate. As Matsumura and Tsui admit, the subjective element of the estimator makes it less attractive than the James-Stein estimator which adjusts toward a point determined by the data itself, i.e. the grand mean, $\bar{\hat{p}}$. Finally, the Tsui estimator is the most similar of the three to the James-Stein estimator given in (1) as it adjusts the MLE point estimates toward their median observation instead of their common mean. However, it requires the number of populations sampled to be at least six while the James-Stein procedure requires only four attribute populations. This is a rather minor operational drawback and the actual behavior of the Tsui estimator will be compared with that of the James-Stein estimator in Section 3.3.

Secondly, the requirement that the estimator guarantees risk dominance over the entire parameter space may be so severe as to rule out estimators which perform quite well in all relevant and realistic circumstances. Matsumura and Tsui [1982, p. 163] comment on the development of James-Stein estimators by stating that "a critical assumption in the exposition is that the sampling distributions follow (at least approximately) a multivariate normal distribution."

While this assumption may be critical in the proof of guaranteed universal risk dominance, it does not necessarily follow that the estimator performs poorly under other circumstances. In fact, as Judge and Bock [1978, p. 310] conclude: "Since the operating characteristics of the Stein-rules depend on the means and variances of the observations and the unknown coefficients, the estimators are robust relative to the normality assumption."

Indeed, the risk performance of the James-Stein estimator in the specific circumstances of error rate estimation is as much an observable item as a debatable one. When Ghosh, et. al. [1983] simulate the behavior of various Poisson based estimators they also investigate a near normal based estimator which is <u>not</u> proven to be universally risk dominant and which adjusts the MLE estimates toward their geometric mean much in the manner of the JS estimator (1). It is interesting to note that this near normal based estimator resulted in risk savings which were approximately three times larger than those obtained by the best guaranteed risk dominant estimator tested. As the results of Section 3.3 show, the James-Stein estimator consistently dominates both the MLE and the Poisson based Tsui estimator over a wide range of low error rate patterns even with sample sizes as low as 50. The concerns of Matsumura and Tsui about the efficiency of the James-Stein estimator when sample sizes and error rates are low appear to be without merit. Furthermore, the analogous Poisson based estimator which they propose is shown to be grossly inefficient relative to the James-Stein estimator.

The third reason that an auditor might wish to use James-Stein estimation in spite of the non-normal sampling distribution is its natural link to Bayesian estimation. As Berger [1983, p. 368] notes:

> The impetus for constructing improved simultaneous estimators has generally come from two directions: (i) the decision-theoretic approach involving production of estimators dominating (in terms of risk) "usual" estimators which are inadmissible in higher dimensions, and (ii) the Bayesian or empirical Bayesian approach of taking advantage of prior information about the unknown parameters (often information concerning a plausible structure or "model" for these parameters) to produce substantially better estimators tailored to the supposed prior information. It is generally the case that the Bayesian or empirical Bayesian approach is of more practical interest, especially when the unknown parameters...are thought to have some common structure, in that very substantial improvement over standard estimators is usually possible only in such situations.

The development of the Poisson based estimators given in Matsumura and Tsui [1982] arises out of the frequentist decision-theoretic approach of strict risk dominance. However, as Section 2.3.1 established, the James-Stein estimator is a reasonable estimator in its own right as a parametric empirical Bayes (PEB) estimator -- apart from its frequentist risk dominance characteristics.

The derivation in Section 2.3.1 linking the PEB and JS estimators assumed that the sampling distribution was normal. Matusura and Tsui reject the normality assumption and model the $\hat{x}_i$'s as Poisson random variables. In the attributes sampling application the sampling distribution is, in fact, neither normal nor Poisson but rather binomial. Morris [1983b] generalizes the PEB estimator to all sampling distributions in the NEF-QVF family. NEF-QVF includes as special cases the normal, Poisson, and binomial distributions. Morris shows that with a binomial sampling distribution using the natural conjugate prior (the beta distribution) for the underlying error rates produces the following PEB estimator:

$$E\{ p_i | \hat{p}_i \} = \hat{B}^{PEB}\, \bar{\bar{p}} + (1 - \hat{B}^{PEB})\, \hat{p}_i$$

$$\text{Where } \hat{B}^{PEB} = \frac{n}{n-1}\hat{B}^{JS} - \frac{k-1}{(n-1)\,k}$$

Thus, the PEB estimator for population error rates asymptotically approaches the traditional James-Stein estimator of equation (1) as the sample size increases. With any reasonable sample size the difference is negligible and the simpler unadjusted James-Stein estimator as initially proposed by Ijiri and Leitch [1980] is used in the following sections.

Interestingly enough, if the auditor chooses to model the $\hat{x}_i$'s under a Poisson distribution as Matsumura and Tsui suggest, then the natural conjugate prior is the gamma distribution and the resulting PEB estimator is identical to the James-Stein estimator of (1).

A fourth drawback to the Poisson based estimators is their dependence on the weighting coefficients of the loss function. The loss function implicit in the formulation of the Poisson based estimator is

$$L(\mathbf{p},\tilde{\mathbf{p}}) = \sum_{i=1}^{k} (p_i - \tilde{p}_i)^2 = (\mathbf{p} - \tilde{\mathbf{p}})' \mathbf{I} (\mathbf{p} - \tilde{\mathbf{p}})$$

Where $\mathbf{I}$ = k x k identity matrix.

The frequentist risk dominance claim of these estimators is that the expected loss for any fixed set of error rates is no greater than the expected loss for the MLE estimator. However, the claim holds only for the unweighted loss function and would not be true for a more general weighted loss function, say

$$L'(\mathbf{p},\tilde{\mathbf{p}}) = (\mathbf{p} - \tilde{\mathbf{p}})' \mathbf{C} (\mathbf{p} - \tilde{\mathbf{p}})$$

Where $\mathbf{C}$ = arbitrary k x k symmetric weighting matrix

As demonstrated in Chapter II, the Bayes (and hence PEB) estimator is independent of the weighting matrix, $\mathbf{C}$, in the loss function. Thus the JS estimator has a Bayesian interpretation (i.e., both $\hat{p}_i$ and p_i random) as the estimator which mininizes the expected loss of any generalized quadratic loss function. The auditor need not assign specific values for these weightings since the resulting estimates are independent of them. This is not true of the frequentist Poisson based estimators which have no interpretation in the Bayesian view. Their risk limiting properties are limited to the frequentist view with a loss function that has equal population weighting and no interaction effects.

The final drawback to the Poisson based estimators is the lack of any method for establishing confidence intervals about the resulting point estimates. Despite their potential risk dominance over MLE the inability to make inferential statements about the population error rates greatly reduces the number of applications for which an auditor would be willing to use them. However, recent results in PEB estimation procedures provide a method for establishing confidence intervals for the James-Stein estimator. These methods are discussed in Section 3.3.3.

3.3 Performance Characteristics of Stein-type Estimators in Attributes Sampling

The purpose of this section is to present the results of a series of tests examining the performance characteristics of the classical MLE estimator, the pure Bayesian estimator, the JS/PEB estimator of equation (1) and the Tsui estimator of equation (2). The behavior of the estimators is first compared under the frequentist view of fixed population error rates. The behavior of the PEB/JS estimator is also examined under the Bayesian view of random population error rates.

Three different operational characteristics are examined in the following three subsections. These are the efficiency, the bias and the reliability of the estimators. These characteristics were identified by Loebbecke and Neter

[1975] as relevant considerations in the choice between competing estimators for audit testing. Efficiency is defined as the ratio of the estimator's expected loss to the expected loss of MLE. Bias is the difference between an estimator's expected value and the true value of the item being estimated. The reliability of an estimator's confidence interval measures the probability that a confidence interval with a stated level of assurance actually contains the true value being estimated.

3.3.1 Tests of Efficiency

The frequentist behavior of the JS/PEB estimator when used to estimate the underlying error rates of four independent attributes populations was examined under a wide range of error rate patterns. Thirty-five cases representing all possible combinations of 2%, 4%, 6% and 8% error rates for the four populations were evaluated for sample sizes of 50 and 100.

An upper bound of 8% on the error rates was chosen for two reasons. First, it is in low error rate situations that Matsumura and Tsui [1982] question the behavior of the James-Stein estimator. Higher error rates imply a sampling distribution which is more nearly normal and hence satisfy the sufficiency conditions for the efficiency of the James-Stein estimator. Secondly, the AICPA audit guide, Audit Sampling [AICPA, 1983] indicates that the range of tolerable error rates within which the auditor may generally place substantial reliance on the internal accounting control is from 2% to 7%. Error rates in excess of this range imply that the auditor may only place moderate or limited reliance on the control.

An upper bound of 100 on the sample size was chosen for two reasons. First, it is only with low sample sizes that the normality of the sampling distribution is questioned. Secondly, these lower sample sizes are a reasonable representation of actual sample sizes used in practice. Sample sizes much in excess of 100 make it unlikely that sampling to determine compliance with an internal control procedure is a cost effective audit procedure.

The actual individual population frequentist mean squared error or risk, R, of the James-Stein estimate with sample sizes of n and conditional on the actual error rates of the four populations, $\mathbf{p} = (p_1,\ldots,p_4)'$ is given by:

$$(3) \quad R(\hat{p}_i^{JS},\mathbf{p},n) = \sum_{j_1=0}^{n} \cdots \sum_{j_4=0}^{n} (\hat{p}_i^{JS} - p_i)^2 \prod_{h=1}^{4} \text{Prob}(J_h = j_h)$$

Where: $$\text{Prob}(J_h = j_h) = \frac{n!}{j_h!\,(n - j_h)!}\, p_h^{j_h} (1 - p_h)^{n - j_h}$$

$$\hat{p}_i^{JS} = \hat{B}^{JS}\,\bar{\hat{p}} + (1 - \hat{B}^{JS})\, j_h/n$$

$$\hat{B}^{JS} = \min \left\{ \frac{\bar{\hat{p}}\,(1 - \bar{\hat{p}})}{n\,S}, 1 \right\}$$

$$\bar{\hat{p}} = 1/4 \sum_{h=1}^{4} j_h/n$$

$$S = \sum_{h=1}^{4} (j_h/n - \bar{\hat{p}})^2$$

The actual MLE risk can be evaluated in a similar manner by substituting $\hat{p}_i$ for $\hat{p}_i^{JS}$ in (3) or by noting that $R(\hat{p}_i, \mathbf{p}, n) = p_i(1 - p_i)/n$.

The ensemble risk is defined as the sum over the four populations of the individual component risks from (3). The efficiency of the James-Stein estimator can be evaluated by computing the ratio of the James-Stein risk to the MLE risk. Thus the efficiency for any individual population, i, is computed as

$$\mathrm{Eff}(\hat{p}_i^{JS}, \mathbf{p}, n) = R(\hat{p}_i^{JS}, \mathbf{p}, n)/R(\hat{p}_i, \mathbf{p}, n).$$

The ensemble or composite efficiency for all populations combined is computed as

$$\mathrm{Eff}(\hat{\mathbf{p}}^{\mathbf{JS}}, \mathbf{p}, n) = \sum_{i=1}^{4} R(\hat{p}_i^{JS}, \mathbf{p}, n) / \sum_{i=1}^{4} R(\hat{p}_i, \mathbf{p}, n).$$

The actual values for these efficiency measures for each of the 35 population error rate patterns were calculated by direct evaluation of (3). The results are given in Tables 3 and 4 for sample sizes of 50 and 100, respectively. As previously noted, despite the lack of normality for the sampling distribution the James-Stein estimator is frequentist ensemble risk efficient relative to MLE in every instance.

James-Stein estimation is most beneficial when the population error rates are close (e.g., case 1 with ensemble efficiency of .638 representing a 36.2% risk reduction over MLE with a sample of 50). The risk savings diminish as the error rates become more disperse (e.g., case 10) or if one

TABLE 3

James-Stein Frequentist Efficiency, n = 50

Case	Population Error Rates 1	2	3	4	Efficiency 1	2	3	4	Ensemble
1	.02	.02	.02	.02	.638	.638	.638	.638	.638
2	.02	.02	.02	.04	.662	.662	.662	.765	.702
3	.02	.02	.02	.06	.721	.721	.721	.896	.806
4	.02	.02	.02	.08	.783	.783	.783	.981	.893
5	.02	.02	.04	.04	.696	.696	.724	.724	.715
6	.02	.02	.04	.06	.759	.759	.726	.846	.786
7	.02	.02	.04	.08	.820	.820	.753	.942	.862
8	.02	.02	.06	.06	.818	.818	.816	.816	.817
9	.02	.02	.06	.08	.872	.872	.809	.904	.865
10	.02	.02	.08	.08	.918	.918	.878	.878	.887
11	.02	.04	.04	.04	.759	.685	.685	.685	.695
12	.02	.04	.04	.06	.839	.691	.691	.792	.748
13	.02	.04	.04	.08	.903	.724	.724	.900	.821
14	.02	.04	.06	.06	.927	.700	.764	.764	.768
15	.02	.04	.06	.08	.992	.732	.762	.862	.819
16	.02	.04	.08	.08	1.055	.760	.837	.837	.844
17	.02	.06	.06	.06	1.037	.739	.739	.739	.770
18	.02	.06	.06	.08	1.111	.740	.740	.825	.806
19	.02	.06	.08	.08	1.192	.742	.804	.804	.822
20	.02	.08	.08	.08	1.287	.787	.787	.787	.827
21	.04	.04	.04	.04	.646	.646	.646	.646	.646
22	.04	.04	.04	.06	.661	.661	.661	.729	.683
23	.04	.04	.04	.08	.703	.703	.703	.849	.760
24	.04	.04	.06	.06	.681	.681	.700	.700	.692
25	.04	.04	.06	.08	.726	.726	.703	.805	.748
26	.04	.04	.08	.08	.769	.769	.780	.780	.777
27	.04	.06	.06	.06	.711	.673	.673	.673	.680
28	.04	.06	.06	.08	.765	.680	.680	.760	.721
29	.04	.06	.08	.08	.822	.687	.737	.737	.739
30	.04	.08	.08	.08	.890	.717	.717	.717	.743
31	.06	.06	.06	.06	.648	.648	.648	.648	.648
32	.06	.06	.06	.08	.659	.659	.659	.711	.675
33	.06	.06	.08	.08	.674	.674	.689	.689	.682
34	.06	.08	.08	.08	.695	.668	.668	.668	.673
35	.08	.08	.08	.08	.649	.649	.649	.649	.649

TABLE 4

James-Stein Frequentist Efficiency, n = 100

Case	Population Error Rates 1	2	3	4	Efficiency 1	2	3	4	Ensemble
1	.02	.02	.02	.02	.646	.646	.646	.646	.646
2	.02	.02	.02	.04	.701	.701	.701	.845	.758
3	.02	.02	.02	.06	.798	.798	.798	1.006	.900
4	.02	.02	.02	.08	.871	.871	.871	1.059	.975
5	.02	.02	.04	.04	.766	.766	.778	.778	.774
6	.02	.02	.04	.06	.854	.854	.782	.941	.870
7	.02	.02	.04	.08	.913	.913	.822	1.027	.945
8	.02	.02	.06	.06	.922	.922	.890	.890	.898
9	.02	.02	.06	.08	.962	.962	.874	.974	.938
10	.02	.02	.08	.08	.992	.992	.936	.936	.948
11	.02	.04	.04	.04	.885	.715	.715	.715	.740
12	.02	.04	.04	.06	.989	.733	.733	.875	.818
13	.02	.04	.04	.08	1.028	.788	.788	.997	.906
14	.02	.04	.06	.06	1.097	.751	.830	.830	.843
15	.02	.04	.06	.08	1.126	.801	.824	.944	.898
16	.02	.04	.08	.08	1.158	.837	.905	.905	.916
17	.02	.06	.06	.06	1.238	.796	.796	.796	.842
18	.02	.06	.06	.08	1.272	.799	.799	.899	.880
19	.02	.06	.08	.08	1.322	.805	.870	.870	.893
20	.02	.08	.08	.08	1.396	.849	.849	.849	.894
21	.04	.04	.04	.04	.648	.648	.648	.648	.648
22	.04	.04	.04	.06	.684	.684	.684	.788	.718
23	.04	.04	.04	.08	.760	.760	.760	.961	.839
24	.04	.04	.06	.06	.725	.725	.738	.738	.733
25	.04	.04	.06	.08	.800	.800	.746	.897	.819
26	.04	.04	.08	.08	.861	.861	.852	.852	.855
27	.04	.06	.06	.06	.789	.694	.694	.694	.711
28	.04	.06	.06	.08	.874	.710	.710	.833	.778
29	.04	.06	.08	.08	.959	.727	.794	.794	.805
30	.04	.08	.08	.08	1.062	.764	.764	.764	.808
31	.06	.06	.06	.06	.649	.649	.649	.649	.649
32	.06	.06	.06	.08	.676	.676	.676	.758	.701
33	.06	.06	.08	.08	.707	.707	.718	.718	.713
34	.06	.08	.08	.08	.751	.683	.683	.683	.697
35	.08	.08	.08	.08	.650	.650	.650	.650	.650

error rate is significantly different from the others (e.g., case 4). The efficiency appears to be closely related to the relative variance (the variance divided by the mean) of the underlying population error rates. For example, cases 12 and 25 represent two differing population error rate patterns both of which have a relative variance of .005. The ensemble efficiencies for these two cases are virtually identical. In general, the James-Stein estimator increases in efficiency relative to MLE as the relative variance of the underlying population error rates decreases.

An alternative interpretation of these efficiency measures can be made in terms of sample sizes. The sample size, n^*, which would be required to generate the same ensemble risk using MLE as with James-Stein estimation using a sample size of n is given by $n^* = n \, / \, \mathrm{Eff}(\hat{\mathbf{p}}^{JS}, \mathbf{p}, n)$. Thus, for example, in case 1 sample sizes of 78 = 50/.638 would be required for the auditor using traditional MLE techniques in order to achieve the same ensemble risk provided by James-Stein estimation using samples of 50.

Tables 3 and 4 give the frequentist efficiency measures for the simultaneous estimation of the error rates in four populations -- the minimum number required by the James-Stein estimator. The estimator increases in efficiency, ceteris paribus, as the number of populations, k, increases. As the number of populations increases much beyond four, the direct evaluation of exact mean squared errors and efficiencies by equation (3) becomes intractable due to the extremely large number of terms in the k summations. However, Table 5 presents the results of a series of Monte Carlo experiments demonstrating the increased efficiency of the James-Stein estimator as the number of populations increases. The Monte Carlo methodology has been employed extensively in research of the behavior and characteristics of statistical auditing techniques when exact results by analytical methods are intractable. Examples include Neter and Loebbecke [1975], Reneau [1978], Duke, Neter and Leitch [1982], Frost and Tamura [1982], and Dworin and Grimlund [1984]. A description of each specific Monte Carlo experiment used in this study accompanies the presentation of the results. For a description of the Monte Carlo methodology in a general setting see Shreider [1966], Smith [1973], or Chambers [1977, pp. 186-191].

Within each comparison group the mean and variance (and hence the relative variance) of the error rates were held constant in order to control for their effect on the efficiency of the estimator. The behavior of the estimator over k = 4, 6, and 8 populations was examined. For each of the 1,000 trials in the Monte Carlo experiment independent random samples were simulated from the error rate populations. These samples were used to construct both the MLE and James-Stein estimates of the population error rates. The ensemble efficiency represents the ratio of the average ensemble squared errors for the two estimates over the 1,000 Monte Carlo trials. The reliability of the Monte Carlo

TABLE 5

Sensitivity of James-Stein Frequentist Efficiency
to the Number of Populations

k	Population Error Rates								Relative Variance	Frequentist Ensemble Efficiency n=50	n=100
4	.02	.02	.02	.02	-	-	-	-	.0000	.635	.647
6	.02	.02	.02	.02	.02	.02	-	-	.0000	.368	.394
8	.02	.02	.02	.02	.02	.02	.02	.02	.0000	.283	.292
4	.02	.02	.04	.04	-	-	-	-	.0033	.717	.769
6	.02	.02	.02	.04	.04	.04	-	-	.0033	.494	.595
8	.02	.02	.02	.02	.04	.04	.04	.04	.0033	.406	.520
4	.02	.02	.06	.06	-	-	-	-	.0100	.819	.895
6	.02	.02	.02	.06	.06	.06	-	-	.0100	.664	.792
8	.02	.02	.02	.02	.06	.06	.06	.06	.0100	.587	.732
4	.02	.02	.08	.08	-	-	-	-	.0180	.883	.945
6	.02	.02	.02	.08	.08	.08	-	-	.0180	.775	.869
8	.02	.02	.02	.02	.08	.08	.08	.08	.0180	.708	.826
4	.04	.04	.04	.04	-	-	-	-	.0000	.639	.654
6	.04	.04	.04	.04	.04	.04	-	-	.0000	.392	.388
8	.04	.04	.04	.04	.04	.04	.04	.04	.0000	.283	.291
4	.04	.04	.06	.06	-	-	-	-	.0020	.694	.731
6	.04	.04	.04	.06	.06	.06	-	-	.0020	.468	.531
8	.04	.04	.04	.04	.06	.06	.06	.06	.0020	.374	.440
4	.04	.04	.08	.08	-	-	-	-	.0067	.781	.862
6	.04	.04	.04	.08	.08	.08	-	-	.0067	.590	.722
8	.04	.04	.04	.04	.08	.08	.08	.08	.0067	.520	.647
4	.06	.06	.06	.06	-	-	-	-	.0000	.649	.655
6	.06	.06	.06	.06	.06	.06	-	-	.0000	.408	.402
8	.06	.06	.06	.06	.06	.06	.06	.06	.0000	.288	.258
4	.06	.06	.08	.08	-	-	-	-	.0014	.677	.721
6	.06	.06	.06	.08	.08	.08	-	-	.0014	.451	.491
8	.06	.06	.06	.06	.08	.08	.08	.08	.0014	.342	.400
4	.08	.08	.08	.08	-	-	-	-	.0000	.657	.657
6	.08	.08	.08	.08	.08	.08	-	-	.0000	.405	.397
8	.08	.08	.08	.08	.08	.08	.08	.08	.0000	.278	.287

technique is shown by the accuracy of the efficiency measures when the number of populations is four. For these cases the efficiency measures in Table 5 should be comparable to the exact values in Tables 3 and 4. The results show the exact and Monte Carlo estimated efficiencies to be consistently close, differing by at most .007.

The results given in Table 5 confirm the fact that the James-Stein estimator increases in efficiency as the number of populations increase. Reductions of mean squared error of up to 70% were obtained with k = 8 populations.

Table 6 presents the results of a series of Monte Carlo experiments comparing the frequentist efficiency of the James-Stein estimator (equation (1)) to the Poisson based Tsui estimator (equation (2)) proposed by Matsumura and Tsui [1982]. The James-Stein estimator shrinks the MLE sample rates toward their common mean while the Tsui estimator shrinks them toward their common median. The results of the Monte Carlo experiments show that although the Tsui estimator is indeed efficient relative to MLE it is grossly inefficient relative to the James-Stein estimator.

In summary, the results presented in Tables 3 through 6 indicate that the concerns of Matsumura and Tsui [1982] regarding the use of the James-Stein estimator in attributes sampling are unwarranted. Not only is the James-Stein estimator consistently frequentist efficient relative to MLE despite low sample sizes and error rates, but the only

TABLE 6

Comparison of James-Stein and Tsui
Ensemble Frequentist Efficiencies

Population Error Rates						n=50 James-Stein	n=50 Tsui	n=100 James-Stein	n=100 Tsui
.02	.02	.02	.02	.02	.02	.368	.919	.394	.919
.02	.02	.02	.02	.04	.04	.496	.925	.584	.934
.02	.02	.02	.02	.06	.06	.611	.939	.809	.959
.02	.02	.02	.04	.04	.04	.494	.929	.595	.940
.02	.02	.02	.06	.06	.06	.664	.933	.792	.961
.02	.02	.04	.04	.04	.04	.483	.920	.562	.933
.02	.02	.04	.04	.06	.06	.591	.929	.709	.949
.02	.02	.06	.06	.06	.06	.613	.933	.740	.954
.04	.04	.04	.04	.04	.04	.392	.916	.391	.927
.04	.04	.04	.04	.06	.06	.466	.922	.520	.939
.04	.04	.04	.06	.06	.06	.457	.922	.526	.938
.04	.04	.06	.06	.06	.06	.456	.920	.511	.939
.06	.06	.06	.06	.06	.06	.382	.915	.408	.933

operationally feasible Poisson estimator they propose is grossly inefficient relative to the JS estimator.

All of the results in Tables 3 through 6 are given in the frequentist setting of fixed population error rates. As previously shown the JS estimator has a second intepretation as a PEB estimator under the Bayesian perspective of random population error rates. Tables 7 and 8 gives the results of a series of Monte Carlo experiments investigating the efficiency of the JS/PEB estimator from this Bayesian viewpoint.

Twelve underlying distributions for the true but unknown error rates were examined. For each of the twelve underlying distributions sampling from k = 4, 6, and 8 populations with sample sizes of n = 50 and 100 was examined. This resulted in a total of 72 Monte Carlo experiments.

For each of the 1,000 Monte Carlo trials in an experiment a set of k population error rates, $\mathbf{p} = (p_1,...,p_k)'$ was "drawn" in accordance with the underlying distribution. For each of the k populations independent samples of n items were "drawn" with each item having the probability of being an observed error based on $\mathbf{p}$. The MLE sample error rates obtained in this manner were then used to construct the PEB/JS estimates of the population error rates using equation (1). For each trial the MLE and PEB/JS squared errors were calculated. This process was repeated 1,000 times. The PEB/JS efficiency

TABLE 7

Monte Carlo Simulation of PEB and Bayes Efficiency — Bayesian Perspective
n = 50

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Normal (.04,.0001)	4	.678	.124	.188	.210	.233	.498	.388	.989	.653	1.682	1.029	.253	.455
	6	.461	.120	.181	.216	.230	.514	.389	1.016	.658	1.720	1.037	.246	.465
	8	.379	.144	.173	.209	.225	.505	.386	1.000	.658	1.694	1.039	.233	.468
(2) Normal (.04,.0004)	4	.763	.384	.327	.432	.343	.676	.466	1.116	.697	1.752	1.036	.364	.547
	6	.583	.370	.315	.437	.348	.698	.491	1.153	.741	1.804	1.100	.332	.542
	8	.522	.370	.319	.438	.350	.703	.490	1.166	.738	1.826	1.095	.338	.560
(3) Normal (.05,.0001)	4	.689	.188	.208	.099	.150	.174	.190	.413	.328	.817	.564	.246	.528
	6	.453	.189	.204	.098	.146	.178	.191	.431	.339	.855	.589	.240	.529
	8	.353	.183	.197	.095	.142	.177	.188	.428	.335	.848	.584	.233	.538
(4) Normal (.05,.0004)	4	.755	.449	.371	.345	.298	.412	.326	.650	.456	1.058	.687	.376	.597
	6	.569	.448	.360	.342	.291	.409	.325	.650	.464	1.063	.707	.362	.591
	8	.503	.437	.352	.340	.289	.416	.332	.665	.479	1.089	.731	.354	.599
(5) Normal (.06,.0001)	4	.671	.395	.333	.164	.181	.086	.126	.159	.167	.383	.305	.333	.616
	6	.463	.368	.308	.153	.168	.078	.117	.143	.156	.349	.285	.309	.605
	8	.340	.391	.323	.161	.173	.084	.121	.160	.167	.390	.311	.321	.623
(6) Normal (.06,.0004)	4	.727	.627	.483	.391	.326	.308	.266	.378	.303	.601	.438	.461	.664
	6	.567	.602	.464	.376	.314	.297	.257	.365	.294	.580	.424	.446	.666
	8	.480	.623	.476	.390	.322	.308	.265	.378	.305	.599	.441	.457	.677

TABLE 7 (Continued)

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(7) Normal (.07,.0001)	4	.679	.676	.529	.345	.302	.144	.162	.074	.110	.135	.145	.494	.723
	6	.441	.669	.511	.339	.286	.139	.149	.069	.100	.130	.138	.474	.717
	8	.339	.684	.520	.347	.290	.143	.151	.074	.102	.138	.143	.479	.736
(8) Normal (.07,.0004)	4	.739	.857	.654	.533	.429	.336	.290	.267	.236	.327	.267	.611	.748
	6	.546	.877	.652	.545	.425	.344	.286	.275	.236	.367	.275	.605	.750
	8	.456	.871	.652	.540	.424	.339	.286	.270	.235	.331	.274	.606	.759
(9) Normal (.08,.0001)	4	.663	1.063	.802	.629	.496	.319	.276	.131	.142	.067	.095	.724	.837
	6	.440	1.018	.769	.604	.476	.307	.266	.128	.138	.066	.092	.699	.830
	8	.331	1.040	.782	.616	.482	.312	.267	.128	.135	.063	.088	.710	.850
(10) Normal (.08,.0004)	4	.728	1.251	.943	.820	.636	.511	.414	.321	.275	.252	.221	.866	.865
	6	.528	1.202	.901	.783	.604	.484	.391	.304	.261	.243	.215	.828	.865
	8	.445	1.203	.902	.785	.605	.486	.392	.306	.263	.244	.216	.831	.870
(11) Uniform (.00, .08)	4	.795	.567	.429	.646	.456	.925	.590	1.402	.830	2.079	1.178	.417	.611
	6	.663	.548	.417	.629	.452	.906	.593	1.379	.841	2.047	1.195	.404	.612
	8	.603	.547	.412	.626	.445	.909	.588	1.395	.843	2.084	1.209	.398	.623
(12) Exponential (.04)	4	.888	1.712	1.180	1.747	1.159	1.978	1.241	2.405	1.429	3.027	1.720	1.176	.743
	6	.824	1.797	1.202	1.854	1.206	2.117	1.322	2.586	1.549	3.262	1.889	1.211	.766
	8	.790	1.807	1.175	1.872	1.189	2.145	1.316	2.624	1.557	3.311	1.910	1.211	.753

TABLE 8

Monte Carlo Simulation of PEB and Bayes Efficiency — Bayesian Perspective
n = 100

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Normal (.04,.0001)	4	.725	.213	.330	.348	.382	.799	.553	1.564	.845	2.645	1.256	.459	.701
	6	.525	.223	.338	.370	.391	.845	.569	1.650	.871	2.784	1.299	.472	.714
	8	.442	.205	.315	.354	.373	.827	.554	1.623	.860	2.743	1.289	.443	.698
(2) Normal (.04,.0004)	4	.822	.627	.484	.702	.500	1.096	.640	1.811	.904	2.845	1.292	.519	.778
	6	.700	.632	.481	.740	.515	1.176	.676	1.938	.963	3.027	1.376	.522	.765
	8	.647	.597	.464	.697	.502	1.103	.658	1.816	.933	2.834	1.328	.500	.766
(3) Normal (.05,.0001)	4	.716	.328	.372	.174	.273	.294	.323	.689	.490	1.358	.776	.413	.788
	6	.496	.327	.339	.176	.274	.314	.333	.741	.518	1.458	.827	.408	.796
	8	.415	.325	.340	.178	.275	.321	.333	.755	.516	1.479	.823	.412	.806
(4) Normal (.05,.0004)	4	.807	.712	.501	.546	.419	.660	.458	1.055	.618	1.730	.898	.508	.808
	6	.676	.726	.519	.561	.442	.676	.485	1.069	.650	1.741	.934	.527	.829
	8	.629	.750	.530	.580	.450	.691	.490	1.083	.653	1.757	.937	.537	.848
(5) Normal (.06,.0001)	4	.701	.687	.489	.289	.299	.149	.231	.265	.284	.637	.459	.466	.895
	6	.489	.674	.492	.288	.306	.154	.239	.271	.291	.639	.461	.474	.907
	8	.405	.650	.468	.272	.288	.142	.227	.259	.282	.624	.455	.448	.891
(6) Normal (.06,.0004)	4	.796	1.085	.693	.675	.491	.523	.410	.628	.451	.991	.614	.648	.914
	6	.655	1.079	.685	.680	.492	.540	.422	.659	.474	1.037	.649	.641	.918
	8	.592	1.064	.675	.665	.481	.524	.410	.641	.460	1.016	.633	.637	.920

TABLE 8 (Continued)

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(7) Normal (.07,.0001)	4	.689	1.207	.759	.614	.455	.256	.270	.133	.205	.244	.258	.652	1.007
	6	.474	1.203	.753	.613	.452	.258	.270	.137	.207	.250	.264	.648	1.007
	8	.386	1.203	.750	.612	.448	.256	.267	.136	.205	.251	.264	.646	1.006
(8) Normal (.07,.0004)	4	.781	1.567	.958	.972	.649	.613	.461	.492	.392	.608	.443	.864	.997
	6	.627	1.516	.907	.937	.609	.588	.428	.468	.365	.578	.418	.827	.981
	8	.576	1.572	.945	.978	.640	.617	.453	.487	.385	.590	.436	.861	1.000
(9) Normal (.08,.0001)	4	.698	1.832	1.113	1.089	.710	.557	.420	.234	.244	.122	.180	.952	1.086
	6	.465	1.785	1.081	1.056	.687	.535	.405	.221	.235	.114	.178	.924	1.084
	8	.368	1.846	1.127	1.096	.720	.558	.427	.233	.250	.121	.187	.965	1.121
(10) Normal (.08,.0004)	4	.777	2.111	1.274	1.382	.876	.858	.589	.541	.413	.429	.348	1.161	1.060
	6	.612	2.108	1.268	1.379	.870	.854	.583	.533	.406	.416	.339	1.155	1.078
	8	.552	2.110	1.268	1.379	.871	.857	.585	.542	.412	.435	.350	1.158	1.093
(11) Uniform (.00, .08)	4	.864	.896	.581	.986	.604	1.382	.746	2.087	1.007	3.099	1.388	.551	.816
	6	.771	.892	.573	1.006	.608	1.437	.766	2.183	1.045	3.245	1.446	.540	.818
	8	.718	.926	.591	1.054	.633	1.510	.801	2.295	1.096	3.407	1.516	.554	.830
(12) Exponential (.04)	4	.936	3.051	1.633	3.051	1.576	3.372	1.644	4.015	1.837	4.978	2.155	1.861	.881
	6	.890	2.763	1.445	2.828	1.439	3.201	1.553	3.884	1.786	4.874	2.139	1.643	.861
	8	.894	2.939	1.483	3.036	1.489	3.474	1.625	4.252	1.892	5.372	2.289	1.792	.896

relative to MLE as shown in column 1 represents the ratio of the PEB/JS and MLE squared errors averaged over the 1,000 trials. The results show the PEB/JS estimator to be efficient relative to MLE for each of the 72 Monte Carlo experiments.

The first six underlying distributions are normal (truncated at zero) so that the use of the natural conjugate prior in the construction of the PEB estimator is in fact a relatively accurate representation of reality. The last two distributions are uniform and exponential, respectively. The uniform distribution implies that every possible outcome within some fixed interval has equal probability. Thus,

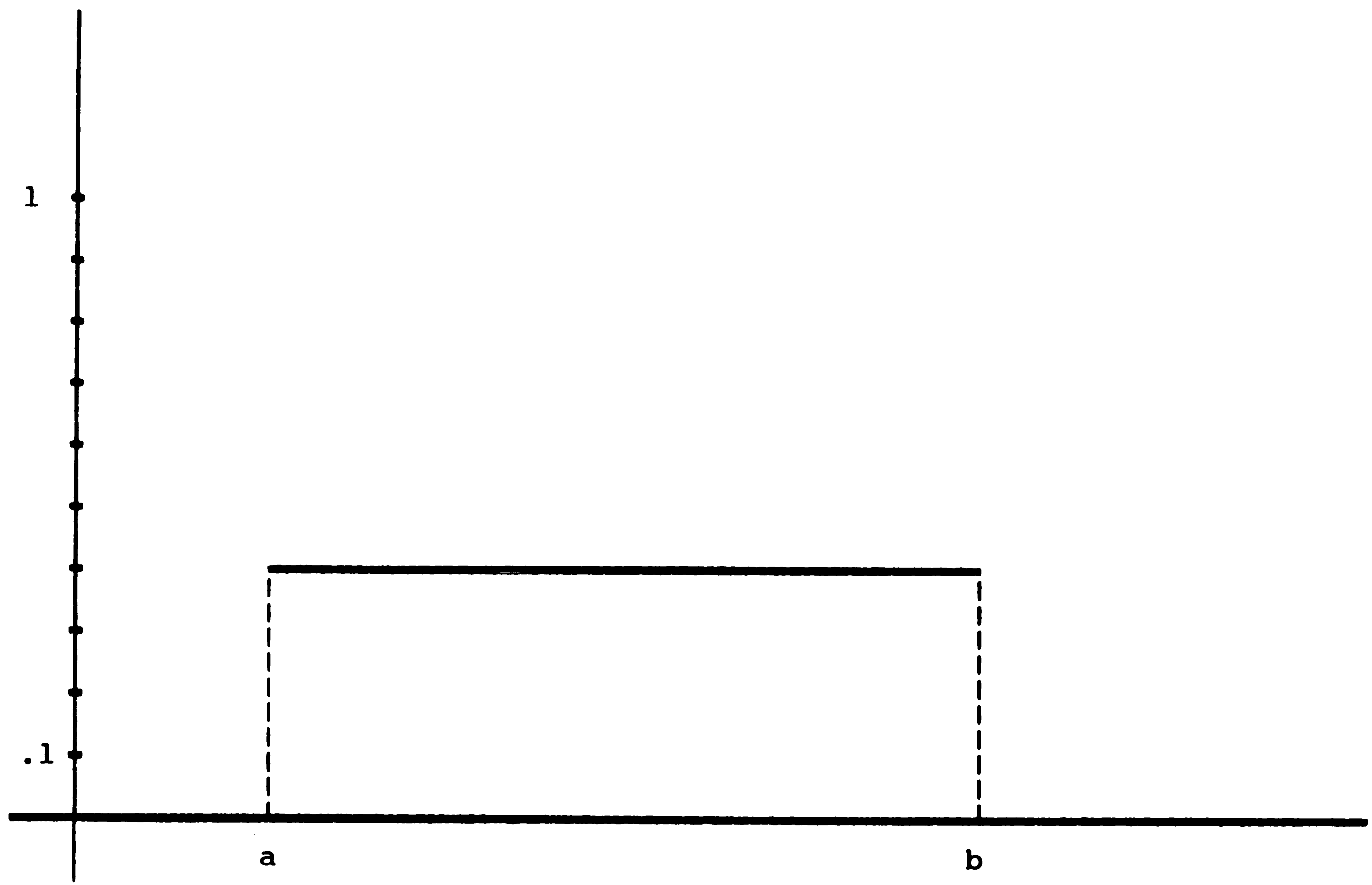
$$\text{Uniform } (a,b) \sim g(\theta) = 1/(b-a) \quad \text{for } \theta \in [a,b]$$
$$= 0 \quad \text{otherwise}$$
$$\text{Mean}(\theta) = (a+b)/2 \quad , \text{Variance}(\theta) = (b-a)^2/12$$

The exponential distribution is highly skewed to the right with functional form:

$$\text{Exponential } (a) \sim g(\theta) = \frac{1}{a} e^{-\theta/a} \quad \text{for } \theta \geq 0$$
$$= 0 \quad \text{otherwise}$$
$$\text{Mean } (\theta) = a \ , \text{Variance } (\theta) = a^2$$

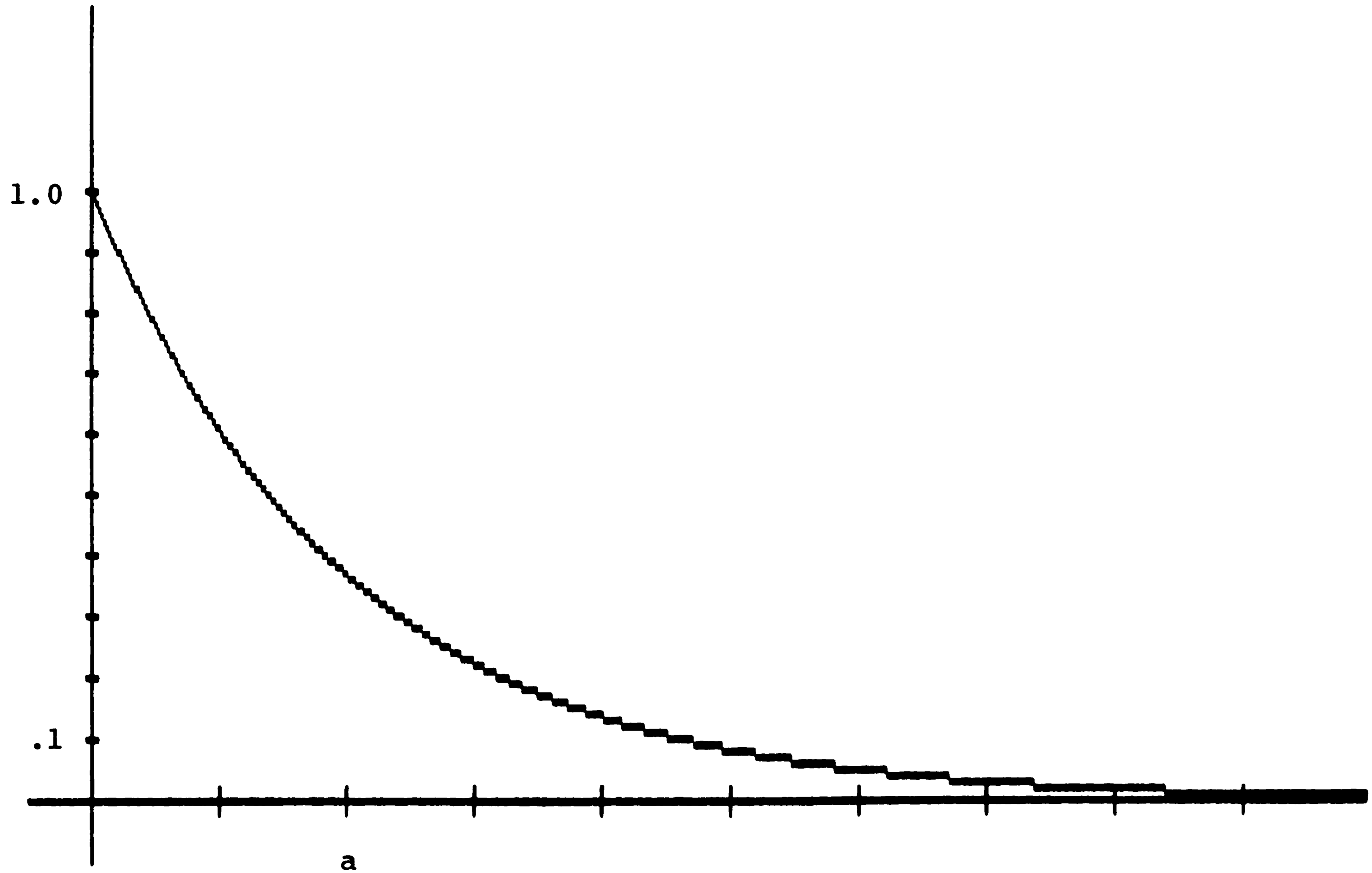
Illustrative graphs of these two distributions are given in Figures 4 and 5.

As can be seen from the graphs, the shapes of both the uniform and the exponential distributions are quite different from the "bell-shaped" curve of the normal distribution. Each is included in the Monte Carlo experiment to demonstrate that the PEB estimator is robust



Uniform (a,b) Density

FIGURE 4



Exponential (a) Density

FIGURE 5

against the actual functional form of the underlying distribution. Recall that the PEB estimator used the normal distribution as the natural conjugate prior. However, the results show the estimator is still efficient relative to MLE even when the actual underlying distribution is quite different from the normal curve.

The last twelve columns of Tables 7 and 8 show the efficiencies of various pure Bayesian estimators relative to MLE. For each of the 1,000 Monte Carlo trials a set of pure Bayesian estimators was calculated using the prior assumptions indicated at the top of each column. The entries represent the efficiencies of the Bayes estimate relative to MLE determined by the ratio of the Bayes mean squared error to the MLE mean squared error when averaged over the 1,000 trials.

These results show that when the prior distribution is properly specified the pure Bayesian estimator is preferred to both MLE and PEB. However, the costs of misspecification can be great, rendering the pure Bayes estimator inefficient relative to the MLE estimator. These costs of misspecification are avoided by the PEB estimator since it estimates the prior parameters using the sample data itself rather than relying on subjective prior specifications. In every instance the PEB estimator was efficient relative to MLE.

3.3.2 Tests of Bias

The preceding results addressed only one characteristic of the estimators, namely their expected mean squared error or loss. A second characteristic which should be analyzed is their bias (Loebbecke and Neter [1975]). One desirable property for an estimator is that it be unbiased or have expected value equal to the unknown parameter being estimated.

Although not explicitly mentioned by either Ijiri and Leitch [1980] or Matsumura and Tsui [1982], one disadvantage of Stein-type estimators the auditor should consider is the fact that estimators of this type may be biased under the frequentist view. Since the James-Stein estimator shrinks the unbiased MLE estimates toward their common mean, it tends to under- (over-) estimate in the populations with higher (lower) error rates. The magnitude of this bias can be evaluated using equation (3) by replacing $(\hat{p}_i^{JS} - p_i)^2$ with $(\hat{p}_i^{JS} - p_i)$. The actual frequentist bias of the James-Stein estimator for each of the 35 error rate patterns is given in Tables 9 and 10 for sample sizes of 50 and 100, respectively.

The PEB/JS estimator is biased only under the frequentist view of fixed population error rates. Under the Bayesian view of random error rates the PEB/JS estimator is unbiased. Tables 11 and 12 show the average bias for the PEB/JS and pure Bayesian estimators for the 72 Monte Carlo

TABLE 9

James-Stein Frequentist Bias, n = 50

Case	Population Error Rates 1	2	3	4	Bias 1	2	3	4
1	.02	.02	.02	.02	.0000	.0000	.0000	.0000
2	.02	.02	.02	.04	.0013	.0013	.0013	-.0039
3	.02	.02	.02	.06	.0021	.0021	.0021	-.0062
4	.02	.02	.02	.08	.0024	.0024	.0024	-.0073
5	.02	.02	.04	.04	.0028	.0028	-.0028	-.0028
6	.02	.02	.04	.06	.0036	.0036	-.0036	-.0036
7	.02	.02	.04	.08	.0039	.0039	-.0006	-.0072
8	.02	.02	.06	.06	.0044	.0044	-.0044	-.0044
9	.02	.02	.06	.08	.0048	.0048	-.0032	-.0063
10	.02	.02	.08	.08	.0052	.0052	-.0052	-.0052
11	.02	.04	.04	.04	.0046	-.0015	-.0015	-.0015
12	.02	.04	.04	.06	.0056	-.0003	-.0003	-.0049
13	.02	.04	.04	.08	.0058	.0006	.0006	-.0071
14	.02	.04	.06	.06	.0067	.0008	-.0038	-.0038
15	.02	.04	.06	.08	.0071	.0017	-.0026	-.0062
16	.02	.04	.08	.08	.0076	.0025	-.0051	-.0051
17	.02	.06	.06	.06	.0081	-.0027	-.0027	-.0027
18	.02	.06	.06	.08	.0086	-.0016	-.0016	-.0055
19	.02	.06	.08	.08	.0093	-.0006	-.0044	-.0044
20	.02	.08	.08	.08	.0102	-.0034	-.0034	-.0034
21	.04	.04	.04	.04	.0000	.0000	.0000	.0000
22	.04	.04	.04	.06	.0014	.0014	.0014	-.0041
23	.04	.04	.04	.08	.0023	.0023	.0023	-.0070
24	.04	.04	.06	.06	.0029	.0029	-.0029	-.0029
25	.04	.04	.06	.08	.0038	.0038	-.0016	-.0061
26	.04	.04	.08	.08	.0049	.0049	-.0049	-.0049
27	.04	.06	.06	.06	.0046	-.0015	-.0015	-.0015
28	.04	.06	.06	.08	.0057	-.0002	-.0002	-.0052
29	.04	.06	.08	.08	.0069	.0010	-.0040	-.0040
30	.04	.08	.08	.08	.0083	-.0028	-.0028	-.0028
31	.06	.06	.06	.06	.0000	.0000	.0000	.0000
32	.06	.06	.06	.08	.0014	.0014	.0014	-.0042
33	.06	.06	.08	.08	.0029	.0029	-.0029	-.0029
34	.06	.08	.08	.08	.0045	-.0015	-.0015	-.0015
35	.08	.08	.08	.08	.0000	.0000	.0000	.0000

TABLE 10

James-Stein Frequentist Bias, n = 100

Case	Population Error Rates				Bias			
	1	2	3	4	1	2	3	4
1	.02	.02	.02	.02	.0000	.0000	.0000	.0000
2	.02	.02	.02	.04	.0012	.0012	.0012	-.0035
3	.02	.02	.02	.06	.0016	.0016	.0016	-.0048
4	.02	.02	.02	.08	.0016	.0016	.0016	-.0049
5	.02	.02	.04	.04	.0025	.0025	-.0025	-.0025
6	.02	.02	.04	.06	.0028	.0028	-.0012	-.0044
7	.02	.02	.04	.08	.0027	.0027	-.0003	-.0050
8	.02	.02	.06	.06	.0033	.0033	-.0033	-.0033
9	.02	.02	.06	.08	.0032	.0032	-.0021	-.0043
10	.02	.02	.08	.08	.0033	.0033	-.0033	-.0033
11	.02	.04	.04	.04	.0042	-.0014	-.0014	-.0014
12	.02	.04	.04	.06	.0047	-.0003	-.0003	-.0042
13	.02	.04	.04	.08	.0043	.0005	.0005	-.0053
14	.02	.04	.06	.06	.0053	.0007	-.0030	-.0030
15	.02	.04	.06	.08	.0051	.0013	-.0018	-.0046
16	.02	.04	.08	.08	.0052	.0018	-.0035	-.0035
17	.02	.06	.06	.06	.0064	-.0021	-.0021	-.0021
18	.02	.06	.06	.08	.0063	-.0011	-.0011	-.0041
19	.02	.06	.08	.08	.0065	-.0003	-.0031	-.0031
20	.02	.08	.08	.08	.0070	-.0023	-.0023	-.0023
21	.04	.04	.04	.04	.0000	.0000	.0000	.0000
22	.04	.04	.04	.06	.0013	.0013	.0013	-.0039
23	.04	.04	.04	.08	.0020	.0020	.0020	-.0059
24	.04	.04	.06	.06	.0027	.0027	-.0027	-.0027
25	.04	.04	.06	.08	.0033	.0033	-.0013	-.0052
26	.04	.04	.08	.08	.0039	.0039	-.0039	-.0039
27	.04	.06	.06	.06	.0043	-.0014	-.0014	-.0014
28	.04	.06	.06	.08	.0050	-.0002	-.0002	-.0047
29	.04	.06	.08	.08	.0059	.0009	-.0034	-.0034
30	.04	.08	.08	.08	.0070	-.0023	-.0023	-.0023
31	.06	.06	.06	.06	.0000	.0000	.0000	.0000
32	.06	.06	.06	.08	.0014	.0014	.0014	-.0041
33	.06	.06	.08	.08	.0027	.0027	-.0027	-.0027
34	.06	.08	.08	.08	.0044	-.0015	-.0015	-.0015
35	.08	.08	.08	.08	.0000	.0000	.0000	.0000

TABLE 11

Monte Carlo Simulation of PEB and Bayes Bias -- Bayesian Perspective
n = 50

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Normal (.04,.0001)	4	.000	-.001	-.001	.008	.005	.017	.011	.025	.018	.034	.024	-.002	-.006
	6	.000	.000	-.001	.008	.006	.017	.012	.026	.018	.035	.025	-.001	-.005
	8	.000	.000	.000	.009	.006	.017	.012	.026	.019	.035	.025	-.001	-.005
(2) Normal (.04,.0004)	4	.000	-.002	-.002	.007	.004	.015	.010	.024	.017	.033	.023	-.033	-.006
	6	.001	-.001	-.001	.007	.005	.016	.012	.025	.018	.034	.025	-.002	-.004
	8	.000	-.001	-.001	.007	.005	.016	.012	.025	.018	.034	.024	-.002	-.005
(3) Normal (.05,.0001)	4	.000	-.009	-.008	.000	-.001	.008	.006	.017	.013	.026	.019	-.008	-.010
	6	.000	-.009	-.008	.000	-.001	.009	.006	.018	.013	.027	.020	-.007	-.010
	8	.000	-.009	-.007	.000	-.001	.009	.006	.018	.013	.027	.020	-.007	-.009
(4) Normal (.05,.0004)	4	.000	-.010	-.008	-.001	-.002	.008	.005	.017	.012	.026	.019	-.008	-.010
	6	.000	-.010	-.008	-.001	-.001	.008	.006	.017	.013	.026	.020	-.008	-.009
	8	.000	-.010	-.008	-.001	-.001	.008	.006	.017	.013	.027	.020	-.007	-.009
(5) Normal (.06,.0001)	4	.000	-.018	-.015	-.009	-.008	.000	-.001	.009	.006	.018	.014	-.014	-.014
	6	.000	-.019	-.015	-.009	-.008	.000	-.001	.009	.007	.018	.014	-.014	-.014
	8	.000	-.018	-.015	-.009	-.008	.000	.000	.009	.007	.018	.014	-.014	-.014
(6) Normal (.06,.0004)	4	.001	-.019	-.015	-.009	-.008	.000	-.001	.009	.006	.018	.014	-.014	-.013
	6	.000	-.019	-.015	-.009	-.008	.000	-.001	.009	.006	.018	.014	-.015	-.013
	8	.000	-.019	-.015	-.009	-.008	.000	-.001	.009	.007	.018	.014	-.014	-.013

TABLE 11 (Continued)

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(7) Normal (.07,.0001)	4	-.001	-.028	-.024	-.019	-.016	-.010	-.009	.000	-.001	.009	.006	-.022	-.019
	6	.000	-.028	-.023	-.019	-.016	-.010	-.008	.000	-.001	.009	.007	-.022	-.018
	8	.000	-.028	-.023	-.019	-.016	-.009	-.008	.000	.000	.009	.007	-.022	-.018
(8) Normal (.07,.0004)	4	.000	-.028	-.024	-.019	-.016	-.009	-.009	.000	-.001	.009	.006	-.022	-.018
	6	.001	-.028	-.023	-.019	-.016	-.009	-.008	.000	-.001	.009	.007	-.022	-.017
	8	.001	-.028	-.023	-.019	-.016	-.009	-.008	.000	-.001	.009	.007	-.022	-.018
(9) Normal (.08,.0001)	4	.001	-.038	-.032	-.028	-.024	-.019	-.016	-.010	-.008	.000	-.001	-.030	-.023
	6	.000	-.038	-.032	-.028	-.024	-.019	-.016	-.010	-.008	.000	-.001	-.030	-.023
	8	.000	-.038	-.032	-.028	-.024	-.019	-.016	-.010	-.008	.000	-.001	-.030	-.023
(10) Normal (.08,.0004)	4	.000	-.038	-.032	-.029	-.025	-.019	-.017	-.010	-.009	-.001	-.001	-.031	-.023
	6	.000	-.037	-.032	-.028	-.024	-.019	-.016	-.009	-.008	.000	.000	-.030	-.022
	8	.000	-.037	-.032	-.028	-.024	-.019	-.016	-.009	-.008	.000	.000	-.030	-.022
(11) Uniform (.00, .08)	4	.000	-.001	-.002	.008	.004	.016	.011	.025	.017	.033	.023	-.002	-.004
	6	.000	-.001	-.001	.008	.005	.017	.012	.026	.018	.034	.024	-.002	-.004
	8	.000	-.001	-.001	.008	.005	.017	.012	.025	.018	.034	.024	-.002	-.004
(12) Exponential (.04)	4	.000	-.002	-.003	.006	.003	.015	.009	.023	.015	.031	.021	-.005	-.003
	6	.000	-.001	-.002	.007	.004	.016	.010	.024	.017	.033	.023	-.004	-.002
	8	.000	-.001	-.002	.008	.005	.016	.011	.025	.017	.034	.024	-.003	-.002

TABLE 12

Monte Carlo Simulation of PEB and Bayes Bias -- Bayesian Perspective
n = 100

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Normal (.04,.0001)	4	.000	-.001	-.001	.007	.004	.015	.009	.023	.013	.031	.018	-.001	-.005
	6	.000	.000	-.001	.007	.004	.015	.009	.023	.014	.031	.019	-.001	-.005
	8	.000	.000	.000	.008	.004	.015	.009	.023	.014	.031	.019	-.001	-.005
(2) Normal (.04,.0004)	4	.000	-.002	-.002	.006	.003	.014	.008	.022	.013	.029	.017	-.002	-.005
	6	.000	-.001	-.001	.007	.004	.015	.008	.022	.013	.030	.018	-.002	-.004
	8	.000	-.001	-.001	.007	.004	.015	.009	.022	.014	.030	.019	-.002	-.004
(3) Normal (.05,.0001)	4	.000	-.009	-.006	-.001	-.001	.008	.004	.016	.010	.024	.015	-.005	-.008
	6	.000	-.009	-.008	.000	.000	.008	.005	.016	.010	.024	.016	-.005	-.008
	8	.000	-.008	-.008	.000	.008	.008	.005	.016	.010	.025	.016	-.005	-.008
(4) Normal (.05,.0004)	4	.000	-.009	-.006	-.001	-.001	.008	.004	.016	.010	.024	.015	-.006	-.007
	6	.000	-.009	-.006	-.001	-.001	.008	.005	.016	.010	.024	.015	-.006	-.007
	8	.000	-.009	-.006	-.001	-.001	.007	.004	.016	.010	.024	.015	-.006	-.007
(5) Normal (.06,.0001)	4	.000	-.017	-.012	-.009	-.006	.000	-.001	.008	.005	.016	.011	-.010	-.011
	6	.000	-.017	-.012	-.009	-.006	.000	-.001	.008	.005	.017	.011	-.010	-.011
	8	.000	-.017	-.012	-.009	-.006	.000	.000	.008	.006	.017	.011	-.010	-.010
(6) Normal (.06,.0004)	4	.000	-.017	-.013	-.009	-.007	-.001	-.001	.008	.005	.016	.011	-.011	-.010
	6	.000	-.017	-.012	-.009	-.006	.000	.000	.008	.005	.017	.011	-.011	-.010
	8	.000	-.017	-.012	-.009	-.006	.000	-.001	.008	.005	.017	.011	-.011	-.010

TABLE 12 (Continued)

Prior Distribution For Error Rates	K	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(7) Normal (.07,.0001)	4	.000	-.026	-.019	-.017	-.013	-.009	-.007	.000	-.001	.008	.006	-.017	-.013
	6	.000	-.026	-.019	-.017	-.013	-.009	-.006	.000	.000	.008	.006	-.017	-.013
	8	.000	-.026	-.019	-.017	-.013	-.009	-.006	.000	.000	.009	.006	-.016	-.013
(8) Normal (.07,.0004)	4	.000	-.026	-.019	-.017	-.013	-.009	-.007	.000	.000	.009	.006	-.017	-.013
	6	.001	-.026	-.019	-.017	-.013	-.009	-.006	.000	.000	.009	.006	-.017	-.012
	8	.000	-.026	-.019	-.018	-.013	-.009	-.007	.000	.000	.008	.006	-.017	-.012
(9) Normal (.08,.0001)	4	.000	-.035	-.026	-.027	-.020	-.018	-.013	-.009	-.007	.000	-.001	-.024	-.016
	6	.000	-.035	-.026	-.026	-.020	-.018	-.013	-.009	-.007	.000	.000	-.024	-.015
	8	.000	-.035	-.026	-.027	-.020	-.018	-.013	-.009	-.007	.000	.000	-.024	-.016
(10) Normal (.08,.0004)	4	.000	-.035	-.026	-.027	-.020	-.018	-.013	-.009	-.007	.000	-.001	-.025	-.015
	6	.000	-.036	-.027	-.027	-.020	-.018	-.014	-.009	-.007	-.001	-.001	-.025	-.015
	8	.000	-.035	-.026	-.026	-.020	-.018	-.013	-.009	-.007	.000	.000	-.024	-.015
(11) Uniform (.00, .08)	4	.000	-.001	-.001	.007	.003	.014	.008	.022	.013	.030	.018	-.002	-.003
	6	.000	-.001	-.001	.007	.004	.015	.009	.023	.013	.030	.018	-.002	-.003
	8	.000	-.001	-.001	.007	.004	.015	.009	.023	.014	.031	.019	-.001	-.003
(12) Exponential (.04)	4	.000	-.002	-.003	.005	.002	.013	.006	.020	.011	.028	.015	-.005	-.002
	6	.000	-.001	-.002	.006	.003	.014	.008	.022	.012	.029	.017	-.004	-.001
	8	.000	-.001	-.001	.007	.003	.014	.008	.022	.013	.030	.017	-.004	-.002

experiments described earlier. Notice that the pure Bayesian estimator is unbiased only when the mean of the prior distribution has been properly specified. When the prior mean has been misspecified the Bayes estimator "shrinks" the MLE toward the false prior value, resulting in a biased estimate. The empirical Bayesian avoids this potential error by refusing to specify any value for the prior mean, choosing instead to use an unbiased estimate for the mean derived from the data itself.

Whether or not the user is willing to accept an estimator which is frequentist biased but produces smaller expected squared error is, of course, a matter of personal taste. If the auditor insists on frequentist unbiased estimators, then he has little choice but to continue with the use of the unbiased MLE estimator since it is the estimator with minimum expected squared error among the class of all unbiased estimators. However, there is no reason to suppose a priori that auditors would insist upon unbiasedness as a required characteristic of an estimator. In fact, there is good reason to assume the contrary since auditors have for years used the classical ratio estimator described in Section 2.1 which is a frequentist biased estimator [Cochran, 1977, pp. 160-162]. Although Frost and Tamura [1982] investigated methods for reducing the bias of the ratio estimator, their primary objective was to improve the estimator's efficiency and reliability, not to seek an unbiased estimator for its own sake. This lends support to the notion that auditors may be willing to adopt estimators which are frequentist biased in exchange for substantially reduced expected squared error. The results of this study give measures for the benefits from risk reduction (Tables 3 and 4) versus the cost of increased bias (Tables 9 and 10) to assist the auditor in his decision of whether or not to adopt James-Stein estimation procedures.

There is, however, a specific concern about the nature of the PEB/JS estimator's potential for bias which is unique to audit sampling as contemplated by SAS 39, Audit Sampling [AICPA, 1985]. This concern centers around SAS 39's view of sampling risk as it pertains to compliance testing. As deomonstrated by the frequentist bias measures in Tables 8 and 9, since the JS estimate shrinks the unbiased MLE values toward their common mean, it tends to overestimate populations with relatively low error rates. Overstating the error rate in a population increases the risk of the auditor's underreliance on internal accounting controls. The result may be an unnecessary expansion in the scope of substantive audit procedures. SAS 39 minimizes the concern for this form of increased risk since the result is an equally effective but inefficient audit (SAS 39, par. 13, AICPA [1985]).

Conversely, by shrinking high MLE values downward the JS estimator increases the likelihood of underestimation in high error rate populations. This increases the risk of the auditor's overreliance on internal accounting controls for high error rate populations. This increased risk of overreliance in high error rate populations and the resultant diminished audit effectiveness (SAS 39, par. 14, AICPA [1985]) may cause auditors to reject James-Stein estimation despite its superior ensemble risk performance.

To address this somewhat disconcerting characteristic of the James-Stein estimator an ad hoc "positive adjustment" estimator is proposed. The estimator is defined as

$$\hat{p}_i^{JS+} = \max\ \{\hat{p}_i^{JS},\ \hat{p}_i\} \text{ for all } i.$$

The positive adjustment estimator adjusts MLE to the JS estimate only if the adjustment is in the upward direction. The estimator is neither PEB nor known to be guaranteed frequentist risk dominant over MLE but is considered as a compromise between these features of JS estimation and the auditor's apparent concerns about overreliance on individual internal controls.

The estimator is, of course, biased toward overstatement in every population and over the entire parameter space (including situations where the pure JS estimator is unbiased or biased toward understatement). This is shown by Tables 13 and 14 which give the results of exact bias computations for the positive adjustment estimator over 35 error rate patterns and sample sizes of n = 50 and 100, respectively. The surprising feature of the estimator is that despite its guaranteed overstatement in every population, it nonetheless continues to dominate the unbiased MLE in ensemble mean squared error. This is

TABLE 13

Positive Adjustment Estimator Operating Characteristics, n = 50

Case	Population Error Rates				Frequentist Efficiency					Frequentist Bias			
	1	2	3	4	1	2	3	4	Ensemble	1	2	3	4
1	.02	.02	.02	.02	.868	.868	.868	.868	.868	.0025	.0025	.0025	.0025
2	.02	.02	.02	.04	.870	.870	.870	.918	.889	.0032	.0032	.0032	.0012
3	.02	.02	.02	.06	.887	.887	.887	.959	.922	.0034	.0034	.0034	.0005
4	.02	.02	.02	.08	.904	.904	.904	1.001	.947	.0033	.0033	.0033	.0002
5	.02	.02	.04	.04	.886	.886	.901	.901	.896	.0043	.0043	.0017	.0017
6	.02	.02	.04	.06	.910	.910	.899	.944	.921	.0046	.0046	.0019	.0008
7	.02	.02	.04	.08	.930	.930	.904	.972	.944	.0046	.0046	.0019	.0004
8	.02	.02	.06	.06	.939	.939	.938	.938	.938	.0052	.0052	.0010	.0010
9	.02	.02	.06	.08	.960	.960	.937	.967	.956	.0053	.0053	.0010	.0005
10	.02	.02	.08	.08	.983	.983	.965	.965	.969	.0055	.0055	.0005	.0005
11	.02	.04	.04	.04	.932	.879	.879	.879	.887	.0058	.0024	.0024	.0024
12	.02	.04	.04	.06	.977	.877	.877	.923	.907	.0064	.0028	.0028	.0013
13	.02	.04	.04	.08	1.005	.884	.884	.958	.930	.0064	.0029	.0029	.0006
14	.02	.04	.06	.06	1.039	.874	.913	.913	.919	.0073	.0034	.0015	.0015
15	.02	.04	.06	.08	1.074	.882	.912	.949	.937	.0075	.0036	.0017	.0008
16	.02	.04	.08	.08	1.115	.888	.945	.945	.951	.0079	.0040	.0009	.0009
17	.02	.06	.06	.06	1.127	.900	.900	.900	.924	.0085	.0020	.0020	.0020
18	.02	.06	.06	.08	1.178	.897	.897	.937	.938	.0089	.0022	.0022	.0011
19	.02	.06	.08	.08	1.241	.893	.931	.931	.948	.0095	.0026	.0012	.0012
20	.02	.08	.08	.08	1.324	.922	.922	.922	.955	.0103	.0015	.0015	.0015
21	.04	.04	.04	.04	.858	.858	.858	.858	.858	.0037	.0037	.0037	.0037
22	.04	.04	.04	.06	.861	.861	.861	.892	.871	.0043	.0043	.0043	.0021
23	.04	.04	.04	.08	.875	.875	.875	.934	.898	.0046	.0046	.0046	.0011
24	.04	.04	.06	.06	.869	.869	.879	.879	.875	.0053	.0053	.0026	.0026
25	.04	.04	.06	.08	.887	.887	.879	.920	.897	.0057	.0057	.0029	.0014
26	.04	.04	.08	.08	.908	.908	.914	.914	.912	.0063	.0063	.0016	.0016
27	.04	.06	.06	.06	.890	.865	.865	.865	.869	.0066	.0034	.0034	.0034
28	.04	.06	.06	.08	.918	.864	.864	.901	.886	.0073	.0038	.0038	.0019
29	.04	.06	.08	.08	.954	.864	.893	.893	.896	.0082	.0044	.0023	.0023
30	.04	.08	.08	.08	1.004	.883	.883	.883	.901	.0093	.0028	.0028	.0028
31	.06	.06	.06	.06	.851	.851	.851	.851	.851	.0045	.0045	.0045	.0045
32	.06	.06	.06	.08	.854	.854	.854	.878	.861	.0052	.0052	.0052	.0027
33	.06	.06	.08	.08	.860	.860	.868	.868	.865	.0061	.0061	.0033	.0033
34	.06	.08	.08	.08	.874	.857	.857	.857	.860	.0074	.0041	.0041	.0041
35	.08	.08	.08	.08	.847	.847	.847	.847	.847	.0051	.0051	.0051	.0051

TABLE 14

Positive Adjustment Estimator Operating Characteristics, n = 100

Case	Population Error Rates				Frequentist Efficiency					Frequentist Bias			
	1	2	3	4	1	2	3	4	Ensemble	1	2	3	4
1	.02	.02	.02	.02	.859	.859	.859	.859	.859	.0019	.0019	.0019	.0019
2	.02	.02	.02	.04	.875	.875	.875	.934	.898	.0023	.0023	.0023	.0006
3	.02	.02	.02	.06	.905	.905	.905	.981	.942	.0022	.0022	.0022	.0001
4	.02	.02	.02	.08	.926	.926	.926	.996	.965	.0018	.0018	.0018	.0000
5	.02	.02	.04	.04	.908	.908	.914	.914	.912	.0032	.0032	.0008	.0008
6	.02	.02	.04	.06	.942	.942	.917	.970	.947	.0032	.0032	.0009	.0002
7	.02	.02	.04	.08	.959	.959	.928	.992	.967	.0028	.0028	.0009	.0001
8	.02	.02	.06	.06	.977	.977	.966	.966	.969	.0035	.0035	.0003	.0003
9	.02	.02	.06	.08	.991	.991	.966	.989	.982	.0033	.0033	.0031	.0001
10	.02	.02	.08	.08	1.008	1.008	.988	.988	.992	.0033	.0033	.0001	.0001
11	.02	.04	.04	.04	1.003	.883	.883	.883	.901	.0047	.0014	.0014	.0014
12	.02	.04	.04	.06	1.063	.886	.886	.947	.931	.0049	.0016	.0016	.0005
13	.02	.04	.04	.08	1.066	.902	.902	.983	.956	.0044	.0016	.0016	.0001
14	.02	.04	.06	.06	1.145	.885	.938	.938	.950	.0055	.0020	.0006	.0006
15	.02	.04	.06	.08	1.151	.899	.939	.976	.968	.0052	.0021	.0007	.0002
16	.02	.04	.08	.08	1.172	.909	.974	.974	.981	.0052	.0023	.0002	.0002
17	.02	.06	.06	.06	1.269	.924	.924	.924	.960	.0064	.0009	.0009	.0009
18	.02	.06	.06	.08	1.289	.922	.922	.967	.973	.0063	.0010	.0010	.0003
19	.02	.06	.08	.08	1.331	.916	.961	.961	.982	.0066	.0012	.0004	.0004
20	.02	.08	.08	.08	1.401	.953	.953	.953	.990	.0070	.0005	.0005	.0005
21	.04	.04	.04	.04	.848	.848	.848	.848	.848	.0026	.0026	.0026	.0026
22	.04	.04	.04	.06	.861	.861	.861	.905	.876	.0031	.0031	.0031	.0011
23	.04	.04	.04	.08	.888	.888	.888	.962	.917	.0031	.0031	.0031	.0003
24	.04	.04	.06	.06	.882	.882	.889	.889	.886	.0040	.0040	.0014	.0014
25	.04	.04	.06	.08	.912	.912	.894	.948	.920	.0041	.0041	.0015	.0005
26	.04	.04	.08	.08	.943	.943	.943	.943	.943	.0044	.0044	.0006	.0006
27	.04	.06	.06	.06	.929	.867	.867	.867	.878	.0054	.0021	.0021	.0021
28	.04	.06	.06	.08	.976	.871	.871	.925	.906	.0057	.0024	.0024	.0009
29	.04	.06	.08	.08	1.035	.873	.916	.916	.925	.0063	.0028	.0011	.0011
30	.04	.08	.08	.08	1.119	.903	.903	.903	.935	.0073	.0014	.0014	.0014
31	.06	.06	.06	.06	.844	.844	.844	.844	.844	.0032	.0032	.0032	.0032
32	.06	.06	.06	.08	.854	.854	.854	.889	.865	.0037	.0037	.0037	.0015
33	.06	.06	.08	.08	.869	.869	.875	.875	.873	.0046	.0046	.0019	.0019
34	.06	.08	.08	.08	.901	.858	.858	.858	.867	.0059	.0026	.0026	.0026
35	.08	.08	.08	.08	.841	.841	.841	.841	.841	.0037	.0037	.0037	.0037

exhibited in Tables 13 and 14 which show the positive adjustment estimator as ensemble risk efficient relative to MLE for all error rate patterns and for both sample sizes. Auditors may be surprised to learn that an estimator exists which is everywhere greater than or equal to their traditional point estimator and yet appears to produce smaller mean squared errors, at least for the representative 35 error rate patterns examined.

As an alternative to the "positive adjustment" estimator, $\hat{p}_i^{JS+}$, the auditor could choose to use a Stein estimator which adjusts the MLE estimates toward some arbitrarily high value, M. Such an estimator is given by James and Stein [1961] and Efron and Morris [1975]:

$$\hat{p}_i^M = M \frac{(k-2)\sigma^2}{S} + \left(1 - \frac{(k-2)\sigma^2}{S}\right) \hat{p}_i$$

$$\text{Where } S = \sum_{i=1}^{k} (\hat{p}_i - M)^2$$

This estimator is guaranteed to be frequentist ensemble risk dominant over the MLE estimator since

$$\sum_{i=1}^{k} R(\hat{p}_i^M, \mathbf{p}) = k\sigma^2 \left(1 - \frac{(k-2)^2}{k - 2 - \Sigma(p_i - M)^2}\right) < k\sigma^2 = \sum_{i=1}^{k} R(\hat{p}_i, \mathbf{p})$$

This estimator shrinks the MLE estimates toward the prespecified value M. If the auditor chooses a value of M large enough, for example M = 1, the estimator will always revise the MLE estimates upward toward M. However, since the error rates to be estimated are low, it follows that S will be quite large. This results in only a minor adjustment to the MLE estimates and negligible risk savings.

Of course, the auditor could subjectively select any M toward which the MLE estimates would be adjusted. However, the subjective element of arbitrarily selecting M and the resulting uncertainty surrounding the ultimate risk savings are elements to be avoided in the design of estimators for audit testing.

More central to the issue of the auditor's apparent greater concern for overstatements than for understatements in error rate estimation is the central assumption of the symmetric quadratic form for the loss function. Scott [1975] investigates the form of the auditor's loss function using a consumption-investment model. In discussing his results he made the following observations about the shape of the loss function [p. 109]. "The general appearance is that of a quadratic. A closer look reveals, however, that the loss functions are not symmetric." He found a pronounced tendency for the loss to be greater for overstatements of net assets than for understatements of the same magnitude. His tests relate only to the loss from errors in estimation of final account balances. However, the differing levels of concern expressed in SAS no. 39 for overestimation and underestimation of error rates would seem to imply an asymmetric loss function for compliance testing as well.

The following is a summary of Scott's findings about the general shape of the loss functions when estimating net assets:

1. The general appearance is that of a quadratic.
2. The loss is zero when the correct amount is equal to the estimated amount.
3. The loss function tends to be negative when the correct amount is slightly less than the estimated amount.
4. The loss function is not symmetric. In particular it tends to rise faster when the actual amount is less than the estimated amount and slower when the actual amount is higher than the estimated amount.

The general shape of a loss function satisfying these four characteristics is given in Figure 6 (see also Scott [1975, Figure 5, p. 114]).

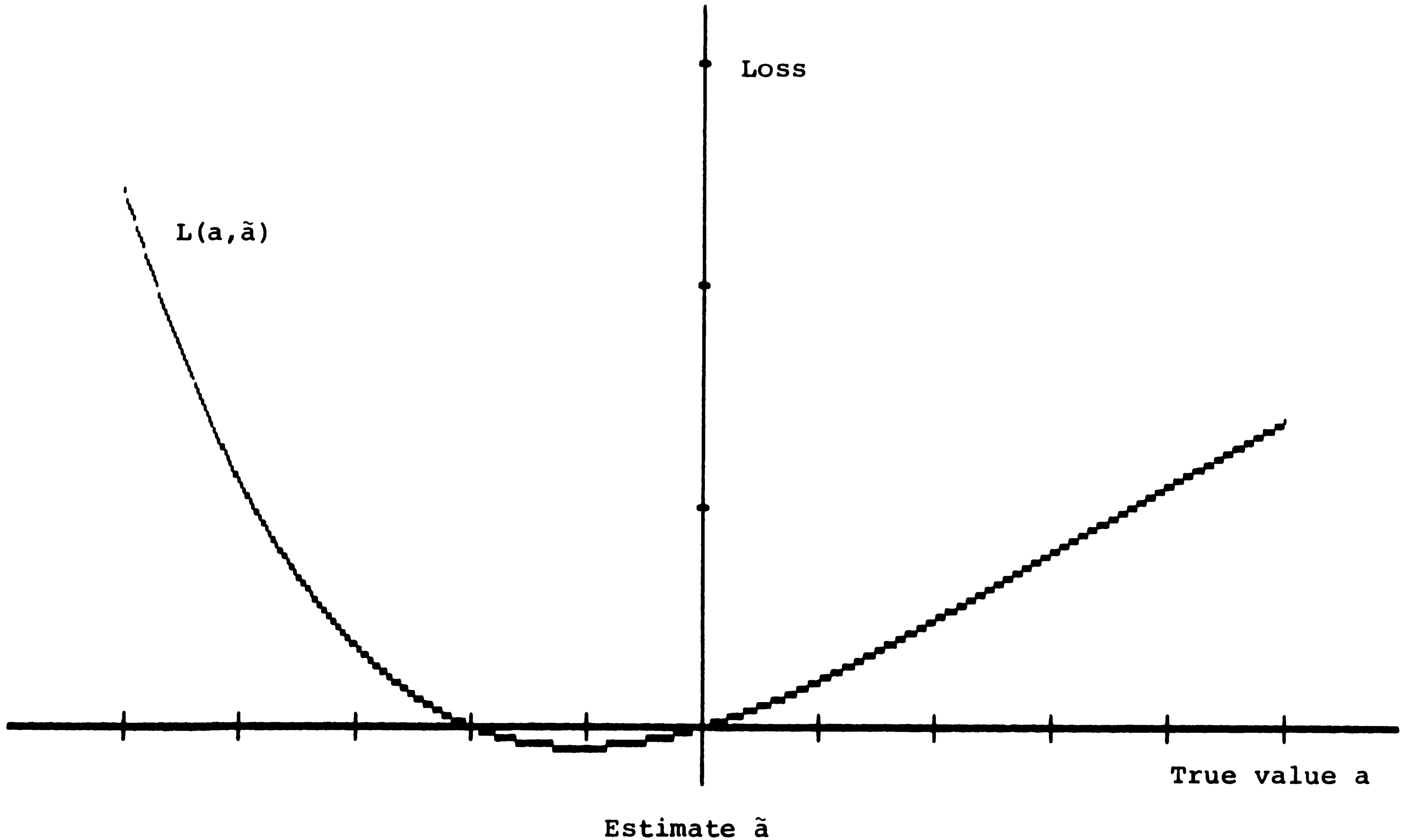
The loss function in Figure 6 relates to errors in the estimation of net assets during substantive testing. During compliance testing the direction of asymmetry would be reversed since the understatement of an error rate is viewed as more costly to the auditor than an overstatement of the same amount. Thus, an asymmetric loss function for compliance testing can be represented by the mirror image of Figure 6. Such an asymmetric compliance test loss function is shown in Figure 7.

Asymmetric loss functions can be incorporated into PEB estimation quite easily. A loss function which has the general shape of the curve given in Figure 6 is

$$L(\theta,\tilde{\theta}) = \delta(\theta - \tilde{\theta})(\theta - \tilde{\theta} + \Delta)\ e^{\alpha(\theta-\tilde{\theta})}$$

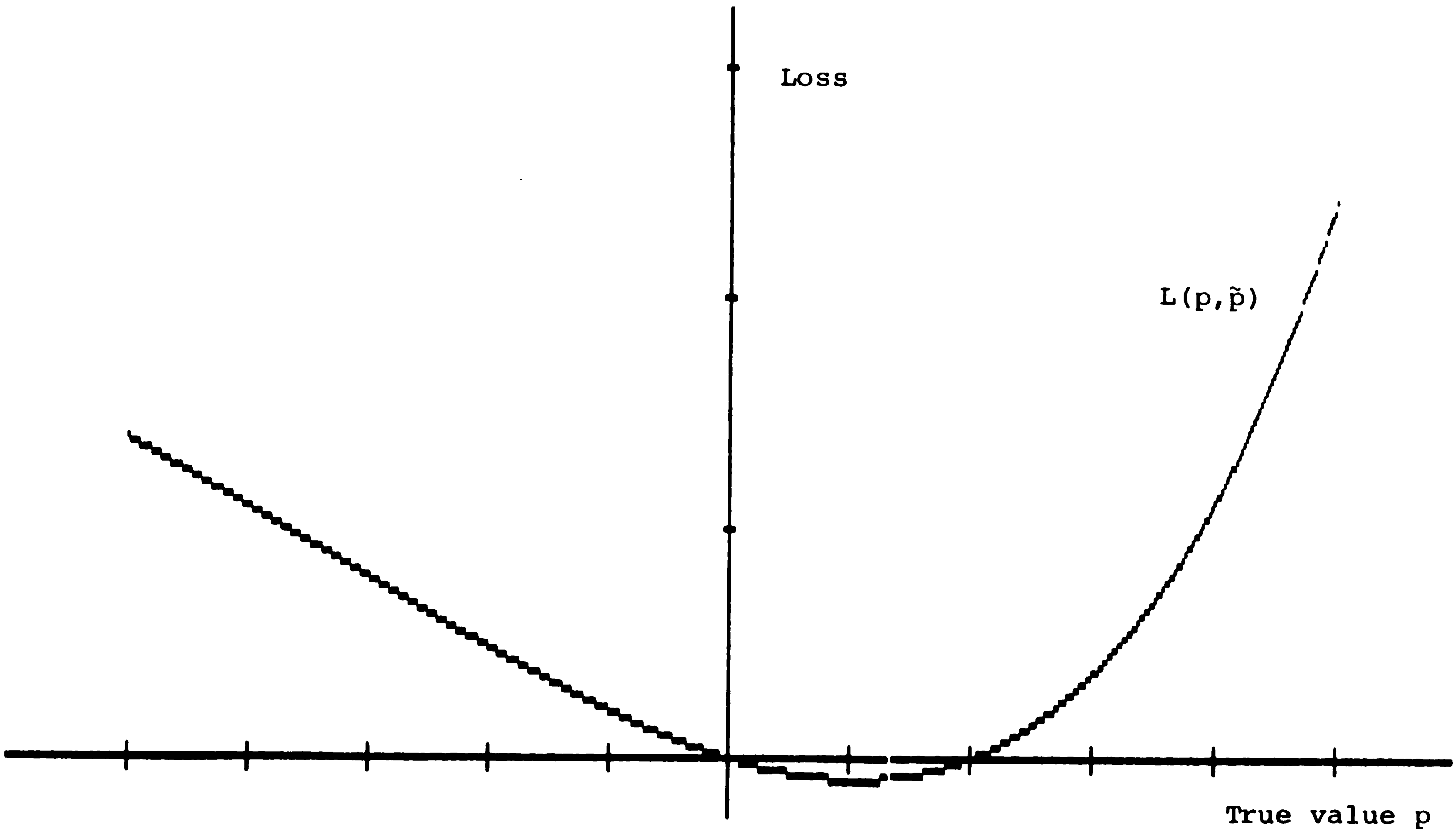
$$\text{for } \delta, \Delta, \alpha > 0.$$

The constant Δ shifts the loss function down and to the left. The results of Scott [1975] show the effects of Δ to



Asymmetric Loss Function in the Estimation of Net Assets

FIGURE 6



Asymmetric Loss Function in the Estimation of Internal Control Error Rates

FIGURE 7

be minimal. Thus, to simplify the analysis and to avoid the somewhat disconcerting notion of a negative loss function it will be assumed that $\Delta = 0$. The constant α controls the asymmetry of the loss function. In general, the larger the value of α the higher the penalty for understatements relative to overstatements with $\alpha = 0$ representing a symmetric loss function.

This asymmetric loss function may be generalized to the multivariate situation of simultaneous estimation of k population parameters by

$$(4) \qquad L(\boldsymbol{\theta},\tilde{\boldsymbol{\theta}}) = \Sigma\, \delta_i\, (\theta_i - \tilde{\theta}_i)^2\, e^{\alpha_i(\theta_i - \tilde{\theta}_i)} .$$

To find the Bayes estimate for this asymmetric loss function assume, as before

$$p_i \sim g(p_i) = \text{Normal } (\mu,\tau)$$

$$\hat{p}_i | p_i \sim f(\hat{p}_i | p_i) = \text{Normal } (p_i, \sigma^2)$$

so that

$$p_i | \hat{p}_i \sim g(p_i | \hat{p}_i) = \text{Normal } \left(\frac{\mu\sigma^2}{\tau + \sigma^2} + \frac{\hat{p}_i \tau}{\tau + \sigma^2}, \frac{\tau\sigma^2}{\tau + \sigma^2}\right).$$

The Bayes estimates are chosen to minimize the expected loss function conditional on the sample results:

$$\min \int \cdot\cdot \int \sum_{i=1}^{k} \delta_i (p_i - \tilde{p}_i)^2\, e^{\alpha_i(p_i - \tilde{p}_i)} \prod_{i=1}^{k} g(p_i | \tilde{p}_i)\, dp_i \cdot\cdot\, dp_k$$

The resulting Bayes estimator is

$$\mu \frac{\sigma^2}{\tau + \sigma^2} + p_i \frac{\tau}{\tau + \sigma^2} + \alpha_i \frac{\tau\sigma^2}{\tau + \sigma^2} + \frac{1}{\alpha_i} \left(1 - \left(1 - \alpha_i^2 \frac{\tau\sigma^2}{\tau + \sigma^2}\right)^{\frac{1}{2}}\right).$$

The asymmetric loss function adds a "conservatism factor" of

$$\alpha_i \frac{\tau\sigma^2}{\tau + \sigma^2} + \frac{1}{\alpha_i} \left(1 - \left(1 - \alpha_i^2 \frac{\tau\sigma^2}{\tau + \sigma^2}\right)^{\frac{1}{2}}\right)$$

to the traditional Bayes estimate. Clearly, the first two terms of the above expression can be evaluated using the same PEB techniques as under a symmetric loss function. The "conservation factor", can be estimated using PEB procedures by noting:

$$\frac{\tau\sigma^2}{\tau + \sigma^2} = \left(1 - \frac{\sigma^2}{\tau + \sigma^2}\right) \sigma^2$$

Since $\hat{B}^{PEB}$ is used in PEB estimation as a substitute for $\frac{\sigma^2}{\tau + \sigma^2}$, a reasonable choice for a PEB estimator with an asymmetric loss function is

$$\hat{p}_i^{PEB} = \hat{B}^{PEB} \bar{\hat{p}} + (1 - \hat{B}^{PEB}) \hat{p}_i + \alpha_i (1 - \hat{B}^{PEB}) \hat{\sigma}^2 + \frac{1}{\alpha_i} \left(1 - \left(1 - \alpha_i^2 (1 - \hat{B}^{PEB}) \hat{\sigma}^2\right)^{\frac{1}{2}}\right).$$

From the frequentist perspective an estimator is sought which minimizes the expected value of the asymmetric loss function conditional on the underlying population error rates and independent of the sampling results from other populations:

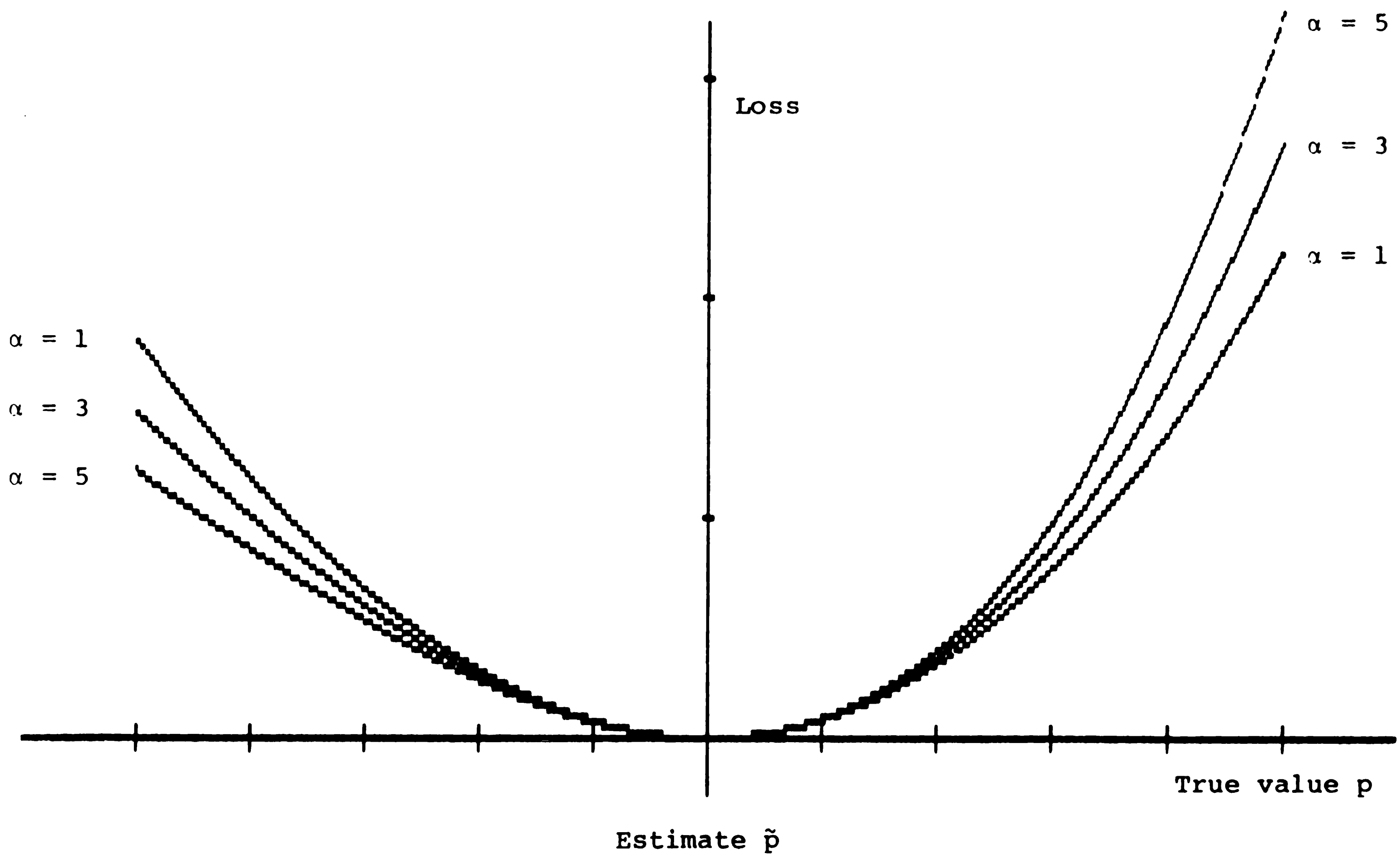
$$\min \int \cdots \int \sum_{i=1}^{k} \delta_i (p_i - \tilde{p}_i)^2 e^{\alpha_i (p_i - \tilde{p}_i)} \prod_{i=1}^{k} f(\tilde{p}_i | p_i) \, dp_i \cdots dp_k$$

The resulting frequentist estimator under the asymmetric loss function is

$$\hat{p}_i^F = \hat{p}_i + \alpha_i \hat{\sigma}^2 + \frac{1}{\alpha_i} (1- (1-\alpha_i^2 \hat{\sigma}^2)^{\frac{1}{2}}).$$

It remains to investigate whether the asymmetric loss PEB estimator constructed under the assumption of random error rates is nonetheless ensemble loss efficient with respect to the frequentist asymmetric estimator even under the frequentist persepctive of fixed population error rates. Such is the case, of course, for the JS/PEB estimator under a symmetric quadratic loss function (see Tables 3 and 4). Toward this end 210 Monte Carlo experiments were conducted computing the average ensemble efficiencies over 1,000 trials under 35 patterns of fixed population error rates, two sample sizes (n=50 and 100) and three sets of asymmetry parameters (α_i = 1, 3 and 5 for all i). Since both the frequentist and PEB estimators are independent of the loss function parameters δ_i, all Monte Carlo experiments were conducted with δ_i= 1 for all i without loss of generality. Figure 8 displays a graph of the three asymmetric loss functions tested. For each of the three asymmetry parameters loss is graphed as a function of the magnitude of over-(under-) estimation. All three loss functions are asymmetric with the penalty for underestimation relatively higher than that for overestimation. The larger the value of α, the larger the degree of asymmetry.

The results of these Monte Carlo experiments are given in Table 15. The results confirm the efficiency of the PEB



Asymmetric Loss Functions Used in Monte Carlo Experiments

FIGURE 8

TABLE 15

PEB Frequentist Efficiency, Asymmetric Loss Function

Case	Population Error Rates 1	2	3	4	Ensemble Efficiency n = 50 $\alpha_i=1$	n = 50 $\alpha_i=3$	n = 50 $\alpha_i=5$	n = 100 $\alpha_i=1$	n = 100 $\alpha_i=3$	n = 100 $\alpha_i=5$
1	.02	.02	.02	.02	.623	.596	.538	.621	.595	.571
2	.02	.02	.02	.04	.675	.629	.592	.740	.714	.691
3	.02	.02	.02	.06	.781	.740	.709	.880	.859	.841
4	.02	.02	.02	.08	.861	.820	.793	.968	.951	.939
5	.02	.02	.04	.04	.685	.638	.602	.758	.732	.709
6	.02	.02	.04	.06	.732	.679	.647	.845	.819	.797
7	.02	.02	.04	.08	.831	.789	.759	.937	.916	.900
8	.02	.02	.06	.06	.787	.741	.706	.882	.855	.833
9	.02	.02	.06	.08	.851	.811	.783	.912	.888	.869
10	.02	.02	.08	.08	.871	.832	.805	.945	.924	.908
11	.02	.04	.04	.04	.670	.624	.586	.730	.703	.678
12	.02	.04	.04	.06	.718	.673	.638	.806	.781	.759
13	.02	.04	.04	.08	.793	.752	.722	.882	.860	.842
14	.02	.04	.06	.06	.741	.694	.657	.824	.798	.775
15	.02	.04	.06	.08	.784	.742	.712	.889	.864	.844
16	.02	.04	.08	.08	.815	.773	.742	.904	.881	.862
17	.02	.06	.06	.06	.740	.696	.661	.820	.793	.770
18	.02	.06	.06	.08	.770	.724	.689	.867	.841	.818
19	.02	.06	.08	.08	.794	.753	.722	.883	.858	.837
20	.02	.08	.08	.08	.795	.750	.716	.878	.854	.834
21	.04	.04	.04	.04	.611	.567	.531	.632	.607	.584
22	.04	.04	.04	.06	.650	.608	.574	.704	.679	.656
23	.04	.04	.04	.08	.739	.697	.666	.828	.805	.787
24	.04	.04	.06	.06	.661	.615	.577	.712	.686	.663
25	.04	.04	.06	.08	.728	.682	.647	.804	.780	.759
26	.04	.04	.08	.08	.750	.707	.674	.845	.821	.800
27	.04	.06	.06	.06	.647	.603	.567	.693	.668	.645
28	.04	.06	.06	.08	.697	.650	.612	.763	.737	.715
29	.04	.06	.08	.08	.716	.671	.637	.786	.763	.744
30	.04	.08	.08	.08	.718	.673	.636	.788	.762	.739
31	.06	.06	.06	.06	.623	.580	.544	.647	.622	.600
32	.06	.06	.06	.08	.646	.605	.571	.686	.661	.640
33	.06	.06	.08	.08	.657	.613	.577	.701	.679	.660
34	.06	.08	.08	.08	.646	.606	.573	.688	.663	.641
35	.08	.08	.08	.08	.616	.573	.537	.639	.616	.595

estimator over the frequentist estimator in each of the 210 cases.

Clearly, the resulting PEB estimator is biased from both the frequentist and Bayesian perspectives. This overstatement has been intentionally built in as a mechanism to minimize the expected value of an asymmetric loss function. Note that under a symmetric quadratic loss function (α_i=0) the PEB estimator is unbiased from the Bayesian perspective (see Tables 11 and 12).

One drawback of the estimator is its dependence on the parameters of the loss function, α_i. More research into the specific nature of the auditor's loss function (along the lines of Scott [1975]) and the auditor's ability to specify parameters of an asymmetric loss function is needed before the estimator could be successfully implemented in practice.

3.3.3 Tests of Reliability

While not explicitly recognized by either Ijiri and Leitch [1980] or Matsumura and Tsui [1982], one major disadvantage of Stein-type estimators has been that until very recently no method existed for establishing confidence intervals about the estimates. Without such a procedure for comparing the estimate against the auditor's "tolerable error rate" as discussed in SAS 39, it is doubtful that any Stein-type estimator would receive wide acceptance by the profession.

Morris [1983a] considers the PEB/JS estimator in an empirical Bayes framework and proposes a technique for establishing empirical Bayes confidence intervals. Morris proposes equation (5) as a 100 x (1 - α)% confidence interval on the true but unknown error rate, p_i.

(5) $$\text{Probability } \{ \hat{p}_i^{PEB} - z_\alpha s_i \le p_i \le \hat{p}_i^{PEB} + z_\alpha s_i \} \ge 1-\alpha$$

Where: $s_i^2 = \frac{\bar{\bar{p}}(1-\bar{\bar{p}})}{n} (1 - \frac{k-1}{k} \hat{B}^{PEB}) + v(\hat{p}_i - \bar{\bar{p}})^2$

$v = \frac{2}{k-3} (\hat{B}^{PEB})^2$

z_α = 100 x (1 - α/2)th percentile of the standard normal distribution

Morris [1983a, 1983c] provide some evidence that the probability coverage is as claimed. In general, the actual probability coverage of (5) is

(6) $$\int\int\cdots\int g(p_1,\ldots,p_k) \sum_{j_1=0}^{n} \cdots \sum_{j_k=0}^{n} I(\hat{p}_i^{PEB} | j_1,\ldots,j_k) \cdot \prod_{h=1}^{k} \text{Prob}(J_h=j_h)\, dp_1 \cdots dp_k.$$

Where: $g(p_1,\cdots,p_k)$ = Actual joint distribution from which the underlying error rates arise, e.g., iid Normal (μ,τ)

$I(\hat{p}_i^{PEB} | j_1,\cdots,j_k)$ = 1 if $\hat{p}_i^{PEB} - z_\alpha s_i \le p_i \le \hat{p}_i^{PEB} + z_\alpha s_i$

= 0 otherwise

$\text{Prob}(J_h=j_h)$, $\hat{p}_i^{PEB}$ = As defined in (3)

z_α, s_i = As defined in (5)

Direct evaluation of (6) for various distributions on the underlying population error rates is intractable. However, Tables 16 and 17 give the results of a series of 72

TABLE 16

Monte Carlo Simulation of Classical, PEB and Bayes 95% Upper Confidence Limits -- Bayesian Perspective
Observed Reliability for 1,000 Trials
n = 50

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Normal (.04,.0001)	4	.998	.924	.931	.966	.989	.983	.998	.993	1.000	.994	1.000	.996
	6	.998	.926	.940	.980	.993	.992	1.000	.998	1.000	.999	1.000	.999
	8	.998	.923	.945	.982	.992	.994	1.000	.997	1.000	.999	1.000	1.000
(2) Normal (.04,.0004)	4	.995	.925	.781	.922	.892	.960	.960	.978	.986	.986	.996	.994
	6	.994	.924	.796	.930	.900	.965	.962	.984	.987	.993	.997	.997
	8	.992	.903	.801	.932	.903	.967	.962	.984	.988	.995	.997	.997
(3) Normal (.05,.0001)	4	.993	.928	.738	.934	.936	.973	.993	.988	.999	.996	.999	.997
	6	.992	.936	.752	.951	.940	.982	.993	.994	.999	.998	1.000	1.000
	8	.992	.934	.752	.957	.946	.987	.994	.996	1.000	.999	1.000	1.000
(4) Normal (.05,.0004)	4	.990	.924	.640	.863	.790	.926	.896	.963	.962	.982	.988	.991
	6	.990	.925	.634	.876	.793	.938	.905	.970	.964	.988	.988	.994
	8	.991	.911	.640	.881	.796	.939	.909	.973	.967	.988	.989	.996
(5) Normal (.06,.0001)	4	.982	.934	.391	.862	.746	.956	.942	.984	.995	.994	1.000	.997
	6	.978	.934	.378	.859	.742	.954	.944	.986	.994	.996	1.000	.999
	8	.980	.934	.397	.871	.752	.958	.947	.990	.995	.999	1.000	1.000
(6) Normal (.06,.0004)	4	.990	.935	.448	.771	.638	.874	.802	.939	.904	.972	.967	.990
	6	.984	.912	.443	.769	.634	.870	.796	.937	.908	.971	.965	.990
	8	.985	.914	.445	.775	.634	.879	.796	.941	.906	.974	.965	.989

TABLE 16 (Continued)

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(7) Normal (.07,.0001)	4	.972	.922	.108	.664	.377	.850	.740	.949	.942	.982	.994	.994
	6	.975	.941	.100	.676	.376	.870	.743	.963	.948	.990	.994	.998
	8	.982	.945	.110	.685	.382	.873	.747	.962	.946	.993	.994	.999
(8) Normal (.07,.0004)	4	.976	.921	.256	.622	.447	.757	.629	.862	.802	.931	.906	.968
	6	.984	.930	.264	.628	.442	.776	.634	.877	.798	.944	.904	.974
	8	.981	.924	.261	.630	.442	.771	.635	.875	.795	.938	.907	.975
(9) Normal (.08,.0001)	4	.980	.940	.014	.406	.106	.676	.376	.863	.748	.958	.949	.989
	6	.978	.942	.014	.413	.108	.674	.385	.864	.730	.960	.944	.990
	8	.979	.946	.011	.398	.099	.674	.371	.874	.742	.969	.949	.993
(10) Normal (.08,.0004)	4	.982	.932	.127	.432	.250	.604	.434	.755	.623	.870	.786	.936
	6	.981	.932	.127	.450	.261	.617	.440	.768	.640	.878	.805	.946
	8	.979	.924	.126	.446	.260	.618	.437	.768	.634	.879	.798	.946
(11) Uniform (.00, .08)	4	.992	.912	.707	.897	.832	.945	.948	.965	.989	.977	.993	.983
	6	.991	.904	.713	.901	.834	.952	.948	.975	.994	.987	.998	.991
	8	.992	.900	.713	.910	.843	.958	.955	.983	.998	.994	1.000	.998
(12) Exponential (.04)	4	.992	.906	.748	.832	.807	.864	.847	.888	.873	.908	.902	.924
	6	.994	.920	.757	.841	.809	.876	.850	.899	.883	.920	.909	.936
	8	.994	.913	.754	.845	.810	.875	.848	.901	.884	.920	.908	.936

TABLE 17

Monte Carlo Simulation of Classical, PEB and Bayes 95% Upper Confidence Limits -- Bayesian Perspective
Observed Reliability for 1,000 Trials
n = 100

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Normal (.04,.0001)	4	.976	.924	.933	.959	.988	.979	.998	.990	1.000	.994	1.000	.996
	6	.980	.917	.932	.968	.991	.986	.998	.994	1.000	.997	1.000	.999
	8	.977	.924	.940	.976	.992	.991	.999	.996	1.000	.998	1.000	.999
(2) Normal (.04,.0004)	4	.981	.920	.794	.926	.907	.959	.962	.979	.988	.987	.997	.992
	6	.980	.915	.810	.933	.912	.961	.964	.981	.989	.990	.997	.996
	8	.981	.918	.803	.934	.905	.963	.963	.980	.987	.989	.996	.994
(3) Normal (.05,.0001)	4	.972	.925	.750	.935	.940	.972	.992	.989	1.000	.997	1.000	.999
	6	.974	.930	.764	.943	.948	.977	.994	.991	1.000	.997	1.000	.999
	8	.976	.925	.768	.945	.947	.981	.994	.994	.999	.998	1.000	1.000
(4) Normal (.05,.0004)	4	.977	.932	.663	.889	.806	.940	.912	.970	.972	.985	.991	.992
	6	.971	.914	.653	.884	.805	.933	.908	.965	.964	.982	.988	.991
	8	.976	.906	.644	.886	.806	.935	.907	.966	.963	.984	.987	.992
(5) Normal (.06,.0001)	4	.976	.936	.409	.869	.754	.941	.943	.979	.995	.992	1.000	.998
	6	.974	.935	.412	.866	.752	.948	.944	.980	.992	.993	1.000	.998
	8	.972	.931	.473	.871	.761	.950	.952	.984	.994	.995	1.000	.999
(6) Normal (.06,.0004)	4	.974	.932	.459	.797	.646	.881	.801	.934	.908	.966	.966	.984
	6	.976	.926	.461	.804	.647	.886	.802	.940	.911	.973	.968	.988
	8	.974	.916	.453	.808	.651	.891	.810	.940	.915	.974	.966	.986

TABLE 17 (Continued)

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(7) Normal (.07,.0001)	4	.975	.942	.117	.728	.403	.869	.763	.949	.952	.984	.995	.996
	6	.974	.945	.125	.735	.408	.875	.759	.954	.947	.984	.993	.995
	8	.974	.931	.121	.730	.418	.877	.763	.954	.942	.986	.994	.997
(8) Normal (.07,.0004)	4	.973	.936	.280	.681	.454	.798	.652	.888	.808	.942	.918	.974
	6	.975	.934	.267	.683	.452	.808	.649	.892	.812	.948	.913	.977
	8	.971	.921	.266	.672	.447	.795	.639	.884	.801	.943	.908	.974
(9) Normal (.08,.0001)	4	.969	.936	.016	.505	.116	.720	.398	.871	.753	.952	.940	.982
	6	.970	.944	.013	.515	.119	.722	.402	.808	.758	.957	.952	.986
	8	.967	.932	.016	.507	.117	.717	.400	.872	.754	.953	.950	.987
(10) Normal (.08,.0004)	4	.968	.929	.136	.506	.268	.660	.442	.794	.636	.886	.801	.939
	6	.970	.930	.130	.505	.266	.660	.433	.787	.631	.882	.801	.943
	8	.968	.920	.134	.517	.271	.668	.450	.794	.646	.888	.805	.943
(11) Uniform (.00, .08)	4	.981	.922	.710	.908	.836	.946	.950	.970	.992	.983	.998	.990
	6	.980	.913	.724	.912	.845	.950	.952	.972	.994	.985	.999	.990
	8	.981	.922	.726	.911	.846	.951	.952	.975	.995	.986	1.000	.994
(12) Exponential (.04)	4	.987	.922	.752	.838	.804	.866	.844	.889	.876	.908	.899	.930
	6	.987	.927	.761	.860	.815	.888	.857	.910	.889	.926	.913	.941
	8	.986	.926	.767	.857	.810	.884	.853	.909	.886	.928	.910	.942

Monte Carlo experiments which simulated the behavior of the suggested confidence interval within the context of the low error rate attributes sampling problem faced by auditors during compliance testing. For each of the 1,000 Monte Carlo trials the classical 95% one-sided upper confidence limits were obtained in accordance with the table given in the AICPA audit guide, Audit Sampling [AICPA, 1983, p. 108]. A 95% one-sided confidence interval was calculated for the PEB/JS estimate using the obvious one-sided analogue to (5). This process was repeated 1,000 times for each of the nine underlying distributions and for sampling from k = 4, 6, and 8 populations with sample sizes of n = 50 and 100 to produce the results given in Tables 16 and 17.

The results of the Monte Carlo simulations show that the proposed confidence interval does appear to produce the desired coverage within the context of low error rate attributes sampling. It is significant to note that the method is able to exploit the efficiency of the PEB/JS estimator by providing the desired reliability with confidence intervals which are 30% to 40% narrower than those currently used by auditors using MLE, as shown in Tables 18 and 19.

Caution should be used in interpreting these confidence intervals. They should not be interpreted as frequentist confidence intervals in the sense of providing $100 \times (1 - \alpha)\%$ coverage for each population under a fixed set of underlying population error rates. Indeed, they may

TABLE 18

Monte Carlo Simulation of Classical, PEB and Bayes 95% Upper Confidence Limits -- Bayesian Perspective
Observed Average Width of Confidence Intervals
n = 50

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Normal (.04,.0001)	4	.079	.036	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.078	.030	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.079	.027	.015	.025	.015	.026	.015	.026	.015	.026	.015	.026
(2) Normal (.04,.0004)	4	.079	.038	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.079	.034	.015	.027	.015	.027	.015	.027	.015	.027	.015	.027
	8	.078	.031	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(3) Normal (.05,.0001)	4	.082	.040	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
	6	.082	.033	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
	8	.082	.030	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
(4) Normal (.05,.0004)	4	.082	.041	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
	6	.082	.036	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
	8	.082	.033	.016	.027	.016	.027	.016	.027	.016	.027	.016	.027
(5) Normal (.06,.0001)	4	.086	.043	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
	6	.086	.037	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
	8	.086	.032	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
(6) Normal (.06,.0004)	4	.086	.045	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
	6	.085	.039	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
	8	.085	.036	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028

TABLE 18 (Continued)

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(7) Normal (.07,.0001)	4	.089	.046	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	6	.089	.039	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	8	.089	.034	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
(8) Normal (.07,.0004)	4	.088	.048	.016	.028	.016	.028	.016	.028	.016	.028	.016	.028
	6	.089	.042	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	8	.089	.038	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
(9) Normal (.08,.0001)	4	.092	.049	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	6	.092	.041	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	8	.092	.037	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
(10) Normal (.08,.0004)	4	.092	.051	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	6	.092	.044	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
	8	.092	.040	.016	.029	.016	.029	.016	.029	.016	.029	.016	.029
(11) Uniform (.00, .08)	4	.077	.038	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.078	.034	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.078	.033	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(12) Exponential (.04)	4	.076	.039	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
	6	.076	.038	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.076	.037	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026

TABLE 19

Monte Carlo Simulation of Classical, PEB and Bayes 95% Upper Confidence Limits
Observed Average Width of Confidence Intervals
n = 100

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Normal (.04,.0001)	4	.048	.026	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
	6	.048	.022	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
	8	.048	.021	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
(2) Normal (.04,.0004)	4	.048	.028	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
	6	.048	.025	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
	8	.048	.025	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
(3) Normal (.05,.0001)	4	.051	.029	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
	6	.051	.024	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
	8	.051	.022	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
(4) Normal (.05,.0004)	4	.051	.030	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
	6	.051	.028	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
	8	.051	.026	.015	.024	.015	.024	.015	.024	.015	.024	.015	.024
(5) Normal (.06,.0001)	4	.054	.031	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
	6	.054	.026	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
	8	.054	.024	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
(6) Normal (.06,.0004)	4	.054	.033	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
	6	.054	.029	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025
	8	.054	.028	.015	.025	.015	.025	.015	.025	.015	.025	.015	.025

TABLE 19 (Continued)

Prior Distribution For Error Rates	K	Classical	PEB	Pure Bayes for Various Priors (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(7) Normal (.07,.0001)	4	.057	.033	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.057	.028	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.057	.025	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(8) Normal (.07,.0004)	4	.056	.035	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.056	.032	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.056	.030	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(9) Normal (.08,.0001)	4	.059	.035	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.059	.030	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.059	.026	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(10) Normal (.08,.0004)	4	.059	.037	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	6	.059	.033	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
	8	.059	.031	.015	.026	.015	.026	.015	.026	.015	.026	.015	.026
(11) Uniform (.00, .08)	4	.047	.028	.014	.023	.014	.023	.014	.023	.014	.023	.014	.023
	6	.047	.027	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
	8	.047	.026	.015	.023	.015	.023	.015	.023	.015	.023	.015	.023
(12) Exponential (.04)	4	.046	.029	.014	.022	.014	.022	.014	.022	.014	.022	.014	.022
	6	.046	.029	.014	.022	.014	.022	.014	.022	.014	.022	.014	.022
	8	.046	.028	.014	.022	.014	.022	.014	.022	.014	.022	.014	.022

perform very badly for certain fixed error rate patterns. Tables 20 and 21 show the results of exact reliability computations over a variety of fixed error rate patterns for sample sizes of 50 and 100, respectively. These exact reliability computations were made using equation (6) by recognizing that the underlying distribution on the error rates, $g(p_1,...,p_k)$, degenerates to one when evaluated at the fixed error rates and zero elsewhere. The results given in Tables 20 and 21 simply reflect the fact that the probability statement given in (5) can not be interpreted in a frequentist manner since the probability is taken over both the conditional sampling distribution of $\hat{p}_i$ and the prior underlying distribution on p_i.

Since the PEB/JS estimator does have a frequentist interpretation as a James-Stein estimator (see Section 2.3.1), it would be useful to have confidence intervals with frequentist interpretations. Of course, the traditional maximum likelihood upper confidence limits given in the audit guide [AICPA, 1983, pp. 108-109] could be used. However, these would not exploit the efficiency of the JS estimator. Accordingly, lower upper confidence limits which still provide the desired frequentist reliability are sought.

Since the exact conditional distribution of the JS estimator is not known, construction of confidence intervals based upon the traditional theory of mathematical statistics is impossible (Judge and Bock [1978, p. 310], Morris

TABLE 20

Frequentist Reliability of PEB 95% Upper Confidence Limit
n = 50

Case	Population Error Rates				Reliability			
	1	2	3	4	1	2	3	4
1	.02	.02	.02	.02	.929	.929	.929	.929
2	.02	.02	.02	.02	.968	.968	.968	.798
3	.02	.02	.02	.06	.986	.986	.986	.808
4	.02	.02	.02	.08	.994	.994	.994	.779
5	.02	.02	.04	.04	.986	.986	.865	.865
6	.02	.02	.04	.06	.994	.994	.911	.834
7	.02	.02	.04	.08	.998	.998	.942	.793
8	.02	.02	.06	.06	.998	.998	.863	.863
9	.02	.02	.06	.08	.999	.999	.890	.817
10	.02	.02	.08	.08	1.000	1.000	.846	.846
11	.02	.04	.04	.04	.994	.913	.913	.913
12	.02	.04	.04	.06	.998	.945	.945	.862
13	.02	.04	.04	.08	.999	.966	.966	.812
14	.02	.04	.06	.06	.999	.968	.891	.891
15	.02	.04	.06	.08	1.000	.980	.915	.839
16	.02	.04	.08	.08	1.000	.989	.867	.867
17	.02	.06	.06	.06	1.000	.916	.916	.916
18	.02	.06	.06	.08	1.000	.935	.935	.865
19	.02	.06	.08	.08	1.000	.950	.890	.890
20	.02	.08	.08	.08	1.000	.911	.911	.911
21	.04	.04	.04	.04	.949	.949	.949	.949
22	.04	.04	.04	.06	.970	.970	.970	.892
23	.04	.04	.04	.08	.982	.982	.982	.836
24	.04	.04	.06	.06	.983	.983	.918	.918
25	.04	.04	.06	.08	.990	.990	.938	.865
26	.04	.04	.08	.08	.995	.995	.891	.891
27	.04	.06	.06	.06	.991	.940	.940	.940
28	.04	.06	.06	.08	.995	.955	.955	.891
29	.04	.06	.08	.08	.998	.967	.914	.914
30	.04	.08	.08	.08	.999	.933	.933	.933
31	.06	.06	.06	.06	.958	.958	.958	.958
32	.06	.06	.06	.08	.970	.970	.970	.916
33	.06	.06	.08	.08	.979	.979	.935	.935
34	.06	.08	.08	.08	.986	.951	.951	.951
35	.08	.08	.08	.08	.963	.963	.963	.963

TABLE 21

Frequentist Reliability of PEB 95% Upper Confidence Limit
n = 100

Case	Population Error Rates 1	2	3	4	Reliability 1	2	3	4
1	.02	.02	.02	.02	.953	.953	.953	.953
2	.02	.02	.02	.04	.984	.984	.984	.836
3	.02	.02	.02	.06	.995	.995	.995	.767
4	.02	.02	.02	.08	.998	.998	.998	.818
5	.02	.02	.04	.04	.996	.996	.891	.891
6	.02	.02	.04	.06	.999	.999	.928	.809
7	.02	.02	.04	.08	1.000	1.000	.950	.823
8	.02	.02	.06	.06	1.000	1.000	.858	.858
9	.02	.02	.06	.08	1.000	1.000	.896	.838
10	.02	.02	.08	.08	1.000	1.000	.863	.863
11	.02	.04	.04	.04	.999	.932	.932	.932
12	.02	.04	.04	.06	1.000	.956	.956	.841
13	.02	.04	.04	.08	1.000	.969	.969	.827
14	.02	.04	.06	.06	1.000	.971	.885	.885
15	.02	.04	.06	.08	1.000	.980	.918	.846
16	.02	.04	.08	.08	1.000	.985	.874	.874
17	.02	.06	.06	.06	1.000	.917	.917	.917
18	.02	.06	.06	.08	1.000	.940	.940	.866
19	.02	.06	.08	.08	1.000	.954	.890	.892
20	.02	.08	.08	.08	1.000	.914	.914	.914
21	.04	.04	.04	.04	.961	.961	.961	.961
22	.04	.04	.04	.06	.975	.975	.975	.874
23	.04	.04	.04	.08	.982	.982	.982	.827
24	.04	.04	.06	.06	.984	.984	.915	.915
25	.04	.04	.06	.08	.989	.989	.941	.847
26	.04	.04	.08	.08	.992	.992	.881	.881
27	.04	.06	.06	.06	.991	.944	.944	.944
28	.04	.06	.06	.08	.994	.960	.960	.870
29	.04	.06	.08	.08	.996	.971	.902	.902
30	.04	.08	.08	.08	.998	.927	.927	.927
31	.06	.06	.06	.06	.964	.964	.964	.964
32	.06	.06	.06	.08	.974	.974	.974	.894
33	.06	.06	.08	.08	.981	.981	.924	.924
34	.06	.08	.08	.08	.986	.946	.946	.946
35	.08	.08	.08	.08	.963	.963	.963	.963

[1983a, p. 51]). It is in such circumstances that the bootstrap methods of Efron [1979, 1982] may be useful. For a non-technical introduction to the topic of bootstrap statistical methods see Diaconis and Efron [1983]. Marais, et.al. [1984] and Marais [1984] represent the first applications of the bootstrap method in accounting research.

The general problem which the bootstrapping method seeks to address is as follows. Given a random sample $\mathbf{X} = (X_1, \ldots, X_n)$ from some unknown probability distribution, F, estimate the sampling distribution of some prespecified random variable $R(\mathbf{X})$ on the basis of one realization of the random sampling procedure $\mathbf{x} = (x_1, \ldots x_n)$.

The bootstrap method consists of the following set of procedures (outlined by Efron [1979, p. 3]):

1. Construct a hypothetical population of size n consisting of the realized sample values $\mathbf{x} = (x_i, \ldots x_n)$. Construct a probability distribution, $\hat{F}$, for the population with each item in the population assigned equal probability, 1/n.

2. Draw a random sample with replacement from the hypothetical population under $\hat{F}$. Call this the bootstrap sample, $\mathbf{x}^* = (x_i^*, \ldots x_n^*)$.

3. Compute the variable at interest for the bootstrap sample, $R^*(\mathbf{x}^*)$.

4. Approximate the sampling distribution of $R(\mathbf{X})$ under F (or some attribute of the distribution, e.g., the mean or variance). Typically this is accomplished by replicating steps two and three a large number of times. The results of these Monte Carlo bootstrapping iterations are used to approximate the underlying distribution.

Efron [1982] proposes the "bootstrap t" method of constructing confidence intervals about statistics with unknown distributions. The following develops its use in constructing upper confidence limits for the JS estimator of population error rates.

Consider the normal based formula for the upper confidence limit for population proportions (Cochran, equation (3.19), [1977, pp. 57-58]):

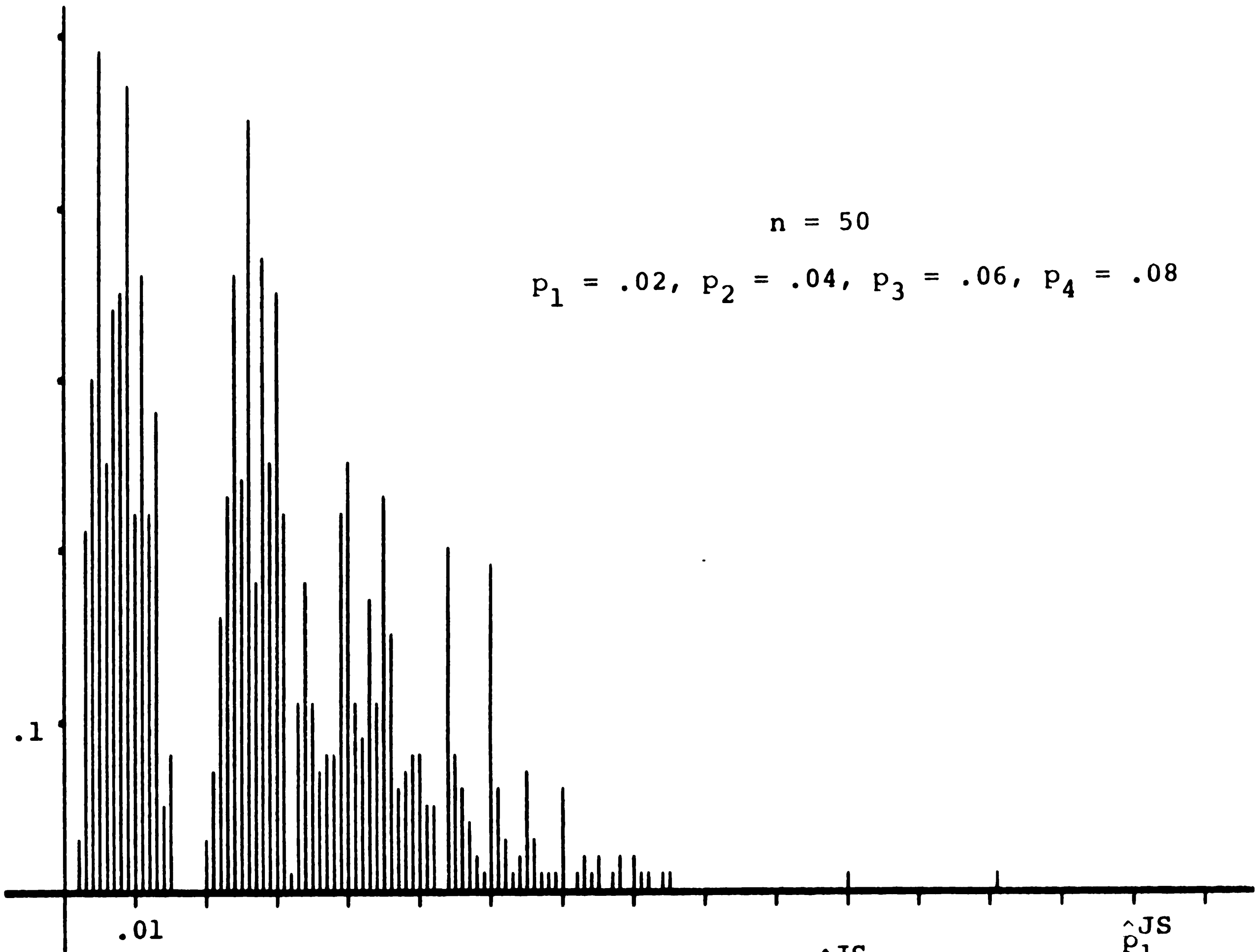
$$1 - \alpha \text{ upper confidence limit} = \hat{p} + z_{2\alpha}(\hat{p}(1-\hat{p})/(n-1))^{1/2} + 1/2n$$

A naive extension of this formula to the James-Stein estimator would be:

$$1 - \alpha \text{ upper confidence limit} = \hat{p}_i^{JS} + z_{2\alpha} V^{1/2} + 1/2n$$

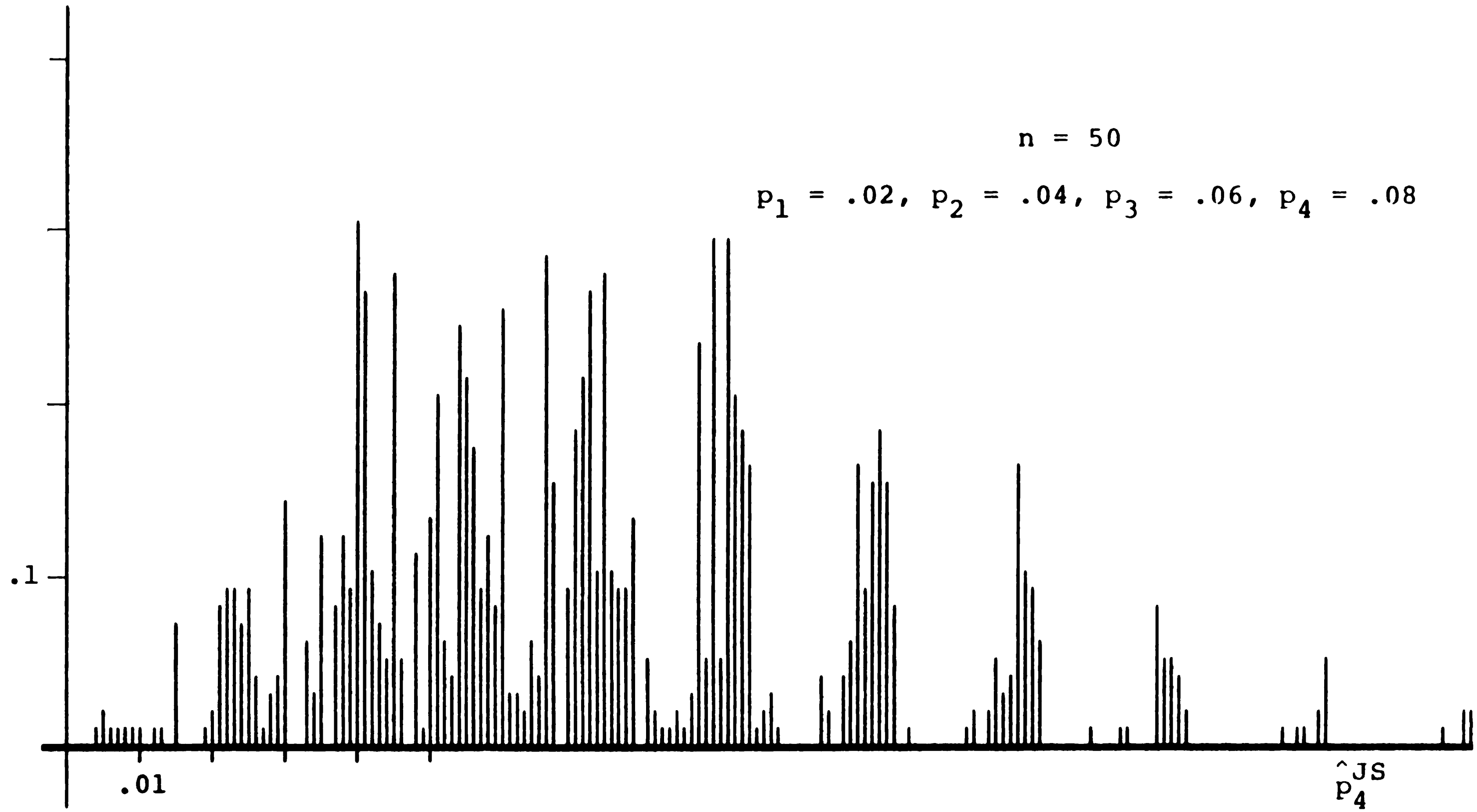
$$\text{Where } V = \hat{\bar{p}}\,(1 - \hat{\bar{p}})/n$$

A maintained assumption behind this formulation is that $\hat{p}_i^{JS}$ is distributed as a normal random variable with mean p_i and variance V. It is easily shown that this is not the case. As noted in Section 3.3.2, $\hat{p}_i^{JS}$ is in most circumstances a biased estimator of p_i. In addition, the distribution of $\hat{p}_i^{JS}$ is not symmetric as shown by Figures 9 and 10 which display the approximate probability density



Example of Sampling Density of $\hat{p}_1^{JS}$

FIGURE 9



Example of Sampling Density of $\hat{p}_4^{JS}$

FIGURE 10

function for $\hat{p}_1^{JS}$ and $\hat{p}_4^{JS}$ under one particular set of underlying population error rates. Note that in this example the distribution of $\hat{p}^{JS}$ is skewed to the right.

Since $T_i = (\hat{p}_i^{JS} - p_i) / V^{\frac{1}{2}}$ is not distributed as a standard normal random variable, it is not true that $\text{Prob} \{ p_i \leq \hat{p}_i^{JS} + z_{2\alpha} V^{\frac{1}{2}} + 1/2n \} = 1 - \alpha$. In lieu of $z_{2\alpha}$ the bootstrap t method seeks the appropriate value $t_i(\alpha)$ such that $\text{Prob} \{ p_i \leq \hat{p}_i^{JS} + t_i(\alpha) V^{\frac{1}{2}} + 1/2n \} = 1 - \alpha$. In general the true distribution of T_i remains unknown but can be estimated using the bootstrapping methods. In particular the 100 x αth percentile of T_i can be estimated by bootstrapping. Let $t_i^*(\alpha)$ be determined from a large number of bootstrap computations of $T_i^* = (p_i^{JS*} - \hat{p}_i) / (V^*)^{\frac{1}{2}}$ such that

$$\frac{\text{Number of bootstrap } T_i^*\text{'s} \leq t_i^*(\alpha)}{\text{Total number of bootstrap iterations}} = \alpha.$$

Thus,

$$\text{Prob} \{T_i = (\hat{p}_i^{JS} - p_i) / V^{\frac{1}{2}} \leq t_i^*(\alpha) \} \simeq \alpha$$

or

$$\text{Prob} \{ p_i \leq \hat{p}_i^{JS} - t_i^*(\alpha) V^{\frac{1}{2}} \} \simeq 1 - \alpha.$$

Incorporating the standard continuity correction term yields the following formula for the bootstrap t upper confidence limit:

(7) $1 - \alpha$ upper confidence limit $= \hat{p}_i^{JS} - t_i^*(\alpha) V^{\frac{1}{2}} + 1/2n$

Three operational difficulties arise when attempting to use (7) for the determination of small error rate upper confidence limits.

1. In the event that no errors are observed in the initial sample, that is $\hat{p}_i = 0$, it is unlikely that bootstapping will provide any meaningful information about the distribution of $\hat{p}_i^{JS}$ or T_i under the true underlying error rate $p_i > 0$. Since the bootstrap population degenerates to n elements, all zero, when $\hat{p}_i = 0$, no variation in $\hat{p}_i^*$ and very little variation in $\hat{p}_i^{JS*}$ will be observed among the repeated bootstrap iterations. Hence, little to no information about the distribution of $\hat{p}_i^{JS}$ can be expected to be extracted from bootstrapping procedures. Thus, when $\hat{p}_i = 0$ the traditional upper confidence limits should be used instead of equation (7).

2. It is possible that $V^* = 0$ on some bootstrap trials so that some T_i^*'s and hence $t_i^*(\alpha)$ may be unbounded. In such instances, equation (7) should again be abandoned in favor of the classical upper confidence limits.

3. It is possible that the upper confidence limit obtained from equation (7) exceeds the classical upper confidence limit. In such instances the lower classical limit can be used to improve efficiency without risking actual reliability levels less than the desired confidence levels.

Finally, the results of a series of Monte Carlo pilot studies showed that defining

$$T_i^* = (\hat{p}_i^{JS*} - \max\{\hat{p}_i , \bar{\hat{p}}\}) / (V^*)^{\frac{1}{2}}$$

results in more reliable upper confidence limits. This definition of T_i^* was used in developing the following results.

A series of 70 Monte Carlo experiments was performed to examine the frequentist behavior of the bootstrap t 95% upper confidence limits. For each of 35 error rate patterns simple random samples of 50 and 100 were simulated 500 times from each of the four populations. Traditional MLE 95% upper confidence limits were obtained using the tables in the AICPA audit guide, <u>Audit Sampling</u> [AICPA, 1983, p. 108]. Bootstrap t 95% upper confidence limits for the James-Stein estimator were constructed using the procedures outlined above with a total of 300 bootstrap iterations used to approximate the bootstrap distribution of T_i^*.

The observed frequentist reliabilities given in Tables 22 and 23 represent the percentage of trials out of the 500 Monte Carlo iterations that the computed 95% upper confidence limit exceeded the true population error rate using simulated sample sizes of n=50 and 100, respectively. Both the MLE and bootstrap t James-Stein upper confidence limits achieve the desired nominal level of 95% confidence. Table 24 gives the ratio of the average bootstrap t upper

TABLE 22

Monte Carlo Simulation of Classical and JS Bootstrap t 95% Upper Confidence Limits
Observed Reliability for 500 trials, n = 50

Case	Population Error Rates				Classical UCL Reliability				JS Bootstrap t UCL Reliability			
	1	2	3	4	1	2	3	4	1	2	3	4
1	.02	.02	.02	.02	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	.02	.02	.02	.04	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	.02	.02	.02	.06	1.000	1.000	1.000	.970	1.000	1.000	1.000	.944
4	.02	.02	.02	.08	1.000	1.000	1.000	.980	1.000	1.000	1.000	.938
5	.02	.02	.04	.04	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
6	.02	.02	.04	.06	1.000	1.000	1.000	.946	1.000	1.000	1.000	.938
7	.02	.02	.04	.08	1.000	1.000	1.000	.986	1.000	1.000	1.000	.944
8	.02	.02	.06	.06	1.000	1.000	.968	.962	1.000	1.000	.964	.958
9	.02	.02	.06	.08	1.000	1.000	.936	.982	1.000	1.000	.936	.944
10	.02	.02	.08	.08	1.000	1.000	.986	.982	1.000	1.000	.956	.942
11	.02	.04	.04	.04	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
12	.02	.04	.04	.06	1.000	1.000	1.000	.952	1.000	1.000	1.000	.948
13	.02	.04	.04	.08	1.000	1.000	1.000	.984	1.000	1.000	1.000	.936
14	.02	.04	.06	.06	1.000	1.000	.966	.950	1.000	1.000	.962	.946
15	.02	.04	.06	.08	1.000	1.000	.958	.980	1.000	1.000	.958	.948
16	.02	.04	.08	.08	1.000	1.000	.982	.990	1.000	1.000	.968	.976
17	.02	.06	.06	.06	1.000	.950	.946	.968	1.000	.950	.946	.968
18	.02	.06	.06	.08	1.000	.962	.964	.992	1.000	.962	.964	.974
19	.02	.06	.08	.08	1.000	.978	.988	.984	1.000	.978	.972	.966
20	.02	.08	.08	.08	1.000	.982	.990	.988	1.000	.968	.984	.976
21	.04	.04	.04	.04	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
22	.04	.04	.04	.06	1.000	1.000	1.000	.948	1.000	1.000	1.000	.946
23	.04	.04	.04	.08	1.000	1.000	1.000	.986	1.000	1.000	1.000	.960
24	.04	.04	.06	.06	1.000	1.000	.962	.940	1.000	1.000	.962	.940
25	.04	.04	.06	.08	1.000	1.000	.946	.992	1.000	1.000	.946	.970
26	.04	.04	.08	.08	1.000	1.000	.970	.984	1.000	1.000	.956	.966
27	.04	.06	.06	.06	1.000	.964	.946	.964	1.000	.964	.946	.964
28	.04	.06	.06	.08	1.000	.956	.944	.984	1.000	.956	.944	.966
29	.04	.06	.08	.08	1.000	.964	.984	.990	1.000	.962	.974	.980
30	.04	.08	.08	.08	1.000	.982	.986	.980	1.000	.980	.978	.972
31	.06	.06	.06	.06	.964	.966	.948	.960	.964	.966	.948	.960
32	.06	.06	.06	.08	.948	.934	.954	.992	.948	.934	.954	.986
33	.06	.06	.08	.08	.946	.972	.986	.988	.946	.972	.980	.984
34	.06	.08	.08	.08	.958	.984	.990	.984	.958	.982	.986	.982
35	.08	.08	.08	.08	.988	.990	.986	.974	.986	.990	.982	.974

TABLE 23

Monte Carlo Simulation of Classical and JS Bootstrap t 95% Upper Confidence Limits
Observed Reliability for 500 trials, n = 100

Case	Population Error Rates				Classical UCL Reliability				JS Bootstrap t UCL Reliability			
	1	2	3	4	1	2	3	4	1	2	3	4
1	.02	.02	.02	.02	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	.02	.02	.02	.04	1.000	1.000	1.000	.992	1.000	1.000	1.000	.962
3	.02	.02	.02	.06	1.000	1.000	1.000	.980	1.000	1.000	1.000	.940
4	.02	.02	.02	.08	1.000	1.000	1.000	.988	1.000	1.000	1.000	.932
5	.02	.02	.04	.04	1.000	1.000	.992	.968	1.000	1.000	.968	.948
6	.02	.02	.04	.06	1.000	1.000	.984	.984	1.000	1.000	.984	.944
7	.02	.02	.04	.08	1.000	1.000	.988	.972	1.000	1.000	.984	.940
8	.02	.02	.06	.06	1.000	1.000	.978	.980	1.000	1.000	.958	.956
9	.02	.02	.06	.08	1.000	1.000	.984	.980	1.000	1.000	.962	.926
10	.02	.02	.08	.08	1.000	1.000	.966	.964	1.000	1.000	.938	.926
11	.02	.04	.04	.04	1.000	.988	.988	.992	1.000	.984	.984	.992
12	.02	.04	.04	.06	1.000	.980	.984	.996	1.000	.980	.984	.968
13	.02	.04	.04	.08	1.000	.972	.992	.976	1.000	.972	.992	.940
14	.02	.04	.06	.06	1.000	.992	.988	.988	1.000	.992	.980	.976
15	.02	.04	.06	.08	1.000	.982	.998	.976	1.000	.982	.984	.950
16	.02	.04	.08	.08	1.000	.988	.968	.956	1.000	.988	.964	.948
17	.02	.06	.06	.06	1.000	.992	.996	.984	1.000	.988	.992	.980
18	.02	.06	.06	.08	1.000	.984	.968	.956	1.000	.968	.968	.932
19	.02	.06	.08	.08	1.000	.988	.948	.960	1.000	.988	.940	.948
20	.02	.08	.08	.08	1.000	.964	.952	.974	1.000	.952	.944	.960
21	.04	.04	.04	.04	.978	.982	.980	.974	.978	.980	.980	.974
22	.04	.04	.04	.06	.992	.980	.988	.996	.992	.980	.988	.976
23	.04	.04	.04	.08	.996	.984	.988	.968	.996	.984	.988	.940
24	.04	.04	.06	.06	.988	.988	.988	.984	.988	.988	.976	.976
25	.04	.04	.06	.08	.984	.984	.992	.964	.984	.984	.988	.936
26	.04	.04	.08	.08	.984	.996	.960	.974	.984	.996	.952	.952
27	.04	.06	.06	.06	.984	.980	.984	.980	.984	.980	.980	.980
28	.04	.06	.06	.08	.976	.976	.988	.964	.976	.972	.980	.948
29	.04	.06	.08	.08	.988	.992	.948	.956	.988	.992	.940	.944
30	.04	.08	.08	.08	.964	.976	.968	.944	.964	.976	.968	.944
31	.06	.06	.06	.06	.988	.996	1.000	.984	.988	.996	1.000	.984
32	.06	.06	.06	.08	.976	.992	.988	.960	.976	.992	.988	.948
33	.06	.06	.08	.08	.984	.984	.964	.960	.984	.984	.960	.960
34	.06	.08	.08	.08	.992	.956	.980	.976	.992	.956	.980	.976
35	.08	.08	.08	.08	.972	.956	.976	.968	.972	.956	.976	.968

TABLE 24

Monte Carlo Simulation of Classical and JS Bootstrap t 95% Upper Confidence Limits
Ratio of Average JS Bootstrap t to Classical Upper Confidence Limit Widths for 500 Trials

Case	Population Error Rates				n = 50				n = 100			
	1	2	3	4	1	2	3	4	1	2	3	4
1	.02	.02	.02	.02	.852	.853	.846	.850	.795	.793	.791	.790
2	.02	.02	.02	.04	.848	.845	.849	.870	.835	.841	.831	.800
3	.02	.02	.02	.06	.834	.851	.839	.860	.867	.885	.885	.864
4	.02	.02	.02	.08	.870	.861	.869	.890	.916	.920	.926	.911
5	.02	.02	.04	.04	.847	.843	.814	.886	.875	.878	.787	.780
6	.02	.02	.04	.06	.855	.860	.805	.814	.921	.909	.820	.845
7	.02	.02	.04	.08	.877	.876	.808	.857	.951	.943	.844	.881
8	.02	.02	.06	.06	.889	.889	.819	.819	.941	.945	.821	.821
9	.02	.02	.06	.08	.908	.908	.810	.844	.966	.965	.823	.860
10	.02	.02	.08	.08	.920	.920	.826	.829	.983	.982	.857	.854
11	.02	.04	.04	.04	.874	.808	.800	.805	.910	.794	.790	.793
12	.02	.04	.04	.06	.885	.817	.813	.817	.945	.832	.827	.810
13	.02	.04	.04	.08	.907	.827	.837	.830	.972	.870	.863	.844
14	.02	.04	.06	.06	.913	.824	.816	.802	.969	.868	.810	.828
15	.02	.04	.06	.08	.855	.846	.805	.817	.983	.900	.832	.839
16	.02	.04	.08	.08	.945	.868	.815	.815	.985	.925	.839	.833
17	.02	.06	.06	.06	.938	.806	.805	.798	.983	.811	.815	.821
18	.02	.06	.06	.08	.946	.826	.812	.800	.988	.850	.849	.838
19	.02	.06	.08	.08	.969	.835	.817	.802	.999	.869	.841	.848
20	.02	.08	.08	.08	.968	.817	.814	.829	.996	.830	.829	.838
21	.04	.04	.04	.04	.806	.802	.802	.809	.813	.810	.817	.808
22	.04	.04	.04	.06	.820	.824	.825	.797	.852	.849	.858	.788
23	.04	.04	.04	.08	.843	.845	.837	.804	.888	.886	.889	.818
24	.04	.04	.06	.06	.842	.844	.800	.796	.878	.880	.785	.777
25	.04	.04	.06	.08	.865	.868	.809	.798	.915	.901	.833	.806
26	.04	.04	.08	.08	.884	.883	.797	.803	.936	.930	.807	.815
27	.04	.06	.06	.06	.861	.806	.802	.810	.898	.812	.827	.809
28	.04	.06	.06	.08	.891	.825	.825	.811	.923	.833	.837	.783
29	.04	.06	.08	.08	.915	.835	.803	.807	.953	.883	.802	.809
30	.04	.08	.08	.08	.928	.824	.810	.812	.968	.836	.837	.819
31	.06	.06	.06	.06	.820	.813	.815	.812	.826	.827	.823	.829
32	.06	.06	.06	.08	.837	.830	.831	.807	.857	.867	.858	.789
33	.06	.06	.08	.08	.847	.855	.803	.806	.882	.874	.816	.814
34	.06	.08	.08	.08	.871	.822	.823	.816	.902	.827	.830	.818
35	.08	.08	.08	.08	.823	.820	.838	.827	.842	.840	.843	.849

confidence limits to the average MLE confidence limit for the 500 Monte Carlo trials. The results show that the bootstrap t method is able to partially exploit the efficiency of the James-Stein estimator by providing upper confidence limits which are consistently lower than the MLE limits while still providing the desired frequentist reliability.

CHAPTER IV

THE BEHAVIOR OF PARAMETRIC EMPIRICAL BAYES ESTIMATORS IN SUBSTANTIVE TESTING

The purpose of this chapter is to consider the application of parametric empirical Bayes estimators within the context of substantive audit testing. Section 4.1 considers the use of parametric empirical Bayes estimators for interitem integration of statistical sampling substantive test results. Section 4.2 considers the use of parametric empirical Bayes estimators for interprocedural integration of analytical review and statistical sampling substantive test procedures.

4.1 Interitem Integration of Variables Sampling Results

As discussed in Section 2.1, classical variables sampling procedures are often used by auditors during substantive testing to estimate the correct or "audited" account balance for comparison with the client's reported book balance. Typically a separate sampling plan is developed for each account balance or company division to be tested. Integration of the results across these account populations is done only subjectively through the auditor's judgmental evaluation of the results. However, as was the case for compliance test applications discussed in Chapter III, pure Bayesian or parametric empirical Bayesian procedures could be used to combine the results of several classical variables sampling estimators for k different account populations.

Define θ_i as the ratio of the true account balance to the client's reported balance for the ith account. The auditor estimates θ_i by $\hat{\theta}_i$ using some classical variables sampling procedure. This estimate is either made directly by use of a ratio estimator or indirectly by estimating the true account balance using a difference or mean-per-unit estimator so that $\hat{\theta}_i$ represents the ratio of the auditor's estimate of the true account balance to the client's reported balance.

Pure Bayesian estimation of θ_i for each of k different populations can be accomplished in the following manner (see also Section 2.2.3):

(8) Auditor's prior belief: $\theta_i \overset{\sim}{iid}$ Normal (μ, τ)

Sampling distribution: $\hat{\theta}_i \sim$ Normal (θ_i, σ_i^2)

Loss function: $L(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})' \mathbf{C} (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})$

Pure Bayes estimator: $\hat{\theta}_i^B = \dfrac{\sigma_i^2}{\tau + \sigma_i^2} \mu + \dfrac{\tau}{\tau + \sigma_i^2} \hat{\theta}_i$

The Bayesian estimator is developed to minimize the expected loss, $E\{L(\theta, \hat{\theta})\}$. The resulting estimator is independent of the loss function's k x k weighting matrix, **C**. Thus, Bayesian estimation behaves as if the loss function was composed of the squared estimation errors of the ratios of true to reported balances. In truth the auditor's loss function may be more properly represented by ratio estimation errors weighted by the size of the accounts, i.e. the dollar amount of the estimation errors. However, since the Bayes estimator is independent of the weighting matrix, **C**, the Bayesian auditor can proceed "as if" his loss function was based on squared errors in the estimates of the ratios themselves rather than on estimation errors of account balances.

As demonstrated in Chapter III, auditor misspecifications in the parameters of the prior distribution, μ and τ, can be quite costly in terms of reduced efficiency and reliability and increased bias. Parametric empirical Bayes estimation procedures avoid these misspecification costs by substituting estimates for the prior distribution parameters which are based on the sampling results.

In developing the PEB estimator for attribute sampling applications a simplifying assumption was utilized. This assumption was that the variances of the sampling distributions, σ_i^2, were nearly equal. This assumption is not unreasonable when performing attributes sampling in populations whose error rates are close. The reasonableness of the assumption is born out by the success of the PEB estimator in the results of the simulation analyses discussed in Chapter III.

However, the assumption of equal sampling variances in variables sampling applications is less reasonable. The distribution of the variables sampling estimator is much more complex than the binomial sampling distribution found in error rate estimation. In particular the distribution depends upon both the proportion of items in the population in error and the shape of the distribution of the dollar size of the error given one exists. For example, Neter and Loebbecke [1975, p. 45] in their landmark study into the behavior of variables sampling estimators found that the popular ratio estimator had sampling variances which differed by up to a factor of 64 for two accounting populations with the same overall error rate. This dramatic difference in variances was attributable to the different shapes of the distributions of dollar errors for those population items whose book and audit values differed.

Morris [1983a] generalizes the PEB estimator discussed in Chapter III to the case of unequal sampling variances. This estimator is given by:

$$(9)\quad \hat{\theta}_i^{PEB} = \hat{B}_i^{PEB}\,\bar{\hat{\theta}}_i + (1 - \hat{B}_i^{PEB})\,\hat{\theta}_i$$

$$\text{Where:}\quad \bar{\hat{\theta}} = \left(\sum_{i=1}^{k} W_i\,\hat{\theta}_i\right) \Big/ \sum_{i=1}^{k} W_i$$

$$W_i = 1 / (\hat{\sigma}_i^2 + \hat{\tau}^+)$$

$$\hat{\tau} = \sum_{i=1}^{k} W_i \{(k/(k-1))(\hat{\theta}_i - \bar{\hat{\theta}})^2 - \hat{\sigma}_i^2\} / \sum_{i=1}^{k} W_i$$

$$\hat{\tau}^+ = \max\{0, \hat{\tau}\}$$

$$\hat{B}_i^{PEB} = ((k-3)/(k-1))\,\hat{\sigma}_i^2 / (\hat{\sigma}_i^2 + \hat{\tau}^+).$$

Since the values of $\hat{\tau}$ and W_i are dependent upon each other, they must be determined by "guessing" at $\hat{\tau}$, computing the weights, W_i, using these weights to improve the estimate of $\hat{\tau}$, and iterating the process to convergence. Convergence to a relative change in $\hat{\tau}$ of less than .001 generally took less than a dozen iterations in the numerous simulations performed in this study.

Morris goes on to propose the following as a 100 x (1 - α)% parametric empirical Bayes confidence interval:

$$(10)\quad \text{Probability}\,\{\hat{\theta}_i^{PEB} - z_\alpha s_i \le \theta_i \le \hat{\theta}_i^{PEB} + z_\alpha s_i\} \ge 1 - \alpha$$

$$\text{Where:}\quad s_i^2 = \hat{\sigma}_i^2\,[(W_i / \Sigma W_i)\,\hat{B}_i^{PEB}] + v_i\,(\hat{\theta}_i - \bar{\hat{\theta}})^2$$

$$v_i = (2/(k-3))(\hat{B}_i^{PEB})^2\,(\bar{\hat{\sigma}}^2 + \hat{\tau}^+) / (\hat{\sigma}_i^2 + \hat{\tau}^+)$$

$$\bar{\hat{\sigma}}^2 = (\Sigma W_i\,\hat{\sigma}_i^2) / \Sigma W_i$$

The generalized PEB estimator and confidence interval, equations (9) and (10), reduce to equations (1) and (5) of Chapter III in the case of equal sampling variances.

In order to test the behavior of the PEB variables sampling estimator and its confidence interval a series of Monte Carlo experiments was performed. These Monte Carlo experiments compared the efficiency and reliability of the PEB estimator to a set of pure Bayes estimators and to a classical variables sampling estimator. A description of these Monte Carlo experiments and their results follows.

The classical variables sampling estimator chosen for study in these Monte Carlo experiments was the probability proportional to size mean-per-unit estimator (Neter and Loebbecke [1975, p. 117], Cochran [1977, p. 252], Roberts [1978, p. 116], Arens and Loebbecke [1981, p. 298] and Bailey [1981, p. 184]). The estimator is defined in the following manner. A sample of n items is drawn from the population with the probability of selection proportional to the recorded value of the item. Define y_j as the recorded value of the jth sample item and x_j as its audited value. The probability proportional to size mean-per-unit (PPS-MPU) estimator is then defined as

$$\hat{\theta} = 1/n \sum_{j=1}^{n} x_j \; / \; y_j .$$

This differs slightly from the traditional ratio estimator where the sample items are drawn using sample random sampling and the estimator is defined as

$$\hat{\theta} = \sum_{j=1}^{n} x_j \Big/ \sum_{j=1}^{n} y_j .$$

The behavior of the PPS-MPU estimator has received only limited study in the auditing literature. Two exceptions are Neter and Loebbecke [1975] and Garstka and Ohlson [1979]. Nonetheless, the estimator was used in this study for several reasons. First, it is an unbiased estimator while the more traditional and widely studied ratio estimator is not. Secondly, the results of Neter and Loebbecke [1975] showed the PPS-MPU estimator to have standard errors which were never substantially greater than those for the traditional ratio estimator and which were less than one half of the ratio estimator's standard error in certain applications. Additionally, the PPS procedure is in itself an attractive sample selection method for the auditor. The procedure is an alternative to stratifying the population by recorded amounts to give greater weight to items with large book values. In fact, Neter, Leitch and Feinberg [1978, p. 78] note that PPS sampling "can be thought of as employing the ultimate stratification by book value." Finally, because of the form of the estimator and its sample selection procedure, it is particularly tractable for Monte Carlo study. This is because only the error rate and error tainting distributions need to be specified and not the distribution of the book values themselves.

As noted, a study into the behavior of a variables sampling estimator requires the specification of the error rate and error tainting distributions. These error characteristics of accounting populations have been studied by Ramage, Krieger and Spero [1979] and Johnson, Leitch and Neter [1981] using substantially the same data base. A summary of their relevant findings and their application to this work follows.

Johnson, Leitch and Neter [1981] studied the distribution of error rates for 55 audits of accounts receivable populations and 26 audits of inventory populations. Table 25 presents the distribution of error rates observed for these two types of populations. An obvious difference between the two distributions is the tendency for inventory populations to have higher error rates than accounts receivable populations. The median inventory error rate is greater than .14 while the median accounts receivable error rate is less than .024. The low error rates observed in accounts receivable populations severly impair the performance of classical statistical estimators without exceedingly large sample sizes. For this reason the behavior of the PPS-MPU estimator and its pure Bayes and PEB revised estimators will only be studied within the context of inventory audits.

Table 26 presents the distribution of the proportion of errors in the inventory populations which were observed to be overstatement errors in the Johnson, Leitch and Neter sample of inventory audits. This split of errors between

TABLE 25

Johnson, Leitch and Neter [1981] Observed Distribution of Error Rates for Accounts Receivable and Inventory Audits

	Error Rate	Percent of Audits
Accounts Receivable		
	.000 - .005	29.1%
	.005 - .024	21.9
	.025 - .088	23.6
	.089 - .120	10.9
	.121 - .163	10.9
	.164 - .266	1.8
	.267 - .861	1.8
Inventory		
	.000 - .074	26.9%
	.074 - .139	19.3
	.140 - .296	23.1
	.297 - .416	11.5
	.417 - .645	7.7
	.646 - .758	11.5

TABLE 26

Johnson, Leitch and Neter [1981] Observed Distribution of Overstatement/Understatement Errors in Inventory Audits

Proportion of Errors Which Were Overstatements	Percent of Audits
.10 - .19	3.8%
.20 - .29	3.8
.30 - .39	7.7
.40 - .49	7.7
.50 - .59	26.9
.60 - .69	19.2
.70 - .79	11.5
.80 - .89	7.7
.90 - .90	3.8
1.0	7.7

overstatements and understatements partially defines the tainting pattern of the errors. The remaining characteristic defining the error tainting pattern is the distribution of the size of the understatements and overstatements. Johnson, Leitch and Neter analyze the range, mean, standard deviation and skewness of the observed error amounts to provide evidence about the shape of the error tainting distribution. In general, their results show the distributions "are frequently highly positively skewed, and have heavy concentrations in a comparatively small portion of the entire range" [Johnson, Leitch and Neter 1981, p. 28]. In subsequent Monte Carlo simulation studies of a dollar unit sampling estimator Leitch, et.al. [1982] chose to model the error tainting pattern using chi-square probability distributions based on the general patterns uncovered by Ramage, Krieger and Spero [1979] and Johnson, Leitch and Neter [1981]. Dworin and Grimlund [1984] recently used these same distributions in their dollar unit sampling simulation studies. In particular, Dworin and Grimlund [1984] modeled the distribution of overstatement percentage errors as .1 times a chi-square random variable with either 1, 2 or 3 degrees of freedom and understatement percentage errors as .1 times a chi-square random variable with either 1 or 2 degrees of freedom. These same tainting patterns (six possible combinations in total) are used in this study.

Within this framework the following parameters completely define the characteristics of the accounting population.

r_t = total error rate

r_o = percentage of errors which are overstatements

r_u = percentage of errors which are understatements = $1 - r_o$

df_o = degrees of freedom for chi-square overstatement distribution (1, 2 or 3)

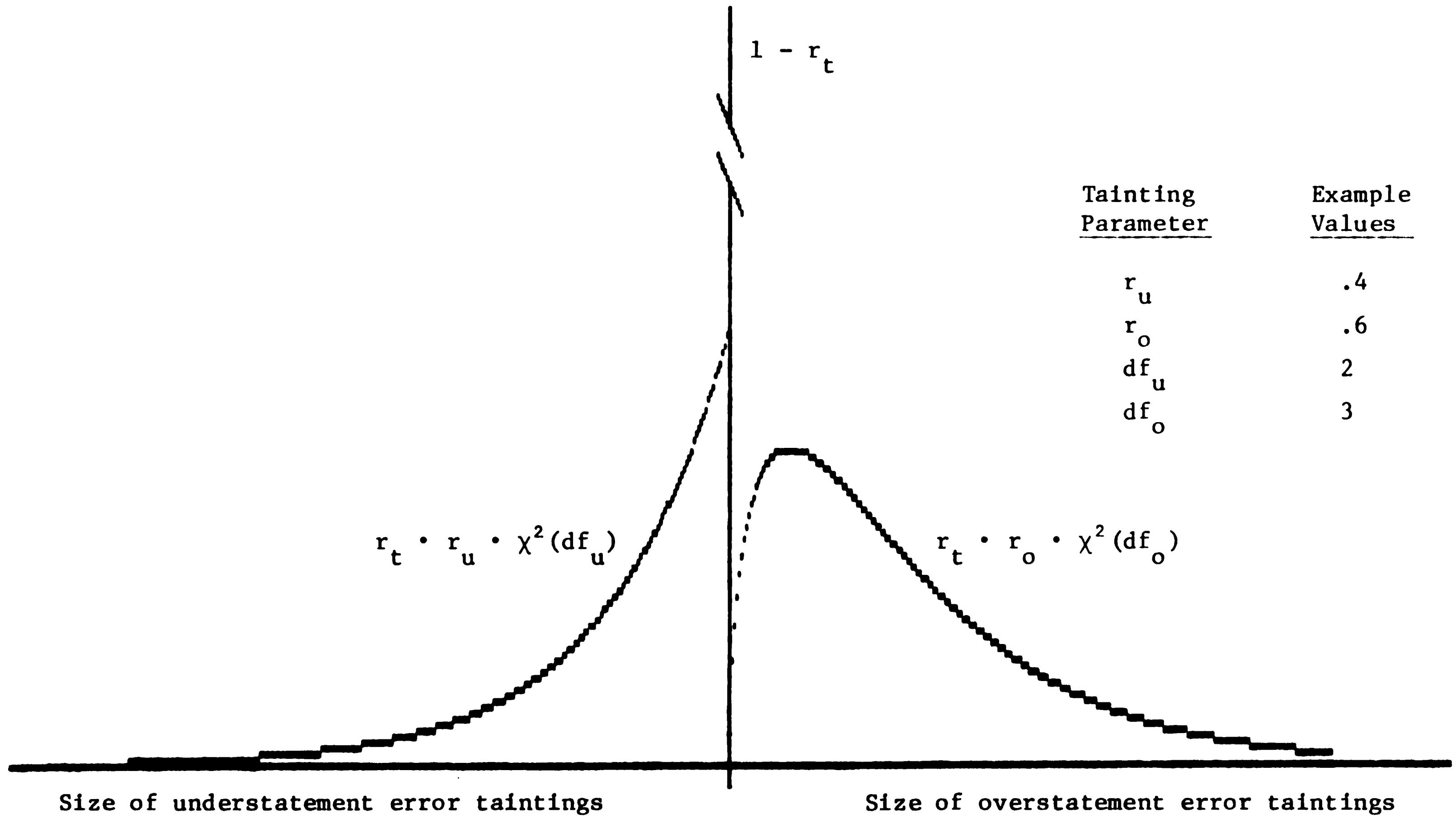
df_u = degrees of freedom for chi-square understatement distribution (1 or 2)

Figure 11 shows one example of an error tainting distribution under this framework (see also Dworin and Grimlund [1984, Figure 1, p. 228]).

Given the identifying parameters of the error tainting distribution, the ratio of the true to reported balance is

$$\theta = 1 - .1r_t\,(r_o\,df_o - r_u\,df_u).$$

The Monte Carlo experiments were conducted in the following manner. For each of k populations total error rates, r_t, and proportions of overstatement errors, r_o, were "drawn" using the underlying distributions given in Tables 25 and 26, respectively. The distribution of these items within the various ranges given in the Tables (e.g., .000 to .074, .074 to .139, etc.) was assumed to be uniform. Separate Monte Carlo experiments were run for each of the six possible combinations of df_u and df_o. From each of the k populations formed in this manner a sample of n dollar



Example of an Error Tainting Distribution

FIGURE 11

units was "drawn". Separate Monte Carlo experiments were performed for sample sizes of n = 50 and n = 100. These samples were constructed in the following manner. Each sample item had probability r_t of being a population error item and probability r_o of being an overstatement if it was an error. The sizes of the overstatement and understatement errors were distributed as $.1\,\chi^2_{df_o}$ and $.1\,\chi^2_{df_u}$ random variables. With a set of n errors, $e_1, \ldots e_n$, "drawn" from a population in this manner the PPS-MPU estimator $\hat{\theta}$, is defined as

$$\hat{\theta} = 1 - \bar{e}$$

$$\text{Where: } \bar{e} = 1/n \sum_{j=1}^{n} e_j$$

$$\hat{\sigma}^2 = 1/(n-1) \sum_{j=1}^{n} (e_j - \bar{e})^2$$

Note that e_j is defined to be zero if the jth sample item is not in error.

The PEB estimators, $\hat{\theta}^{PEB}$, were calculated in accordance with equation (9) and a set of nine pure Bayes estimators were calculated using equation (8) for each Monte Carlo trial. The sensitivity of the Bayes estimator to misspecifications in the parameters of the prior probability distribution was examined. The values for the true expectation and variance of the ratio, θ, of the true account balance to the recorded book balance are:

$$E\{\theta\} = E\{1 - .1\ r_t\ (r_o\ df_o - r_u\ df_u)\}$$
$$= 1 - .1\ E\{r_t\}\ [E\{r_o\}\ df_o - E\{r_u\}\ df_u]$$
$$= 1 - .0243211\ [.61435\ df_o - .38565\ df_u]$$
$$Var\{\theta\} = E\{\theta^2\} - [E\{\theta\}]^2$$
$$= E\{r_t^2\}\ (.01)\ [(df_u + df_o)^2\ E\{r_o^2\}$$
$$- 2(df_u + df_o)\ df_u\ E\{r_o\} + df_u]$$
$$- [E\{r_t\}]^2\ (.01)\ [(df_u + df_o)\ E\{r_o\} - df_u^2]$$
$$= .00107583652\ [(df_u + df_o)^2\ (.423131667)$$
$$- 2(df_u + df_o)\ df_u\ (.61435) + df_u^2]$$
$$- (.243211)^2\ (.01)\ [(df_u + df_o)\ .61435 - df_u]^2$$

Thus, for each Monte Carlo experiment the true values of $E\{\theta\}$ and $Var\{\theta\}$ depend only on the degrees of freedom in the distributions of the size of overstatement and understatement errors, df_o and df_u. Table 27 gives the values of the expectation and variance of θ for the six combinations of df_o and df_u tested.

In each Monte Carlo experiment nine pure Bayes estimators were tested. These nine estimators result from all possible combinations of three prior specifications on both the mean and variance of the prior distribution. The prior specifications for $E\{\theta\}$, represented by μ in equation (8), were:

Perfect specification:	$\mu = E\{\theta\}$
Underestimation of prior:	$\mu = E\{\theta\} - .02$
Overestimation of prior:	$\mu = E\{\theta\} + .02$

TABLE 27

Expectation and Variance of True to Reported Account Balance in Monte Carlo Experiments

Degrees of Freedom in Understatement Distribution df_u	Degrees of Freedom in Overstatement Distribution df_o	Expectation of Ratio of True to Reported Account Balance $E\{\theta\}$	Variance of Ratio of True to Reported Account Balance $Var\{\theta\}$
1	1	.993488	.00022202
1	2	.979496	.00078677
1	3	.964554	.00181546
2	1	1.003817	.00045448
2	2	.980876	.00088808
2	3	.973934	.00178561

The prior specifications for $Var\{\theta\}$, represented by τ in equation (8), were:

Perfect specification:	$\tau = Var\{\theta\}$
Overly tight specification:	$\tau = \frac{1}{4} Var\{\theta\}$
Overly diffuse specification:	$\tau = 4 Var\{\theta\}$

Thus, the nine prior distributions used in each Monte Carlo experiment were:

(1) Normal $(E\{\theta\} + .02 , \frac{1}{4} Var\{\theta\})$
(2) Normal $(E\{\theta\} + .02 , Var\{\theta\})$
(3) Normal $(E\{\theta\} + .02 , 4 Var\{\theta\})$
(4) Normal $(E\{\theta\} , \frac{1}{4} Var\{\theta\})$
(5) Normal $(E\{\theta\} , Var\{\theta\})$
(6) Normal $(E\{\theta\} , 4 Var\{\theta\})$
(7) Normal $(E\{\theta\} - .02 , \frac{1}{4} Var\{\theta\})$
(8) Normal $(E\{\theta\} - .02 , Var\{\theta\})$
(9) Normal $(E\{\theta\} - .02 , 4 Var\{\theta\})$

Tables 28 and 29 give the average efficiencies of the PEB estimator and the nine pure Bayes estimators for the 36 Monte Carlo experiments. The efficiency of an estimator is the ratio of its average mean squared error to the average mean squared error of the classical PPS-MPU estimator. Each Monte Carlo experiment consisted of 2,000 iterations of the population formation, sample selection and estimator calculation process. Table 28 presents the efficiency results for sample sizes of n = 50 and Table 29 gives the results for sample sizes of n = 100.

The results show the PEB estimator is always efficient relative to the classical PPS-MPU estimator. Reductions in mean squared error by as much as 40% were obtained.

TABLE 28

Monte Carlo Simulation of PEB and Bayes Efficiency in Variables Sampling
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions				Pure Bayes for Various Priors								
df_u	df_o	k	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.739	2.310	1.183	.705	1.031	.660	.607	1.685	.902	.673
		6	.624	2.434	1.253	.738	1.111	.714	.638	1.742	.950	.704
		8	.594	2.274	1.176	.710	1.043	.672	.617	1.673	.904	.680
1	2	4	.849	2.544	1.188	.780	1.564	.884	.763	1.339	.801	.779
		6	.813	2.547	1.201	.799	1.579	.903	.781	1.381	.828	.797
		8	.849	2.500	1.203	.803	1.551	.901	.782	1.341	.815	.792
1	3	4	.904	2.612	1.177	.850	1.891	.997	.847	1.542	.907	.854
		6	.889	2.544	1.131	.828	1.847	.966	.830	1.518	.888	.843
		8	.913	2.538	1.125	.826	1.854	.963	.828	1.528	.887	.841
2	1	4	.781	1.476	.800	.709	1.236	.767	.693	2.078	1.121	.749
		6	.662	1.514	.812	.709	1.228	.765	.691	2.209	1.102	.743
		8	.674	1.546	.825	.713	1.280	.788	.698	2.101	1.137	.754
2	2	4	.807	1.853	.943	.760	1.346	.781	.741	1.462	.814	.752
		6	.700	1.877	.944	.751	1.359	.788	.732	1.472	.812	.745
		8	.708	1.843	.937	.750	1.356	.785	.736	1.498	.832	.753
2	3	4	.863	2.251	1.080	.823	1.709	.921	.812	1.532	.860	.805
		6	.793	2.121	1.017	.805	1.608	.871	.797	1.445	.818	.807
		8	.828	2.177	1.041	.807	1.642	.886	.798	1.464	.828	.802

TABLE 29

Monte Carlo Simulation of PEB and Bayes Efficiency in Variables Sampling
n = 100

Degrees of Freedom in the Understatement/ Overstatement Distributions				Pure Bayes for various Priors								
df_u	df_o	k	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.841	3.682	1.455	.844	1.736	.888	.772	2.540	1.082	.806
		6	.744	3.536	1.402	.841	1.638	.869	.781	2.470	1.087	.825
		8	.745	3.568	1.405	.831	1.631	.870	.778	2.513	1.113	.834
1	2	4	.913	3.378	1.237	.875	2.096	.982	.868	1.700	.901	.880
		6	.907	3.407	1.253	.885	2.143	1.004	.878	1.785	.930	.889
		8	.926	3.418	1.247	.878	2.151	1.002	.873	1.796	.934	.888
1	3	4	.943	3.123	1.130	.902	2.280	1.008	.906	1.832	.950	.915
		6	.944	3.118	1.129	.903	2.279	1.008	.906	1.838	.950	.915
		8	.979	3.117	1.127	.900	2.272	1.005	.904	1.822	.945	.913
2	1	4	.855	2.018	.933	.828	1.813	.932	.811	3.037	1.282	.839
		6	.804	2.165	.960	.833	1.941	.957	.820	3.214	1.320	.853
		8	.841	2.090	.958	.838	1.863	.962	.831	3.084	1.321	.869
2	2	4	.855	2.753	1.111	.879	2.042	.960	.867	2.129	.977	.872
		6	.833	2.547	1.058	.869	1.873	.910	.855	1.988	.930	.860
		8	.857	2.666	1.088	.869	1.970	.936	.856	2.057	.951	.861
2	3	4	.926	2.925	1.122	.897	2.234	.996	.893	1.961	.945	.896
		6	.908	2.765	1.095	.900	2.106	.975	.895	1.855	.927	.897
		8	.902	2.721	1.054	.881	2.075	.942	.880	1.829	.902	.886

The use of a pure Bayesian estimator results in additional efficiency gains with a correct prior specification of the mean of the underlying distribution. As was the case in the attributes sampling model, misspecifications in the mean or variance of the prior distribution can be costly during pure Bayesian estimation. The penalties are most acute with any misspecification of the mean (priors 1, 2, 3, 7, 8 and 9) or an overly tight variance specification (priors 1, 4 and 7). In these circumstances the gains from Bayesian estimation can be eroded to such an extent that the estimator is inefficient (at times grossly inefficient) relative to traditional non-Bayesian estimation. These costs of incorrect specification of the prior distribution parameters are avoided in the PEB estimator by estimating these parameters from the sample data itself. In each of the 36 Monte Carlo trials reported in Tables 28 and 29 the PEB estimator was efficient relative to the classical PPS-MPU estimator.

However, three anomalous results were obtained. First, there was a persistent tendency for the PEB estimator to have greater efficiency when sampling from 6 populations than when sampling from 8 populations. One would expect, ceteris paribus, greater efficiency when sampling from a larger number of populations since more precise estimates of the underlying parameters $E\{\theta\}$ and $Var\{\theta\}$ can be obtained from the sampling results.

A second anomaly arose in the behavior of the pure Bayes estimators. The efficiency for the Bayes estimator with proper mean specification but overly diffuse variance specification (prior 6) was consistently superior to that for the Bayes estimator with perfect mean and variance specifications (prior 5). Under ideal conditions one would expect perfectly specified Bayesian priors to outperform any misspecified prior.

Finally, the observed reliability of the confidence intervals was substantially lower than the desired 95% probability coverage. This is shown in Tables 30 and 31 which present the Monte Carlo results for PPS-MPU, PEB and pure Bayesian 95% confidence interval observed reliability levels. In all instances the observed reliability is substantially less than the desired 95% probability coverage.

The unreliability of classical sampling techniques when testing accounting populations has long been known and is well documented in the literature. See, for example, Neter and Loebbecke [1975]. The major reason for this result is the lack of normality exhibited by the sampling estimator principally due to the large proportion of population items which are not in error as well as the skewed distribution of population errors.

One method of increasing the reliability of the estimator is, of course, to sufficiently increase the sample size to insure the selection of a larger number of

TABLE 30

Monte Carlo Simulation of PPS-MPU, PEB and Bayes Reliability in Variables Sampling
Observed Reliability of 95% Confidence Intervals in 2,000 Trials
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions					Pure Bayes for Various Priors								
df_u	df_o	k	PPS-MPU	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.828	.829	.435	.736	.823	.695	.792	.829	.429	.770	.830
		6	.819	.814	.429	.720	.813	.681	.782	.822	.418	.755	.822
		8	.828	.809	.436	.728	.818	.688	.790	.830	.422	.765	.829
1	2	4	.812	.811	.600	.756	.809	.701	.788	.815	.637	.804	.818
		6	.818	.804	.606	.762	.816	.702	.794	.820	.642	.809	.823
		8	.815	.789	.606	.758	.812	.702	.795	.817	.639	.806	.821
1	3	4	.806	.805	.642	.768	.804	.699	.787	.806	.689	.799	.810
		6	.814	.797	.646	.776	.813	.703	.795	.816	.691	.808	.817
		8	.811	.795	.647	.773	.810	.704	.793	.812	.694	.804	.814
2	1	4	.828	.830	.589	.790	.833	.690	.794	.829	.551	.764	.823
		6	.831	.824	.595	.796	.834	.696	.800	.832	.553	.768	.826
		8	.830	.811	.587	.792	.834	.692	.797	.832	.555	.764	.826
2	2	4	.851	.851	.658	.803	.849	.717	.821	.852	.658	.820	.853
		6	.854	.848	.664	.810	.855	.728	.828	.857	.666	.829	.858
		8	.855	.830	.667	.810	.854	.724	.825	.857	.662	.823	.857
2	3	4	.847	.848	.687	.804	.846	.736	.820	.848	.713	.829	.850
		6	.848	.836	.681	.808	.848	.729	.827	.851	.711	.833	.853
		8	.846	.824	.677	.807	.846	.726	.824	.848	.711	.833	.850

TABLE 31

Monte Carlo Simulation of PPS-MPU, PEB and Bayes Reliability in Variables Sampling
Observed Reliability of 95% Confidence Intervals in 2,000 Trials
n = 100

Degrees of Freedom in the Understatement/Overstatement Distributions					Pure Bayes for Various Priors								
df_u	df_o	k	PPS-MPU	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.882	.884	.523	.801	.876	.746	.852	.883	.522	.834	.884
		6	.882	.876	.533	.803	.875	.758	.852	.883	.516	.835	.882
		8	.881	.868	.533	.802	.877	.751	.856	.883	.517	.834	.881
1	2	4	.873	.873	.680	.834	.873	.764	.859	.876	.728	.869	.876
		6	.871	.861	.678	.828	.869	.764	.854	.872	.726	.865	.874
		8	.871	.858	.682	.832	.869	.764	.855	.872	.731	.865	.875
1	3	4	.869	.865	.720	.843	.869	.769	.856	.870	.774	.863	.872
		6	.864	.858	.724	.843	.864	.773	.854	.866	.779	.862	.867
		8	.867	.858	.717	.843	.869	.768	.857	.870	.770	.864	.870
2	1	4	.879	.882	.682	.857	.881	.757	.857	.880	.634	.829	.877
		6	.886	.876	.678	.858	.887	.756	.860	.886	.631	.837	.882
		8	.882	.863	.674	.857	.883	.755	.858	.882	.639	.831	.879
2	2	4	.893	.898	.710	.856	.895	.768	.872	.896	.735	.876	.897
		6	.897	.892	.728	.866	.898	.783	.879	.899	.739	.858	.900
		8	.896	.883	.718	.860	.898	.774	.876	.898	.735	.876	.898
2	3	4	.897	.898	.746	.871	.897	.786	.883	.899	.780	.889	.900
		6	.890	.881	.742	.864	.889	.785	.877	.891	.780	.884	.892
		8	.887	.874	.745	.862	.887	.787	.875	.888	.773	.880	.889

population error items. However, the sample sizes required to obtain the desired reliability are often too high for auditors to be able to employ the procedure in many practical circumstances. A second alternative is to abandon classical variables sampling for populations suspected to have a low rate of errors in favor of the combined attributes-variables or dollar unit sampling method which does not rely on the normality of a sampling estimator. For populations with higher expected error rates the classical sampling techniques can be retained.

Tables 32 through 35 present the results of a series of Monte Carlo experiments analyzing the behavior of the classical PPS-MPU, the PEB and pure Bayesian estimators in populations whose total error rates are 10% or more. With the single exception of this lower bound on the total error rate all other aspects of these Monte Carlo experiments are identical to the previously described trials whose results are given in Tables 28 through 31.

Limiting the use of classical estimators to populations with error rates greater than 10% is consistent with two recent studies which employed the same error tainting patterns used here. Leitch, et.al [1982] limit their analysis to populations with error rates of 6% or more. Similarly, Dworin and Grimlund [1984] limit their study to populations with error rates of 10% or more. The results of Johnson, Leitch and Neter [1981] given in Table 25 indicate that somewhere between 50 to 75% of all inventory

TABLE 32

Monte Carlo Simulation of PEB and Bayes Efficiency in Variables Sampling
Total Error Rates Greater Than 10%
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions				Pure Bayes for Various Priors								
df_u	df_o	k	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.772	1.851	.923	.694	.937	.598	.649	1.581	.820	.701
		6	.636	1.898	.940	.692	.968	.614	.651	1.629	.847	.711
		8	.598	1.838	.911	.689	.933	.600	.652	1.588	.836	.714
1	2	4	.858	1.877	.919	.767	1.234	.749	.770	1.227	.752	.797
		6	.782	1.954	.962	.780	1.290	.779	.780	1.269	.771	.804
		8	.782	1.959	.963	.783	1.282	.776	.781	1.266	.767	.803
1	3	4	.897	1.875	.921	.823	1.401	.820	.829	1.252	.793	.844
		6	.848	2.015	.963	.828	1.514	.857	.836	1.352	.829	.853
		8	.841	1.965	.951	.823	1.487	.849	.831	1.336	.823	.848
2	1	4	.809	1.327	.727	.742	1.117	.694	.725	1.758	.923	.748
		6	.700	1.323	.724	.740	1.139	.701	.724	1.801	.938	.748
		8	.677	1.298	.707	.732	1.102	.677	.714	1.743	.904	.736
2	2	4	.837	1.514	.803	.784	1.193	.714	.773	1.368	.759	.781
		6	.742	1.474	.771	.765	1.176	.699	.763	1.366	.760	.778
		8	.713	1.565	.800	.775	1.235	.715	.769	1.405	.766	.781
2	3	4	.864	1.587	.824	.809	1.283	.754	.811	1.261	.751	.821
		6	.810	1.652	.850	.817	1.336	.779	.819	1.301	.774	.829
		8	.793	1.696	.868	.820	1.368	.789	.820	1.331	.781	.828

TABLE 33

Monte Carlo Simulation of PEB and Bayes Efficiency in Variables Sampling
Total Error Rates Greater Than 10%
n = 100

Degrees of Freedom in the Understatement/Overstatement Distributions				Pure Bayes for Various Priors								
df_u	df_o	k	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.853	2.552	.998	.801	1.336	.740	.794	2.206	.980	.846
		6	.775	2.683	1.059	.823	1.412	.771	.803	2.252	.981	.843
		8	.758	2.651	1.056	.828	1.382	.766	.805	2.251	.981	.843
1	2	4	.908	2.463	1.003	.873	1.645	.873	.879	1.581	.877	.898
		6	.874	2.450	1.011	.878	1.631	.878	.883	1.560	.879	.901
		8	.872	2.546	1.025	.876	1.694	.883	.879	1.602	.877	.896
1	3	4	.939	2.308	.986	.905	1.743	.914	.910	1.527	.895	.920
		6	.934	2.379	1.017	.915	1.809	.944	.920	1.591	.924	.930
		8	.916	2.341	.981	.900	1.769	.908	.905	1.539	.887	.915
2	1	4	.886	1.648	.843	.862	1.522	.838	.851	2.415	1.040	.862
		6	.827	1.727	.859	.862	1.573	.847	.851	2.472	1.052	.863
		8	.822	1.743	.856	.862	1.588	.844	.850	2.485	1.047	.861
2	2	4	.899	2.025	.921	.883	1.610	.854	.881	1.782	.893	.890
		6	.842	1.953	.887	.872	1.558	.825	.871	1.728	.865	.880
		8	.837	1.954	.898	.877	1.569	.838	.876	1.755	.881	.885
2	3	4	.927	2.003	.919	.898	1.631	.870	.901	1.571	.871	.909
		6	.890	1.998	.931	.903	1.624	.878	.905	1.558	.874	.910
		8	.877	2.030	.934	.903	1.648	.879	.903	1.580	.873	.909

TABLE 34

Monte Carlo Simulation of PPS-MPU, PEB and Bayes Reliability in Variables Sampling
Observed Reliability of 95% Confidence Intervals in 2,000 Trials
Total Error Rates Greater Than 10%
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions					Pure Bayes for Various Priors								
df_u	df_o	k	PPS-MPU	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.922	.926	.559	.847	.919	.764	.895	.926	.519	.874	.926
		6	.920	.916	.562	.843	.917	.760	.892	.923	.511	.869	.923
		8	.920	.909	.568	.844	.917	.764	.894	.926	.509	.868	.924
1	2	4	.910	.912	.694	.866	.909	.774	.896	.914	.704	.904	.916
		6	.907	.900	.694	.858	.908	.776	.890	.912	.699	.903	.915
		8	.914	.900	.697	.868	.913	.779	.898	.916	.710	.907	.918
1	3	4	.901	.901	.729	.870	.902	.780	.890	.904	.762	.902	.906
		6	.907	.901	.732	.877	.908	.783	.893	.910	.760	.904	.911
		8	.902	.894	.733	.870	.903	.783	.891	.906	.749	.900	.908
2	1	4	.918	.922	.691	.893	.923	.772	.896	.922	.647	.867	.917
		6	.918	.915	.690	.895	.924	.768	.895	.921	.648	.864	.916
		8	.921	.906	.692	.897	.925	.770	.898	.923	.654	.869	.919
2	2	4	.933	.935	.741	.903	.935	.785	.920	.935	.730	.917	.936
		6	.933	.930	.741	.903	.935	.782	.917	.937	.721	.913	.937
		8	.932	.926	.737	.902	.936	.783	.915	.937	.731	.914	.938
2	3	4	.923	.926	.759	.898	.925	.794	.911	.926	.770	.917	.927
		6	.920	.918	.754	.897	.923	.791	.911	.926	.764	.914	.928
		8	.928	.920	.760	.900	.930	.791	.914	.932	.765	.920	.933

TABLE 35

Monte Carlo Simulation of PPS-MPU, PEB and Bayes Reliability in Variables Sampling
Observed Reliability of 95% Confidence Intervals in 2,000 Trials
Total Error Rates Greater Than 10%
n = 100

Degrees of Freedom in the Understatement/Overstatement Distributions					Pure Bayes for Various Priors								
df_u	df_o	k	PPS-MPU	PEB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	1	4	.937	.941	.621	.887	.939	.802	.926	.941	.584	.907	.940
		6	.935	.933	.615	.881	.935	.793	.918	.939	.591	.599	.938
		8	.936	.928	.621	.883	.935	.795	.923	.939	.585	.900	.937
1	2	4	.927	.931	.746	.902	.929	.817	.922	.930	.763	.925	.930
		6	.930	.926	.745	.901	.931	.812	.921	.933	.763	.924	.934
		8	.930	.921	.743	.901	.928	.817	.922	.931	.763	.926	.934
1	3	4	.927	.928	.781	.909	.928	.823	.921	.929	.816	.929	.930
		6	.929	.922	.779	.909	.929	.820	.920	.929	.806	.924	.931
		8	.926	.921	.774	.909	.928	.820	.920	.928	.811	.926	.929
2	1	4	.939	.941	.747	.925	.942	.805	.927	.942	.707	.904	.938
		6	.936	.931	.737	.923	.939	.804	.921	.938	.701	.897	.935
		8	.934	.925	.734	.919	.937	.799	.920	.936	.695	.899	.933
2	2	4	.947	.949	.780	.926	.946	.816	.934	.948	.776	.934	.948
		6	.941	.941	.776	.923	.943	.816	.934	.944	.777	.931	.945
		8	.943	.939	.774	.926	.945	.809	.934	.946	.771	.931	.946
2	3	4	.936	.939	.796	.928	.941	.824	.934	.941	.805	.936	.942
		6	.938	.935	.800	.924	.941	.825	.931	.942	.807	.934	.942
		8	.939	.945	.795	.924	.940	.823	.933	.941	.808	.936	.942

populations can be expected to have total error rates in excess of 10%. Thus, the implication of limiting the study to these accounts is that the auditor will use the results of interim attribute tests of the internal control system and knowledge obtained from previous audits to form an assessment about the total error rate of the population prior to selecting a sampling plan. For those populations whose total error rate is determined to be less than 10% an alternative estimation procedure would be employed.

The results given in Tables 32 to 35 show that limiting the use of the sampling procedures to populations with errors in excess of 10% successfully avoids the three anomalies previously noted.

First, PEB estimation was in every instance efficient relative to classical PPS-MPU estimation. The efficiency improves when increasing the number of populations, k, used in the construction of the PEB estimator. Reductions of mean squared error by as much as 40% were obtained using PEB estimation in lieu of the classical PPS-MPU estimator.

Secondly, the best pure Bayesian estimator results when the parametric specifications are accurate for both the mean and variance of the prior distribution. Misspecifications of either of these parameters can be costly, resulting in inefficient estimation relative to both PEB and classical PPS-MPU estimation.

It is significant to note that PEB efficiency levels approach those of optimal Bayesian estimation with perfect prior parameter specification (prior 5) with as few as eight populations used in the construction of the PEB estimator.

Finally, as expected, the reliability of the estimators is greatly enhanced by limiting their use to populations with error rates of 10% or more. The observed reliability levels of 95% nominal confidence intervals were similar for the PPS-MPU and the PEB estimators. In general they fell within the range of 90 to 93% with a sample size of n = 50 and 92.5 to 94% for the larger sample size of n = 100. Misspecifications in the parameters of the prior distribution can be extremely costly in pure Bayesian estimation with observed reliability levels of as low as 51% for 95% desired confidence intervals being noted.

Table 36 displays the ratios of the average widths of the PEB 95% confidence intervals to those for classical PPS-MPU confidence intervals. The results show that the PEB intervals are able to exploit the efficiency of the estimator by establishing confidence intervals which are from 2 to 12% shorter than the classical intervals.

The results of Tables 32 through 36 demonstrate the effectiveness of PEB estimation in substantive audit testing when population error rates are 10% or more. In these circumstances PEB estimation produces confidence intervals with observed reliability only slightly less than the desired level and with confidence intervals somewhat narrower than those yielded by classical procedures. Additionally, the PEB point estimator was shown to have

TABLE 36

Ratio of PEB to Classical 95% Confidence Interval Widths
Total Error Rates Greater Than 10%

Degrees of Freedom in Understatement/ Overstatement Distributions			Ratio of Average PEB to PPS-MPU Confidence Interval Widths	
df_u	df_o	k	n = 50	n = 100
1	1	4	.947	.963
		6	.873	.918
		8	.829	.891
1	2	4	.961	.976
		6	.906	.944
		8	.876	.928
1	3	4	.969	.981
		6	.925	.955
		8	.901	.947
2	1	4	.953	.970
		6	.888	.930
		8	.854	.913
2	2	4	.956	.972
		6	.900	.938
		8	.872	.922
2	3	4	.964	.977
		6	.916	.950
		8	.891	.937

significantly smaller mean squared error than the classical estimator with risk reductions of 6 to 40% obtained. Reduction of mean squared error toward the optimal level of pure Bayesian estimation with perfect prior parameter estimation was rapid. Only k = 8 populations are required to reduce the PEB mean squared error sufficiently to approach optimal Bayesian estimation. This fact combined with the considerable risk additions which arise from imperfect prior parameter specifications in pure Bayesian estimation make PEB techniques an attractive option for auditors.

Finally, the anomalous behavior noted in Tables 28 through 31 when populations with error rates of less than 10% are admitted does not appear to be a result of PEB techniques per se, but rather an indirect effect from the reliance on a nonnormally distributed classical estimator. The PEB estimator was consistently efficient relative to classical estimation in these circumstances as well. In fact the mean squared error reduction is greater when low error rate populations are admitted into the analysis. Thus, if the auditor is committed to using a classical variables sampling approach rather than a combined attributes-variables technique even in the face of low error rates, then the benefits in terms of risk reduction from PEB estimation are still present. However, reliance on PEB confidence intervals in these circumstances is unwise, as it would be for classical estimation as well.

4.2 Interprocedural Integration of Variables Sampling and Analytical Review Procedures

The purpose of this section is to illustrate the use of PEB estimation techniques for the integration of classical variables sampling estimators with auxiliary analytical review information. The purpose of this integration is to obtain a single estimate of the true balance of an accounting population which is more efficient than either the sampling information or the auxiliary analytical review information used in isolation could provide.

Morris [1983a] generalizes the PEB estimator given by equation (9) in Section 4.1 to allow for the integration of auxiliary information in the estimation process. This auxiliary information PEB estimator is developed in the following manner.

As before let $\hat{\theta}_i^S$ be a classical sampling estimate of the ratio, θ_i, of the true account balance to the client's reported balance for the *ith* account. Thus,

$$\hat{\theta}_i^S \sim \text{Normal}\ (\theta_i,\ \sigma_i^2).$$

Employing the notation of Section 2.1 by representing the true and recorded balances of the *ith* population as X_i and Y_i, respectively, yields the classical sampling estimate of the true account balance:

$$\hat{X}_i^S = Y_i\ \hat{\theta}_i^S \sim \text{Normal}\ (X_i,\ Y_i^2\ \sigma_i^2)$$

Let $\mathbf{z}_i$ be an r-dimensional column vector of observations for a set of auxiliary information variables.

These concomitant information variables could include any of the numerous items typically employed by auditors in formal or informal analytical review models. Examples might include prior year's balances, balances in other related accounts, or general economic and industry indexes.

If it was known that the true balance, X_i, in each population was perfectly correlated with these auxiliary variables through some regression equation

$$X_i = \mathbf{z}_i' \beta$$

then the ideal estimator would be obtained through a weighted least squares process:

$$\hat{X}_i^{AR} = \mathbf{z}_i' \hat{\beta}$$

Where: $\hat{\beta} = (\mathbf{Z}'\mathbf{D}\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{D}\hat{\mathbf{X}}^S$

$\mathbf{Z}$ = k x r matrix with rows $\mathbf{z}_i'$

$\mathbf{D}$ = k x k diagonal matrix with $1/Y_i^2\sigma_i^2$ diagonal elements

$\hat{\mathbf{X}}^S = (X_1^S, \ldots, X_k^S)'$

In general we know that the regression equation

$$X_i = \mathbf{z}_i' \beta$$

does not hold exactly. However, numerous studies using actual accounting and economic data have shown the validity of such an analytical review regression model as a potential audit tool. Examples include Stringer [1975], Albrecht and McKeown [1977], Kaplan [1978], Kinney [1978], Akresh and Wallace [1980], Neter [1980], and Lev [1980]. The general analytical review regression model with an error term is

$$X_i = \mathbf{z}_i' \beta + \varepsilon_i$$

$$\text{Where } \varepsilon_i \sim \text{Normal } (0, \tau).$$

Typically, the noise in an analytical review regression model as represented by the variance of the disturbance term, τ, is of such a magnitude as to make analytical review procedures inefficient relative to classical sampling when each procedure is used in isolation. However, Bayesian estimation provides a compromise estimator formed by a linear combination of the sampling and analytical review estimators. If $\boldsymbol{\beta}$ and τ were known, then the pure Bayes estimator

$$\hat{X}_i^B = B_i \, \mathbf{z}_i'\beta + (1 - B_i) \, \hat{X}_i^S$$

$$B_i = Y_i^2\sigma_i^2/(Y_i^2\sigma_i^2 + \tau)$$

would serve to minimize the expected mean squared error.

Since $\boldsymbol{\beta}$ and τ are not known, in order to operationalize pure Bayesian estimation these parameters would have to be subjectively prespecified by the auditor. Misspecification of one or both of these parameters will prove to be costly, of course, perhaps to the extent of making pure Bayesian estimation inefficient relative to classical sampling.

However, estimates of both $\boldsymbol{\beta}$ and τ are available from the sample data. This leads to a generalized parametric empirical Bayes estimator (Morris [1983a], p. 53):

$$(11) \qquad \hat{X}_i^{PEB} = \hat{B}_i^{PEB}\, \mathbf{z}_i'\hat{\boldsymbol\beta} + (1 - \hat{B}_i^{PEB})\, \hat{X}_i^S$$

Where:

$$\hat{\boldsymbol\beta} = (\mathbf{Z}'\mathbf{D}\mathbf{Z})^{-1}\, \mathbf{Z}'\mathbf{D}\hat{\mathbf{X}}^S$$

$$\mathbf{D} = k \times k \text{ diagonal matrix of weights, } W_i$$

$$W_i = 1/(Y_i^2\sigma_i^2 + \hat{\tau}^+)$$

$$\hat{\tau} = \Sigma W_i[(k/(k-r))(\hat{X}_i^S - \mathbf{z}_i'\hat{\boldsymbol\beta})^2 - Y_i^2\hat{\sigma}_i^2]/\Sigma W_i$$

$$\hat{\tau}^+ = \max\{0, \hat{\tau}\}$$

$$\hat{B}_i^{PEB} = [(k-r-2)/(k-r)]\; Y_i^2\hat{\sigma}_i^2\,/(Y_i^2\hat{\sigma}_i^2 + \hat{\tau}^+)$$

When $r = 1$ and each $z_i = 1$ represents a constant term, then equation (11) reduces to equation (9), the PEB estimator analyzed in the previous section, 4.1.

The PEB estimator can be interpreted as a compromise between a pure sampling estimator, $\hat{X}_i^S$, and an estimate obtained from a noisy analytical review regression model, $\mathbf{z}_i'\hat{\boldsymbol\beta}$. The goal of this compromise estimator is to obtain one estimate of the true account balance which will outperform either extreme choice of total reliance on analytical review or pure sampling procedures. Thus, the PEB estimator can be viewed as a mechanism for integrating across two types of substantive audit procedures to obtain one single efficient estimator.

The PEB coefficient, $\hat{B}_i^{PEB}$, can be viewed as a shrinkage coefficient. If the analytical review model is particularly noisy, yielding little information about the true account balance, then τ is large and relatively little shrinkage from the sampling estimator results.

Finally, it should be noted that no prior information about the analytical review regression model is assumed to be known by the auditor. The estimates $\hat{\beta}$ and $\hat{\tau}^+$ are derived strictly from the current year's sampling information and observations on the regression variables, z_i. This differs from the usual analytical review estimator which requires intertemporal or crossectional estimation outside the current audit engagement. Typically the auditor must assume that the regression equation for the current audit client is unchanged from the model which generated the estimation observations. Thus, one major advantage of this form of PEB estimation is the fact that auditors may exploit the information contained in concomitant variables without prior observations on those variables. This is particularly valuable for newly obtained clients or circumstances in which the auditor believes past relationships may no longer be reflective of the current economic circumstances.

The advantages of PEB integration of analytical review evidence with classical sampling results are brought sharply into focus by the following comments of Kinney [1979, p. 460] regarding traditional analytical review regression techniques:

> To justify a precise account balance probability statement based on the regression results, two conditions must be met. First, the auditor must be able to show that the base period model seems to be valid for the audit period. Second, if the model appears valid, the regression results must be sufficiently precise to allow adequate confidence that the maximum accounting error is less than a material amount.

Both of these disadvantages of pure analytical review regression analysis are mitigated through PEB integration. Since the regression model parameters are estimated using current sampling results and not through the use of prior observations, the stability of the analytical review model from the base observations to the audit period is not a required assumption. Secondly, since the analytical review model is integrated with the sampling estimator to obtain a single more efficient estimator, the analytical review estimator in isolation need not be sufficiently precise for the auditor's substantive test. Even a relatively noisy analytical review model can be exploited to improve the efficiency or precision of a sampling estimator.

In order to evaluate the performance of the PEB estimator given in equation (11), a particular regression model must be posited. To maintain consistency with the prior literature, a model similar to ones employed by Kinney and Salamon [1978, 1982] and Kinney [1979] is used in this study:

$$X_i = 4 + 2\ z_i + \varepsilon_i$$

$$\text{Where:}\quad z_i \sim \text{Normal}\ (48,\ \sigma_z^2)$$

$$\varepsilon_i \sim \text{Normal}\ (0,\ \sigma_\varepsilon^2)$$

Four combinations of σ_z^2 and σ_ε^2 were analyzed. These four combinations represent varying degrees of informativeness of the analytical review regression model. The four combinations are given in Table 37. In general, the higher the variance of the disturbance terms and the

TABLE 37

Analytical Review Models Chosen for Study

Model	Relative Informativeness of the Analytical Review Model	σ^2_z	σ^2_ε	Total Variance of True Balance	Expected Correlation of True Balance and Auxiliary Variable
1	High	20.83	9.00	92.31	.95
2	Medium	16.67	25.61	92.31	.85
3	Medium	83.31	36.00	369.22	.95
4	Low	66.69	102.46	369.22	.85

lower the correlation between the true account balances and the auxiliary explanatory variable, the less informative is the analytical review model. Stringer [1975] reports that the correlation coefficient in Deloitte, Haskins and Sells' applications of analytical review regression models is less than .85 in about one-third of the applications and above .95 in about one-third. Thus, the range chosen for this study reflects a large portion of practical applications.

A series of Monte Carlo experiments was performed investigating the efficiency of the equation (11) PEB estimator. For each Monte Carlo trial values for the true account balance and the auxiliary variable were generated in accordance with the underlying analytical review regression model. Simulated samples of n = 50 and 100 were drawn using the same error rate and error size distributions of Section 4.1. The PEB estimator was computed in accordance with equation (11) based upon the simulated sampling results and auxiliary variable values. Each Monte Carlo experiment consisted of 2,000 of such trials.

The results of these Monte Carlo experiments are given in Table 38 and 39. In each of the 144 Monte Carlo experiments the PEB estimator was efficient relative to the classical sampling estimator. This efficiency was obtained despite the fact that the pure analytical review estimator, $z_i' \hat{\beta}$, is inefficient relative to pure sampling. Thus, PEB techniques were successful in integrating a relatively noisy concomitant variable with a classical sampling estimator to

TABLE 38

Monte Carlo Simulation of the Efficiency of the PEB Estimator Integrating
Classical Sampling and Analytical Review (AR) Estimators
Total Error Rates Greater Than 10%
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions			Model 1		Model 2		Model 3		Model 4	
df_u	df_o	k	PEB	AR	PEB	AR	PEB	AR	PEB	AR
1	1	5	.905	2.916	.964	7.884	.975	11.198	.991	29.330
		7	.852	3.347	.931	8.995	.949	12.029	.981	35.717
		9	.836	3.672	.932	10.206	.940	13.449	.986	37.639
1	2	5	.880	1.683	.939	4.031	.960	5.593	.982	14.712
		7	.777	1.735	.905	4.564	.918	6.101	.982	16.678
		9	.742	1.876	.881	4.964	.900	6.466	.973	18.415
1	3	5	.899	1.729	.916	2.486	.931	3.316	.980	8.570
		7	.763	1.283	.849	2.721	.877	3.713	.960	9.944
		9	.657	1.185	.823	2.921	.848	4.004	.950	10.842
2	1	5	.885	2.020	.952	5.071	.959	6.993	.984	19.442
		7	.793	2.270	.906	5.892	.920	8.033	.963	22.170
		9	.758	2.277	.897	6.507	.907	8.432	.963	24.549
2	2	5	.858	1.324	.932	3.121	.941	4.199	.978	10.887
		7	.740	1.383	.885	3.576	.896	4.729	.973	13.176
		9	.704	1.453	.844	3.628	.877	5.185	.946	14.097
2	3	5	.836	1.067	.900	2.134	.924	2.795	.964	7.150
		7	.688	.991	.836	2.368	.873	3.138	.942	8.524
		9	.644	1.054	.796	2.411	.827	3.259	.934	9.036

TABLE 39

Monte Carlo Simulation of the Efficiency of the PEB Estimator Integrating Classical Sampling and Analytical Review (AR) Estimators
Total Error Rates Greater Than 10%
n = 100

Degrees of Freedom in the Understatement/ Overstatement Distributions			Model 1		Model 2		Model 3		Model 4	
df_u	df_o	k	PEB	AR	PEB	AR	PEB	AR	PEB	AR
1	1	5	.957	5.612	.987	15.414	.991	21.953	.998	55.567
		7	.928	6.417	.975	17.613	.986	25.602	1.000	67.643
		9	.904	6.867	.970	19.381	.981	26.809	.994	75.819
1	2	5	.925	2.907	.977	7.471	.985	10.778	.997	28.162
		7	.857	3.298	.946	8.899	.964	11.978	.996	33.772
		9	.856	3.649	.936	9.101	.965	13.328	.993	35.759
1	3	5	.884	1.747	.959	4.648	.961	5.995	.988	16.021
		7	.807	1.968	.922	5.201	.942	7.201	.982	19.291
		9	.769	2.099	.900	5.506	.921	7.828	.981	21.156
2	1	5	.939	3.843	.976	9.708	.983	13.184	.995	39.017
		7	.902	4.329	.959	11.573	.965	15.796	.987	43.615
		9	.869	4.591	.947	12.263	.963	17.158	.978	49.382
2	2	5	.915	2.393	.956	6.020	.975	8.023	.990	22.261
		7	.850	2.570	.943	7.015	.957	9.113	.990	25.971
		9	.820	2.809	.923	7.286	.936	10.140	.984	29.108
2	3	5	.883	1.652	.947	3.965	.957	5.164	.982	13.797
		7	.791	1.730	.906	4.362	.938	5.997	.974	16.214
		9	.751	1.821	.871	4.653	.905	6.471	.977	17.718

obtain a single estimate which was more efficient than either of the two used in isolation.

In general the efficiency of the PEB estimator improves with the informativeness of the analytical review model and the number of populations used to compute the estimator.

Morris [1983a, p. 53] proposes a confidence interval for the generalized PEB estimator of equation (11):

(12) Probablility $\{\hat{X}_i^{PEB} - z_\alpha s_i \leq X_i \leq \hat{X}_i^{PEB} + z_\alpha s_i\} \geq 1 - \alpha$

Where:
$$s_i^2 = Y_i^2 \sigma_i^2 \,[1-(1-\hat{r}_i)\hat{B}_i^{PEB}] + v_i(\hat{X}_i - \mathbf{z}_i'\hat{\beta})^2$$
$$\hat{r}_i = W_i \,[\, \mathbf{Z}(\mathbf{Z}'\mathbf{D}\mathbf{Z})^{-1}\mathbf{Z}'\,]_{ii}$$
$$v_i = (2/(k-r-2))(\hat{B}_i^{PEB})^2 (\overline{Y_i^2\hat{\sigma}_i^2} + \hat{\tau}^+)/(Y_i^2\hat{\sigma}_i^2 + \hat{\tau}^+)$$
$$\overline{Y_i^2\hat{\sigma}_i^2} = \Sigma W_i Y_i^2 \hat{\sigma}_i^2 \;/\; \Sigma W_i$$

When r = 1 and each z_i = 1 represents a constant term, then equation (12) reduces to equation (10), the PEB confidence interval analyzed in the previous section, 4.1.

Tables 40 and 41 present the results of a series of Monte Carlo experiments investigating the reliability of the PEB confidence interval for each of the analytical review models discussed above. The results show that that the 95% nominal PEB confidence intervals provide actual reliability levels of 90 to 93% for samples sizes of n = 50 and 92.5 to 94.5% for n=100. While the PEB reliability is slightly less than the desired amount, it is never significantly less than the classical reliability level shown in the PPS-MPU columns of Tables 40 and 41.

Table 42 exhibits the average ratio of the width of the PEB confidence intervals to the width of the classical

TABLE 40

Monte Carlo Simulation of the Reliability of 95% Confidence Intervals for the PEB Estimator Integrating Classical (PPS-MPU) Sampling and Analytical Review (AR) Estimators
Total Error Rates Greater Than 10%
n = 50

Degrees of Freedom in the Understatement/ Overstatement Distributions			Model 1		Model 2		Model 3		Model 4	
df_u	df_o	k	PPS-MPU	PEB	PPS-MPU	PEB	PPS-MPU	PEB	PPS-MPU	PEB
1	1	5	.920	.920	.916	.917	.918	.918	.919	.920
		7	.915	.912	.916	.916	.922	.921	.915	.916
		9	.919	.916	.916	.916	.920	.918	.917	.918
1	2	5	.902	.907	.911	.912	.908	.908	.914	.914
		7	.909	.908	.909	.908	.908	.905	.907	.908
		9	.911	.909	.909	.906	.905	.905	.909	.910
1	3	5	.907	.913	.904	.904	.902	.905	.901	.901
		7	.905	.907	.904	.904	.908	.906	.904	.904
		9	.905	.904	.903	.899	.906	.904	.904	.902
2	1	5	.915	.921	.916	.916	.920	.919	.917	.917
		7	.924	.923	.920	.919	.915	.914	.921	.921
		9	.919	.916	.919	.917	.914	.912	.920	.917
2	2	5	.931	.936	.932	.934	.935	.934	.932	.934
		7	.931	.931	.930	.929	.930	.929	.934	.932
		9	.932	.929	.929	.927	.936	.936	.929	.930
2	3	5	.925	.934	.928	.931	.925	.928	.924	.926
		7	.926	.932	.927	.926	.924	.925	.926	.927
		9	.928	.928	.925	.924	.927	.926	.928	.927

TABLE 41

TABLE 41

Monte Carlo Simulation of the Reliability of 95% Confidence Intervals for the PEB Estimator Integrating Classical (PPS-MPU) Sampling and Analytical Review (AR) Estimators
Total Error Rates Greater Than 10%
n = 100

Degrees of Freedom in the Understatement/ Overstatement Distributions			Model 1		Model 2		Model 3		Model 4	
df_u	df_o	k	PPS-MPU	PEB	PPS-MPU	PEB	PPS-MPU	PEB	PPS-MPU	PEB
1	1	5	.936	.936	.935	.935	.943	.942	.932	.933
		7	.935	.935	.934	.933	.939	.938	.936	.938
		9	.939	.937	.937	.936	.939	.938	.942	.942
1	2	5	.927	.929	.931	.930	.933	.932	.930	.930
		7	.930	.931	.929	.930	.929	.929	.938	.934
		9	.932	.928	.932	.933	.930	.928	.930	.929
1	3	5	.933	.936	.928	.929	.925	.925	.929	.929
		7	.925	.926	.926	.925	.926	.925	.928	.927
		9	.928	.927	.926	.924	.926	.925	.929	.930
2	1	5	.937	.939	.934	.935	.938	.939	.940	.940
		7	.932	.931	.938	.938	.934	.934	.937	.936
		9	.938	.936	.935	.934	.935	.932	.935	.936
2	2	5	.941	.944	.943	.945	.945	.946	.941	.941
		7	.940	.941	.946	.945	.940	.938	.939	.940
		9	.943	.942	.942	.941	.942	.942	.942	.941
2	3	5	.938	.943	.936	.940	.942	.942	.939	.940
		7	.941	.943	.938	.939	.940	.940	.941	.941
		9	.939	.938	.938	.939	.939	.939	.939	.937

TABLE 42

Ratio of PEB to Classical 95% Confidence Interval Widths
with Integration of Auxiliary Analytical Review Information
Total Error Rates Greater Than 10%

Degrees of Freedom in the Understatement/Overstatement Distributions			Model 1		Model 2		Model 3		Model 4	
df_u	df_o	k	n = 50	n = 100	n = 50	n = 100	n = 50	n = 100	n = 50	n = 100
1	1	5	.976	.986	.988	.995	.991	.996	.997	.999
		7	.943	.969	.977	.988	.982	.991	.993	.997
		9	.928	.961	.970	.985	.976	.989	.991	.996
1	2	5	.964	.977	.981	.989	.985	.992	.994	.997
		7	.915	.949	.957	.978	.966	.983	.987	.994
		9	.890	.934	.947	.972	.958	.979	.983	.992
1	3	5	.957	.968	.972	.984	.976	.987	.990	.995
		7	.889	.926	.938	.965	.948	.972	.978	.989
		9	.852	.900	.919	.955	.935	.965	.973	.986
2	1	5	.969	.982	.985	.992	.988	.993	.995	.998
		7	.928	.958	.967	.984	.973	.988	.989	.995
		9	.906	.946	.956	.978	.966	.984	.988	.994
2	2	5	.960	.974	.976	.988	.982	.990	.992	.997
		7	.903	.940	.950	.973	.960	.979	.984	.992
		9	.870	.921	.934	.966	.950	.974	.980	.990
2	3	5	.952	.967	.971	.982	.974	.985	.988	.994
		7	.884	.919	.934	.960	.948	.968	.977	.987
		9	.841	.892	.910	.951	.926	.959	.968	.985

confidence intervals observed during the Monte Carlo experiments. The PEB confidence intervals are able to exploit the efficiency of the estimator, generating narrower intervals than under classical sampling procedures without sacrificing reliability.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this dissertation was to examine the use of empirical Bayes estimation techniques as an audit tool. Applications of parametric empirical Bayes estimation procedures were considered for both the intenal control compliance testing and account balance substantive testing phases of the audit. The following two sections summarize the results of this study for each of these applications. Section 5.3 considers issues raised by this study which indicate a need for future research.

5.1 Compliance Test Applications

The behavior of the PEB estimator of population error rates for internal control compliance testing was examined from both the frequentist perspective of fixed population error rates and the Bayesian perspective of random underlying population error rates. The results for each of these two perspectives are summarized separately in the following discussion.

The PEB error rate estimator proposed in this study has an interpretation from the frequentist perspective as the

James-Stein estimator of a set of fixed population error rates. Under ideal conditions, for any set of fixed population error rates the James-Stein estimator has smaller expected ensemble squared error than the classical maximum likelihood estimator currently used by auditors. However, with the small error rates encountered by auditors these ideal conditions may not be met in practice. Thus, the actual behavior of the JS/PEB error rate estimator was examined for various sample sizes, error rate patterns and numbers of populations likely to be encountered by auditors. In total 113 exact numerical computations or Monte Carlo simulations were performed computing the PEB/JS frequentist mean squared error under these various sample size and population error rate specifications. In each instance the PEB/JS estimator was ensemble risk efficient relative to MLE estimation.

Matsumura and Tsui [1982] are critical of the use of the PEB/JS error rate estimator in the presence of low error rates and low sample sizes often present in attributes sampling. They propose using a Stein-type estimator based on the Poisson distribution in lieu of the PEB/JS estimator which is based on the normal distribution. However, this study presented five reasons why the auditor might not wish to abandon the PEB/JS estimator in favor of a Poisson based Stein-type estimator. Additionally, the results of this study show that their criticism of the PEB/JS estimator is substantially without merit since under the wide variety of scenarios examined the PEB/JS estimator was consistently efficient relative to MLE, even in the face of relatively small sample sizes and low population error rates. Finally, in a series of 26 Monte Carlo simulations over various low error rate patterns and two sample sizes the most operationally feasible Poisson based Stein-type estimator proposed by Matsumura and Tsui was found to be grossly inefficient relative to the PEB/JS estimator.

In general, the results of this study confirm that the PEB/JS estimator is consistently ensemble risk efficient relative to classical MLE procedures. This efficiency tends to increase with the number of populations used to construct the estimators and the closer the population error rates are to each other.

The results also confirm that the PEB/JS estimator is frequentist biased. This bias is in the direction of overstatements in low error rate populations and understatements in high error rate populations. The authoritative auditing literature indicates the auditor is more concerned with the audit risks of overreliance on internal controls which may be induced by underestimating error rates than with the inefficiency of underreliance on internal controls induced by overrestimating error rates. This study proposed a "positive adjustment" estimator equal to the greater of the PEB/JS and MLE estimators. This estimator is everywhere biased toward error rate overestimation and is never less than the classical MLE estimator used by auditors. Nevertheless, computations of the expected ensemble squared errors for the two estimators demonstrated that the positive adjustment estimator was efficient relative to MLE over each of the 70 error rate pattern and sample size combinations considered.

An examination into the efficiency of the PEB/JS error rate estimator was also made from the Bayesian perspective of random error rates. A series of 72 Monte Carlo experiments over various underlying prior distributions for the population error rates and two sample sizes were performed. These investigated the efficiency of the PEB/JS and pure Bayes estimators relative to classical estimation. The results showed that in every instance the PEB/JS estimator was efficient relative to MLE. Additionally, the pure Bayes estimator was found to be preferred to both MLE and PEB/JS when the parameters of the prior distribution were accurately specified ex ante by the auditor. However, it was shown that the costs of parameter misspecification can be great, rendering the pure Bayes estimator inefficient relative to both PEB/JS and MLE estimation. One of the major benefits of PEB estimation demonstrated repeatedly in this study is the avoidance of these misspecification costs by estimating the parameters of the prior distribution from the sample data itself.

The PEB estimator was constructed using the normal distribution as the assumed functional form for the prior distribution on the error rates. However, the behavior of the estimator appears to be robust against this functional form specification as it was found to be efficient relative to MLE when the prior error rates were generated under both the uniform and the exponential distributions as well as the normal distribution.

Finally, the construction of upper confidence limits for the true population error rates based on the PEB/JS estimator was examined. Upper confidence limits from the frequentist perspective were proposed using bootstrapping methods. PEB confidence intervals proposed by Morris [1983a, 1983c] were examined from the Bayesian perspective of random error rates. The results of a series of Monte Carlo experiments showed that the proposed confidence limits provide actual reliability which is only slightly less than the desired nominal level. This reliability was obtainedwith confidence intervals which were consistently narrower than under classical estimation.

In addition to being inefficient, the pure Bayes estimator also generated confidence intervals which were grossly unreliable when the prior distribution parameters were misspecified. This cost of missspecification is avoided by PEB estimation by estimating these parameters using the sample data itself.

The implications of these results to the auditor are that more precise estimates of population error rates are obtainable without increasing sample sizes or engaging in risky subjective Bayesian prior probability parameter specifications. These more efficient estimates are obtained through parametric empirical Bayes estimation procedures. This increased efficiency can be exploited to provide confidence intervals for the true error rates which are narrower than under classical estimation. These results hold for both the frequentist and Bayesian perspective of the auditor's sampling problem.

5.2 Substantive Test Applications

The behavior of PEB techniques when estimating the true account balance during the auditor's substantive testing was also examined in this study. PEB estimators were proposed for the integration of classical sampling results across populations as well as with auxiliary analytical review model information. The results for each of these two uses of PEB procedures are summarized separately in the following discussion.

A PEB variables sampling estimator was proposed as a technique for combining the results of classical statistical sampling procedures across several account balance populations. The behavior of the PEB estimator was examined in the context of sampling for the true balances of inventory accounts. The error rate and error size distributions used in the Monte Carlo simulations of the estimator's behavior were based on published results of studies into these error characteristics for actual inventory populations. These distributions are consistent with those used by several recent studies of other statistical sampling estimators.

In total 72 Monte Carlo experiments were performed using various sample sizes, numbers of populations, and error rate and error size distributions. In every instance the PEB estimator was efficient relative to the classical probability proportional to size mean-per-unit estimator upon which it was based.

The efficiency of various pure Bayesian estimators combining the results of classical sampling with the auditor's subjective prior beliefs was also examined. As expected, the Bayes estimator was efficient relative to both classical and PEB estimation when the subjective ex ante specification of the parameters of the underlying error distribution was accurate. However, the costs of misspecification were shown to be potentially great, rendering the Bayes estimator inefficient relative to both the PEB and classical techniques. These misspecification costs are avoided by the PEB estimator which estimates the parameters of the prior distribution from the sample data itself rather than relying on risky ex ante auditor beliefs about the parameters. With as few as eight populations used in the construction of the estimator, PEB techniques approach the optimal efficiency of pure Bayesian estimation with proper prior parameter specification.

The reliability of PEB confidence intervals for the true population balance was also examined. These confidence intervals were shown to be somewhat unreliable, providing less than the desired level of probability coverage, when populations with error rates of less than 10% were admitted into the analysis. This result stems from the fact that the PEB estimator is based on a classical estimator which is itself unreliable in these circumstances.

When estimation was limited to populations with error rates in excess of 10%, the reliability of the PEB confidence intervals increased dramatically to just slightly less than the desired nominal level. These confidence intervals were consistently narrower than those provided by classical procedures and yielded the same level of reliability. Pure Bayesian confidence intervals provided the desired reliability when the ex ante subjective specifications were accurate. However, they were grossly unreliable when the subjective specifications were inaccurate. This unreliability due to prior misspecification is avoided by PEB techniques which estimate the prior parameters from the sample data itself.

The implications of these results for the auditor are that more precise estimates of the true account balance are obtainable without increasing sample sizes or engaging in risky subjective Bayesian prior specifications. These more efficient estimates are obtained through parametric empirical Bayes estimation procedures. This increased efficiency holds regardless of the total error rates present in the populations. However, the reliability of the estimator using realistic sample sizes is reasonable only when the procedures are applied to populations with error rates greater than 10%. The auditor should use the results of alternative audit procedures and knowledge from prior years' examinations in order to apply the techniques only when error rates are expected to exceed 10% if he intends to obtain reliable confidence intervals. However, it is also true that classical estimation should not be used in situations where the error rate is expected to be low. Nonetheless, if the auditor chooses to employ classical procedures even in the face of low error rates, then PEB techniques can still be used to obtain more precise estimates. However, reliance on PEB confidence intervals in these circumstances is unwise, but no worse than reliance on classical confidence intervals.

PEB techniques which integrated the results of classical sampling estimators with auxiliary analytical review information were also examined. There are two advantages to this technique over the current practice. First, the PEB integration provides a single estimate which is more efficient than either the sampling or analytical review estimate used in isolation. Currently any integration of the two procedures into one estimate must be done subjectively, if at all. Secondly, the analytical review auxiliary information can be for the current audit period only. Typically, when employing analytical review estimation procedures, the auditor obtains a set of observations on audited account balances and auxiliary analytical review information variables from outside the period or client under audit. He then establishes a model linking the true account balance to the auxiliary information using the estimation period observations. The auditor must assume that the model is also valid for the audit observations. This assumption is not required under the PEB technque which can be constructed using observations from the audit period only. The sample information is used both as a direct estimate of the true account balance and in estimating the parameters of the analytical review model.

In total 144 Monte Carlo simulations examined the behavior of this type of PEB estimator. The performance of the estimator was analyzed over various population error distributions, underlying auxiliary information models, sample sizes, and numbers of populations used in the construction of the estimator. For the reasons discussed above this analysis was limited to populations with error rates in excess of 10%.

In every instance the PEB estimator was efficient relative to classical estimation. This efficiency was obtained despite the fact that the pure analytical review estimator was inefficient relative to classical sampling. PEB techniques were successful in integrating a relatively noisy auxiliary analytical review variable to obtain a more efficient estimate of the true account balance. In general, these efficiency gains were greater with the more informative auxiliary information models. A more informative auxiliary information analytical review model is one with lower total variance and/or higher correlation between the true account balance and the auxiliary information variables.

The reliability of PEB confidence intervals about these estimates was also examined. The observed reliability was the same as under classical estimation and just slightly less than the desired nominal level. Additionally, the PEB confidence interval widths were somewhat narrower than under classical estimation.

5.3 Implications for Future Research

Several issues were raised by this study which indicate the need for future research. Three of the most important topics are discussed in the following.

The efficiency investigations made in this study were, for the most part, made with respect to a quadratic loss function. However, some evidence was cited in the authoritative literature (SAS No. 39) and prior academic research (Scott [1975]) which indicates that the auditor's loss function may be asymmetric. While squared error appears to be a close approximation to the auditor's loss, additional research into the exact form and parameters of the auditor's loss function could lead to the development of improved PEB estimators which reduce the auditor's actual expected loss.

Secondly, this dissertation investigated the behavior of parametric empirical Bayes estimators in audit testing. PEB procedures assume a given functional form for the underlying distribution of the item at interest. However, the identifying parameters of this prior distribution are left unspecified and are estimated from the data itself. The results of this dissertation show that the behavior of the PEB estimator in auditing applications is robust against the normal functional form specification. Nonetheless, an interesting extension of this line of research would investigate the behavior of nonparametric empirical Bayes procedures which estimate both the functional form and the identifying parameters of the underlying distribution from the sample data.

Finally, the results of this dissertation show that PEB confidence intervals for the true account balance are unreliable in the presence of low error rate populations. This stems from the PEB estimator's reliance on a classical estimator which is itself unreliable in these circumstances. For this reason a series of dollar-unit sampling procedures have been developed which construct conservative upper confidence bounds for the true account balance in the presence of low error rates. These dollar-unit sampling or combined attributes-variables sampling techniques do not rely upon a classical statistical sampling point estimator. Recent attempts have been made to introduce Bayesian methods to these techniques (e.g., McCray [1984], Godfrey and Neter [1984], and Dworin and Grimlund [1986]). Future research might examine the potential of introducing empirical Bayes procedures to the dollar-unit-sampling methods.

BIBLIOGRAPHY

BIBLIOGRAPHY

Akresh, A., and W. Wallace. "The Application of Regression Analysis for Limited Review and Audit Planning." Symposium on Auditing Research IV. Urbana: University of Illinois, 1980:69-128.

Albrecht, W., and J. McKeown. "Toward an Extended Use of Statistical Analytical Reviews in the Audit." Symposium on Auditing Research II. Urbana: University of Illinois, 1977:53-69.

American Institute of Certified Public Accountants. Codification of Statements on Auditing Standards: Numbers 1 to 49. New York: AICPA, 1985.

________. Audit Sampling. New York: AICPA, 1983.

Arens, A.A., and J.K. Loebbecke. Applications of Statistical Sampling to Auditing. Englewood Cliffs: Prentice-Hall, Inc., 1981.

Bailey, A.D., Jr. Statistical Auditing: Review, Concepts, and Problems. New York: Harcourt Brace Jovanovich, Inc., 1981.

Beck, P.J., I. Solomon, and L.A. Tomassini. "Subjective Prior Probability Distributions and Audit Risk." Journal of Accounting Research (Spring 1985):37-56.

Berger, J.O. Discussion of "Construction of Improved Estimators in Multiparameter Estimation for Discrete Exponential Families." Annals of Statistics (June 1983):368-369.

Birnberg, J.G. "Bayesian Statistics: A Review." Journal of Accounting Research (Spring 1964):108-116.

Blocher, E. "Assessment of Prior Distributions: The Effect on Required Sample Size in Bayesian Audit Sampling." Accounting and Business Research (Winter 1981):11-20.

Carter, G.M., and J.E. Rolph. "Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities." Journal of the American Statistical Association (December 1974):880-885.

Chambers, J.M. Computational Methods for Data Analysis. New York: John Wiley & Sons, 1977.

Chesley, G.R. "Elicitation of Subjective Probabilities: A Review." The Accounting Review (April 1975):325-337.

Clemmer, B.A., and R.G. Krutchkoff. "The Use of Empirical Bayes Estimators in a Linear Regression Model." Biometrika (November 1968):525-539.

Cochran, W.G. Sampling Techniques. New York: John Wiley & Sons, 1977.

Committee on Basic Auditing Concepts. A Statement of Basic Auditing Concepts, Studies in Accounting Research No. 6. Sarasota: American Accounting Association, 1973.

Corless, J.C. "Assessing Prior Distributions for Applying Bayesian Statistics in Auditing." The Accounting Review (July 1972):556-566.

Crosby, M.A. "Implications of Prior Probability Elicitation on Auditor Sample Size Decisions." Journal of Accounting Research (Autumn 1980):585-593.

________. "Bayesian Statistics in Auditing: A Comparison of Probability Elicitation Techniques." The Accounting Review (April 1981):355-365.

________. "The Development of Bayesian Decision - Theoretic Concepts in Attribute Sampling." Auditing: A Journal of Practice and Theory (Spring 1985):118-132.

Diaconis, P., and B. Efron. "Computer - Intensive Methods in Statistics." Scientific American (May 1983):116-130.

Dworin, L., and R.A. Grimlund. "Dollar Unit Sampling for Accounts Reveivable and Inventory." The Accounting Review (April 1984):218-241.

________. "Dollar-Unit Sampling: A Comparison of the Quasi-Bayesian and Moment Bounds." The Accounting Review (January 1986):36-57.

Duke, G.L., J. Neter, and R.A. Leitch. "Power Characteristics of Test Statistics in the Auditing Environment: An Empirical Study." Journal of Accounting Research (Spring 1982):42-67.

Efron, B. "Bootstrap Methods: Another Look at the Jackknife." The Annals of Statistics (January 1979): 1-26.

________. The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia: Society for Industrial and Applied Mathematics, 1982.

Efron, B., and C. Morris. "Limiting the Risk of Bayes and Empirical Bayes Estimators -- Part II: The Empirical Bayes Case." Journal of the American Statistical Association (March 1972):130-139.

________. "Stein's Estimation Rule and Its Competitors -- An Empirical Bayes Approach." Journal of the American Statistical Association (March 1973):117-130.

________. "Data Analysis Using Stein's Estimator and Its Generalizations." Journal of the American Statistical Association (June 1975):311-319.

________. "Stein's Paradox in Statistics." Scientific American (May 1977):119-127.

Fay, R.E. III, and R.A. Herriot. "Estimates of Income for Small Places: An Application of James - Stein Procedures to Census Data." Journal of the American Statistical Association (June 1979):269-277.

Felix, W.L., Jr. "Evidence on Alternative Means of Assessing Prior Probability Distributions for Audit Decision Making." The Accounting Review (October 1976):800-807.

Felix, W.L., Jr., and R.A. Grimlund. "A Sampling Model for Audit Tests of Composite Accounts." Journal of Accounting Research (Spring 1977):23-41

Frost, P.A., and H. Tamura. "Jackknifed Ratio Estimation in Statistical Auditing." Journal of Accounting Research (Spring 1982):103-120.

Garstka, S.J., and P.A. Ohlson. "Ratio Estimation in Accounting Populations with Probabilities of Sample Selection Proportional to Size of Book Values." Journal of Accounting Research (Spring 1979):23-59.

Ghosh,M., J.T. Hwang, and K. Tsui. "Construction of Improved Estimators in Multiparameter Estimation for Discrete Exponential Families." Annals of Statistics (June 1983):351-367.

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

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

Grimlund, R.A. "A Framework for the Integration of Auditing Evidence." unpublished Ph.D. dissertation, University of Washington, 1977.

Hoadley, B. "The Quality Measurement Plan (QMP)." Bell System Technical Journal (February 1981):215-273.

Ijiri, Y., and R.W. Leitch. "Stein's Paradox and Audit Sampling." Journal of Accounting Research (Spring 1980):91-108.

Jackson, D.A., T.M. O'Donovan, W.J. Zimmer, and J.J. Deeley. "Minimax Estimators in the Exponential Family." Biometrika (August 1970):439-443.

James, W., and C. Stein. "Estimation with Quadratic Loss," in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. Berkeley: University of California Press, 1961.

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

Judge, G.G., and M.E. Bock. The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics Amsterdam: North-Holland, 1978.

Kaplan, R. "Developing a Financial Planning Model for Analytical Review." Symposium on Auditing Research III. Urbana: University of Illinois, 1978:3-30.

Kinney, W.R., Jr. "ARIMA and Regression in Analytical Review: An Empirical Test." The Accounting Review (January 1978):48-60.

________. "Integrating Audit Tests: Regression Analysis and Partitioned Dollar Unit Sampling." Journal of Accounting Research (Autumn 1979):456-475.

Kinney, W.R., Jr., and G.L. Salamon. "The Effect of Measurement Error on Regression Results in Analytical Review." Symposium on Auditing Research III. Urbana: University of Illinois, 1978:49-64.

________. "Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules." Journal of Accounting Research (Autumn 1982):350-366.

Kraft, W.H., Jr. "Statistical Sampling for Auditors: A New Look." Journal of Accountancy (August 1968):49-56.

Krutchkoff, R.G. "An Introduction to Empirical Bayes," in Proceedings of the Symposium on Empirical Bayes Estimation and Computing in Statistics, Texas Tech University Mathematics Series no. 6, edited by T.A. Atcheson and H.F. Martz, Jr. Lubbock: Texas Tech University, 1969.

Leitch, R.A., J. Neter, R. Plante, and P. Sinha. "Modified Multinomial Bounds for Larger Numbers of Errors in Audits." The Accounting Review (April 1982):384-400.

Lev, B. "On the Use of Index Models in Analytical Review by Auditors." Journal of Accounting Research (Autumn 1980):524-550.

Lindley, D.V. Discussion of C.M. Stein, "Confidence Sets for the Mean of a Multivariate Normal Distribution." Journal of the Royal Statistical Society, Series B, (1962):285-287.

Loebbecke, J.K., and J. Neter. "Considerations in Choosing Statistical Sampling Procedures in Auditing." Studies on Statistical Methodology in Auditing. Supplement to Journal of Accounting Research (1975):38-52.

Maier, S.F., D.W. Peterson, and J.H. Vander Weide. "An Empirical Bayes Estimate of Market Risk." Management Science (July 1982):728-737.

Marais, M.L. "An Application of the Bootstrap Method to the Analysis of Squared, Standardized Market Model Prediction Errors." Studies on Current Econometric Issues in Accounting Research. Supplement to Journal of Accounting Research (1984):34-54.

Marais, M.L., J.M. Patell, and M.A. Wolfson. "The Experimental Design of Classification Models: An Application of Recursive Partitioning and Bootstrapping to Commercial Bank Loan Classifications." Studies on Current Econometric Issues in Accounting Research. Supplement to Journal of Accounting Research (1984): 87-114.

Matsumura, E.A., and K. Tsui. "Stein-type Poisson Estimators in Audit Sampling." Journal of Accounting Research (Spring 1982):162-170.

Mautz, R.K., and H.A. Sharaf. The Philosophy of Auditing. Sarasota: American Accounting Association, 1961.

McCray, J.H. "A Quasi-Bayesian Audit Risk Model for Dollar Unit Sampling." The Accounting Review (January 1984): 35-51.

Morris, C.N. "Parametric Empirical Bayes Inference: Theory and Applications." Journal of the American Statistical Association (March 1983a):47-55.

________. "Natural Exponential Families with Quadratic Variance Functions: Statistical Theory." Annals of Statistics (June 1983b):515-529

________. "Parametric Empirical Bayes Confidence Intervals." Scientific Inference, Data Analysis, and Robustness: Publication no. 48 of the Mathematics Research Center of the University of Wisconsin -- Madison, edited by G.E.P. Box, T. Leonard, and C. Wu. New York: Academic Press, (1983c):25-50.

Morris, C.N., and L. Van Slyke. "Empirical Bayes Methods for Pricing Insurance Classes." Proceedings of the Business and Economics Statistics Section, American Statistical Association (1978):579-582.

Neter, J. "Two Case Studies on the Use of Regression for Analytical Review." Symposium on Auditing Research IV. Urbana: University of Illinois, 1980:293-337.

Neter, J., R.A. Leitch, and S.E. Feinberg. "Dollar Unit Sampling: Multinomial Bounds for Total Overstatement and Understatement Errors." The Accounting Review (January 1978):77-93.

Neter, J., and J.K. Loebbecke. Behavior of Major Statistical Estimators in Sampling Accounting Populations: An Empirical Study, Auditing Research Monograph no. 2. New York: AICPA, 1975.

Newman, D.P., and L.A. Tomassini. "Calibration of Subjective Probability Assessments: A Methodological Perspective," in Proceedings of the Conference on Decision Making in Accounting (1983).

Park, S.H. "A Model of Audit Process with Explicit Consideration of the Interrelationship Among Account Balances," unpublished Ph.D. dissertation, University of Iowa, 1977.

Parzen, E. "On Estimation of a Probability Density Function and Mode." Annals of Mathematical Statistics (1962):1065-1076.

Ramage. J.G., A.M. Krieger and L.L. Spero. "An Empirical Study of Error Characteristics in Audit Populations." Studies on Auditing -- Selections from the "Research Opportunities in Auditing" Program. Supplement to Journal of Accounting Research (1979):72-102.

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

Robbins, H. "An Empirical Bayes Approach to Statistics," in Proceedings of the Third Berkeley Symposium on Mathematics, Statistics and Probability, vol. 1. Berkeley: University of California Press, 1956.

________. "The Empirical Bayes Approach to Statistical Decision Problems." The Annals of Mathematical Statistics (1964):1-20.

Roberts, D.M. Statistical Auditing. New York: AICPA, 1978.

Rubin, D.B. "Using Empirical Bayes Techniques in the Law School Validity Studies." Journal of the American Statistical Association (December 1981):801-816.

Rutherford, J.R., and R.G. Krutchkoff. "Some Empirical Bayes Techniques in Point Estimation." Biometrika (March 1969):133-137.

Scott, W.R. "A Bayesian Approach to Asset Valuation and Audit Size." Journal of Accounting Research (Autumn 1973):304-330.

________. "Auditor's Loss Functions Implicit in Consumption-Investment Models." Studies on Statistical Methodology in Auditing. Supplement to Journal of Accounting Research (1975):98-117.

Shreider, Y.A. The Monte Carlo Method: The Method of Statistical Trials. Oxford: Pergamon Press, 1966.

Smith, V.K. Monte Carlo Methods: Their Role for Econometrics. Lexington, Mass.: Lexington Books, 1973.

Soloman, I., J.L. Krogstad, M.B. Romney, and L.A. Tomassini. "Auditors' Prior Probability Distributions for Account Balances." Accounting, Organizations and Society (January 1982):27-41.

Soloman, I., L.A. Tomassini, J.L. Krogstad, and M.B. Romney. "On the Use of the Equivalent Prior Sample Probability Elicitation Method by Auditors." Advances in Accounting (1984):267-290.

Sorenson, J.E. "Bayesian Analysis in Auditing." The Accounting Review (July 1969):555-561.

Stein, C. "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. Berkeley: University of California Press, 1956.

Stringer, K.W. "A Statistical Technique for Analytical Review." Studies on Statistical Methodology in Auditing. Supplement to Journal of Accounting Research (1975):1-13.

Tomassini, L.A., I. Soloman, M.B. Romney, and J.L. Krogstad. "Calibration of Auditors' Probabilistic Judgments: Some Empirical Evidence." Organizational Behavior and Human Performance (December 1982):391-406.

Tracy, J.A. "Bayesian Statistical Methods in Auditing." The Accounting Review (January 1969a):90-98.

Tracy, J.A. "Bayesian Statistical Confidence Intervals for Auditors." Journal of Accountancy (July 1969b):41-47.

Winkler, R.L. "The Assessment of Prior Distributions in Bayesian Analysis." Journal of the American Statistical Association (September 1967):776-800.