BAYESIAN INFERENCE AND PORTFOLIO EFFICIENCY

by

Shmuel Kandel
Robert McCulloch
Robert F. Stambaugh

8-91

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

Bayesian Inference and Portfolio Efficiency

by*

Shmuel Kandel
Recanati Graduate School of Business Administration
Tel-Aviv University

Robert McCulloch
Graduate School of Business
University of Chicago

Robert F. Stambaugh
The Wharton School
University of Pennsylvania
and
National Bureau of Economic Research

March 1991

Abstract

Bayesian posterior distributions allow one to investigate the approximate efficiency of a portfolio without specifying the maximum degree of inefficiency *a priori*. The difference in expected returns between the value-weighted equity portfolio and an efficient portfolio of equal variance has a disperse posterior distribution, and our experiments confirm that such uncertainty is inherent in the sample sizes typically encountered in empirical studies. The maximum correlation between the value-weighted portfolio and an efficient portfolio has a posterior that is concentrated, often around low values, but this result appears to reflect nonlinearity in the function rather than information in the sample.

---

*We are grateful to Craig MacKinlay, Peter Rossi, and workshop participants at the University of Pennsylvania for helpful discussions and comments. Shmuel Kandel was a visiting associate professor at the Wharton School while portions of this research were conducted. Robert McCulloch acknowledges support from the Graduate School of Business and the IBM Faculty Research Fund at the University of Chicago.

1. Introduction

A portfolio is *efficient* if it offers the highest expected return for a given variance of return. The concept of portfolio efficiency, pioneered by Markowitz (1952, 1959), offers obvious normative content as well as a framework for representing modern theories of financial asset pricing. In the latter context, for example, the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Black (1972) maintains that the market portfolio is efficient, and the consumption-based model of Breeden (1979) requires the efficiency of the portfolio whose return has the maximum correlation with consumption.

In the traditional empirical approach, efficiency is tested as a simple (point) hypothesis describing an observed "test" portfolio, and classical methods of statistical inference are used to accept or reject the hypothesis at a given significance level.[1] Two departures from this traditional approach have emerged. The first of these stems from the recognition that, in tests of asset pricing models, the test is an imperfect "proxy" for the true portfolio of interest, such as the market portfolio. This imprecision suggests a composite hypothesis that includes some amount of inefficiency in test portfolio. Kandel and Stambaugh (1987) and Shanken (1987a) have explored such approaches. A second major departure from the traditional approach relies on Bayesian inference instead of classical sampling theory. Shanken (1987b), McCulloch and Rossi (1990a, 1990b), and Harvey and Zhou (1990) develop and apply Bayesian approaches to drawing inferences about portfolio efficiency and asset pricing models.

This study develops and analyzes a framework that combines both of the above departures from the traditional approach. The inefficiency of the test portfolio can be described by univariate measures with simple intuitive appeal. We analyze two measures: Δ, the difference in expected returns between the test portfolio and an efficient portfolio of equal variance, and ρ, the maximum correlation between the returns on the test portfolio and an efficient portfolio. These definitions of Δ and ρ apply whether or not a riskless asset exists, and both measures incorporate the restriction that an efficient portfolio must lie on the *positively* sloped portion of the minimum-variance boundary.[2] Our framework yields posterior distributions for these inefficiency measures, and the posterior distributions reflect the prior uncertainty

[1]The earliest examples of this approach include Douglas (1969), Black, Jensen and Scholes (1972), and Fama and MacBeth (1973). For recent discussions and interpretations of such methods see Gibbons, Ross, and Shanken (1989), Kandel and Stambaugh (1989), and Shanken (1989).

[2]The latter restriction is not incorporated, for example, in the Bayesian approaches of Shanken (1987b), Harvey and Zhou (1990), and McCulloch and Rossi (1990b) or in the maximum-likelihood tests summarized by Kandel and Stambaugh (1989).

about the parameters of the joint distribution of all assets in the sample. Previous studies that have investigated portfolio efficiency in a Bayesian context have either conditioned on sample moments of the test portfolio or analyzed less interpretable measures, such as the difference in squared Sharpe measures between the test portfolio and an efficient portfolio. The previous Bayesian studies have analyzed the case in which a riskless asset is included; we analyze that case as well as the zero-beta case (no riskless asset).

It seems reasonable to investigate the composite hypothesis of "approximate efficiency" by obtaining a posterior distribution for a measure of inefficiency. This allows us to describe the portfolio of interest without specifying a given degree of inefficiency *a priori*. Such an approach offers an alternative to the tests proposed by Kandel and Stambaugh (1987) and Shanken (1987a), wherein the hypothesis of approximate efficiency is specified as a given minimum correlation between the test portfolio and an efficient portfolio.

The remainder of the paper proceeds as follows. Section 2 defines the measures of portfolio inefficiency and computes their posterior distributions for the value-weighted portfolio of the New York and American Stock Exchanges. Prior beliefs are represented as diffuse priors on the mean vector (E) and the covariance matrix (V) of the joint distribution of asset returns. We find that the posterior distribution of Δ is rather disperse. In contrast, most of the posterior probability for ρ is confined to a narrow range, concentrating on values well away from $\rho = 1$ when a riskless asset is included. To gain some perspective on these seemingly disparate results, we conduct experiments in section 3 designed to reveal the behavior of the posterior distributions for various sample sizes. With small sample sizes, as typically encountered in practice, posterior beliefs about the model's fundamental parameters E and V are fairly disperse. Nevertheless, even when the test portfolio is efficient *ex post*, these disperse posterior beliefs about E and V can imply posterior beliefs about ρ that are concentrated well away from unity, due to the nonlinear relations governing ρ. Considerably larger samples are required to have the posterior of ρ reflect substantial information in the data. Thus, although ρ has simple intuitive appeal as a measure of inefficiency, casual inspection of its posterior distribution can be misleading. We conclude that evidence favoring inefficiency of the value-weighted equity portfolio is weaker than other approaches suggest.

2. Posterior Distributions of Inefficiency Measures

2.1. Definitions and Notation

Let the vector R_t contain returns in period t on n risky assets, where one of the n assets is the test portfolio whose efficiency is to be investigated. If a riskless asset exists, then R_t contains excess returns on these assets, i.e., returns in excess of the riskless rate. Let $\mu(p)$ and $\sigma(p)$ denote the mean and the standard deviation of a given portfolio p, where moments of "portfolios" should be interpreted throughout as moments of returns on the portfolios. A "bar" over a portfolio, e.g. $\bar{p}$, denotes a *minimum-variance portfolio*, a portfolio with minimum variance among those having the same mean. Portfolio $\bar{g}$ has the global minimum variance among all portfolios composed solely of risky assets. Let $\bar{p}_\mu$ denote the minimum-variance portfolio with the same mean as portfolio p, and let $\bar{p}_\sigma$ denote the efficient portfolio with the same standard deviation as p.

When a riskless asset exists, we assume that $\mu(\bar{g}) > 0$ (recall that R_t then contains excess returns). The Sharpe measure of portfolio p, $S(p)$, is defined, in mean-standard-deviation space, as the slope of the ray connecting p to the origin. That is $S(p) \equiv \mu(p)/\sigma(p)$. The Sharpe-Lintner tangent portfolio $\bar{s}$ is the (efficient) portfolio of risky assets having the maximum Sharpe measure.

For a given portfolio p, the two measures of inefficiency, Δ and ρ, are easily computed from the above quantities. The first measure is computed as

$$\Delta = \mu(\bar{p}_\sigma) - \mu(p) \ . \tag{1}$$

This measure is simply the average loss in rate of return associated with holding an inefficient portfolio instead of an efficient portfolio of equal risk. For the computation of ρ, the maximum correlation between p and an efficient portfolio, we rely on the results of Kandel and Stambaugh (1987). In the presence of a riskless asset,

$$\rho = S(p)/S(\bar{s}) \ . \tag{2}$$

That is, ρ is simply the ratio of the portfolio's Sharpe measure to the maximum Sharpe measure.[3] When no riskless asset exists, then

$$\rho = \begin{cases} \sigma(\bar{p}_\mu)/\sigma(p) & \text{if } \mu(p) > \mu(\bar{g}) \\ \sigma(\bar{g})/\sigma(p) & \text{otherwise.} \end{cases} \tag{3}$$

[3]Shanken (1987) also provides this result.

The inefficiency measures Δ and ρ for the test portfolio are computed as nonlinear functions of E and V, the parameters of the joint distribution of returns on the n risky assets. If the test portfolio is efficient, then $\Delta = 0$ and $\rho = 1$.

2.2. Constructing Marginal Posterior Distributions for Δ and ρ

We assume that the n-vector of returns R_t is distributed multivariate normal, independently across t, with mean E and variance-covariance matrix V. For the prior distribution of E and V in the case without a riskless asset, we use the usual non-informative prior of the multivariate normal distribution [e.g., Zellner (1971)]:

$$\mathrm{p}(E, V) \propto |V|^{-(n+1)/2} \tag{4}$$

For the case with a riskless asset, we truncate this prior distribution so that positive density is assigned to the parameters (E, V) only if they satisfy the condition that $\mu(\bar{g}) > 0$. Technically, we augment the prior in (4) by an indicator function that receives the value of unity if this condition is satisfied and zero otherwise.

Note that the return on the test portfolio is included as an element of the n-vector R_t. In the case where the riskless asset is assumed to exist, an alternative approach is to consider the distribution of returns on the other $n-1$ assets conditional on the return on the test portfolio. It is well known that this conditional distribution will be in the form of a multivariate regression model, and the test portfolio is on the minimum-variance boundary of the $n-1$ assets if and only if the regression intercepts equal zero. The latter hypothesis about the regression intercepts has been considered in the Bayesian framework by Shanken (1987b), McCulloch and Rossi (1990), and Harvey and Zhou (1990). In this paper we investigate the efficiency of a given portfolio by computing the posterior distribution of measures of inefficiency based on that portfolio's location in mean-standard-deviation space. With this approach there is an important difference between the conditional model and the unconditional model, which includes the given portfolio in the joint multivariate normal distribution. The inefficiency measures Δ and ρ depend on the mean and the standard deviation of the test portfolio. In the conditional approach, estimated values for the test portfolio's mean and standard deviation are used to construct the desired posterior distributions, but uncertainty about those estimated values is ignored. Our posterior distributions in the unconditional approach reflect uncertainty about the parameters of the conditional distribution of the $n-1$ assets given the test portfolio as well as the parameters of the marginal distribution of the test portfolio.

For a given sample of T observations of R_t, define the sample estimates

$$\hat{E} = \frac{1}{T}\sum_{t=1}^{T} R_t , \quad (5)$$

$$\hat{V} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{E})(R_t - \hat{E})' . \quad (6)$$

Given a sample of T observations and the prior distribution (4), the matrix V^{-1} has a Wishart posterior distribution with $(T-1)$ degrees of freedom and parameter matrix $(T\hat{V})^{-1}$. Given V, the posterior distribution of E is normal with mean $\hat{E}$ and covariance matrix V/T. These posterior distributions apply for the case without a riskless asset. When a riskless asset exists, the joint posterior distribution of (E, V) is adjusted to reflect the zero prior probability of values that are inconsistent with $\mu(\bar{g}) > 0$.

Although the inefficiency measures Δ and ρ are functions of E and V, we do not obtain the marginal posterior distributions of Δ and ρ analytically given the joint posterior distribution of E and V. Instead, we follow a Monte Carlo method and repeatedly draw values for E and V from the posterior distribution given above. For each draw of E and V, we compute the corresponding values of Δ and ρ. For a large number of draws, the empirical distribution obtained from the computed values of each measure will provide a good approximation to the posterior marginal distribution. The results presented below are based on 5000 draws.

Note that with the use of the non-informative prior, some aspects of our Bayesian posterior distribution correspond to confidence regions obtained by standard frequentist analysis. For example, the quantity $(\hat{E} - E)'\hat{V}^{-1}(\hat{E} - E)$ has a posterior distribution (where E is random and $\hat{E}$ and $\hat{V}$ are considered fixed) exactly the same as the frequentist distribution (where E and V are considered fixed but $\hat{E}$ and $\hat{V}$ are random). The posterior and frequentist distributions are both $[n/(T-n)]\mathrm{F}_{n,T-n}$. The usual elliptical confidence region for E has posterior probability equal to the frequentist confidence level. Thus, our Bayesian results provide inferences about basic model parameters with a level of uncertainty equivalent to that obtained through the frequentist approach. This suggests that while the prior in (4) is sufficiently diffuse to represent ignorance, it does not inject an inordinate amount of uncertainty into our inference process. The Bayesian analysis allows us to compute marginal posterior distributions for the measures Δ and ρ, while there is no easy way to compute the corresponding frequentist confidence intervals.

2.3. Sample Results

Our sample includes weekly data for the period January 1, 1963, through December 31, 1987. The daily CRSP master is used to compute weekly raw returns on stocks of the New York Stock Exchange (NYSE) and the American Stock Exchange (ASE). Weekly returns on treasury bills are computed from quotations of T-bills obtained from the *Wall Street Journal*. The T-bill returns were originally compiled by Oldfield and Rogalsky (1987) and were subsequently extended by Ferson, Kandel, and Stambaugh (1987) and McCulloch and Rossi (1990).

We consider a set of twelve risky assets: ten size-ranked portfolios and two indices of the stock market. The formulation and rebalancing of the size-ranked portfolios are described in a greater detail by McCulloch and Rossi (1990a). Every four weeks, firms are sorted by market value and placed into decile portfolios. Portfolios' raw returns are computed by value-weighting of the individual firms returns. The two market indices are the NYSE–ASE value-weighted and equally weighted portfolios. Excess returns are computed by subtracting the weekly returns on the T-bills from the raw returns. The test portfolio is the value-weighted market.

Figure 1a presents the marginal posterior of ρ and figure 1b presents the marginal posterior of Δ when the riskless asset is included. The values of Δ are multiplied by 52 to put them roughly on an annual basis. The histogram for ρ seems to give clear evidence against the hypothesis that the value-weighted market portfolio is efficient. Virtually all of the mass lies between the values of $\rho = 0$ and $\rho = .25$. The posterior distribution of Δ assigns appreciable mass to differences ranging from 0 up to values well in excess of 100% per year, indicating a very high level of uncertainty. Thus, the two inefficiency measures tell different stories: the posterior of ρ is concentrated on values clearly inconsistent with efficiency, while the posterior of Δ is spread over a wide range of values. Figure 2 presents the results obtained when the riskless asset is excluded. The results are different from those of figure 1 in that the posterior of ρ lies closer to unity, so neither measure provides clear evidence against the model. Nevertheless, the posterior for Δ still indicates substantial uncertainty about Δ while the posterior for ρ is again more concentrated. Which measure should we believe? Can a sample provide information about one measure and not the other? Is the sample sufficiently informative to allow us to make a reliable inference? These questions are pursued in section 3. We show there how the posterior of ρ, a bounded and non-linear function of E and V, can appear to be tight, even though the posteriors of E and V are not.

We conclude this section by presenting the posterior distribution of the mean return of the value-weighted market. The histogram in figure 3a gives the posterior of the mean of value-weighted market returns in excess of the risk-free rate, and the histogram in figure 3b gives the posterior of the mean for total returns. Although the marginal posterior of the mean return can be obtained analytically (it is Student t), we present the Monte-Carlo distribution for comparison with the earlier figures. Clearly, there is substantial posterior uncertainty about the mean returns. Both distributions assign appreciable mass to intervals having ranges of about 15%. Since the joint distribution of the 12 risky assets has an additional 89 parameters, the overall level of uncertainty about E and V given the sample would seem to be high. The posterior distribution of Δ appears to reflect this uncertainty.

3. Interpretation of the Posteriors

3.1. Tight Posteriors and Uninformative Samples: A Simple Example

We first present a simple example illustrating how the marginal distribution of a function of parameters can appear to be tight even though the distribution of the underlying parameters is not. If f is a non-linear function and μ is a random variable, we know that $f(E\{\mu\})$ need not equal $E\{f(\mu)\}$. In order to interpret the results obtained in section 2, it is also important to note that the variance of $f(\mu)$ need not decrease monotonically as the variance of μ decreases.

Our underlying parameter, μ, will be the mean excess return on the value-weighted market. To assess the effect of the level of uncertainty, we consider the marginal posterior distribution of μ that would be obtained given artificial multivariate samples of various sizes having the same sample mean $\hat{E}$ and sample covariance matrix $\hat{V}$ as our actual observations. Since the sample mean vector and covariance matrix are sufficient statistics for the multivariate normal model given the number of observations, these are the only sample characteristics that we must specify in order to compute the posterior for μ.

Figure 4a displays the posteriors for μ obtained from samples corresponding to 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data. For each sample size, the 1%, 25%, 50%, 75%, and 99% quantiles of the posterior distribution are indicated. A solid line connects the 75% and 25% quantiles, and dotted lines extend out to the 1% and 99% quantiles. The sample size increases from left to right on the horizontal scale. One can read from the figure, for example, that the 99% quantile of the posterior for μ given 25 years of data is about 10% (the returns have been multiplied by 52). As expected, the distributions tighten significantly as the sample size grows—there is relatively little uncertainty about the market mean after 8,000 years. As the distribution tightens, the mean stays constant at about .034.

Figure 4b displays the corresponding posteriors of the nonlinear function $f(\mu)$ given by,

$$f(\mu) = \left(1 + [500 \cdot (\mu - .034)]^2\right)^{-1} . \tag{7}$$

Note that $0 \leq f \leq 1$, so that f is bounded both above and below and $f(.034) = 1$. The distribution of $f(\mu)$ behaves quite differently from that of μ. Even though the median of μ is constant, the median of $f(\mu)$ moves from the smallest possible value of 0 to nearly the maximum value of 1. If we measure the variability of the distribution of $f(\mu)$ by the interquartile range (the difference between the 75% and the 25% quantiles) the variability of $f(\mu)$ is actually smallest when the variability of μ is largest. As the sample size increases and the distribution of μ concentrates around .034, the distribution of $f(\mu)$ initially spreads out before finally tightening around $f(E\{\mu\}) = 1$.

Of course, this peculiar behavior of the posterior of $f(\mu)$ arises because f maps all values sufficiently far from .034 to values close to 0. When the variance of μ is large, we get values far from the mean of .034 with high probability and they are all essentially mapped to 0. Only when the distribution of μ is tight around the mean value of .034 does the distribution of $f(\mu)$ concentrate around 1.

The important practical consequence of this exercise is that, for a measure such as ρ, which is nonlinear in E and V and bounded, a seemingly tight posterior need not imply that the sample is informative. If the posterior distribution of E and V is diffuse, then most of the mass is on (E, V) pairs for which the test portfolio is far from efficient and thus the corresponding ρ is close to 0. This situation must not be confused with that in which the joint posterior of (E, V) is concentrated about a set of values corresponding to $\rho \approx 0$. In both cases we obtain a marginal posterior distribution for ρ that is concentrated around 0, but only in the latter case do the data provide evidence against efficiency of the tested portfolio.

3.2. Sample Size and the Posteriors for Δ and ρ

In this section we show that the posterior of ρ behaves much like the posterior of the function f in our simple example above. We perform similar sample-size experiments to study the effect of the level of information on the posterior distributions of both ρ and Δ for the case in which a riskless asset is included.

In order to compute the posterior distribution of E and V, we need only specify the sample size and the sufficient statistics. In section 3.1 we used $\hat{E}$ and $\hat{V}$, which are the maximum likelihood estimates of E and V for the multivariate normal model, in combination with various sample sizes. In this section we will use the maximum likelihood estimates subject to the restriction that the value-weighted market be the Sharpe-Lintner tangent portfolio $\bar{s}$. Let $\hat{E}_r$ and $\hat{V}_r$ denote the *restricted* maximum likelihood estimates in our sample. We compute the marginal posterior distributions of ρ and Δ that would be obtained given a sample having *unrestricted* maximum likelihood estimates equal to $\hat{E}_r$ and $\hat{V}_r$ and sample sizes of 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data. The choices of $\hat{E}_r$ and $\hat{V}_r$ for the sufficient statistics gives us samples that provide the greatest possible support for the hypothesis that the value-weighted market is efficient given the sample size.[4] The sample size determines the quantity of information. Clearly, these artificial samples cannot provide evidence against the hypothesis. They may however, fail to provide conclusive evidence in favor of the hypothesis when the sample sizes are small. When the sample sizes are large, the posterior distribution of (E, V) will collapse around $\hat{E}_r$ and $\hat{V}_r$. Since ρ and Δ equal 1 and 0 respectively when E and V equal $\hat{E}_r$ and $\hat{V}_r$, the posterior distributions of ρ and Δ will collapse around 1 and 0.

Figures 5a and 5b display the posteriors of ρ and Δ for the eight artificial sample sizes. The posterior distributions are presented as in figure 4, with the sample size increasing from left to right. The marginal posteriors for ρ behave much like those of $f(\mu)$ in the example discussed previously in section 3.1. When the uncertainty about E and V is at its highest level, 98% of the mass is on values between 0 and .2 even though the unrestricted maximum likelihood estimate of ρ is 1. As the sample size increases and the uncertainty about E and V decreases, the posterior of ρ increases toward 1. The spread of the distribution first increases and then decreases as the posterior finally begins to tighten up around 1. The obvious interpretation of the posterior of ρ based on 25 years of data is that there is clear evidence against the efficiency of the test portfolio. By construction of the hypothetical data, we know that this is not the case.

Harvey and Zhou (1990) use a different sample and follow a somewhat different approach to constructing posterior distributions of ρ for the value-weighted NYSE portfolio. They use

[4] We repeated our experiments with the restricted maximum likelihood estimates based on subsamples of our data and obtained results similar to those reported.

the conditional (regression) form of the model and construct several posteriors of ρ, each conditioned on a set of prespecified values for the mean and variance of the value-weighted portfolio. As discussed in section 2.2, this approach does not include prior uncertainty about the parameters of the test portfolio. Nevertheless, Harvey and Zhou's posterior distributions for ρ are also concentrated away from 1, and the authors conclude from this evidence that "it is unlikely that the [NYSE] value-weighted portfolio is efficient." Given the experiments reported here, these conclusions may warrant reconsideration.

The interpretation of the posteriors for Δ in figure 5b is more straightforward. As the sample size increases, the spread of the distribution decreases and the distribution collapses around 0. Because Δ must always be positive, however, a higher variance of the posterior distribution tends to be accompanied by an overall shift to larger values Δ. With only 25 years of data, for example, about 98% of the mass occurs for values of Δ greater than 25%. While the posterior distribution for Δ better captures the level of parameter uncertainty than does the posterior for ρ, care is still needed in the interpretation of its location when the level of uncertainty is high.

Our results provide information about the amount of data that an uninformed person would require in order to make sharp inferences about the efficiency of the test portfolio. For example, in order to infer with .95 probability that Δ is less than 2% (annualized), figure 5b indicates that we would require more than 8000 years of weekly data. Evidently, only our distant descendants will have the opportunity to make reach such inferences. These results seem to parallel those of MacKinlay (1987), who concludes that multivariate tests of efficiency in the classical framework are likely to possess low power in samples of sizes encountered in practice. Note also that, since we set the sample's sufficient statistics equal to $\hat{E}_r$ and $\hat{V}_r$, the implied sample sizes required to infer a given degree of efficiency are actually *conservative*. That is, even if the true parameters E and V were such that the test portfolio were efficient, we would not observe sample moments $\hat{E}$ and $\hat{V}$ corresponding to a maximum likelihood estimate of ρ exactly equal to 1.

3.3. Informative Priors

Our experiments also allow us to analyze the roles of informative priors. In section 2.2 we noted that the general appeal of the improper prior in (4) stems from its non-informative property and the close correspondence between the resulting posterior and classical results. One might prefer, however, to use a proper prior distribution reflecting an informative set of beliefs about the model. For example, we might believe *a priori* that ρ lies between 0.5

and 1.0, so we would attempt to construct a prior distribution for E and V that places high prior probability on ρ between those values.

A standard choice for a proper prior in the multivariate normal model is the conjugate prior were E is multivariate normal given V, and V has a Wishart distribution. The conjugate prior has the property that the posterior will be of the same type. An important property of the conjugate prior is that it can always be viewed as the posterior distribution arising from the non-informative prior in (4) and a data set with T_p observations having sufficient statistics $\hat{E}_p$ and $\hat{V}_p$. The parameters of the conjugate prior simply represent choices of T_p, $\hat{E}_p$ and $\hat{V}_p$. With this choice of prior, the posterior obtained from a data set of size T is then exactly the same as the posterior obtained using the non-informative prior and an augmented data set of size $T + T_p$ consisting of the actual data combined with any data set with T_p observations and sufficient statistics $\hat{E}_p$ and $\hat{V}_p$.

We can construct a conjugate prior that puts the desired mass on large values of ρ by choosing $\hat{E}_p = \hat{E}_r$, $\hat{V}_p = \hat{V}_r$, and T_p large enough. Our experiments can then be interpreted as an investigation designed to find the appropriate value of T_p. To place about 50% of the prior probability for ρ between 0.5 and 1.0, for example, figure 5a indicates that T_p would have to correspond to almost 4000 years of weekly data. With such a large value of T_p underlying the informative prior, the posterior distribution resulting from only 25 years of data would result in a posterior virtually identical to the prior. Thus, a sample of the size encountered in practice could do little to influence our posterior beliefs about the efficiency of the test portfolio.

3.4. The Number of Assets

With n assets, the inefficiency measures Δ and ρ are nonlinear functions of the $\frac{n(n+3)}{2}$ underlying parameters in E and V. For a given number of time-series observations, intuition suggests that reducing the number of assets would reduce the uncertainty about the composition of efficient portfolios and thereby reduce the uncertainty about the inefficiency measures.

In this section we conduct additional experiments designed to investigate the role that the number of assets plays in determining the posterior uncertainty about Δ and ρ. Rather than use all 10 of the size-based portfolios, we now include only the first and tenth size based portfolios along with the equally and value-weighted market indexes. With only these 4 assets, the total number of parameters in the multivariate normal distribution is 14, as compared to 90 with all 12 assets. As in the previous section, we consider the case in which a riskless asset is included, and each *artificial* sample is assumed to generate *unrestricted* maximum likelihood estimates equal to $\hat{E}_r$ and $\hat{V}_r$, the *restricted* estimates in the *actual* sample. The sizes of the artificial samples are the same as those used previously: 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data.

Figures 6a and 6b display, for the smaller set of assets, the posterior distributions analogous to those in figures 5a and 5b for the larger set of assets. (The same scales are used to facilitate comparison.) The posteriors for ρ in figure 6a exhibit some similarities to those in figure 5a for the larger set of assets: in both figures the posteriors center around low values of ρ for the smaller sample sizes and then concentrate around 1 as the sample size increases. On the other hand, some interesting differences between the two figures also appear. With a 25-year sample size, for example, the posterior distribution for ρ in the 4-asset case in figure 6a is more spread out than in the 12-asset case in figure 5a. Comparing only these two cases might suggest that the uncertainty about ρ is greater in the 4-asset case. A more thorough comparison indicates otherwise, however. As noted previously, the posteriors for ρ in the 12-asset case actually spread out as the sample size initially increases beyond 25 years. The 25-year posterior in the 4-asset case is more appropriately compared to the posteriors produced by larger samples in the 12-asset case. This interpretation is supported by observing that, for the larger sample sizes, the posteriors concentrate around 1 more rapidly in the 4-asset case. Thus, reducing the number of assets does indeed appear to increase the information about ρ provided by a given number of time-series observations, but this conclusion is reached only after examining the behavior of the posterior distributions for both small and large samples.

As in the 12-asset case, the interpretation of the posterior distribution of Δ is more straightforward. For both small and large samples, the posterior of Δ tightens around 0 as sample size increases. Comparing figures 5b and 6b illustrates the substantial reduction in uncertainty obtained by reducing the number of assets. With a 25-year sample, for example, 98% of the posterior mass for Δ lies between 25% and 150% in the 12-asset case, but the same mass lies between 1% and 65% in the 4-asset case. In both cases, however, a large number of time-series observations is required to obtain a reasonably tight posterior for Δ. In order for 99% of the mass to lie below a 10% annual return, for example, one would require over 1000 years of weekly data in the 4-asset case and roughly 4000 years of data in the 12-asset case. It is instructive to compare this uncertainty about Δ to that of the market mean in figure 4a. In the latter case, a 1000-year sample produces a posterior in which 98% of the mass lies within a 2% range of returns.

4. Conclusions

The approximate efficiency of a given "test" portfolio can be investigated in a Bayesian framework by examining posterior distributions of scalar measures of portfolio inefficiency. This study pursues such an approach to investigating the efficiency of the value-weighted NYSE–ASE portfolio using a 25-year sample of weekly returns on ten size-ranked portfolios and two market indexes.

The posterior distribution of ρ, the maximum correlation between the test portfolio and an efficient portfolio, appears to be concentrated away from 1 when a riskless asset is included. We conclude, however, that this result should not be interpreted as strong evidence against efficiency of the test portfolio. A disperse distribution for the underlying fundamental parameters E and V can be accompanied by tight distributions of bounded nonlinear functions of those parameters, such as ρ. Our investigation indicates that much larger samples would be required in order for the data to significantly reduce the uncertainty about the test portfolio's efficiency.

An alternative measure, Δ, is the difference in expected returns between the test portfolio and an efficient portfolio of equal variance. The posterior distribution for Δ obtained in our sample is spread over a wide range of values and reveals substantial uncertainty about the degree of the test portfolio's inefficiency. Our experiments indicate that the dispersion in this posterior distribution provides a reasonable representation of the true uncertainty accompanying a sample of this size.

Even when sample point estimates correspond to exact portfolio efficiency, very large samples are required to obtain posterior distributions of ρ and Δ that are concentrated close to the values of $\rho = 1$ and $\Delta = 0$. We conclude that sample sizes encountered in practice are incapable of providing tight posterior distributions of a portfolio's degree of inefficiency.

References

Black, Fischer, 1972, "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business* 45, 444–454.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, "The Capital Asset Pricing Model: Some Empirical Tests," in Michael C. Jensen, ed., *Studies in the Theory of Capital Markets*, Praeger, New York.

Breeden, Douglas T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics* 7, 265–296.

Douglas, George W., 1969, "Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency," *Yale Economic Essays* 9, 3–45.

Fama, Eugene F. and James D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy* 81, 607–636.

Ferson, Wayne, Shmuel Kandel, and Robert F. Stambaugh, 1987, "Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas," *Journal of Finance* 42, 201–220.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica* 57, 1121–1152.

Harvey, Campbell R. and Goufu Zhou, 1990, "Bayesian Inference in Asset Pricing Tests," forthcoming in *Journal of Financial Economics.*

Kandel, Shmuel and Robert F. Stambaugh, 1987, "On Correlations and Inferences about Mean-Variance Efficiency," *Journal of Financial Economics* 18, 61–90.

Kandel, Shmuel and Robert F. Stambaugh, 1989, "A Mean-Variance Framework for Tests of Asset Pricing Models," *Review of Financial Studies* 2, 125–156.

Lintner, John, 1965, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics* 47, 13–27.

MacKinlay, A. Craig, 1987, "On Multivariate Tests of the CAPM," *Journal of Financial Economics* 18, 341–371.

Markowitz, Harry, 1952, "Portfolio Selection," *Journal of Finance* 7, 77–91.

Markowitz, Harry, 1959, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York.

McCulloch, Robert and Peter E. Rossi, 1990a, "Posterior, Predictive, and Utility-Based Approaches to Testing the Arbitrage Pricing Theory," forthcoming in *Journal of Financial Economics.*

McCulloch, Robert and Peter E. Rossi, 1990b, "A Bayesian Approach to Testing the Arbitrage Pricing Theory," forthcoming in *Journal of Econometrics.*

Oldfield, George S. and Richard J. Rogalsky, 1987, "The Stochastic Properties of Term

Structure Movements," *Journal of Monetary Economics* 19, 229–254.

Shanken, Jay, 1987a, "Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique," *Journal of Financial Economics* 18, 91–110.

Shanken, Jay, 1987b, "A Bayesian Approach to Testing Portfolio Efficiency," *Journal of Financial Economics* 19, 195–215.

Shanken, Jay, 1989, "On the Estimation of Beta-Pricing Models," working paper, University of Rochester, Rochester, NY.

Sharpe, William F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance* 19, 425–442.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.

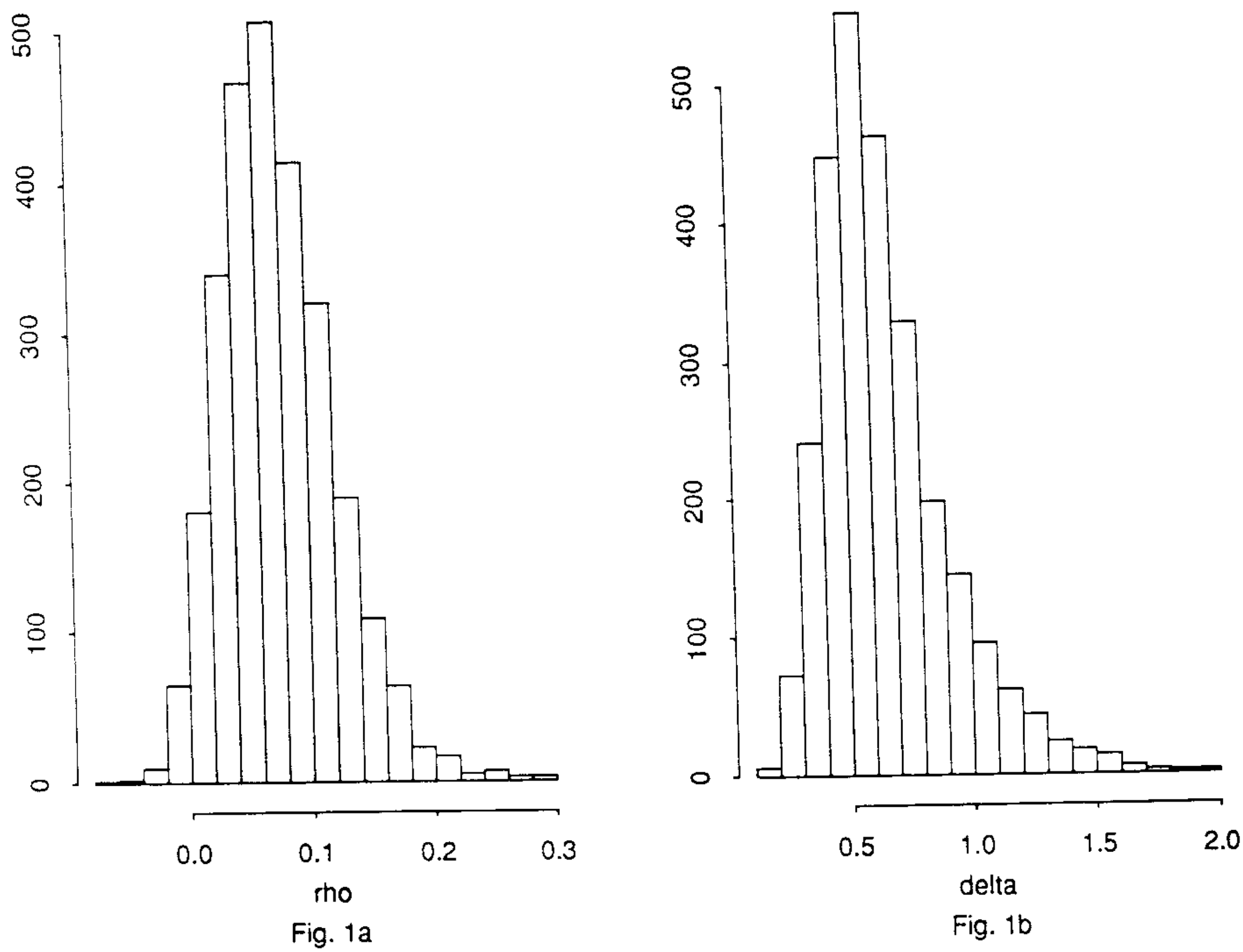

Figure 1. Marginal posterior distributions of inefficiency measures with the riskless asset included.

Figure 1a is the histogram based on 5000 draws from the posterior distribution of the inefficiency measure ρ, which is the maximum correlation between the value-weighted portfolio and efficient portfolios. Figure 1b is the histogram based on 5000 draws from the posterior of the inefficiency measure Δ, which is the difference between the mean of the value-weighted portfolio and the efficient portfolio having the same variance. Efficient portfolios are computed using weekly excess returns from January 1963 through December 1987 for a set of assets consisting of ten size-ranked portfolios and the equally and value-weighted NYSE–ASE indexes. Returns are in excess of the return on a one-week Treasury Bill, and all return values are multiplied by 52.

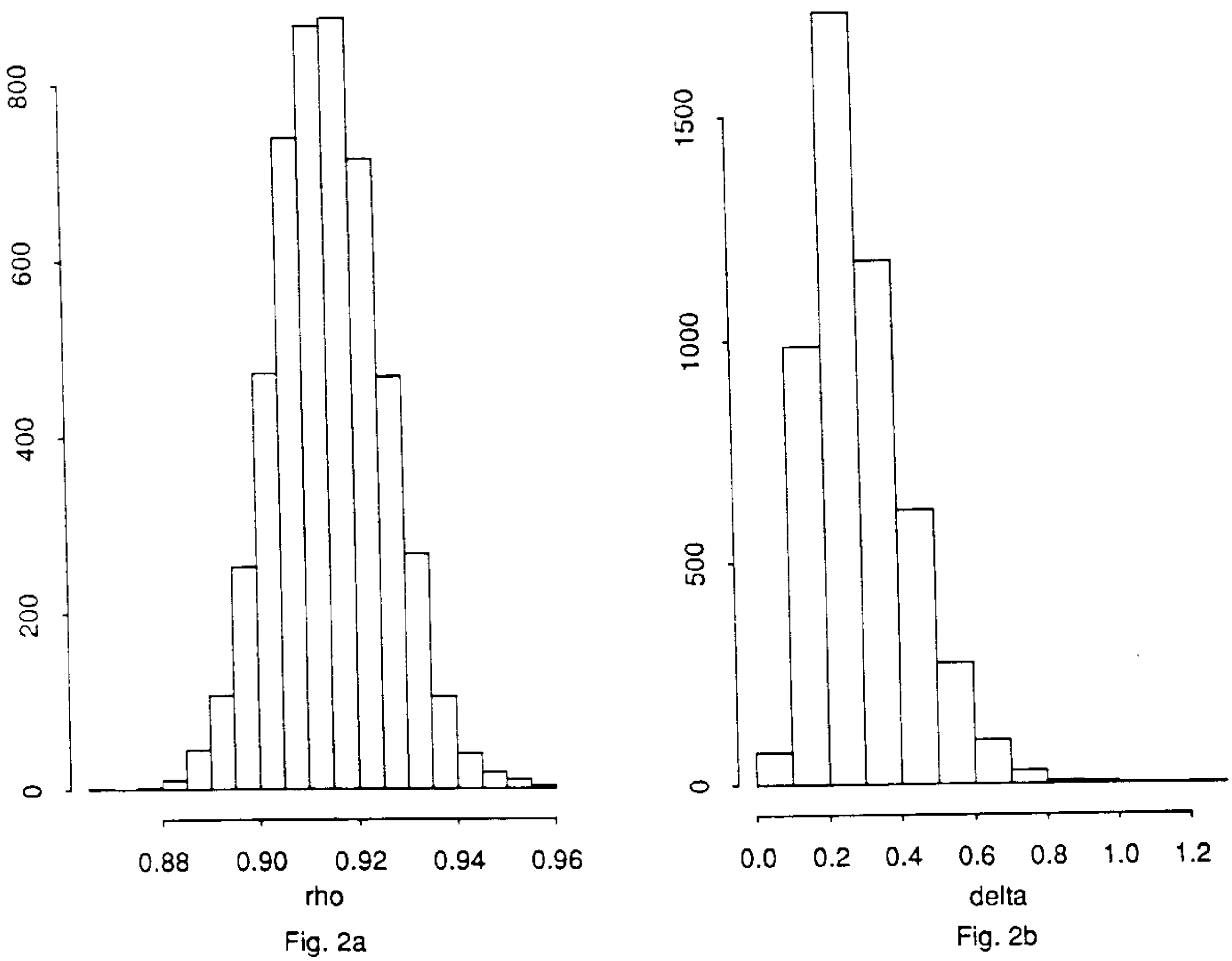

Figure 2. Marginal posterior distributions of inefficiency measures without the riskless asset.

Figure 2a is the histogram based on 5000 draws from the posterior of the inefficiency measure ρ, which is the maximum correlation between the value-weighted portfolio and efficient portfolios. Figure 2b is the histogram based on 5000 draws from the posterior of the inefficiency measure Δ, which is the difference between the mean of the value-weighted portfolio and the efficient portfolio having the same variance. Efficient portfolios are computed using weekly returns from January 1963 through December 1987 for a set of assets consisting of ten size-ranked portfolios and the equally and value-weighted NYSE–ASE indexes. All return values are multiplied by 52.

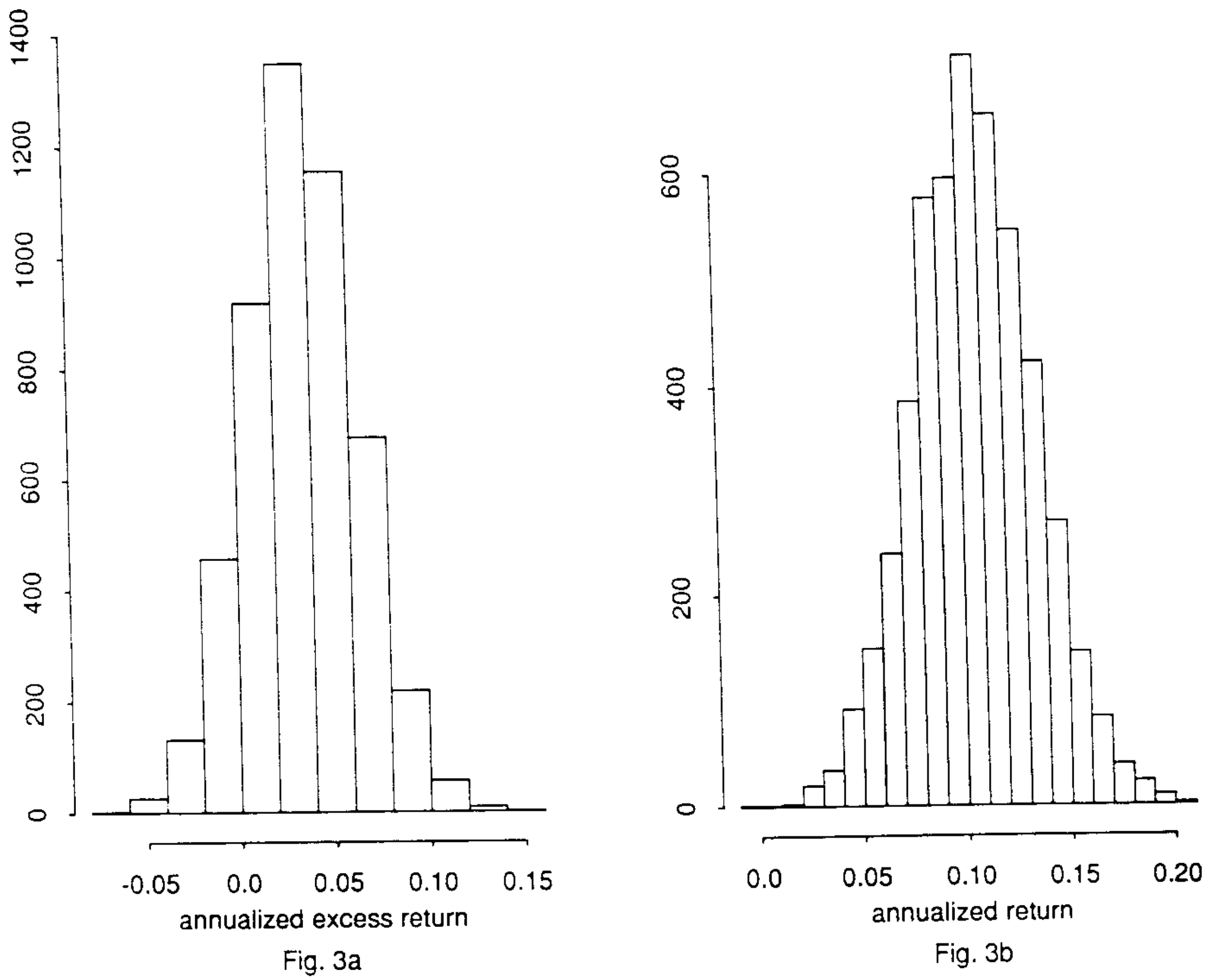

Figure 3. Marginal posterior distributions of the market return.

Figure 3a is the histogram based on 5000 draws from the posterior distribution of the mean excess return of the value-weighted portfolio based on weekly returns from January 1963 through December 1987. Figure 3b is the histogram based on 5000 draws from the posterior distribution of the mean return of the value-weighted portfolio. All return values are multiplied by 52.

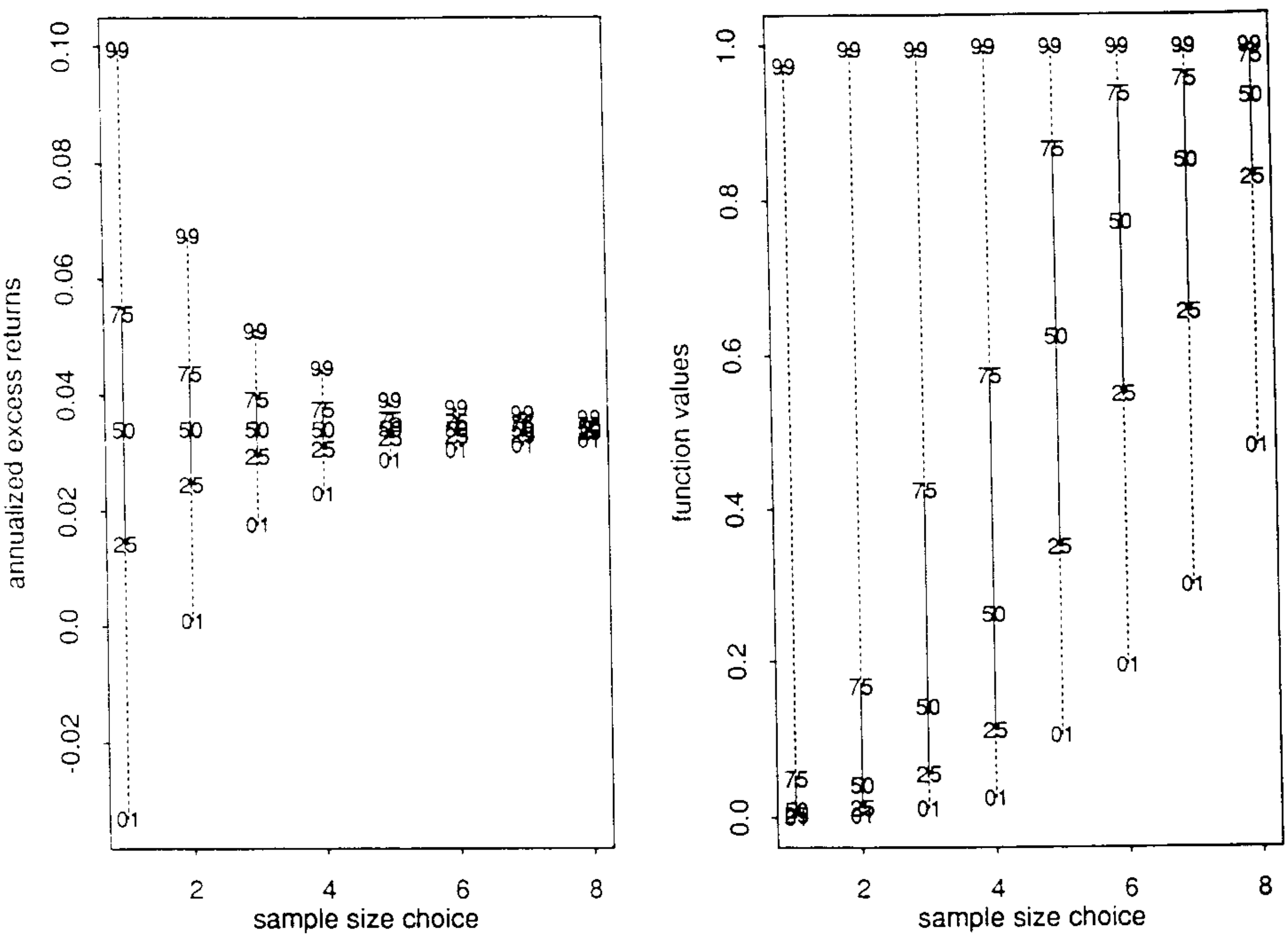

Fig. 4a

Fig. 4b

Figure 4. Marginal posterior distributions of the mean excess return on the value-weighted portfolio and a bounded nonlinear function for artificial data sets of various sample sizes.

We consider posteriors based on artificial data sets having maximum likelihood estimates of the mean vector E and the covariance matrix V equal to the estimates from our actual data. The eight sample sizes, increasing from left to right, take values corresponding to 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data. For each sample size, the figure displays the 1%, 25%, 50%, 75%, and 99% quantiles of 5000 draws from the posterior distribution. The 25% and 75% quantiles are connected by solid lines, and dotted lines extend to the 1% and 99% quantiles. Figure 4a displays the posterior distribution of μ, the mean excess return on the value-weighted portfolio. Figure 4b displays the posterior distribution of the nonlinear function $f(\mu) = (1 + [500 \cdot (\mu - .034)]^2)^{-1}$. All return values are multiplied by 52.

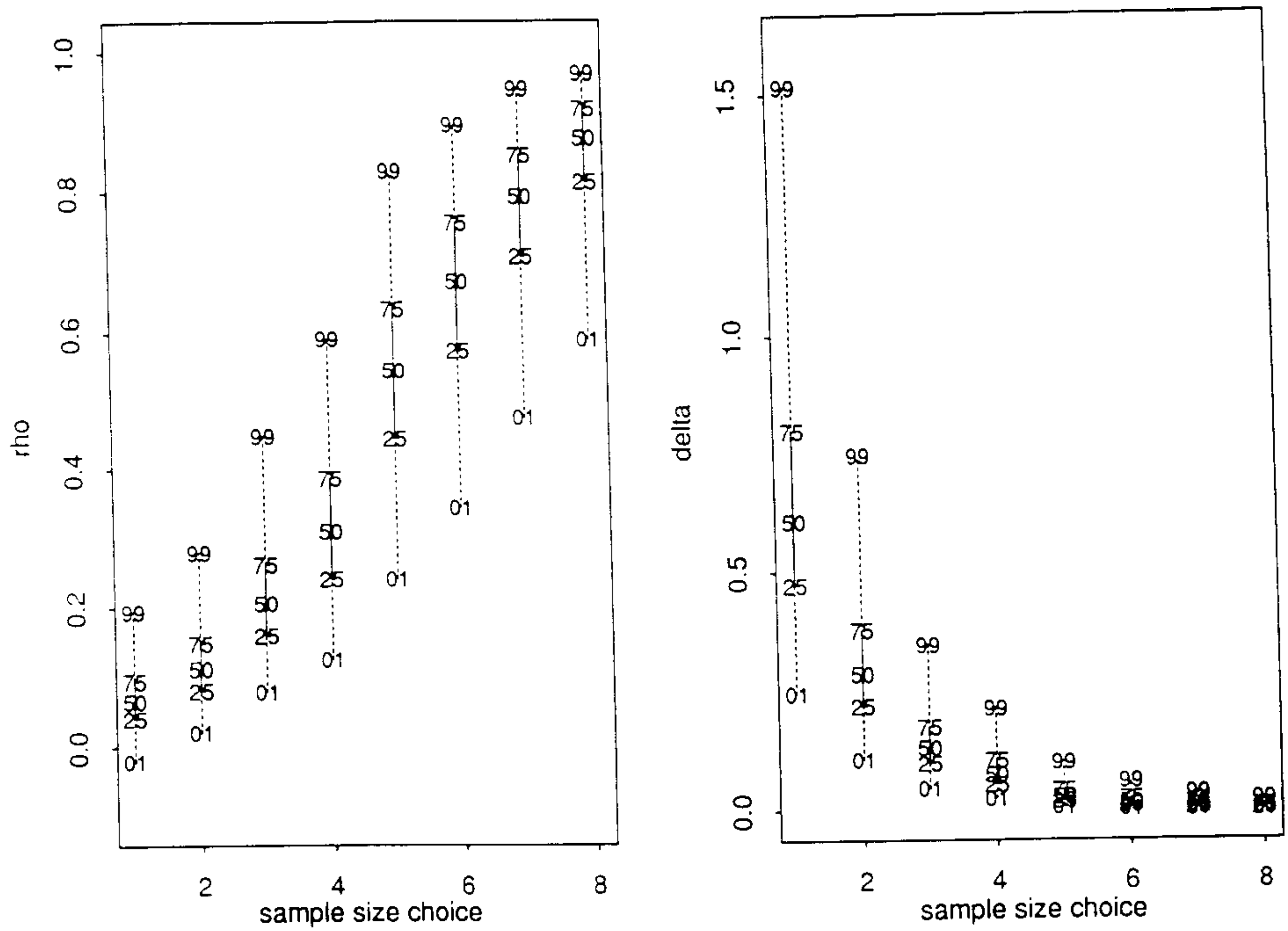

Fig. 5a

Fig. 5b

Figure 5. Marginal posterior distributions of inefficiency measures with twelve risky assets and a riskless asset for artificial data sets of various sample sizes.

We consider posteriors based on artificial data sets having unrestricted maximum likelihood estimates of the mean vector E and covariance matrix V equal to the maximum likelihood estimates based on our actual data satisfying the restriction that the value-weighted portfolio be the Sharpe-Lintner tangent portfolio. The eight sample sizes, increasing from left to right, take values corresponding to 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data. For each sample size, the figure displays the 1%, 25%, 50%, 75%, and 99% quantiles of 5000 draws from the posterior distribution. The 25% and 75% quantiles are connected by solid lines, and dotted lines extend to the 1% and 99% quantiles. Figure 5a displays the posterior distribution of the inefficiency measure ρ. Figure 5b displays the posterior distribution of the inefficiency measure Δ. All return values are multiplied by 52.

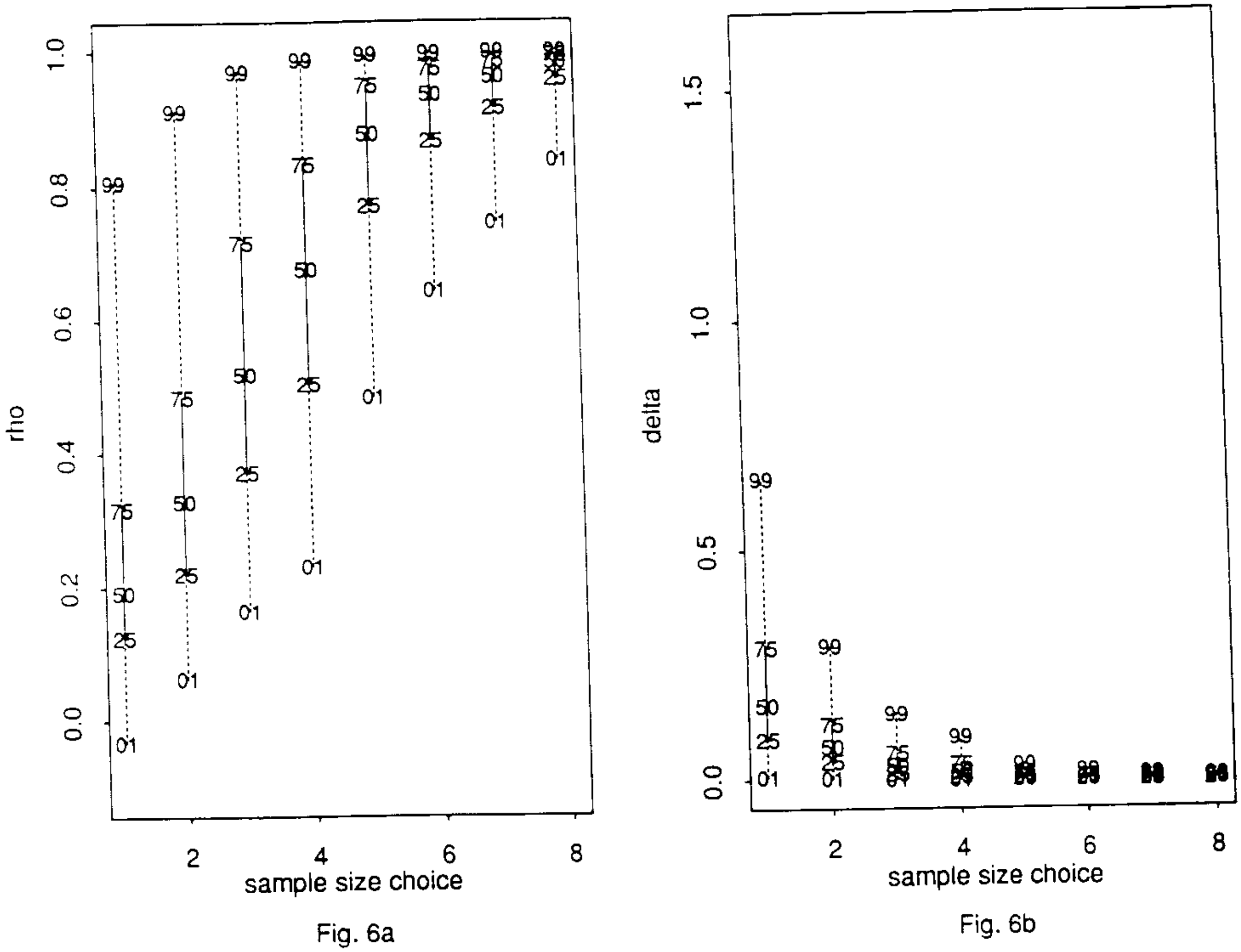

Figure 6. Marginal posterior distributions of inefficiency measures with four risky assets and a riskless asset for artificial data sets of various sample sizes.

The results presented in this figure use returns on only the value-weighted portfolio, the equally weighted portfolio, and the first and tenth size-ranked portfolios. We consider posteriors based on artificial data sets having unrestricted maximum likelihood estimates of the mean vector E and covariance matrix V equal to the maximum likelihood estimates based on our actual data satisfying the restriction that the value-weighted portfolio be the Sharpe-Lintner tangent portfolio. The eight sample sizes, increasing from left to right, take values corresponding to 25, 100, 400, 1000, 4000, 8000, 16,000, and 32,000 years of weekly data. For each sample size, the figure displays the 1%, 25%, 50%, 75%, and 99% quantiles of 5000 draws from the posterior distribution. The 25% and 75% quantiles are connected by solid lines, and dotted lines extend to the 1% and 99% quantiles. Figure 5a displays the posterior distribution of the inefficiency measure ρ. Figure 5b displays the posterior distribution of the inefficiency measure Δ. All return values are multiplied by 52.