JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS VOL 21, NO. 3, SEPTEMBER 1986

An Empirical Bayes Approach to Efficient Portfolio Selection

Peter A. Frost and James E. Savarino*

Abstract

When portfolio optimization is implemented using the historical characteristics of security returns, estimation error can degrade the desirable properties of the investment portfolio that is selected. Given the problem of estimation risk, it is natural to formulate rules of portfolio selection within a Bayesian framework. In this framework, portfolio selection is based on maximization of expected utility conditioned on the predictive distribution of security returns. Most researchers have addressed the problem of estimation risk by asserting a noninformative diffuse prior that reduces the detrimental effect of estimation risk, but does not directly reduce estimation error. Portfolio performance can be improved by specifying an informative prior that reduces estimation error. An informative prior that all securities have identical expected returns, variances, and pairwise correlation coefficients is asserted. This informative prior reduces estimation error by drawing the posterior estimates of each security's expected return, variance, and pairwise correlation coefficients toward the average return, average variance, and average correlation coefficient, respectively, of all the securities in the population. The amount that each of these parameters is drawn toward its grand mean depends upon the degree to which the sample is consistent with the informative prior. This empirical Bayes method is shown to select portfolios whose performance is superior to that achieved, given the assumption of a noninformative prior or by using classical sample estimates.

I. Introduction

Modern portfolio theory provides a realistic paradigm for optimal portfolio selection when risk-averse investors prefer investment portfolios that are mean-variance efficient. Markowitz [16] provides a normative framework through which such optimal portfolios may be identified. Optimal portfolio selection requires knowledge of each security's expected return, variance, and covariance with other security returns. In practice, each security's expected return, variance, and covariance with other security returns are unknown and must be estimated from available historical or subjective information. When portfolio optimization is implemented using the historical characteristics of security returns, estimation

* Graduate School of Business Administration, University of Washington, Seattle, WA 98195, and School of Business Administration, Pacific Lutheran University, Tacoma, WA 98447, respectively. The authors wish to thank W. W. Alberts, Chan-Fu Chen, Kaun-Pin Lin, Ed Rice, Paul Sampson, Phil Smith, and Rex Thompson for their helpful comments and suggestions. The authors are especially indebted to an anonymous *JFQA* referee for inciteful and thorough comments.

error can degrade the desirable properties of the investment portfolio that is selected. This problem is referred to as the problem of estimation risk in portfolio selection. Bawa, Brown, and Klein [5] provide an excellent overview of this problem.

Once estimation risk in portfolio selection is acknowledged, it is natural to formulate rules of portfolio selection within a Bayesian framework. In this framework, portfolio selection is based on maximization of expected utility derived from the joint predictive distribution of security returns. Most researchers have used a noninformative diffuse prior that accounts for estimation risk, but does not reduce estimation error. It should be possible to obtain superior portfolio performance by specifying an informative prior that reduces estimation error. This paper suggests an informative prior that is based on prior knowledge that the more a sample estimate of a parameter for a particular security differs from the mean of that parameter for all securities, the more likely it is that the sample estimate is estimated with error. This informative prior asserts that all stocks have identical means and variances, and that the correlation coefficient between any two stock's returns is the same. This prior also effectively reduces estimation error, on average, by drawing the sample estimates of each individual security's parameters toward the grand mean of those parameters. The Bayesian method for combining this ''all stocks are identical'' informative prior with the sample estimate of each security's expected return provides a predictive expected return that is a weighted average of the security's historically estimated expected return and the historical average return for all stocks. The predictive estimates of each security's variance and covariance with other security returns are determined in a similar fashion. In other words, to the extent that each security's historical characteristics differ from the average of all stocks, those characteristics are drawn towards the grand average by a Bayesian adjustment factor. This paper presents a straightforward method for calculating this adjustment factor, which essentially amounts to determining the goodness of fit of the prior assumption with available sample information. This empirical Bayes procedure adjusts all relevant stock characteristics towards their historical grand average more when the historical sample size is small, and when the historical data are consistent with the prior assumption. The objective is to reduce estimation error thereby improving the risk return characteristics of the investment portfolio selected.

To evaluate this empirical Bayes procedure, a sample of 25 stocks is constructed. This paper asserts a true multivariate normal distribution of security returns that is utilized to both simulate sample observations and to evaluate the performance of alternative investment rules. Fifty replications of portfolio optimization are carried out, positing three alternative approaches to constructing an optimal portfolio. The first is the Classical Estimated investment rule, which applies historical estimates of each security's expected return, variance, and covariance with other security returns in portfolio selection. The second is the noninformative Bayes Diffuse investment rule, which is suggested by previous research. The third is this paper's proposed Empirical Bayes investment rule, which allows the specification of an informative prior. The Empirical Bayes investment rule selects investment portfolios whose performance is superior to that provided by either the Classical Estimated or Bayes Diffuse investment rules.

Section II describes in more detail the nature of the estimation risk and portfolio choice problem being addressed. Section III develops the traditional Bayesian approach to specifying this paper's particular informative prior. Section IV describes an empirical Bayes approach to estimating the adjustment factors associated with this informative prior. In section V, the simulation model and results are presented. Section VI contains concluding remarks.

II. Estimation Risk in Portfolio Selection

Identification of the set of investment portfolios that are mean-variance efficient requires knowledge of the multivariate normal density $f_n(\underline{R}|\underline{\mu},\underline{\Sigma})$ of the $N \times 1$ vector $\underline{R}$ of security returns, where N represents the number of securities in the population. Vector $\underline{\mu}$ and matrix $\underline{\Sigma}$ represent, respectively, the $N \times 1$ vector of security expected returns and the $N \times N$ variance-covariance matrix of security returns. The set $\underline{P}$ is defined as the mean-variance efficient set of investment portfolios identified, given knowledge of the true multivariate normal distribution of security returns.

The corresponding admissible set $\underline{X} \in \underline{P}$ of investment weights, where vector $\underline{X}$ represents the percentage of one's wealth invested in each of N securities, is given by the solution to

$$\text{(1)} \qquad \min_{\underline{X} \in \underline{P}} \sigma_p^2 \qquad \text{s.t.} \quad \begin{aligned} \sigma_p^2 &= \underline{X}'\underline{\Sigma}\underline{X} \\ \mu_p &= \underline{X}'\underline{\mu} \\ 1.0 &= \underline{1}'\underline{X} \end{aligned}$$

for all levels of expected portfolio return μ_p above that of the minimum variance portfolio. In equation (1) above, $\underline{1}$ represents the unity vector. Equation (1) allows unrestricted short sales. Any investor, in maximizing expected utility, prefers a particular investment portfolio on Markowitz's efficient frontier. This portfolio is identified as optimal portfolio P^*, with optimal investment weights $\underline{X}^* \in \underline{P}$, expected return μ_p^*, and variance σ_p^{2*}.

The traditional or Classical approach to estimating Markowitz's efficient frontier utilizes the sample estimate $f_n(\underline{R}|\hat{\underline{\mu}},\hat{\underline{\Sigma}})$ in portfolio selection. An individual is assumed to select an investment portfolio from the sample estimated mean-variance efficient frontier. Here the sample estimated efficient set $\underline{C}$ of investment portfolios is provided by the solution to equation (1), given that sample estimates $\hat{\underline{\mu}}$ and $\hat{\underline{\Sigma}}$ of mean return vector $\underline{\mu}$ and variance-covariance matrix $\underline{\Sigma}$ are substituted for their true values. Portfolio selection from the Classical sample estimated efficient set $\underline{C}$ is referred to as the Classical Estimated investment rule.

III. A Traditional Bayesian Approach to Efficient Portfolio Selection

Barry [3], Barry and Winkler [4], Brown [6], Kalymon [12], Klein and Bawa [13], [14], Mao and Sarndal [15], Winkler [20], and Winkler and Barry [21] provide a Bayesian approach to efficient portfolio selection. Within a Bayesian framework, portfolio selection is based on maximization of expected utility derived from knowledge of the joint predictive distribution $g(\underline{R})$ of security returns. There are an infinite number of possible Bayesian posterior efficient frontiers as each posterior estimated efficient frontier is a function of the particular prior distribution $f(\underline{\mu}, \underline{\Sigma})$ assumed for multivariate normal density $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$. Barry [3], Brown [6], and Klein and Bawa [13] have suggested the assumption of a Bayesian noninformative diffuse prior in portfolio selection. Below, the Bayesian efficient frontier is defined given the assumption of a diffuse prior

$$f(\underline{\mu}, \underline{\Sigma}) \propto |\underline{\Sigma}|^{-(N+1)/2}$$

on multivariate normal distribution $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$. Given this assumption, predictive density $g(\underline{R})$ can be shown to be student-t distributed according to $f_s(\underline{R}|\underline{\mu}'', \underline{\Sigma}'', T-N)$ where T equals the number of sample observations on each of N security returns. Posterior estimate $\underline{\mu}''$ is identical to the sample estimate $\hat{\underline{\mu}}$. Posterior estimate $\underline{\Sigma}''$ is a scalar multiple k,

$$k = (1 + 1/T)(T - 1)/(T - N - 2) > 1$$

of sample estimate $\hat{\underline{\Sigma}}$. A two-parameter efficient set $\underline{D}$ of investment weights exists given that posterior estimates $\underline{\mu}''$ and $\underline{\Sigma}''$ are used as input to equation (1). Portfolio selection from the Bayesian efficient set $\underline{D}$ is referred to as the Bayes Diffuse investment rule.

Brown [6] shows that the Bayes Diffuse investment rule provides superior security selection relative to that of the Classical Estimated investment rule. However, Barry [3] demonstrates that the set of investment weights $\underline{D}$ of portfolios on the Bayes Diffuse efficient frontier are identical to the set of investment weights $\underline{C}$ of portfolios on the Classical Estimated efficient frontier. This equivalence severely limits the Bayes Diffuse investment rule's ability to reduce estimation risk in portfolio selection. To improve on the Bayes Diffuse investment rule, an informative prior that reduces estimation error can be specified. The objective is to provide a less biased and more precise estimate of the investment weights associated with the investor's optimal portfolio.[1]

The primary method of introducing prior information into any Bayesian decision rule is through specification of a natural conjugate prior. The Normal-Wishart probability distribution is the conjugate prior to the multivariate normal probability distribution. Following Ando and Kaufman [2], this paper asserts a Normal-Wishart conjugate prior with prior parameters $\underline{\Omega}^{-1}$, ν, $\underline{\mu}_o$, and τ,

$$(2) \qquad f_{nw}\left(\underline{\mu}, \underline{\Lambda} \mid \underline{\Omega}^{-1}, \nu, \underline{\mu}_o, \tau\right) = f_n\left(\underline{\mu} \mid \underline{\mu}_o, (\underline{\Lambda}\tau)^{-1}\right) \cdot f_w\left(\underline{\Lambda} \mid (\nu\underline{\Omega})^{-1}, \nu\right),$$

where $\underline{\Lambda} \equiv \underline{\Sigma}^{-1}$. Prior parameters ν and τ determine the strength of belief in prior values $\underline{\Omega}^{-1}$ and $\underline{\mu}_o$. Given the assumption of a Normal-Wishart conjugate prior, predictive density $g(\underline{R})$ is multivariate student-t distributed according to

[1] Klein and Bawa [13] have suggested the application of two particular Bayesian informative priors in portfolio selection. From a practical perspective, neither informative prior suggested in [13] can be implemented in portfolio selection. Each informative prior suggested by Klein and Bawa requires either that specific parameters of multivariate normal density $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$ are known or that a specific functional relationship between the parameters of $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$ is known.

$f_s(\underline{R}|\underline{\mu}'', \underline{\Sigma}'', v + T)$ with $v + T$ degrees of freedom. Posterior estimates $\underline{\mu}''$ and $\underline{\Sigma}''$ follow from the posterior distribution of $\underline{\mu}$ and $\underline{\Sigma}$ and are given by[2]

$$\underline{\mu}'' = \underline{\mu}_o(\omega_\tau) + \hat{\underline{\mu}}(1 - \omega_\tau) \tag{3a}$$

$$\underline{\Sigma}'' = \left[\frac{v + T}{v + T - 2}\right]\left[1 + \frac{1}{\tau + T}\right]\left[\underline{\Omega}\omega_v + \hat{\underline{\Sigma}}(1 - \omega_v) + \omega_\tau\left(\frac{T}{v + T}\right)(\hat{\underline{\mu}} - \underline{\mu}_o)(\hat{\underline{\mu}} - \underline{\mu}_o)'\right], \tag{3b}$$

where

$$\omega_\tau = \tau/(\tau + T)$$
$$\omega_v = v/(v + T) .$$

For any given amount of sample information T, strength of belief weights ω_τ and ω_v are determined, respectively, by prior parameters τ and v. From equation (3a), posterior estimate $\underline{\mu}''$ of mean return vector $\underline{\mu}$ will more closely approximate prior value $\underline{\mu}_o$, the larger the strength of belief weight ω_τ. Similarly, from equation (3b), for any given value of T and ω_τ, the inverse of the prior $\underline{\Omega}^{-1}$ will be provided a larger weight in determining posterior estimate $\underline{\Sigma}''$, the larger the value of strength of belief weight ω_v.

The distribution of $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$ needs to be inferred from existing sample and prior information. The problem is to specify an informative prior that will provide a less biased and more precise estimate of the investment weights associated with the investor's optimal portfolio. It is known that the more the sample estimate of a parameter for a particular security differs from the average of that parameter for all securities in the populaton, the more likely that sample estimate is estimated with error. Consequently, a better estimate of that parameter may be obtained by using a weighted average of the sample estimate and a prior on that parameter's average population value. This suggests the informative prior assumption that the prior $\underline{\Omega}^{-1}$ on $\underline{\Sigma}^{-1}$ is of intraclass form

$$\underline{\Omega}^{-1} \equiv \begin{bmatrix} H & G & \cdot\cdot & G \\ G & H & \cdot\cdot & G \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ G & G & \cdot\cdot & H \end{bmatrix}$$

where

$$H = \frac{1 + (N-2)\gamma/\delta}{\delta(1-\gamma/\delta)(1+(N-1)\gamma/\delta)};$$
$$G = \frac{-\gamma/\delta}{\delta(1-\gamma/\delta)(1+(N-1)\gamma/\delta)}.$$

[2] See [6], p. 145.

By definition,

$$\underline{\Omega} \equiv \begin{bmatrix} \delta & \gamma & \cdot & \cdot & \gamma \\ \gamma & \delta & \cdot & \cdot & \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \gamma & \gamma & \cdot & \cdot & \delta \end{bmatrix}$$

The prior $\underline{\mu}_o$ on mean return vector $\underline{\mu}$ is such that $\underline{\mu}_o = \alpha \cdot \underline{1}$. Parameters α, δ, and γ represent, respectively, the Bayesian's prior on the unknown average expected return, variance, and pairwise covariance of the securities in the population. Inspection of equations (3a) and (3b) will verify that the above prior assumption is such that for given values of ω_τ, ω_v, and T, the posterior estimate of each security's mean, variance, and covariance will be adjusted more towards its corresponding prior value α, δ, or γ, the more that sample estimate $\hat{\mu}_i$, $\hat{\sigma}_i^2$, or $\hat{\sigma}_{ij}$ differs from this value.

IV. An Empirical Bayes Approach to Efficient Portfolio Selection

A full Bayesian approach to efficient portfolio selection requires assertion of all the prior parameters $\underline{\Omega}^{-1}$, $\underline{\mu}_o$, τ, and v associated with prior density $f_{nw}(\underline{\mu}, \underline{A}|\underline{\Omega}^{-1}, \underline{\mu}_o, \tau, v)$. A more practical alternative to the above fully Bayesian approach entails point estimation of prior parameters $\underline{\Omega}^{-1}$, $\underline{\mu}_o$, τ, and v. Such an empirical Bayes approach to maximum likelihood estimation of the prior parameters of a prior density function is suggested in [8]. Chen [7] applies such an empirical Bayes estimation technique when estimating a Normal dispersion matrix. In the context of our portfolio selection problem, prior parameters $\underline{\Omega}^{-1}$, $\underline{\mu}_o$, τ, and v could be estimated simultaneously through maximum likelihood estimation. However, for computational convenience, prior values $\underline{\mu}_o$ and $\underline{\Omega}^{-1}$ are equated with their maximum likelihood sample estimates $\hat{\underline{\mu}}_o$ and $\hat{\underline{\Omega}}^{-1}$, given the assumption of equality of the means, variances, and covariances.[3] Empirical Bayes estimates, $\hat{\tau}$ and $\hat{v}$ and v of τ are then obtained given $\hat{\underline{\mu}}_o$ and $\hat{\underline{\Omega}}^{-1}$.

[3] The maximum likelihood estimate $\hat{\underline{\mu}}_o$ of $\underline{\mu}$ is equal to $\hat{\underline{\mu}}_o = \hat{\alpha} \cdot \underline{1}$, where $\hat{\alpha} = \sum_{i=1}^{i=N} \hat{\mu}_i / N$ ([19], p. 271). From Anderson [1], pp. 48-49, the maximum likelihood estimate $\hat{\underline{\Omega}}^{-1}$ equals the inverse of the maximum likelihood estimate $\hat{\underline{\Omega}}$ of $\underline{\Omega}$. The maximum likelihood estimate $\hat{\underline{\Omega}}$ of $\underline{\Omega}$ is given by

$$\hat{\underline{\Omega}} = \begin{bmatrix} \hat{\delta} & \hat{\gamma} \\ \hat{\gamma} & \hat{\delta} \end{bmatrix} \quad \text{where}$$

$$\hat{\delta} = \sum_{i=1}^{i=N} \left[s_i^2 + \left(\hat{\mu}_i - \hat{\alpha}\right)^2 \right] \Big/ N$$

$$\hat{\gamma} = \hat{\rho} \cdot \hat{\delta}$$

$$\hat{\rho} = \left[\sum_{\substack{i=1 \\ i \neq j}}^{i=N} \sum_{j=1}^{j=N} s_{ij} + \left(\hat{\mu}_i - \hat{\alpha}\right)\left(\hat{\mu}_j - \hat{\alpha}\right) \right] \Big/ N(N-1)$$

and s_i^2 and s_{ij} are the maximum likelihood estimates of each security's variance and covariance with other security returns. See Wilks [19], p. 270-271.

The objective of the empirical Bayes estimation technique is to determine the values for prior parameters τ and v, which maximize the likelihood function

$$
(4) \quad \begin{aligned}
g\left(\underline{S}, \hat{\underline{\mu}} \mid \hat{\underline{\Omega}}^{-1}, \hat{\underline{\mu}}_o, v, \tau\right) &= \int_{\underline{\mu}} \int_{\underline{\Lambda}} f\left(\underline{S}, \hat{\underline{\mu}}, \underline{\Lambda}, \underline{\mu} \mid \underline{\Omega}^{-1}, \underline{\mu}_o, \tau, v\right) d\underline{\Lambda}\, d\underline{\mu} \\
&= \int_{\underline{\mu}} \int_{\underline{\Lambda}} f_{nw}(\underline{S}, \hat{\underline{\mu}} \mid \underline{\Lambda}, \underline{\mu}) \\
&\qquad \cdot f_{nw}\left(\underline{\Lambda}, \underline{\mu} \mid \hat{\underline{\Omega}}^{-1}, \hat{\underline{\mu}}_o, v, \tau\right) d\underline{\Lambda}\, d\underline{\mu}\,.
\end{aligned}
$$

Matrix $\underline{S}$ represents the sum of the squared errors, $\underline{S} \equiv (T - 1)\hat{\underline{\Sigma}}$. The likelihood function in equation (4) reduces to

$$
(5) \quad \begin{aligned}
g\left(\underline{S}, \hat{\underline{\mu}} / \hat{\underline{\Omega}}^{-1}, \underline{\mu}_0, v, \tau\right) &= k_1 \cdot k_2 \cdot |\underline{S}|^{((T-N)/2)-1} \\
&\quad \cdot |v\underline{\Omega}|^{(v+N-1)/2} \cdot |\underline{Q}|^{-(v+T+N-1)/2} \\
k_1 &= 2^{-N/2} \cdot [\tau T/(\tau + T)]^{N/2} \\
k_2 &= w(N, T-1) \cdot w(N, v) \cdot [w(N, v+T)]^{-1} \\
w(N, p) &= \left[2^{Np/2} \cdot \pi^{N(N-1)/4} \cdot \prod_{j=1}^{j=N} \Gamma\left(\frac{p+1-j}{2}\right)\right]^{-1} \\
\underline{Q} &= \underline{S} + v\hat{\underline{\Omega}} + \left(\frac{\tau T}{\tau + T}\right)\left(\hat{\underline{\mu}} - \underline{\mu}_o\right)\left(\hat{\underline{\mu}} - \underline{\mu}_o\right)'
\end{aligned}
$$

where $\Gamma(z)$ represents the gamma function evaluated at z.[4] Maximum likelihood estimates of v and τ, given $\underline{S}$, $\hat{\underline{\mu}}$, $\hat{\underline{\Omega}}^{-1}$ and $\hat{\underline{\mu}}_o$, are obtained by maximizing equation (5) w.r.t. v and τ. Such estimates may be numerically approximated via any conventionally accepted nonlinear optimization technique. Point estimates $\hat{\tau}$ and $\hat{v}$ of prior parameters τ and v, given $\hat{\underline{\Omega}}$ and $\hat{\underline{\mu}}_o$, determine posterior estimates $\underline{\Sigma}''$ and $\underline{\mu}''$ from equations (3a) and (3b). An empirical Bayes efficient set $\underline{B}$ of investment weights is provided by using posterior estimates $\underline{\mu}''$ and $\underline{\Sigma}''$ as input to equation (1). Portfolio selection from set $\underline{B}$ is referred to as the Empirical Bayes investment rule.

Before concluding the discussion on both specification of an informative prior and application of this paper's suggested empirical Bayes estimation procedure in portfolio selection, two aspects of the estimation risk and portfolio choice problem being addressed should be emphasized. First, as with any Bayesian decision rule, there is no definitive approach to specification of a particular informative prior. Specification of alternative informative priors that similarly consider prior information regarding estimation error in portfolio selection may provide superior security selection relative to that provided by the informative

[4] Ando and Kaufman [2], p. 353 demonstrate that

$$g\left(\hat{\underline{\mu}}, \underline{S} \mid \hat{\Omega}^{-1}, \hat{\underline{\mu}}_o, v, \tau\right) \propto \underline{S}^{1/2((T-N)-1)} \cdot |\underline{Q}|^{-(v+T+N-1)/2}.$$

Equation (5) may be derived by adding the appropriate normalizing constants while following the general logic of Ando and Kaufman's derivation.

prior suggested in Section III. Second, an empirical Bayes approach to estimation of prior parameters can be used in portfolio selection for alternative informative conjugate priors.

V. Simulation Results

Portfolio performance is analyzed in 25 sample trials of the Classical Estimated, the Bayes Diffuse, and the Empirical Bayes investment rules. In order to compare the performance of these three investment rules, both a population of securities and utility functions for different investors are specified. A population of 25 securities was randomly selected from all common stocks listed on the New York Stock Exchange for which data were consistently available during the January 1, 1953 to August 1, 1971 sample period.[5] Historical observations on each security's real monthly rate of return for the above sample period provide sample estimate $f_n(\underline{R}|\hat{\underline{\mu}}, \hat{\underline{\Sigma}})$.[6] This estimate is asserted to represent the hypothetical true multivariate normal distribution of security returns. This hypothetical distribution is then utilized to both simulate sample observations on security returns and to evaluate the actual performance of investment porfolios selected on the basis of historical information.[7]

Sample observations on security returns are simulated for sample sizes of $T = 50, 75, 125, 225, 425$, and 825 months. These sample sizes are chosen so that $T - N$ doubles with each increment in T. $T - N$ is a more appropriate measure of sample information since more observations than securities are required in order to estimate the inverse of variance-covariance matrix $\underline{\Sigma}$. For all investment rules, portfolio performance achieved is reported relative to $T - N$, the degree to which sample information exceeds the minimum amount required. Sample observations are simulated on security returns in each of 50 independent trials.[8] For

[5] The January 1953 through July 1971 sample period (via Fama [10]) is specified in order to both provide a sufficient number of sample observations for security selection and avoid the problem of measuring inflation, given price controls initiated by the U.S. Government in late 1971 and maintained on certain petroleum products throughout the entire decade. In order to help ensure that a representative population universe of N securities is selected, a stratified random sampling procedure is followed. All securities for which data are consistently available during our sample period are ranked (stratified) into quintiles by their estimated historical betas. An equal number of $N/5$ securities are then randomly selected from each representative quintile in which each security has an equal probability of being selected.

[6] The sample information on nominal monthly returns on all securities is provided by the Center for Research in Security Prices Stock Files. All required historical real rates of return R_t are obtained from

$$1 + R_t = \left(1 + R_t^N\right)/(1 + \dot{p}) ,$$

where R_t^N is the nominal security return and $\dot{p}$ is the rate of inflation. We use the rate of change in the Consumer Price Index as our measure of inflation.

[7] The simulation results are strictly valid only for this paper's assumed joint distribution of security returns. Brown [6] evaluates the relative performance of alternatively specified investment rules from an assumed multivariate normal distribution based on historical sample information on security returns from, alternatively, the 1951 to 1958 and 1961 to 1968 sample periods. The relative performance of the different investment rules evaluated together with the observed impact of estimation risk on portfolio selection appeared insensitive to the different values for $\underline{\mu}$ and $\underline{\Sigma}$ assumed. There is no reason to expect that the simulation results are highly sensitive to the specific values for $\underline{\mu}$ and $\underline{\Sigma}$ assumed.

[8] IMSL mathematical and statistical subroutine GGNSM is used to simulate sample observations.

each sample trial, sample estimates of the security means, variances, and covariances are obtained from these sample observations and used as input to the alternative portfolio selection techniques.

Investor i has the negative exponential utility function

$$U\left(W_{ip}\right) = \exp\left[- h_i W_{i0}\left(1 + R_{ip}\right)\right] = - \exp\left[- \lambda_i\left(1 + R_{ip}\right)\right]$$

where W_{ip} is investor i's terminal wealth, W_{i0} is investor i's initial wealth, and R_{ip} is the real rate of return on investor i's portfolio. The parameter $\lambda_i = h_i W_{i0}$ completely summarizes individual i's risk aversion. Investor i will choose the portfolio that maximizes his expected utility or, equivalently, that maximizes his certainty equivalent return.[9] Utility functions are specified for three different investors. Information regarding λ_i, the expected real return, and the variance of the real return on each investor's optimal portfolio is provided in Table 1. Investor 2 is specified to hold an optimal portfolio whose expected return of 1.1063 percent equals the average mean return of the 25-security population.

TABLE 1
Optimal Portfolios for Individual Investors
(Returns in Percent)

Investor	λ_i	μ_{p^*}	$\sigma^2_{p^*}$
First	0.4964268	0.7733	7.2032
Second	0.1654756	1.1063	9.8859
Third	0.0992853	1.4392	15.2514

As the simulation results are virtually identical for each of the three investors, only the results for the second investor are discussed.[10] The monthly certainty equivalent return associated with the second investor's optimal portfolio is 0.29 percent. The average certainty equivalent returns reported in Table 2 summarize the results of the 50 independent trials. These results are portrayed graphically in Figure 1. The simulation results substantiate Brown's [6] finding that the Bayes Diffuse investment rule is superior to the Classical Estimated investment rule. The average certainty equivalent return is greater for the Bayes Diffuse investment rule for all sample sizes. As would be expected, the absolute magnitude of this difference declines as the amount of sample information increases. Using the Wilcoxon rank sign test, the null hypothesis can be rejected that the median certainty equivalent return for the Bayes Diffuse and the Classical Estimated investment rules are equivalent at a 0.0005 level of significance for each sample size evaluated.

The basic premise of this paper is that the Empirical Bayes investment rule will select investment portfolios whose portfolio performance is superior to that provided by either the Bayes Diffuse or Classical Estimated investment rules.

[9] Expected utility is given by $E[U(W_p)] = -\exp[-\lambda_i (1 + \mu_p - ½\lambda_i \sigma_p^2)]$. The certainty equivalent return R_c associated with the expected utility that is provided by the expected return and variance of any particular portfolio is given by

$$R_c = \left[\mathrm{Ln}\left(-1/E\left[U_i\left(W_p\right)\right]\right)/\lambda_i\right] - 1 = \mu_p - \frac{1}{2}\lambda_i\sigma_p^2 .$$

[10] The results for the first and third investors are available from the authors upon request.

TABLE 2

Average Monthly Certainty Equivalent Return (Returns in Percent)

$T-N$	Bayes Diffuse Versus Classical Estimated			Empirical Bayes Versus Bayes Diffuse			Empirical Bayes Versus Classical Estimated		
	Bayes Diffuse	Classical Estimated	Difference	Empirical Bayes	Bayes Diffuse	Difference	Empirical Bayes	Classical Estimated	Difference
25	−3.25	−15.40	12 15	−0.11	−3 25	3.14	−0 11	−15.40	15 29
50	−1 35	− 3.44	2 09	−0.06	−1 35	1.29	−0 06	− 3 44	3.38
100	−0 66	− 1 17	0.51	−0 01	−0.66	0 65	−0.01	− 1.17	1.16
200	−0 18	− 0.30	0 12	0.07	−0.18	0.25	0 07	− 0.30	0 37
400	0 05	0 03	0.02	0 14	0.05	0 09	0.14	0 03	0.11
800	0.17	0 16	0.01	0 19	0.17	0 02	0.19	0 16	0 03
∞	0 29	0.29		0.29	0 29		0 29	0.29	

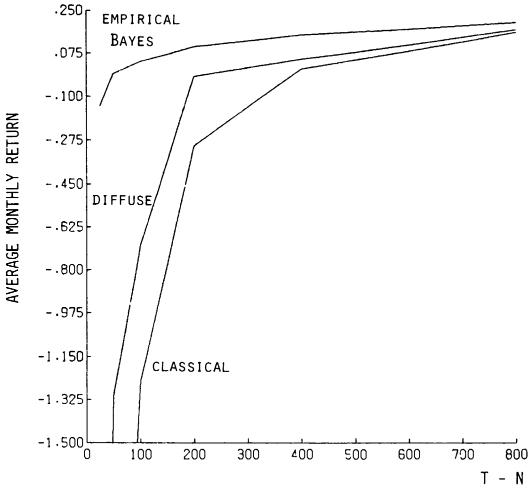

FIGURE 1

Average Monthly Certainty Equivalent Returns

This contention is strongly supported by the simulation results. For each sample size, the Empirical Bayes investment rule provides an average certainty equivalent return greater than that of either the Classical Estimated or Bayes Diffuse investment rules. A Wilcoxon rank sign test rejects the null hypothesis that the median certainty equivalent return for the Empirical Bayes and the Bayes Diffuse investment rules are identical at a 0.0005 level of significance for each sample size evaluated. When $T - N$ is 400 or less, the Empirical Bayes investment rule provides the higher certainty equivalent return in at least 47 of the 50 trials. For the largest sample size, given 825 sample observations, the Empirical Bayes investment rule provides the higher certainty equivalent return in 42 of the 50 trials. In small sample sizes, the economic gain in certainty equivalent return is substantial. When $T - N$ is 25, the Empirical Bayes investment rule provides an average monthly certainty equivalent return that is 314 basis points greater than that provided by the Bayes Diffuse investment rule. For $T - N$ equal to 100, the Empirical Bayes investment rule provides an average certainty equivalent return that is 65 basis points higher than that provided by the Bayes Diffuse investment rule. The economic gain in certainty equivalent return decreases as sample size increases. When $T - N$ is 800, the difference between the average certainty equivalent return of the Empirical Bayes and Bayes Diffuse investment rules decreases to 2 basis points. All three investment rules approach optimal portfolio selection and provide a monthly certainty equivalent return of 0.29 percent as the amount of sample information approaches infinity.

Insight can be gained into how the empirical Bayes procedure works by observing how the strength of belief weights ω_τ and ω_v vary with the number of observations in the sample. In this particular application, it is not presumed that the informative prior structural restriction on $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$ is correct. That is, it is not really believed that all security means, variances, and covariances are identical. The prior structural restriction is imposed as a means of reducing estimation risk in portfolio selection. Given increasing amounts of sample information, it should become increasingly apparent that the prior structural restriction is incorrect.[11] In Table 3, the average values of strength of belief weights ω_τ and ω_v are presented with their standard deviations s_{ω_τ} and s_{ω_v} in the 50 trials. Inspection of Table 3 will verify that the empirical Bayes estimates of prior parameters τ and v provide average strength of belief weights, ω_τ and ω_v, which decline monotonically with the amount of available sample information.

TABLE 3

Average Values and Standard Deviations of ω_τ and ω_v

$T-N$	$\overline{\omega}_\tau$	s_{ω_τ}	$\overline{\omega}_v$	s_{ω_v}
25	0.8238	0.1703	0.3735	0.0361
50	0.8009	0.2043	0.2771	0.0299
100	0.7304	0.2054	0.1844	0.0134
200	0.5679	0.1774	0.1107	0.0057
400	0.4202	0.1386	0.0592	0.0024
800	0.2671	0.0730	0.0308	0.0010

VI. Concluding Remarks

When portfolio optimization is carried out using estimates of security expected returns, variances, and covariances with other security returns, estimation error can degrade the desirable properties of the investment portfolio that is selected. Portfolio performance can be improved by specifying an informative prior that reduces estimation error. This paper suggests an informative prior that effectively reduces estimation error by drawing the sample estimates of each security's parameters toward the historical grand average of those parameters. An empirical Bayes method is utilized to estimate the strength of belief in this informative prior. The simulation results imply that application of this empirical Bayes procedure, given specification of this study's ''all stocks are identical'' informative prior, selects portfolios whose *ex ante* performance is substantially superior to that provided by either the Classical Estimated or the Bayes Diffuse investment rules.

[11] If the prior assumption regarding $f_n(\underline{R}|\underline{\mu}, \underline{\Sigma})$ is correct, strength of belief parameters τ and v approach infinity as T approaches infinity (see [7], p. 239).

In conclusion, two important aspects of applying a Bayesian procedure in portfolio selection can be emphasized. First, as with any Bayesian decision rule, there is no definitive approach to specification of an informative prior. Specification of alternative informative priors that reduce estimation error may provide superior security selection relative to this paper's "all stocks are identical" informative prior. Second, the empirical Bayes procedure that is presented is universally applicable to alternatively specified informative conjugate priors. Portfolio performance can be expected to improve the more that the specified informative prior reduces estimation error.

References

[1] Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley (1958).

[2] Ando, A., and G. M. Kaufman. "Bayesian Analysis of the Independent Multinomial Process—Neither Mean nor Precision Known." *Journal of the American Statistical Association,* 60 (March 1965), 347-353.

[3] Barry, C. B. "Portfolio Analysis under Uncertain Means, Variances, and Covariances." *Journal of Finance,* 29 (May 1974), 515-522.

[4] Barry, C. B., and R. L. Winkler. "Nonstationarity and Portfolio Choice." *Journal of Financial and Quantitative Analysis,* 11 (June 1976), 217-235.

[5] Bawa, V. S.; S. J. Brown; and R. W. Klein. *Estimation Risk and Optimal Portfolio Choice,* Studies in Bayesian Econometrics Bell Laboratories Series. New York: North Holland (1979).

[6] Brown, S. J. "Optimal Portfolio Choice under Uncertainty: A Bayesian Approach." Unpublished Ph.D. diss., Univ. of Chicago, (1976).

[7] Chen, C.-F. "Bayesian Inference for a Normal Dispersion Matrix and Its Application to Stochastic Multiple Regression Analysis." *Journal of the Royal Statistical Society,* 41 (1979), 235-248.

[8] Dempster, A. P.; N. M. Laird; and D. B. Rubin. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society,* 39 (1977), 1-38.

[9] Dempster, A. P. *Elements of Continuous Multivariate Analysis*. Cambridge, MA: Addison-Wesley (1969).

[10] Fama, E. F. *Foundations of Finance: Portfolio Decisions and Security Prices*. New York: Basic Books (1976).

[11] Frankfurter, G.; H. Phillips; and J. Seagle. "Bias in Estimating Portfolio Alpha and Beta Scores." *Review of Economics and Statistics,* 56 (Aug. 1976), 412-414.

[12] Kalymon, B. A. "Estimation Risk in the Portfolio Selection Model." *Journal of Financial and Quantitative Analysis,* 6 (Jan. 1971), 559-582.

[13] Klein, R. W., and V. S. Bawa. "The Effect of Estimation Risk on Optimal Portfolio Choice." *Journal of Financial Economics,* 3 (June 1976), 215-231.

[14] ________. "The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification." *Journal of Financial Economics,* 5 (Aug. 1977), 89-111.

[15] Mao, J., and C. Sarndal. "A Decision Theory Approach to Portfolio Selection." *Management Science,* 12 (April 1966), 323-339.

[16] Markowitz, H. "Portfolio Selection." *Journal of Finance,* 7 (March 1952), 77-91.

[17] Press, S. James. *Applied Multivariate Analysis*. Malabar, FL: Robert E. Krieger Publ. (1982).

[18] Sharpe, W. F. "A Simplified Model for Portfolio Analysis." *Management Science,* 9 (Jan. 1963), 277-293.

[19] Wilks, S. S. "Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Distribution." *Annals of Mathematical Statistics,* 17 (Sept. 1946), 257-281.

[20] Winkler, R. L. "Bayesian Models for Forecasting Future Security Prices." *Journal of Financial and Quantitative Analysis,* 8 (June 1973), 387-406.

[21] Winkler, R. L., and C. B. Barry. ''A Bayesian Model for Portfolio Selection and Revision.'' *Journal of Finance*, 30 (March 1975), 179-192.