Annals of Operations Research 45(1993)21–58

Computing efficient frontiers using estimated parameters

Mark Broadie
Columbia University, Graduate School of Business, New York, NY 10027, USA

The mean-variance model for portfolio selection requires estimates of many parameters. This paper investigates the effect of errors in parameter estimates on the results of mean-variance analysis. Using a small amount of historical data to estimate parameters exposes the model to estimation errors. However, using a long time horizon to estimate parameters increases the possibility of nonstationarity in the parameters. This paper investigates the tradeoff between estimation error and stationarity. A simulation study shows that the effects of estimation error can be surprisingly large. The magnitude of the errors increase with the number of securities in the analysis. Due to the error maximization property of mean-variance analysis, estimates of portfolio performance are optimistically biased predictors of actual portfolio performance. It is important for users of mean-variance analysis to recognize and correct for this phenomenon in order to develop more realistic expectations of the future performance of a portfolio. This paper suggests a method for adjusting for the bias. A statistical test is proposed to check for nonstationarity in historical data.

Keywords: Portfolio selection, mean-variance analysis, efficient frontier, estimation error.

1. Introduction

This paper investigates the effect of estimation error in the Markowitz mean-variance model for portfolio selection. For a description of this model see Markowitz [11]. The tradeoff between estimation error and the stationarity of model parameters is also addressed. This paper provides a framework for providing quantitative results in these areas. For example, how much error in the recommendations of the model is caused by errors in the input parameters? How far away from the efficient frontier is the recommended portfolio? How different is the composition of a recommended portfolio from the true optimal portfolio?

The parameters of the Markowitz model are the means, standard deviations, and correlations of security returns. These parameters can be estimated using historical data, analytical models, analysts' forecasts, or other methods. When historical data is used to estimate model parameters, there are at least two main areas of concern: *stationarity* of the model parameters and *estimation errors.*

Stationarity means that the model parameters do not change over time.[1] Return data from forty years ago might have little bearing on returns this year, so parameters are not likely to be stationary over long periods of time. However, using only very recent data to estimate model parameters exposes the model to estimation errors. Even if the process that generates returns is stationary, using a limited amount of recent data means that estimates of the unknown parameters will differ from their true values. The estimation errors decrease when more data is used, but more data requires a longer time horizon. There is a tradeoff between stationarity and estimation error when deciding *how much data* to use in estimating parameters.

The tendency of mean-variance analysis to exaggerate errors in the input parameters has been noted by previous researchers and practitioners. Jobson and Korkie [7] and Jobson [6] investigate the effect of errors in the input parameters, but without restrictions on short-selling. Michaud [12] discusses the implications of estimation error for portfolio managers. Chopra [2] and Jorion [9] examine the composition of optimal and near-optimal portfolios. Best and Grauer [1] analyze the effect of changes in the mean returns of assets on the mean-variance efficient frontier and the composition of optimal portfolios. Chopra and Ziemba [4] and Kallberg and Ziemba [10] study the relative importance of errors in means, variances, and covariances on the investor's utility function.

In addition to investigating the effects of estimation error, this paper suggests a method for *adjusting for the bias inherent in mean-variance analysis*. This could be an important tool to develop more realistic expectations of the future performance of a portfolio.

In the next section, the effect of estimation error on the results of mean-variance analysis is investigated. In order to isolate the effect of estimation error, it is assumed that stationarity holds. Section 3 provides intuitive explanations for the results. The effect of using improved estimates of parameters is explored in section 4. Section 5 focuses on errors in the estimated frontier. The effect of the number of securities on errors in mean-variance analysis is examined in section 6. A method for adjusting for the bias inherent in mean-variance analysis is illustrated in section 7. A test for stationarity of historical data and results are presented in section 8. Conclusions are given in section 9.

[1] It is important to distinguish constant *parameters* (i.e., stationarity) from constant *outcomes*. Suppose that the annual returns of a security over three years are −10%, 5%, and 18%. These returns could represent normal statistical variation consistent with a model that has constant parameters. For example, the returns are consistent with a normal distribution that has a constant mean of 8% and a constant standard deviation of 13%. The observed returns could also be consistent with a model that has nonstationary (i.e., time-varying) parameters. Three observations are not nearly enough data to reject either possibility.

2. Investigating estimation error using simulation

The effect of estimation error on the computation of mean-variance efficient frontiers is investigated in this section. A *simulation* procedure is used because it can directly show the effect of estimation error on the results of the mean-variance analysis. An analytical approach would be more difficult because of the following reasons. There are simultaneous and correlated estimation errors in a large number of parameters. While it is easy to determine the degree of estimation error for a single parameter, it is more difficult for a large number of parameters. More importantly, investors are not directly interested in the magnitude of the estimation errors in the parameters, but in their effect on the outcome of the mean-variance analysis. Analytical approaches to the latter problem would be more complicated than the simulation procedure used here. In spite of these difficulties, Jobson [6] develops confidence regions for the efficient frontier when there are no restrictions on short-selling. Best and Grauer [1] develop analytical results for the sensitivity of the optimal portfolio to changes in the mean returns of the assets.

Details of the mean-variance analysis and the simulation procedure used in this paper are given in the appendix. The approach will be illustrated with an example. Suppose that there are five securities with true parameters given in table 1. The parameters were chosen to be in the average range of monthly returns for many securities.[2] A monthly time interval was chosen as a reasonable intermediate point between weekly and yearly intervals. A weekly time horizon is too short for most portfolio planning problems and an annual time horizon is too long for gathering relevant historical data. However, the same analysis can be applied to any chosen time horizon.

Using the simulation procedure described in the appendix, 24 monthly returns of the five securities are generated. Using this simulated data, means, standard deviations, and correlations of monthly returns are estimated by the usual sample estimates. The estimated parameters are given in table 2. Comparing the numbers in tables 1 and 2 shows that the estimated standard deviations are quite close to the true values. The estimated correlations are not as accurate, e.g., $\rho_{13} = 0.30$ while $\hat{\rho}_{13} = 0.52$. The security with the largest estimated mean error is security 3, for which $\mu_3 = 0.014$ while $\hat{\mu}_3 = 0.042$.

The effect of using estimated parameters instead of true parameters on the computation of the mean-variance efficient frontier is graphed in fig. 1. Using the true parameters from table 1, the "true" efficient frontier is computed. Next, using the estimated parameters from table 2, a similar frontier, the "estimated" frontier is computed. Finally, the "actual" frontier is computed. The actual frontier is defined as follows. Each point on the estimated frontier corresponds to a portfolio of the five securities. Using the portfolio weights derived from the estimated parameters,

[2] Average monthly returns can be stated in approximate annual terms by multiplying by 12. Monthly standard deviations of return can be stated in approximate annual terms by multiplying by $\sqrt{12}$.

Table 1
True monthly parameters.

Means, μ	0.006	0.010	0.014	0.018	0.022
Std. Devs., σ	0.085	0.080	0.095	0.090	0.100
Correlations, ρ	1.00	0.30	0.30	0.30	0.30
	0.30	1.00	0.30	0.30	0.30
	0.30	0.30	1.00	0.30	0.30
	0.30	0.30	0.30	1.00	0.30
	0.30	0.30	0.30	0.30	1.00

and then applying the true parameters, the true portfolio expected return and standard deviation are computed. This defines a point on the "actual" frontier. Definitions of these terms are given in the appendix.

To summarize, the *true efficient frontier* is the efficient frontier based on the true (but unknown parameters). The *estimated frontier* is the frontier based on the estimated (and hence incorrect) parameters. The *actual frontier* consists of the true portfolio mean and variance points corresponding to portfolios on the estimated frontier. In short, the estimated frontier is what *appears* to be the case based on the data and the estimated parameters, but the actual frontier is what *really* occurs based on the true parameters. The estimated frontier is the only frontier that is observable in practice. The true and actual frontiers are unobservable because the true parameters are unknown.

In fig. 1, the point on the estimated frontier with *maximum* estimated expected return (labeled B) corresponds to the portfolio with *minimum* expected return (labeled F) on the actual frontier. Similarly, the point on the estimated frontier with minimum variance, labeled A, corresponds to point E on the actual frontier. Point E is quite close to the true minimum variance portfolio, labeled C, on the true efficient frontier.

Point B on the estimated frontier corresponds to a portfolio entirely invested in security 3. Recall that security 3 is the security with the largest positive error in the

Table 2
Estimated monthly parameters.

Means, $\hat{\mu}$	−0.009	0.002	0.042	0.022	0.028
Std. Devs., $\hat{\sigma}$	0.090	0.085	0.113	0.097	0.085
Correlations, $\hat{\rho}$	1.00	0.51	0.52	0.08	0.11
	0.51	1.00	0.49	0.05	0.12
	0.52	0.49	1.00	0.54	0.30
	0.08	0.05	0.54	1.00	0.45
	0.11	0.12	0.30	0.45	1.00

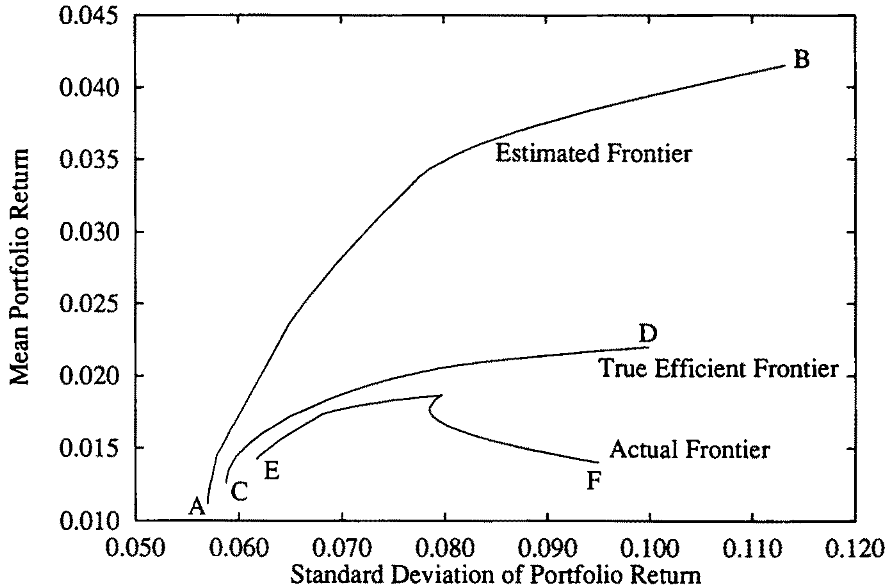

Fig. 1. Mean-variance frontiers using 24 months of simulated data.

estimation of mean returns. This illustrates the *error maximization* property of mean-variance analysis. That is, points on the estimated frontier correspond to portfolios that overweight securities with large positive errors in mean returns, large negative errors in standard deviations, and large negative errors in correlations. Hence the estimated frontier tends to exaggerate certain errors in the input parameters, resulting in optimistically biased estimates of portfolio performance.

What seems most striking about fig. 1 is that the actual frontier lies far from the true efficient frontier and even farther from the estimated frontier. The distance of the actual frontier from the true frontier is not uniform along the curves, e.g., point F is relatively far from point D while point E is close to point C. This small example suggests that with a limited amount of data minimum variance portfolios can be estimated more accurately than maximum return portfolios.

Is the result shown in fig. 1 a typical outcome using 24 data points or does it represent a low probability outcome of a simulation? As the number of data points increases, the estimates of the model parameters approach their true values. The estimated frontier and the actual frontier converge to the true efficient frontier. How fast is this convergence? How much data is necessary to achieve a desired level of convergence? The simulation procedure can be repeated to gain some understanding about these issues.

Each *trial* of the simulation involves generating random monthly returns (based on the true parameters from table 1), estimating model parameters from the simulated returns, and then computing mean-variance frontiers. The results from 25 trials using 24 months of data are shown in fig. 2. Similar results from 25

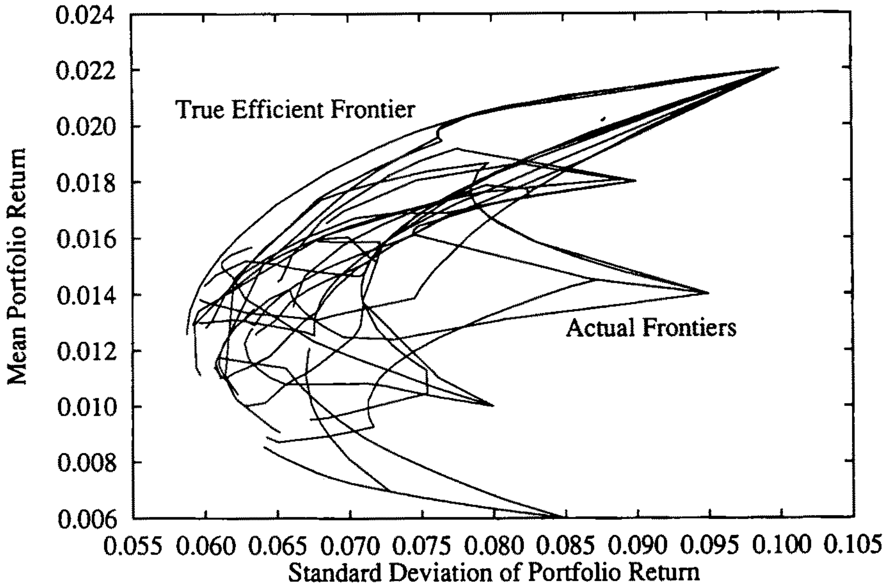

Fig. 2. 24 months of simulated data.

trials using 120 months of data are shown in fig. 3. These figures illustrate the range of uncertainty in the actual frontiers based on a limited amount of data. The results improve with more data, i.e., the actual frontiers lie closer to the true efficient frontier in fig. 3 compared to fig. 2. However, the result in fig. 1 is fairly typical.

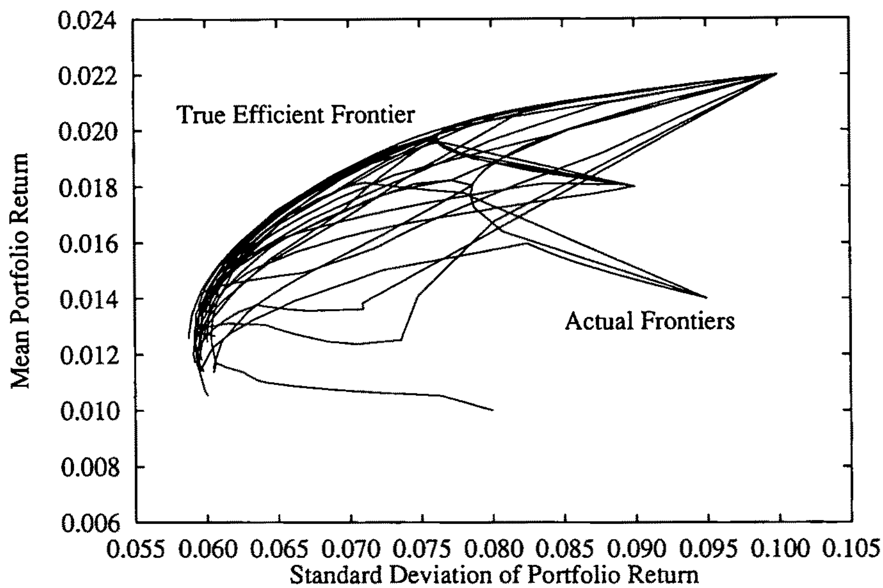

Fig. 3. 120 months of simulated data.

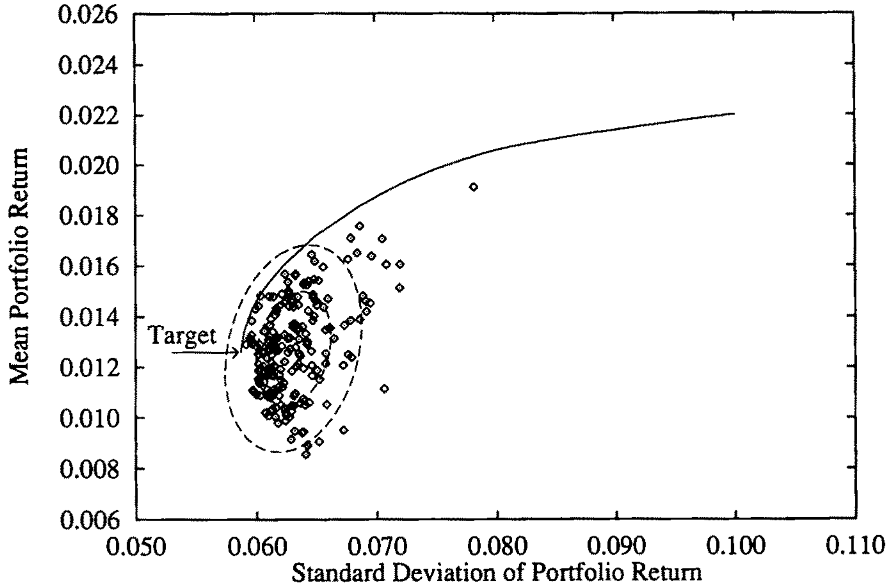

Fig. 4. Actual frontier points; $n = 24$, $t = 10,000$.

The results in figs. 2 and 3 can be clarified by focusing on individual target points on the true efficient frontier. The target point is defined as the point in the mean-variance plane that maximizes $\mu_P - t\sigma_P^2$ for a given value of t. The parameter t represents the tradeoff in utility for expected portfolio return (μ_P) versus risk (measured by variance in portfolio return, σ_P^2). For $t = 0$ the solution is the portfolio with the maximum return. For t equal to infinity, the solution is the minimum variance portfolio. For practical purposes, $t = 10,000$ is more than sufficient to give the minimum variance portfolio. Results are given in figs. 4–11.

In figs. 4–11 the scattered dots represent points on the actual frontier corresponding to a given t value. Each dot is the result of a single simulation trial. The ellipses are a visual aid to summarize the location of the multitude of dots. The ellipses represent 50% and 90% confidence regions, i.e., each smaller ellipse contains approximately 50% of the dots and each larger ellipse contains approximately 90% of the dots.[3] Figures 4–11 illustrate how close a point on the actual frontier is to a target optimal point on the true efficient frontier.

[3] The ellipses are computed under the assumption that the points follow a bivariate normal distribution. In fact the points from the actual frontiers do not follow a bivariate normal distribution. However, the ellipses are still a useful visual aid for summarizing the location and the spread of the points. Also, many points are repeated and repeated points cannot be distinguished from single points in the figures. So the ellipses are also useful because they are computed with the proper weight given to repeated points. The ellipses are computed based on 10,000 points, but to maintain the clarity of the figures, only 200 points are plotted in each figure.

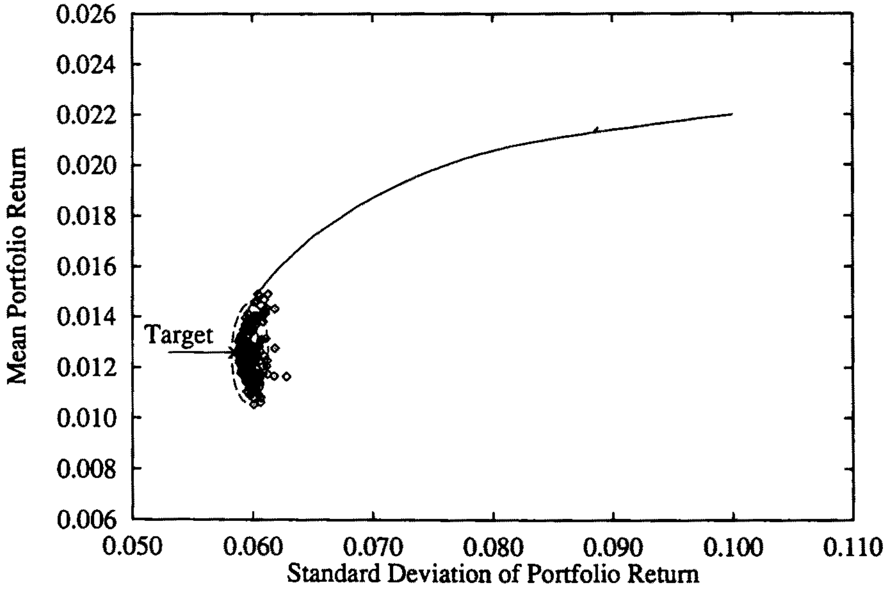

Fig. 5. Actual frontier points; $n = 120$, $t = 10,000$.

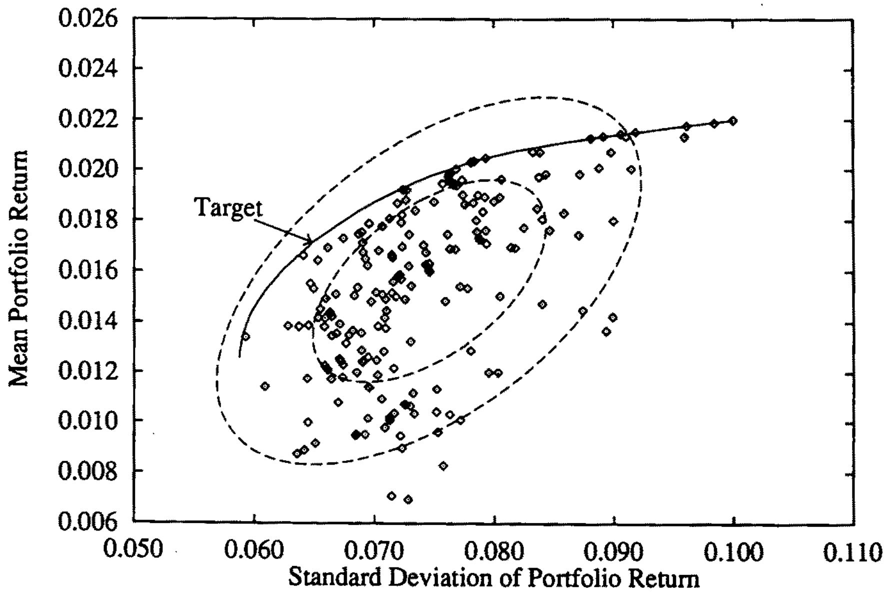

Fig. 6. Actual frontier points; $n = 24$, $t = 3$.

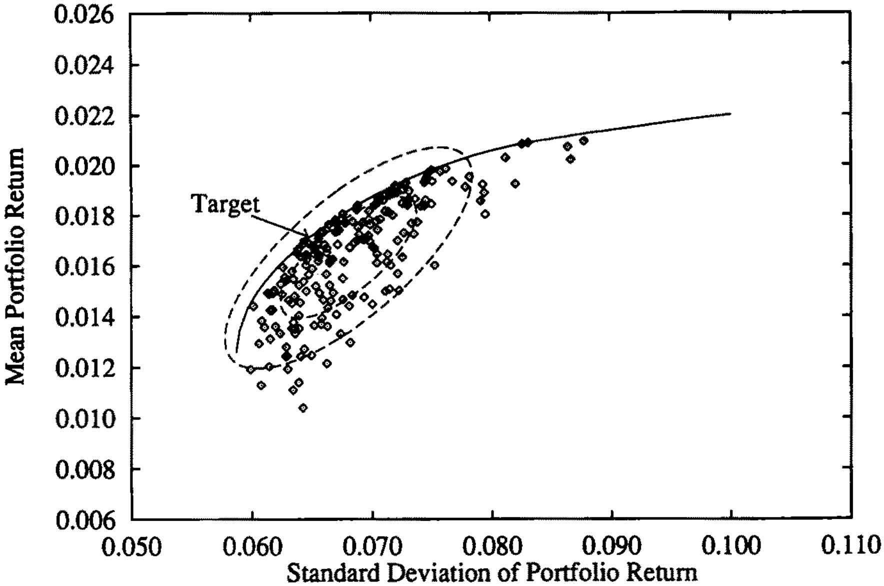

Fig. 7. Actual frontier points; $n = 120$, $t = 3$.

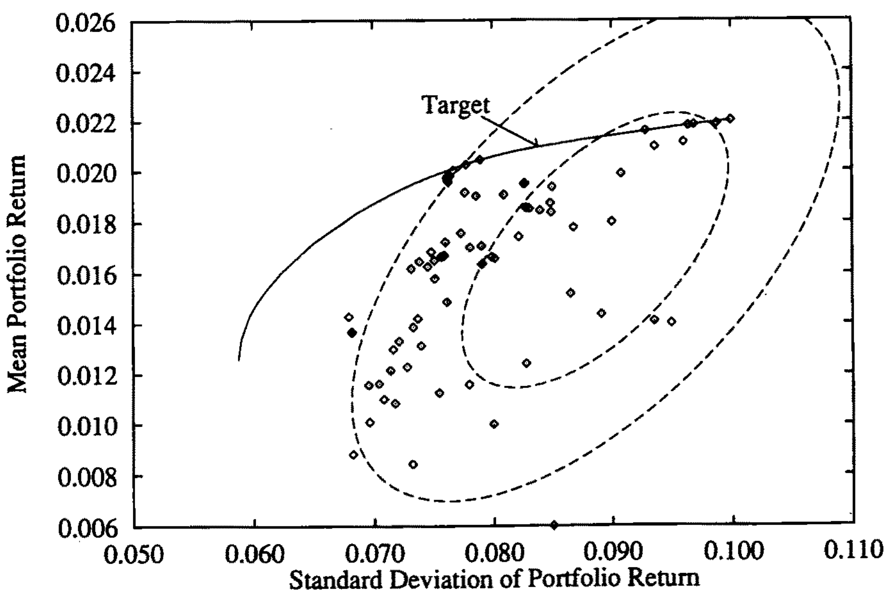

Fig. 8. Actual frontier points; $n = 24$, $t = 0.5$.

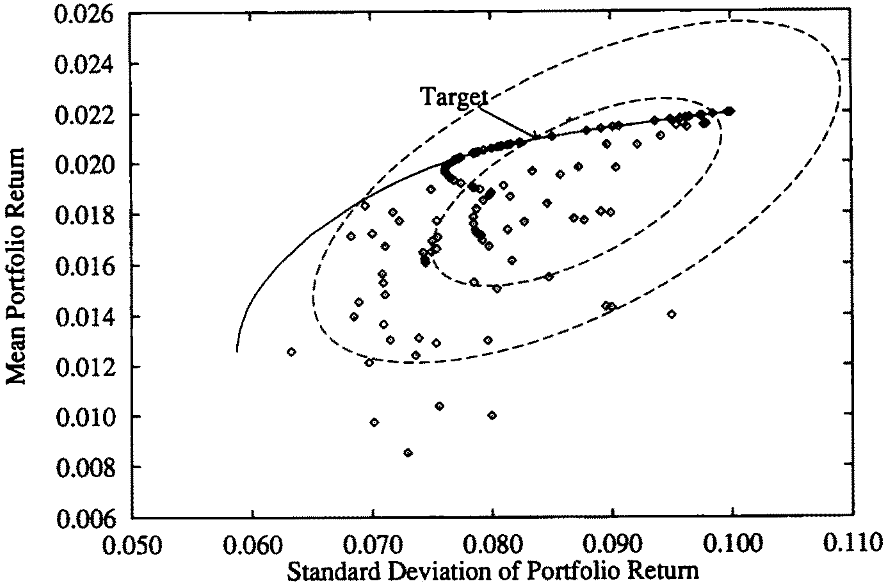

Fig. 9. Actual frontier points; $n = 120$, $t = 0.5$.

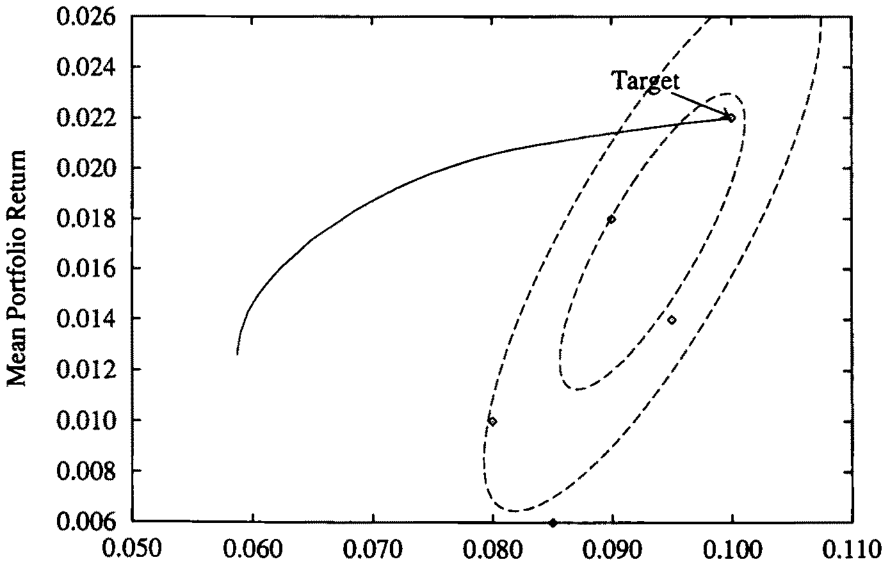

Fig. 10. Actual frontier points; $n = 24$, $t = 0$.

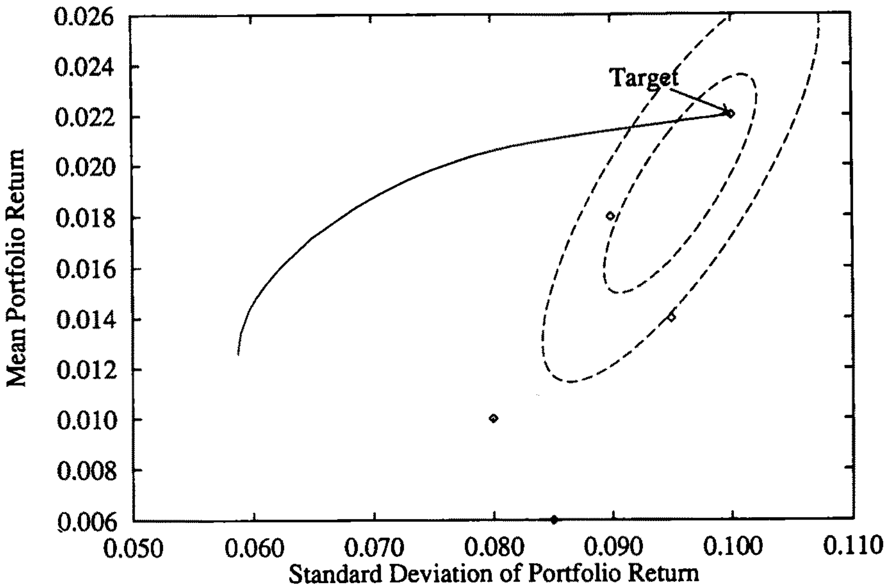

Fig. 11. Actual frontier points; $n = 120$, $t = 0$.

2.1. QUANTITATIVE MEASURES OF ESTIMATION ERROR

In order to determine an appropriate amount of data to use for parameter estimation, it is useful to have a quantitative measure of the error caused by using estimated parameters. The distance between a point on the true efficient frontier and the actual frontier is one way to measure the error.[4] It measures the actual performance of the portfolio relative to the true optimal portfolio. One way to quantify this measure is described next. Let $(\mu_P(t), \sigma_P(t))$ denote the target point on the true efficient frontier for a fixed value of t. Let $(\tilde{\mu}_P^s(t), \tilde{\sigma}_P^s(t))$ represent the mean and standard deviation of portfolio return for a corresponding point on the actual frontier for simulation trial s (where s ranges from 1 to S). Then the root-mean-squared (RMS) μ-error, denoted $f_\mu(t)$, and the RMS σ-error, denoted $f_\sigma(t)$, are given by

$$f_\mu(t) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}(\mu_P(t) - \tilde{\mu}_P^s(t))^2} \quad \text{and} \quad f_\sigma(t) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}(\sigma_P(t) - \tilde{\sigma}_P^s(t))^2}. \tag{1}$$

Table 3 shows how these error measures decline as the number of data points,

[4] Another way is to compare values of the objective function $\mu_P - t\sigma_P^2$ for the true optimal portfolio and for the estimated optimal portfolio. Other useful measures are the cash equivalent loss described in Chopra and Ziemba [4] and the utility loss used in Jorion [8].

Table 3
RMS-error measures.[a]

t	Target		$n=24$		$n=72$		$n=120$		$n=360$		$n=600$	
	$\mu_P(t)$	$\sigma_P(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$
10,000	1.26	5.88	0.19	0.51	0.12	0.20	0.09	0.12	0.05	0.04	0.04	0.02
5	1.53	6.11	0.29	1.11	0.20	0.59	0.16	0.43	0.10	0.21	0.08	0.16
3	1.72	6.49	0.37	1.27	0.26	0.75	0.22	0.58	0.13	0.33	0.10	0.26
1	2.02	7.73	0.56	1.25	0.41	1.09	0.34	1.01	0.19	0.76	0.13	0.62
0.5	2.10	8.40	0.62	1.06	0.46	1.07	0.38	1.06	0.21	1.02	0.15	0.97
0.3	2.18	9.65	0.68	1.06	0.51	1.11	0.43	1.13	0.25	1.11	0.19	1.12
0	2.20	10.00	0.70	0.93	0.53	0.76	0.46	0.69	0.28	0.50	0.21	0.44

[a] All values in the table are expressed in *percent* (except for the t column).
"Target" is the optimal point on the true efficient frontier.
n is the number of data points (i.e., number of months) used to estimate the parameters.
The RMS-error measures $f_\mu(t)$ and $f_\sigma(t)$ are defined in equation (1) and computed with $S = 10,000$.

n, used to estimate the parameters' increases. Suppose the target point on the true efficient frontier corresponds to $t=3$ and $n=24$ months of data are used to estimate parameters. Then, in the sense of the RMS-error defined above, the actual mean portfolio return will be 0.37% from the true optimal mean return of 1.72%. Also, on average, the actual standard deviation of portfolio return will be 1.27% from the true optimal standard deviation of 6.49%.

Compared to the size of the true efficient frontier, these errors are large. For example, the error in the actual portfolio mean as a percent of the difference between the largest and smallest means on the true efficient frontier is 40% $(=0.0037/(0.022-0.0126))$. Similarly, the error in the actual standard deviation as a percent of the difference between the largest and smallest standard deviations on the true efficient frontier is 31% $(=0.0127/(0.1-0.0588))$.

What is a reasonable error level to set, i.e., how close, on average, should the actual frontier point be to the target point on the true efficient frontier? Clearly, the answer differs according to the preferences of the investor. For example, suppose an investor uses $t=1$ and desires RMS-errors in the mean of less than 0.20% per month and RMS-errors in standard deviation of less than 0.75% per month (i.e., $f_\mu(1) < 0.20\%$ and $f_\sigma(1) < 0.75\%$). Then table 3 shows that n should be approximately 360. In other words, 30 years of monthly data are needed to achieve the desired level of accuracy.

3. Intuition behind the results

As the results in the previous section illustrate, the effect of estimation error can be quite large. Along the efficient frontier, it appears to be easier to estimate minimum variance portfolios than to estimate maximum return portfolios. These results can be understood by separately considering the steps in the construction of the efficient frontier. First, how large is the error in each type of input parameter? Second, how do errors in the input parameters translate into errors in the efficient frontier and optimal portfolios? We start by considering estimation errors in expected returns of securities.

3.1. FINDING THE SECURITY WITH MAXIMUM EXPECTED RETURN

The point on the true efficient frontier with maximum expected return corresponds to a portfolio that is 100% invested in the security with the largest expected return. Hence, to correctly identify this point requires only that the security with the largest expected return is correctly identified. The estimates of standard deviations and correlations do not matter for this purpose. If expected returns are estimated with historical data, how much data is necessary to identify the security with the largest expected return?

In order to gain some intuition about this problem, consider the following example consisting of only two securities. Suppose security A has an expected monthly return of 1.5% and security B has an expected monthly return of 1.0%. Furthermore, suppose the monthly returns each have a standard deviation of 7%, the returns have a correlation of 0.5, and both returns are normally distributed. It is not known which returns belong to security A and which belong to

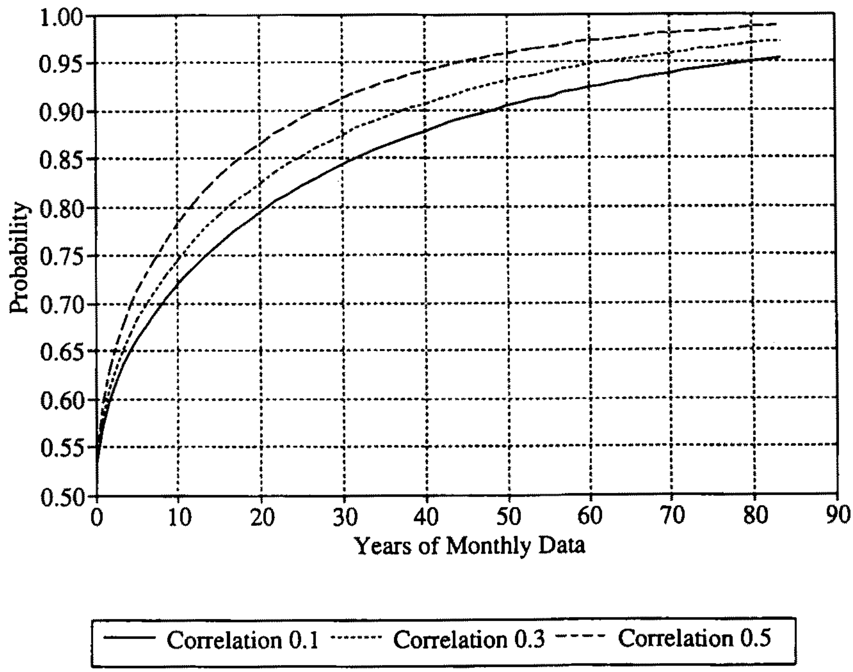

Fig. 12. Probability of distinguishing securities with different mean returns.

security B. *How many months of data would be necessary to correctly distinguish the two securities with a probability of 90%?*

The answer is illustrated in fig. 12. Even though the expected return of security A is 50% larger than security B, more than *26 years* of monthly returns are necessary. Figure 12 shows results for three levels of correlations of returns. If the correlation is 0.1, then approximately 50 years of monthly data are necessary to correctly distinguish the securities with probability 0.9. The 90% probability level is therefore too stringent a requirement to be achieved in practice. Figure 12 shows that to correctly distinguish the securities with probability 0.65 requires less than 5 years of monthly data.[5]

These results occur because the standard deviations of return are large relative to the expected mean returns. However, the values chosen for illustration are typical of monthly returns for many securities.

3.2. FINDING THE SECURITY WITH MINIMUM VARIANCE OF RETURN

Suppose security A has monthly returns that have a standard deviation of 7.0% and security B has monthly returns with a standard deviation of 6.0%. Furthermore, suppose the monthly returns each have a mean of 1%, the returns have a correlation of 0.5, and both returns are normally distributed. It is not known which returns are security A's and which are security B's. How many months of data would be necessary to correctly distinguish the two securities with a probability of 90%?

[5] Denote the random monthly returns of securities A and B by R_A and R_B, respectively. Suppose the returns are normally distributed, i.e., $R_A \sim N(\mu_A, \sigma_A^2)$, $R_B \sim N(\mu_B, \sigma_B^2)$, with correlation ρ. Without loss of generality, suppose that $\mu_A > \mu_B$. Let $R = R_A - R_B$. Then $R \sim N(\mu_A - \mu_B, \sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B)$. Let $\bar{R}$ denote the average of n monthly observations of R. The probability that the securities are correctly identified is $P(\bar{R} > 0)$, namely:

$$P(\bar{R} > 0) = P\left(\frac{\bar{R} - (\mu_A - \mu_B)}{\sigma(\bar{R})} > \frac{-(\mu_A - \mu_B)}{\sigma(\bar{R})}\right) = P\left(Z < \frac{\mu_A - \mu_B}{\sigma(\bar{R})}\right)$$

$$= P\left(Z < \frac{\mu_A - \mu_B}{\sqrt{\frac{1}{n}(\sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B)}}\right).$$

Since Z is a standard normal random variable, the last probability is easily evaluated. If the desired probability is 0.9, the corresponding standard normal z-value is 1.28. Solving for n using the parameter values from the text gives

$$\frac{\mu_A - \mu_B}{\sqrt{\frac{1}{n}(\sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B)}} \geq 1.28 \Rightarrow n \geq \left(\frac{1.28\sqrt{(\sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B)}}{\mu_A - \mu_B}\right)^2 = \left(\frac{1.28(0.07)}{0.005}\right)^2 \approx 321.$$

Hence to correctly identify the securities with 90% probability requires at least 321 months of data.

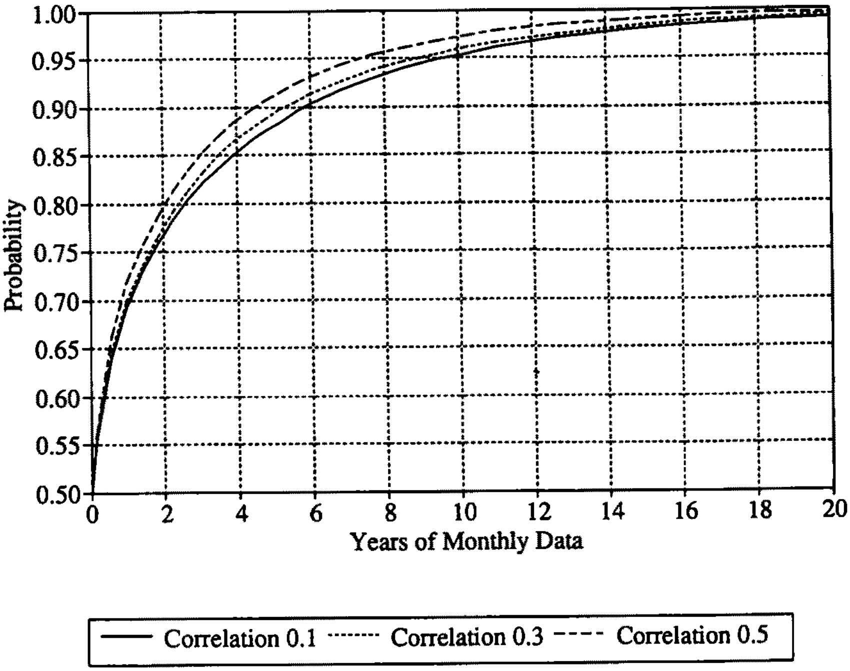

Fig. 13. Probability of distinguishing securities with different std. devs. of return.

The answer is illustrated in fig. 13.[6] Less than five years of monthly data are necessary, even though the standard deviations are quite close. To distinguish the securities with probability 0.65 requires less than one year of data.

These results illustrate that distinguishing securities with different standard deviations of return is much easier (i.e., requires fewer observations) than distinguishing securities with different mean returns. Alternatively, for a given number of observations, the error in the estimate of mean returns is typically much larger than the error in the estimate of standard deviation of returns.

3.3. INTERMEDIATE CASES

Points on the efficient frontier are obtained by maximizing $\mu_P - t\sigma_P^2$ for various values of t. For small values of t, mean returns drive the optimization. For larger values of t, variances become increasingly important in the optimization. In most cases, the investor's objective will lie between the extremes of $t = 0$ and $t = \infty$. Figure 14 shows an investor's indifference curve for $t = 0.5$.

[6] The values in fig. 13 were determined by simulation. Using the procedure outlined in the appendix, returns were simulated and sample standard deviations were computed. The security with the larger sample standard deviation was assumed to be security A. The simulated fraction of the time that this conclusion is correct is an estimate of the true probability of correctly identifying the securities.

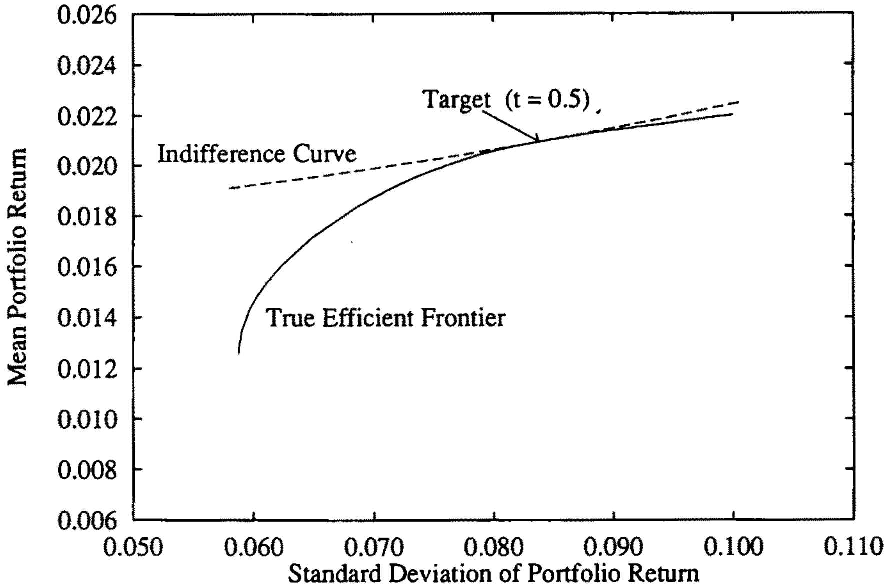

Fig. 14. Indifference curve for $t = 0.5$.

The indifference curve, or iso-utility curve, is a plot of equal values of the objective function $\mu_P - t\sigma_P^2$. Figure 14 shows that the indifference curve and true efficient frontier are both nearly linear in a large region around the optimal solution. Hence a small change in the parameters could lead to a large change in the optimal solution. This explains, in part, why many of the scattered dots in fig. 9 lie on the true efficient frontier, but far from the target point. The points on the actual frontier can be close to the true efficient frontier, but not particularly close to the target point.

3.4. ERROR MAXIMIZATION

Suppose that two securities each have a mean return of 1% per month, but the estimated mean returns are 1.4% and 0.4% per month. A portfolio that is 50% invested in each security will have a true mean return of 1% and an estimated mean return of 0.9%. The error, i.e., the difference between the true and estimated returns, for this fixed portfolio is only 0.1%.[7] For fixed weight portfolios estimation errors tend to cancel. However, if an investor chooses the portfolio which maximizes (estimated) expected return, the error is 0.4%. Optimizing expected return leads the investor to choose securities with the largest (positive) estimation errors. The magnification of estimation error induced by the optimization step in forming a

[7] This notion of error, the difference between *true* and *estimated* portfolio returns, is considered in section 5.

portfolio is commonly termed *error maximization*. The effect becomes more pronounced as the number of securities increases, as detailed in section 6. A similar effect happens when the investor is trying to minimize portfolio variance. In this case, the optimization will overweight securities with the most underestimated standard deviations. The relative importance of errors in means, variances, and covariances is documented in Chopra and Ziemba [4].

The problem of finding the maximum return portfolio involves maximizing a linear function over a convex region. Optimal solutions for this problem are found at extreme points of the feasible region. That is, the maximum return portfolio will be entirely invested in a single security. Any error in estimating returns will tend to be magnified because of this portfolio specialization. Conversely, the problem of finding the minimum variance portfolio involves minimizing a convex function over a convex region. Optimal solutions for this problem are found in the interior of the feasible region, i.e., where the portfolio weights on many securities are positive. The error maximization effect of the optimization will be somewhat mitigated because of this portfolio diversification.

To summarize, using historical data to estimate parameters leads to greater errors in means than in standard deviations. The error maximization property of the optimization step implies that optimal portfolios favor securities with overestimated expected returns and underestimated standard deviations. These remarks indicate why using estimated parameters produces greater errors in the maximum return extreme of the efficient frontier than in the minimum variance extreme.

4. Estimation error with improved parameter estimates

Significant improvements in the results of mean-variance analysis are possible if better methods are used to estimate parameters. Suppose that a security has a mean monthly return of 1%. If historical data is used to estimate the mean, it is quite possible that the estimate will be negative. However, it is unlikely that investors would hold positive amounts of the security if it truly had a negative mean return. Clearly, then it is possible to improve historical estimates of mean returns.

Parameter estimates based solely on historical data can be improved upon in several ways. Chopra [2], Chopra et al. [3], and Jorion [8] investigate techniques for improving historical estimates, including the use of Stein estimators. Historical estimates can be combined with analysts' forecasts. Mathematical models can be used to estimate mean returns and standard deviations. For example, dividend discount models are often used to estimate stock parameters, and yield-based models are often used to estimate parameters for bonds. Several authors have investigated methods for adjusting historical estimates of a security's beta. These methods correct for the tendency of betas to regress towards one. See chapter 5 of Elton and Gruber [5] for a summary and bibliography. These methods can be applied to adjusting historical estimates of a security's mean return and also to estimating the correlation matrix of returns. Theories of efficient markets can be used to determine the market consensus estimate of mean returns. See, for example, the discussion in chapter 7 of Sharpe [13]. This approach is appropriate when the securities represent very large asset classes. Combining implied volatility from option values with historical estimates might give better forecasts of future standard deviations of security returns.

The effect of using improved estimates of mean returns on estimating mean-variance efficient frontiers is investigated next. The method for improving the estimates will remain unspecified, but it will be assumed that some method for adjusting historical estimates of mean returns is used that improves the estimates by 50%. That is, the adjusted estimates lie halfway between the historical estimates and the true values for every security. The parameters from table 1 were used to analyze the distance of points on the actual frontier (AF) from the corresponding point on the true frontier with improved estimates of the mean. The results are presented in figs. 15–20 and in table 4. Figure 15 is directly comparable to fig. 6, fig. 16 is directly comparable to fig. 7, etc. The minimum variance portfolio results (corresponding to $t = 10,000$) are not repeated, because they are identical to figs. 4 and 5.

The results show that *significant reductions in errors are possible with improved estimates of mean returns*. For example, suppose that an investor uses $t = 1$ and desires RMS-errors in the mean of less than 0.20% per month and RMS-errors and RMS-errors in standard deviation of less than 0.75% per month. Recall that the RMS-values are a measure of the distance of the target point on the true efficient frontier from the corresponding point on the actual frontier. If the investor uses a method that improves all of the estimates of mean returns by 50% over

Table 4
RMS-error measures with improved mean estimates.[a]

t	Target		$n = 24$		$n = 72$		$n = 120$		$n = 360$		$n = 600$	
	$\mu_P(t)$	$\sigma_P(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$
10,000	1.26	5.88	0.19	0.51	0.12	0.20	0.09	0.12	0.05	0.04	0.04	0.02
5	1.53	6.11	0.23	0.79	0.15	0.38	0.12	0.27	0.07	0.14	0.06	0.10
3	1.72	6.49	0.27	0.87	0.17	0.47	0.14	0.36	0.08	0.20	0.06	0.16
1	2.02	7.73	0.38	1.09	0.22	0.84	0.17	0.71	0.08	0.44	0.06	0.34
0.5	2.10	8.40	0.42	1.07	0.24	1.03	0.18	0.99	0.09	0.87	0.07	0.76
0.3	2.18	9.65	0.47	1.11	0.28	1.12	0.22	1.12	0.12	1.04	0.09	0.97
0	2.20	10.00	0.49	0.72	0.31	0.53	0.24	0.47	0.12	0.30	0.08	0.21

[a] All values in the table are expressed in *percent* (except for the t column).
"Target" is the optimal point on the true efficient frontier.
n is the number of data points (i.e., number of months) used to estimate the parameters.
The RMS-error measures $f_\mu(t)$ and $f_\sigma(t)$ are defined in equation (1) and computed with $S = 10,000$.

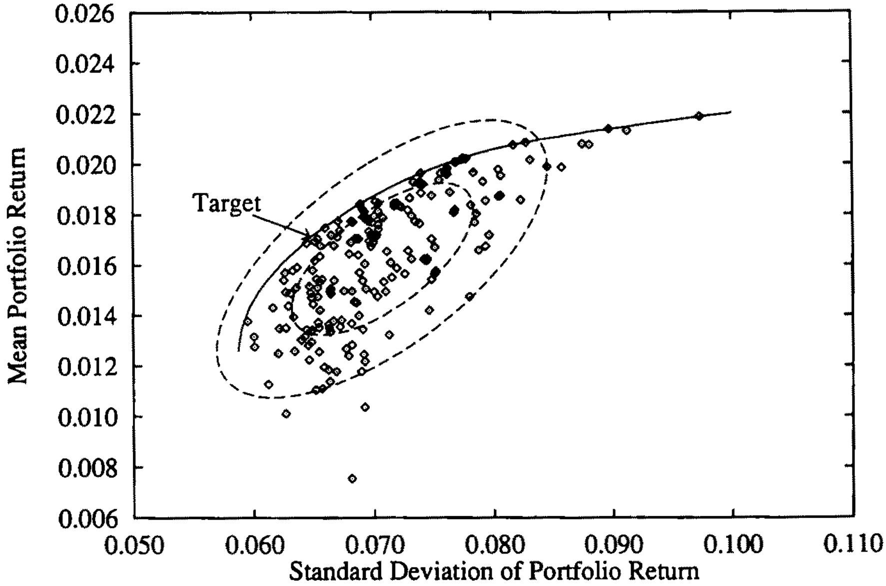

Fig. 15. AF points with improved means; $n = 24$, $t = 3$.

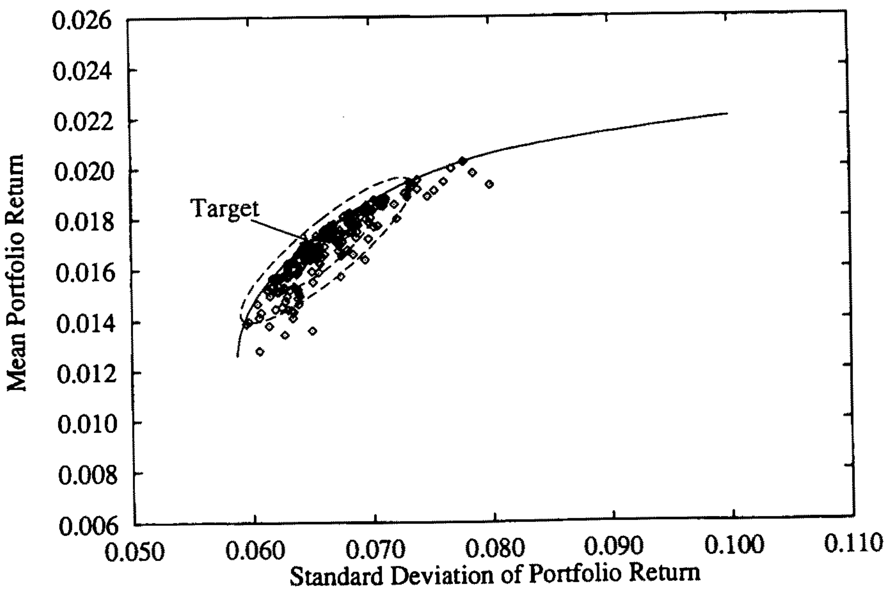

Fig. 16. AF points with improved means; $n = 120$, $t = 3$.

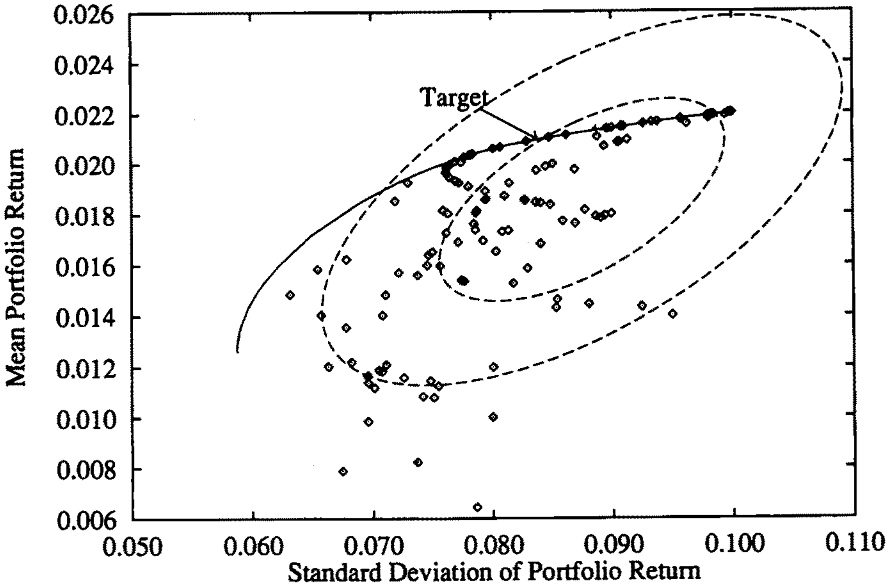

Fig. 17. AF points with improved means; $n = 24$, $t = 0.5$.

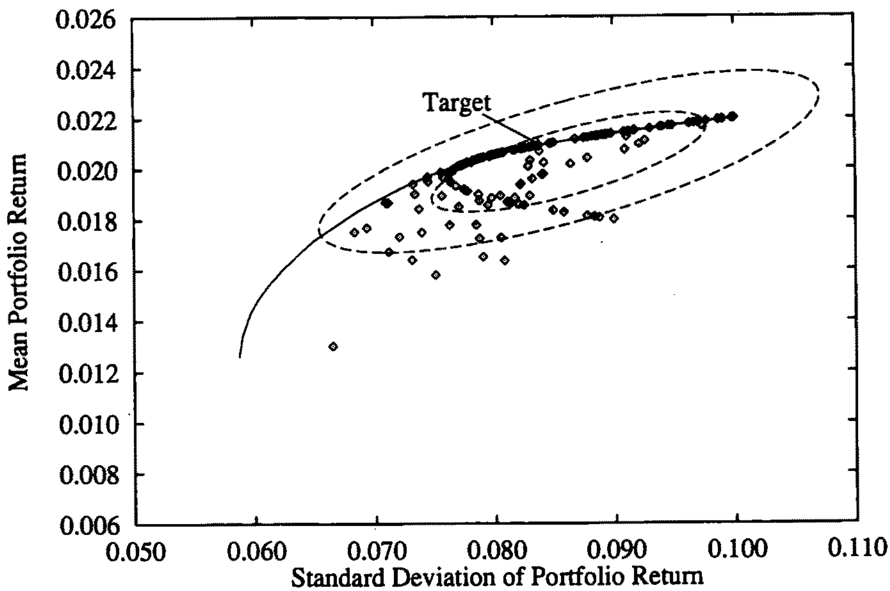

Fig. 18. AF points with improved means; $n = 120$, $t = 0.5$.

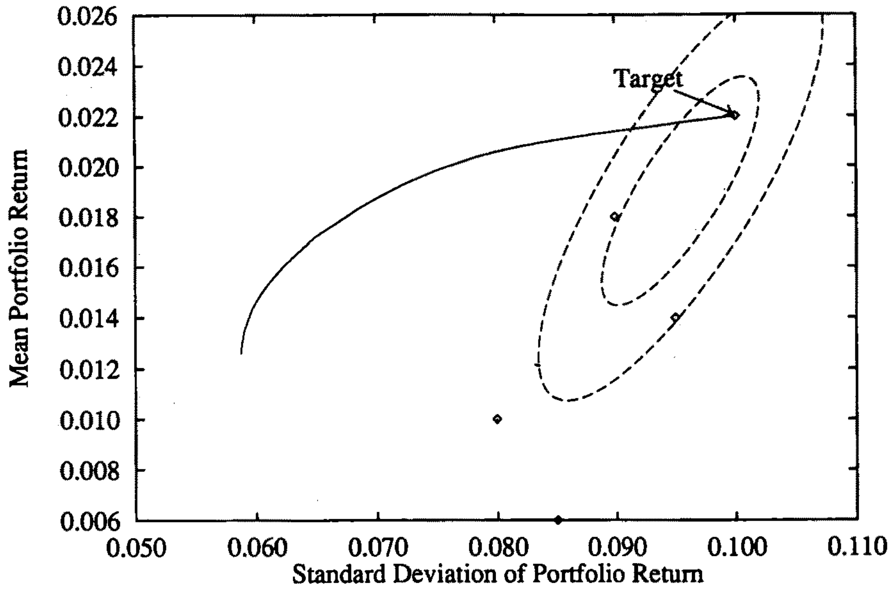

Fig. 19. AF points with improved means; $n = 24$, $t = 0$.

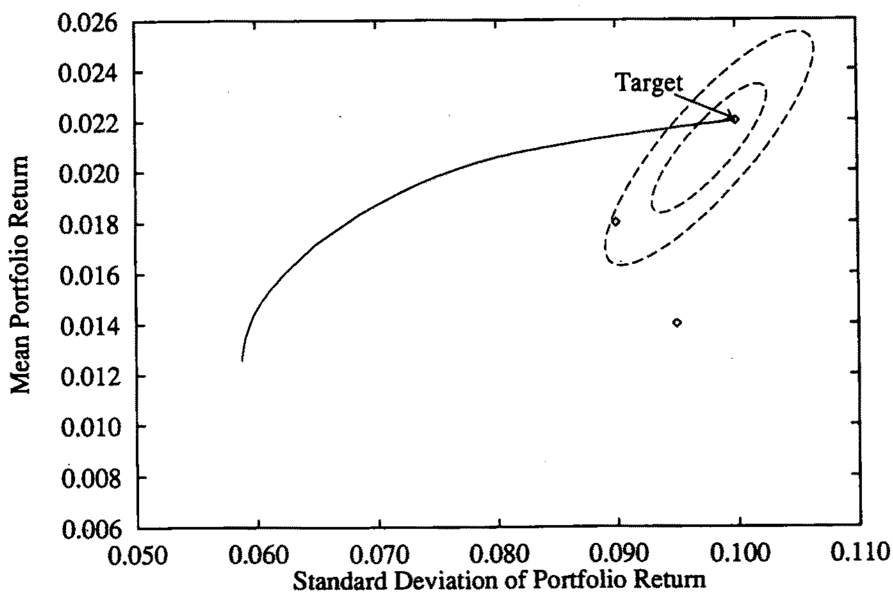

Fig. 20. AF points with improved means; $n = 120$, $t = 0$.

historical estimates, then table 4 shows that n should be approximately 96. Only about *one-fourth* as much historical data is needed to obtain the same level of accuracy. Comparing the results in tables 3 and 4 shows that improving the estimates of mean returns reduces the RMS-μ error *and* the RMS-σ error. Examination of fig. 18 shows that even with a large amount of data, improved estimates of mean returns, and stationary parameters, *the points on the actual frontier can be close to efficient, but still not be close to the target point on the true efficient frontier.* A complementary phenomenon is that portfolios with quite different compositions can be close to the same target point on the true efficient frontier. See Chopra [2] for a discussion of this point.

The same approach can be used to investigate the effect of improved estimates of standard deviations and correlations. Rather than reproduce several tables analogous to table 4, results for $n = 48$ are presented in table 5. In table 5, the improved parameter estimates refer to 50% improvements in all of the parameters of a given type. While improved estimates of mean returns dramatically reduce RMS-errors, improvements in estimates of standard deviations and

Table 5
RMS-error measures with improved parameter estimates.[a]

t	Target		Historical estimates		Improved μ_j		Improved σ_j		Improved ρ_{jk}	
	$\mu_P(t)$	$\sigma_P(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$
10,000	1.26	5.88	0.14	0.29	0.14	0.29	0.11	0.18	0.11	0.19
5	1.53	6.11	0.23	0.75	0.18	0.50	0.22	0.66	0.23	0.73
3	1.72	6.49	0.31	0.92	0.21	0.59	0.30	0.85	0.30	0.93
1	2.02	7.73	0.48	1.17	0.28	0.94	0.47	1.16	0.47	1.17
0.5	2.10	8.40	0.53	1.07	0.31	1.05	0.53	1.07	0.53	1.07
0.3	2.18	9.65	0.59	1.10	0.35	1.13	0.58	1.10	0.58	1.10
0	2.20	10.00	0.61	0.84	0.38	0.61	0.60	0.84	0.60	0.84

t	Target		Historical estimates		Improved μ_j and σ_j		Improved μ_j and ρ_{jk}		Improved μ_j, σ_j, and ρ_{jk}	
	$\mu_P(t)$	$\sigma_P(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$	$f_\mu(t)$	$f_\sigma(t)$
10,000	1.26	5.88	0.14	0.29	0.11	0.18	0.11	0.19	0.07	0.08
5	1.53	6.11	0.23	0.75	0.15	0.39	0.16	0.45	0.13	0.32
3	1.72	6.49	0.31	0.92	0.19	0.49	0.20	0.58	0.18	0.46
1	2.02	7.73	0.48	1.17	0.27	0.92	0.27	0.96	0.27	0.92
0.5	2.10	8.40	0.53	1.07	0.30	1.05	0.30	1.06	0.30	1.05
0.3	2.18	9.65	0.59	1.10	0.35	1.12	0.35	1.12	0.35	1.12
0	2.20	10.00	0.61	0.84	0.37	0.60	0.37	0.60	0.37	0.60

[a] All values in the table are computed with $n = 48$ and $S = 10,000$.

correlations have very little impact. Small additional reductions in RMS-errors are possible if means and standard deviations are improved. Except if the investor is solely interested in minimizing variance, table 5 indicates that historical estimates of mean returns should be used with caution. If possible, *effort should first be focused on improving the historical estimates of mean returns.*

5. Errors in the estimated frontier

The previous sections focused on one type of error, namely the distance between points on the actual frontier and the corresponding points on the true efficient frontier. However, neither of these frontiers is observable in practice. Hence it is also of interest to ask how far is the estimated frontier, which is observable, from the true and actual frontiers.

Figures 21 and 22 show points on the estimated frontier (EF) corresponding to $t = 0.5$. The points are computed assuming that estimates of mean returns are improved by 50% over historical data alone. Figures 21 and 22 are directly comparable to figs. 17 and 18. The estimated frontier has much more error (relative to the true efficient frontier) than the actual frontier.

The centers of the ellipses in figs. 21 and 22 lie beyond the true efficient frontier. This is another manifestation of the *error maximization* property of mean-variance analysis. The estimated frontier, on average, overstates means and understates standard deviations of portfolio return.

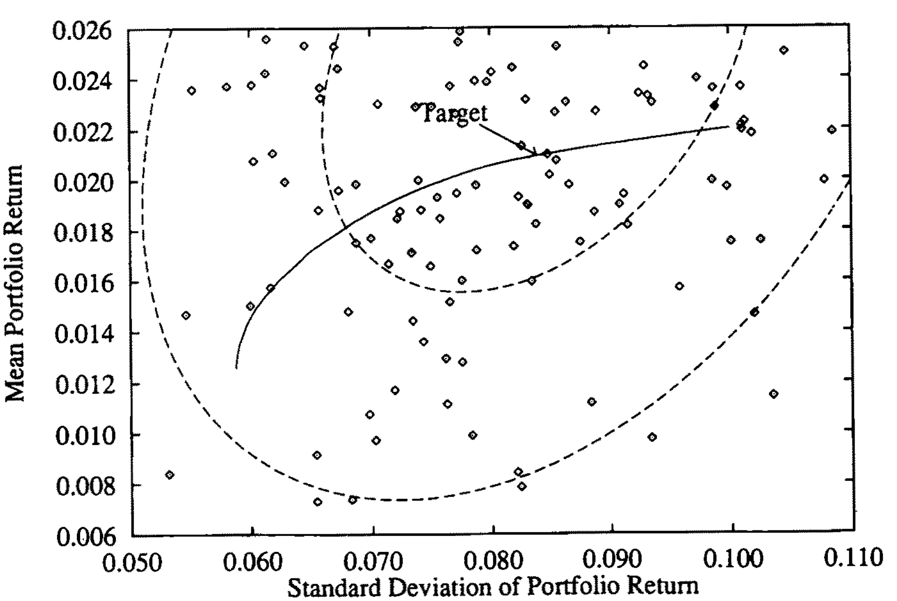

Fig. 21. EF points with improved means; $n = 24$, $t = 0.5$.

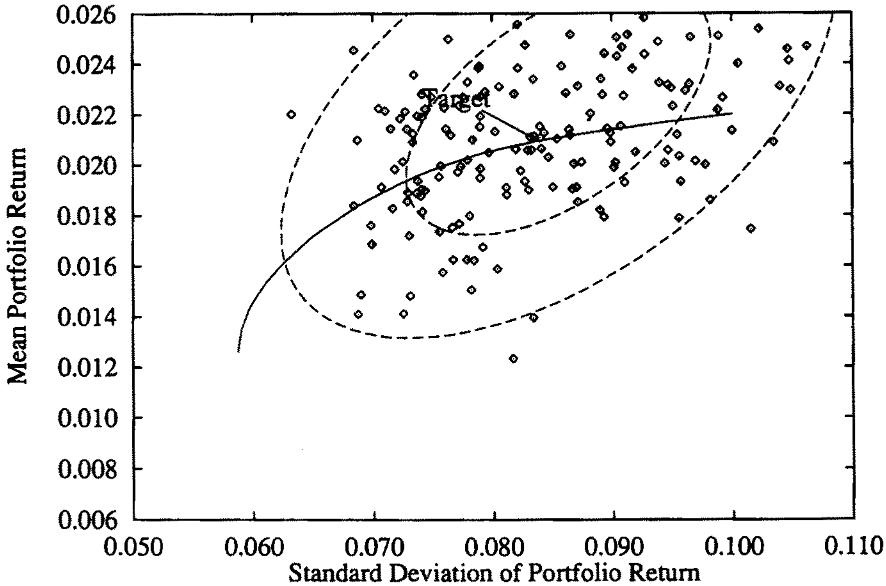

Fig. 22. EF points with improved means; $n = 120$, $t = 0.5$.

5.1. ESTIMATED FRONTIER RELATIVE TO THE ACTUAL FRONTIER

Investors observe the estimated frontier but portfolio performance is given by the actual frontier. How far is the estimated frontier from the actual frontier? Let $(\hat{\mu}_P^s(t), \hat{\sigma}_P^s(t))$ denote the mean and standard deviation of portfolio return for a point on the estimated frontier for simulation trial s and let $(\tilde{\mu}_P^s(t), \tilde{\sigma}_P^s(t))$ denote the corresponding point on the actual frontier. Then the average difference of the μ values, denoted $g_\mu(t)$, and the average difference of the σ values, denoted $g_\sigma(t)$, are given by

$$g_\mu(t) = \frac{1}{S}\sum_{s=1}^{S}(\hat{\mu}_P^s(t) - \tilde{\mu}_P^s(t)) \quad \text{and} \quad g_\sigma(t) = \frac{1}{S}\sum_{s=1}^{S}(\hat{\sigma}_P^s(t) - \tilde{\sigma}_P^s(t)). \tag{2}$$

Results are given in table 6. The g-values summarize the average position of the estimated frontier relative to the actual frontier. The RMS-error measures previously used combined relative position and spread.

The results in table 6 quantify the amount of error maximization in mean-variance analysis. For example, suppose that $n = 48$ months of data are used to estimate parameters, and that estimated means can be improved by 50%. If an investor chooses portfolios with $t = 1$, then the mean portfolio return on the estimated frontier will be 0.39% greater than the actual mean portfolio return, on average. Similarly, the standard deviation of return on the estimated frontier will be 0.26% less than the actual standard deviation of return on average. To be more

Table 6
Relative position of the estimated and actual frontiers with improved mean estimates.[a]

t	True eff. frontier		$n = 24$		$n = 72$		$n = 120$		$n = 360$		$n = 600$	
	$\mu_P(t)$	$\sigma_P(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$
10,000	1.26	5.88	0.00	−1.01	0.00	−0.36	0.00	−0.22	0.00	−0.07	0.00	−0.04
5	1.53	6.11	0.35	−0.92	0.13	−0.35	0.08	−0.22	0.02	−0.07	0.02	−0.04
3	1.72	6.49	0.47	−0.81	0.18	−0.31	0.12	−0.20	0.04	−0.07	0.03	−0.04
1	2.02	7.73	0.66	−0.49	0.27	−0.19	0.17	−0.12	0.05	−0.04	0.04	−0.02
0.5	2.10	8.40	0.70	−0.32	0.30	−0.12	0.20	−0.08	0.07	−0.03	0.04	−0.02
0.3	2.18	9.65	0.70	−0.24	0.30	−0.09	0.20	−0.06	0.06	−0.02	0.04	−0.01
0	2.20	10.00	0.69	−0.10	0.29	−0.03	0.19	−0.02	0.05	−0.01	0.02	0.00

[a] All values in the table are expressed in *percent* (except for the t column).
n is the number of data points (i.e., number of months) used to estimate the parameters.
The error measures $g_\mu(t)$ and $g_\sigma(t)$ are defined in equation (2) and computed with $S = 10,000$.

explicit, suppose the point on the estimated frontier has an average annual return of 25.0% and an annual standard deviation on 24.2%. Then an unbiased estimate of the actual portfolio mean return is 20.3% (= 25.0% − 12(0.39)%) and standard deviation is 25.1% (= 24.2% + $\sqrt{12}$(0.26)%).

These results are important because they indicate the *large degree to which estimated frontiers are optimistically biased predictors of actual portfolio performance*. To get better estimates of actual portfolio performance, the portfolio means and standard deviations reported by the estimated frontier should be adjusted to account for the bias caused by the error maximization effect. An example of this adjustment process is given in section 7.

6. Results with additional securities

As the number of securities increases, the error maximization property of mean-variance analysis becomes even more pronounced. *Points on the estimated frontier tend to increasingly overstate actual portfolio performance as the number of securities increases.* With more securities in the analysis, it is likely that some security will have a large positive error in the estimation of its mean return. It is also likely that some security will have a large negative error in the estimation of its standard deviation of return.

The average magnitude of these effects is illustrated in table 7. Additional securities were created by duplicating the mean and standard deviation parameters from table 1. All correlations p_{jk} for $j \neq k$ were set to 0.30. Table 7 illustrates how the average position of the estimated frontier relative to the actual frontier increases with the number of securities. For example, suppose an investor uses $t = 1$ to choose

Table 7
Effect of the number of securities on the relative position of the estimated and actual frontiers with improved mean estimates.[a]

T	$m = 5$		$m = 10$		$m = 20$		$m = 40$	
	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$	$g_\mu(t)$	$g_\sigma(t)$
10,000	0.00	−0.52	0.00	−0.84	0.00	−1.18	0.00	−1.50
5	0.19	−0.49	0.33	−0.73	0.45	−0.97	0.57	−1.17
3	0.27	−0.44	0.42	−0.62	0.55	−0.79	0.68	−0.94
1	0.39	−0.26	0.57	−0.36	0.72	−0.44	0.88	−0.52
0.5	0.42	−0.17	0.62	−0.24	0.79	−0.28	0.95	−0.33
0.3	0.42	−0.12	0.63	−0.17	0.81	−0.20	0.98	−0.23
0	0.42	−0.04	0.64	−0.05	0.82	−0.05	0.99	−0.05

[a] All values in the table are expressed in *percent* (except for the t column).
m is the number of securities.
All values in the table are computed for $n = 48$.
The error measures $g_\mu(t)$ and $g_\sigma(t)$ are defined in equation (2) and computed with $S = 10,000$.

a portfolio from $m = 5$ securities. From table 7, the estimated mean portfolio return will, on average, exceed the actual portfolio mean return by 4.7% per year (i.e., (12)0.39%). If $m = 40$ securities are used in the analysis, the estimated mean portfolio return will, on average, exceed the actual portfolio mean return by 10.6% per year (i.e., (12)0.88%). Since the largest possible expected return with these parameters is only 26.4%, the optimistic bias in the estimated frontier can be very large. These results are based on the investor using $n = 48$ months of data and improving estimates of mean returns by 50%.

Suppose that the true parameters of securities are known. Then as the number of securities included in mean-variance analysis increases, the resulting portfolio performance must improve. Now suppose that the parameters of the securities are estimated. Table 7 illustrates that the *performance improvement measured by estimated parameters will tend to overstate the gains from additional securities*. The amount of the overstatement can be estimated by procedures similar to the ones used to develop table 7. As noted by Jorion [9] the amount of improvement with additional securities should be tested in this way for statistical significance.

7. Adjusting for the bias in the estimated frontier

The results of the previous sections were all derived using true underlying parameters. In practice the true parameters are rarely, if ever, known. Even so, the effects of estimation error can be estimated using the same simulation methodology, by using the estimated parameters in place of the true parameters.

Suppose an investor wishes to correct for the optimistic bias inherent in the estimated frontier. The same simulation methodology described in the appendix and applied in section 5 can be used with the estimated parameters taking the place of the true parameters. As done in table 6, the average position of the estimated frontier relative to the actual frontier can be estimated. The results can then be used to estimate the position of the actual frontier. More specifically, the simulation procedure can be used to estimate the values $g_\mu(t)$ and $g_\sigma(t)$ defined in equation (2). Denote a point on the estimated frontier by $(\hat{\mu}_P(t), \hat{\sigma}_P(t))$. Then an estimate of the corresponding point on the actual frontier is $(\hat{\mu}_P(t) - g_\mu(t), \hat{\sigma}_P(t) - g_\sigma(t))$.

To illustrate the procedure, the following steps were performed. Using parameters from table 1, $n = 48$ months of historical data were generated. Parameters were estimated and then estimates of mean returns were improved by 50%. Then the estimated frontier was computed. Using the procedure described above, the position of the actual frontier was estimated. The results are plotted in fig. 23. The estimated position of the actual frontier lies beneath the estimated frontier. For reference, fig. 23 also shows the real position of the actual frontier. As expected, the adjustment process is not perfect, but still the estimated position of the actual frontier gives a better indication of portfolio performance than the original estimated frontier. Figure 24 displays 50% and 90% confidence regions for the location of a point on the actual frontier corresponding to a target point on the estimated frontier.[8] Notice that the actual frontier point corresponding to $t = 3$ lies within both the 50% and 90% confidence regions determined by the simulation.

The procedure illustrated in this section gives a method for adjusting for the bias inherent in mean-variance analysis. The procedure can be used to give both point estimates and confidence regions for the bias adjustment.

8. Investigating stationarity using historical data

Nonstationarity means that parameters change over time. If parameters are not stationary, then parameter estimates based on simple sample estimates can be biased. Nonstationary parameters can increase the errors described in the previous sections. In this section a test for nonstationarity in historical data is described.

In order to illustrate the ideas, the following data will be used for analysis. There are five asset classes, with monthly returns from January, 1926 through December, 1988. The asset classes are: (1) common stocks (based on Standard and Poor's Composite index), (2) small stocks, (3) long-term corporate bonds, (4) long-term government bonds, and (5) US Treasury bills.[9]

An obvious approach to investigating the stationarity of data is to plot time

[8] The confidence regions displayed in fig. 24 are approximate, because they are determined from estimated parameters rather than the true parameters.

[9] This data is from Ibbotson Associates.

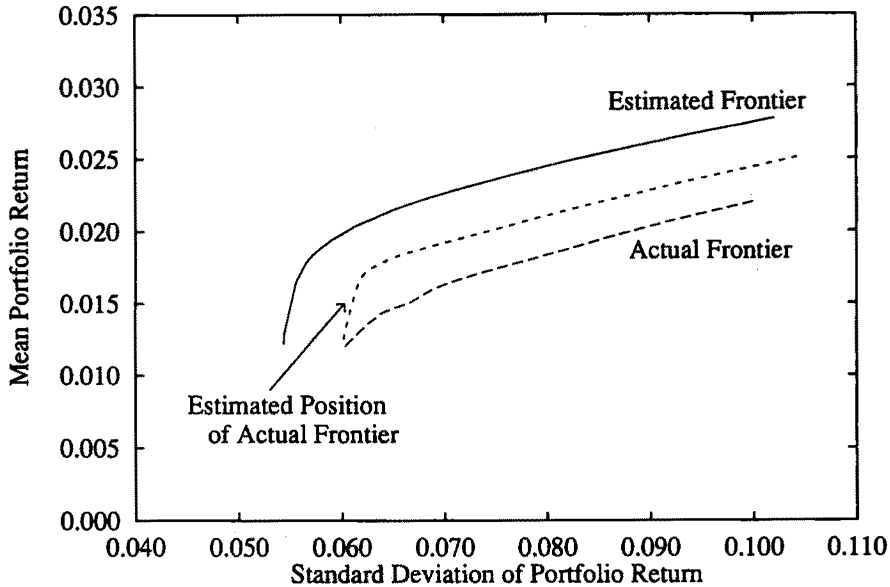

Fig. 23. Estimated position of actual frontier; $n = 48$, improved means.

series of returns and check for trends or patterns. While this is a reasonable starting point, several problems immediately arise. First, if there are m securities, then there are m time series to analyze. In addition to investigating the stationarity of m mean returns, there are also m standard deviations and $(m^2 - m)/2$ correlations to be

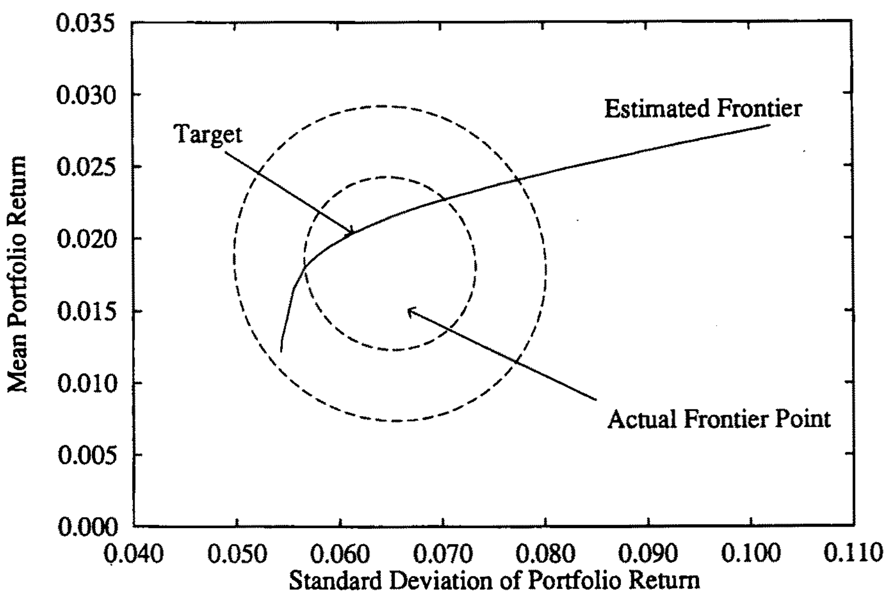

Fig. 24. Confidence regions for AF point; $n = 48$, improved means, $t = 3$.

investigated. Visual evidence of trends might suggest nonstationarity, but without quantitative analysis it is difficult to determine whether the trends are statistically significant.

Figures 25 and 26 plot 2-year moving average estimates of mean returns and standard deviations for common stocks and long-term government bonds, respectively. The horizontal axis scale counts the number of months from the end of the data. For example, $n = 1$ corresponds to December, 1988, $n = 2$ is November, 1988, $n = 24$ is January, 1987, etc. This numbering scheme is used to compare results with values of n from the previous sections. The plots in figs. 25 and 26 go back to $n = 240$ corresponding to January, 1969. The market break of October, 1987 is evidenced by a drop in the moving average return and a sharp rise in the moving average standard deviation at $n = 15$ in fig. 25. Both graphs show evidence of nonstationarity in the parameters.

Stationarity of any correlation parameter can be investigated in the same way. Figure 27 plots 2-year moving average correlation of returns of common stocks and long-term government bonds. From fig. 27 it appears that the correlation between these two asset classes is almost always positive. However, figure 28, which shows results going back to January, 1926, contradicts this conclusion. This is also evidence of nonstationarity of this parameter.

The graphs give evidence of nonstationarity of the return parameters, but are

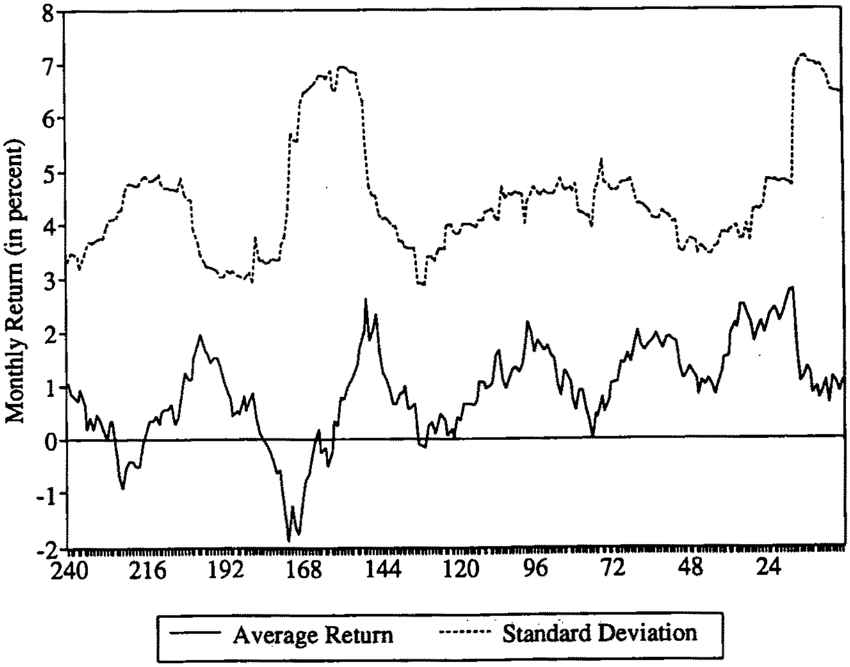

Fig. 25. Moving average results for common stocks.

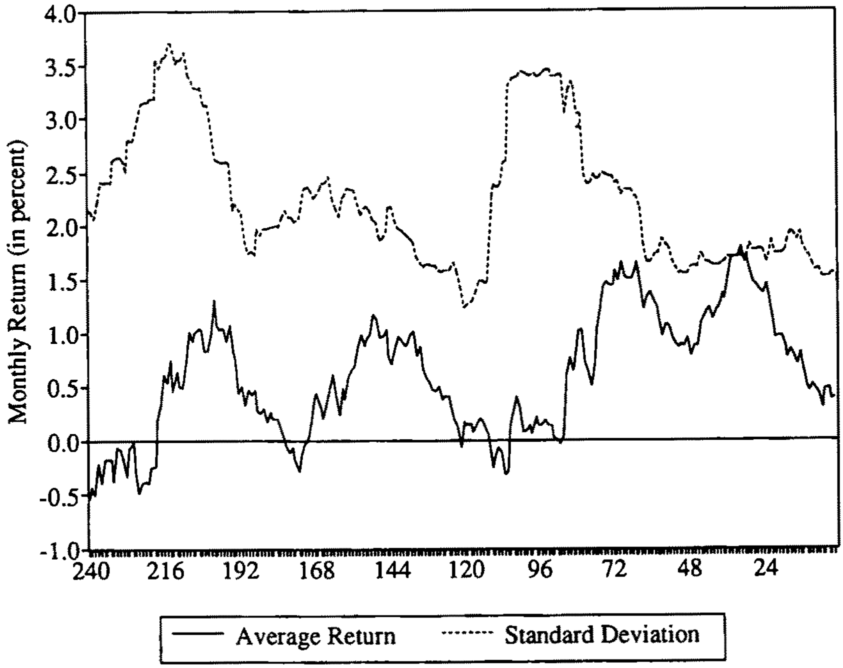

Fig. 26. Moving average results for government bonds.

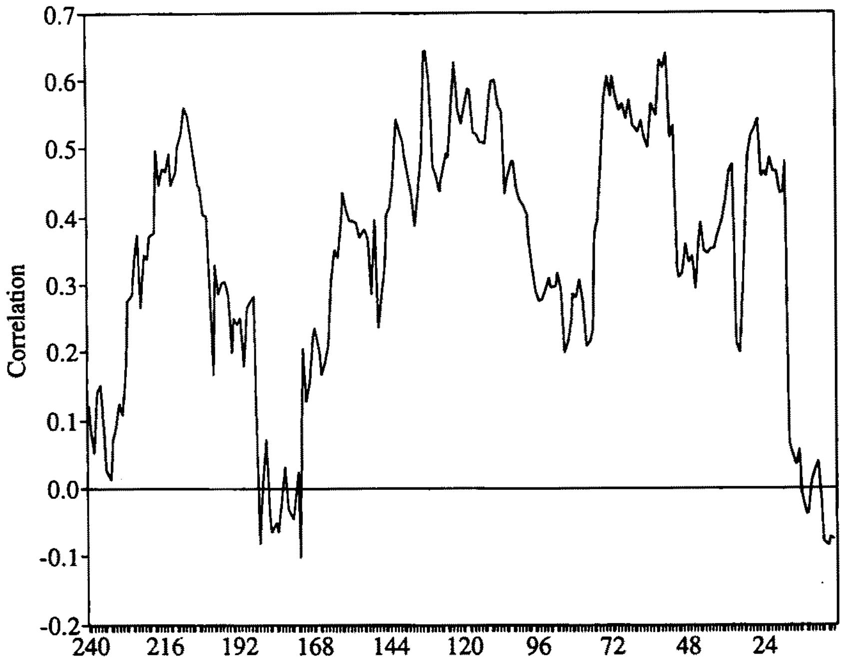

Fig. 27. Moving average correlation of stock and bond returns.

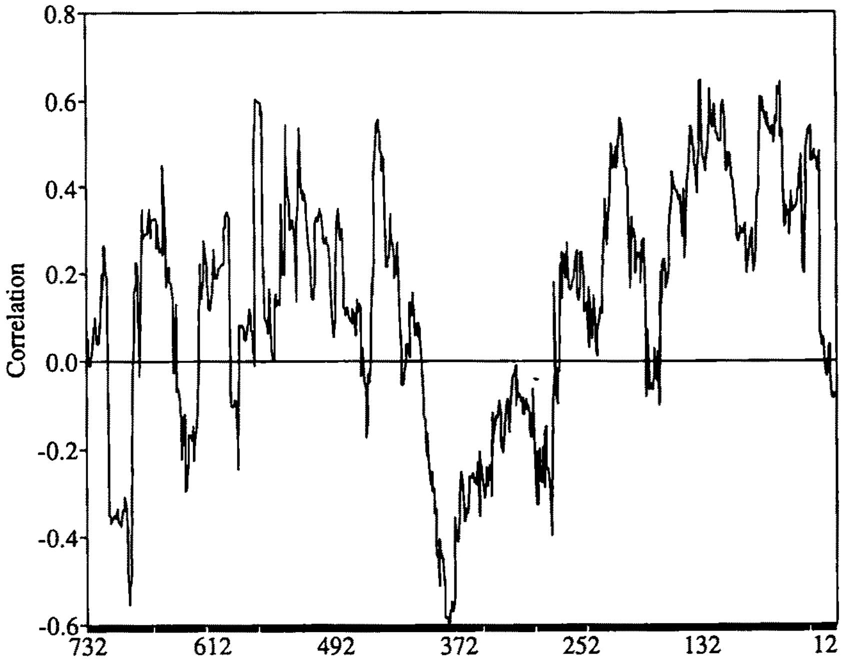

Fig. 28. Moving average correlation of returns (complete data).

the patterns statistically significant? A nonparametric test for detecting nonstationary data is described in the appendix. The test attempts to measure the degree of nonstationarity between the data in the period indexed from $1, \ldots, n/2$ and the data in the period $n/2+1, \ldots, n$. The results in table 8 are presented by type of parameter (μ, σ, and ρ) and an overall measure that combines all parameters is also given. The p-values are a quantitative measure of the nonstationarity of the data, where large p-values (i.e., values close to one) are evidence that the returns are not generated by stationary parameters.

For this data set, the results in table 8 give strong evidence of overall nonstationarity in the data if more than 360 months of data are used. Evidence of nonstationarity in the standard deviations is strong for n at least 240.

The p-values are not exceptionally large for $n = 36$, which includes the market break at roughly the midpoint of the data at $n = 15$. There are several reasons why the p-values are not larger. First, the measures of nonstationarity combine all five asset classes. While common stocks and small stocks experienced an exceptionally large negative return at $n = 15$, the other assets did not. Also, the spike at $n = 15$ was only a one-time occurrence. The test for nonstationarity given in this paper is more powerful at detecting longer term consistent changes in parameters. Also, the reasoning that showed it is difficult to distinguish two securities with different mean returns also implies that it is difficult to detect nonstationarity in mean returns.

Table 8
Results of the nonparametric test for nonstationary data.

n	p-values			
	Mean	Std. Dev.	Correlation	Overall
12	0.89	0.21	0.40	0.51
24	0.21	0.88	0.75	0.69
36	0.59	0.13	0.56	0.44
48	0.51	0.42	0.83	0.76
60	0.61	0.74	0.13	0.25
72	0.28	0.69	0.04	0.04
84	0.39	0.46	0.22	0.16
96	0.21	0.57	0.61	0.47
108	0.49	0.60	0.73	0.71
120	0.78	0.64	0.57	0.71
180	0.89	0.80	0.05	0.44
240	0.82	0.97	0.48	0.89
300	0.73	0.90	0.40	0.72
360	0.88	0.99	0.44	0.93
420	0.73	1.00	0.78	0.99
480	0.89	1.00	1.00	1.00
540	0.78	1.00	1.00	1.00
600	0.68	0.99	1.00	1.00

While nonstationarity appears to be evident from figs. 25 and 26, it does not appear to be statistically significant until $n = 240$. The test for nonstationarity divides the data into two equal parts. Roughly this means that parameters generating returns in the 1970's were different than parameters in the 1980's. Although these results apply to the particular data set being analyzed, the same procedure can be applied to any data set. Clearly, the results for other data sets could be very different than these results.

The amount of historical data to use in estimating parameters depends on many factors, including the number of securities in the analysis, whether or not parameter estimates are improved over historical estimates, and the preferences of the investor. The results of this section indicate the likelihood of nonstationarity over different time horizons in one data set. The results of the previous sections indicate how estimation error affects mean-variance results over different time horizons.

9. Conclusions

The mean-variance model for portfolio selection requires the specification of many input parameters. The problem is to determine the parameters that apply for the *next* planning period. Using a limited amount of historical data exposes the model to *estimation errors* in the parameters. Using a large amount of historical data increases the likelihood of nonstationarity in the parameters. In this case, the estimated parameters may not be representative of the true parameters for the next period.

This paper analyzed the effect of estimation error through the use of *simulation*. An important distinction was made between the *true* efficient frontier, the *estimated* frontier, and the *actual* frontier. The problem of stationarity in historical data was investigated with a nonparametric statistical test.

Although the exposition in this paper focused on specific examples, several conclusions appear to hold generally. The error in computing efficient frontiers caused by using parameters estimated with historical data can be *very large*, even if the underlying parameters are stationary. The results vary depending on the utility parameter t, but errors are generally larger for maximum return portfolios than for minimum variance portfolios.

The *greatest reduction in errors* in mean-variance analysis can be obtained by *improving historical estimates of mean returns* of the securities. The errors in estimates of mean returns using historical data are so large that these parameter estimates should always be used with caution. Improvements in estimates of standard deviations and correlations have much less effect. One recommendation for practitioners is to use historical data to estimate standard deviations and correlations but use a model to estimate mean returns.

Errors in the estimated frontier, relative to the true efficient frontier, are larger than are errors in the actual frontier. Due to the error maximization property of mean-variance analysis, *points on the estimated frontier are optimistically biased predictors of actual portfolio performance*. A point on the estimated frontier will, on average, have a larger mean and a smaller standard deviation than the corresponding point on the actual frontier. These effects become more pronounced as the number of securities increases. It is important for practitioners to recognize and correct for this phenomenon in order to develop more realistic expectations of the future performance of a portfolio. This paper suggested a method for adjusting for this bias.

Appendix

Define

m is the number of securities.

μ is the vector of true mean returns, i.e., $\mu = (\mu_1, \ldots, \mu_m)$.

σ is the vector of true standard deviations of returns, i.e., $\sigma = (\sigma_1, \ldots, \sigma_m)$.

ρ is the matrix of true correlations of returns of the securities, i.e., ρ_{jk} is the true correlation of returns of securities j and k, for $j, k = 1, \ldots, m$.

Mean-variance efficient portfolios are computed by solving the following quadratic program:

$$\text{(MV)} \qquad \max_{x_1,\ldots,x_m} \sum_{j=1}^{m} x_j\mu_j - t\left(\sum_{j=1}^{m} x_j^2\sigma_j^2 + 2\sum_{j=1}^{m}\sum_{k=j+1}^{m} x_j x_k \rho_{jk}\sigma_j\sigma_k\right)$$

$$\text{subject to: } \sum_{j=1}^{m} x_j = 1,$$

$$x_j \geq 0, \quad j = 1, \ldots, m.$$

Denote the optimal solution to (MV) by $x(t) = MV(\mu, \sigma, \rho, t)$, where $x(t) = (x_1(t), \ldots, x_m(t))$. That is, $x(t)$ is the vector of optimal portfolio weights. Corresponding to this optimal solution is an expected portfolio return of $\mu_P(t) = \sum_{j=1}^{m} x_j(t)\mu_j$ and a portfolio variance of $\sigma_P^2(t) = \sum_{i=1}^{m} x_j^2(t)\sigma_j^2 + 2\sum_{j=1}^{m}\sum_{k=j+1}^{m} x_j(t)x_k(t)\rho_{jk}\sigma_j\sigma_k$.

TRUE, ESTIMATED, AND ACTUAL FRONTIERS DEFINED

The *true efficient frontier* is the curve $(\mu_P(t), \sigma_P(t))$ for values of t ranging from zero to infinity. In practice, $t = 10,000$ is sufficient to solve for the minimum variance portfolio.

Let $r_1, \ldots, r_n$ represent vectors of security returns over n time periods. For example, $r_i = (r_{i1}, \ldots, r_{im})$ are the returns of the m securities at time period i. These returns could represent real historical data or simulated data. Expected returns, standard deviations, and correlations can be estimated by the usual sample estimates:

$$\hat{\mu}_j = \frac{1}{n}\sum_{i=1}^{n} r_{ij}, \quad \hat{\sigma}_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(r_{ij} - \hat{\mu}_j)^2, \quad \hat{\rho}_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(r_{ij} - \hat{\mu}_j)(r_{ik} - \hat{\mu}_k).$$

Estimated parameters are distinguished with hats. Denote the solution of (MV) with estimated parameters by $\hat{x}(t) = MV(\hat{\mu}, \hat{\sigma}, \hat{\rho}, t)$, where $\hat{x}(t) = (\hat{x}_1(t), \ldots, \hat{x}_m(t))$. The corresponding estimated expected portfolio return is $\hat{\mu}_P(t) = \sum_{j=1}^{m} \hat{x}_j(t)\hat{\mu}_j$ and the estimated portfolio variance is $\hat{\sigma}_P^2(t) = \sum_{j=1}^{m} \hat{x}_j^2(t)\hat{\sigma}_j^2 + 2\sum_{j=1}^{m}\sum_{k=j+1}^{m} \hat{x}_j(t)\hat{x}_k(t)\hat{\rho}_{jk}\hat{\sigma}_j\hat{\sigma}_k$. The *estimated frontier* is the curve $(\hat{\mu}_P(t), \hat{\sigma}_P(t))$ for values of t ranging from zero to infinity.

For each value of t, the actual portfolio expected return is $\tilde{\mu}_P(t) = \sum_{j=1}^{m} \hat{x}_j(t)\mu_j$ and the actual portfolio variance is given by $\tilde{\sigma}_P^2(t) = \sum_{j=1}^{m} \hat{x}_j^2(t)\sigma_j^2 + 2\sum_{j=1}^{m}\sum_{k=j+1}^{m} \hat{x}_j(t)\hat{x}_k(t)\rho_{jk}\sigma_j\sigma_k$. The actual portfolio mean and variance involve portfolio weights derived from estimated parameters, together with the true security means, standard deviations, and correlations. The *actual frontier* is the curve $(\tilde{\mu}_P(t), \tilde{\sigma}_P(t))$ for values of t ranging from zero to infinity.

The true optimal portfolio, $x(t)$, will differ from the estimated optimal portfolio, $\hat{x}(t)$. Jorion [9] essentially studies the differences in these portfolios to define "statistically equivalent portfolios".

DETAILS OF THE SIMULATION PROCEDURE

Suppose that the returns of m securities follow a multivariate normal distribution with a vector of mean returns μ and a covariance matrix C, where $C = \sigma^{\mathrm{T}} \rho \sigma$. The main logic of the simulation is summarized next. The total number of trials in the simulation is S and the trials are indexed by $s = 1, \ldots, S$.

Step 1. (*Main loop*) FOR $s = 1$ to S:
Step 2. Simulate n independent returns from the multivariate normal distribution $N(\mu, C)$. Denote these returns by $r_1^s, \ldots, r_n^s$.
Step 3. Estimate the parameters $\hat{\mu}^s$, $\hat{\sigma}^s$, and $\hat{\rho}^s$.
Step 4. Compute the estimated and actual frontiers for simulation trial s. Denote these frontiers by $(\hat{\mu}_P^s(t), \hat{\sigma}_P^s(t))$ and $(\tilde{\mu}_P^s(t), \tilde{\sigma}_P^s(t))$, respectively.
Step 5. Goto step 1. (*End of main loop*)

Details of step 2, simulating returns that follow a multivariate normal distribution, are given next. First, compute the eigenvalues and eigenvectors of C and then write C as

$$C = VDV^{\mathrm{T}},$$

where V is the matrix whose columns are eigenvectors of C, D is the diagonal matrix of eigenvalues of C, and V^{T} is the transpose of V. The eigenvalue matrix is normalized so that the length of each eigenvector is one, i.e., $VV^{\mathrm{T}} = I$, where I is the identity matrix.

Let $D^{1/2}$ denote the diagonal matrix whose elements are the square-roots of the eigenvalues. Define the linear transformation from the vector of random variables z to r by

$$r = VD^{1/2}z = Fz + \mu,$$

where $F = VD^{1/2}$. Suppose that the random variables z satisfy $E(z) = 0$ and $Cov(z) = I$. The covariance matrix of z equal to the identity matrix means that z_j has standard deviation 1 for all j and z_j and z_k are uncorrelated for all $j \neq k$. Under these assumptions, $E(r) = \mu$ and

$$Cov(r) = Cov(Fz) = FCov(z)F^{\mathrm{T}} = (VD^{1/2})I(VD^{1/2})^{\mathrm{T}} = VDV^{\mathrm{T}} = C.$$

That is, if returns r are generated by $Fz + \mu$ where $E(z) = 0$ and $Cov(z) = I$, then the resulting returns will satisfy $E(r) = \mu$ and $Cov(r) = C$. Furthermore, if the z's are normally distributed, then r will also be normally distributed.

This reduces the problem of simulating multivariate normally distributed returns r to the problem of simulating independent and identically distributed standard normal random variables z. Many programming languages have built-in routines for the latter problem.

DESCRIPTION OF THE NONPARAMETRIC TEST FOR NONSTATIONARY DATA

To begin, measures of nonstationarity are defined. Let $r_1, \ldots, r_n$ represent vectors of historical security returns over n time periods. As before, $r_i = (r_{i1}, \ldots, r_{im})$ are the returns of the m securities at time period i. Divide the returns into two equal groups. To simplify the discussion, suppose that n is even. Then group A consists of $r_1, \ldots, r_{n/2}$ and group B consists of $r_{n/2+1}, \ldots, r_n$. Denote the sample estimates of means, standard deviations, and correlations of each group by $\hat{\mu}_j(A)$, $\hat{\mu}_j(B)$, $\hat{\sigma}_j(A)$, $\hat{\sigma}_j(B)$, $\hat{\rho}_{jk}(A)$, and $\hat{\rho}_{jk}(B)$. If the returns are stationary, then $\hat{\mu}_j(A)$ should not be significantly different than $\hat{\mu}_j(B)$, $\hat{\sigma}_j(A)$ should not be significantly different than $\hat{\sigma}_j(B)$, etc. Define the following measures of nonstationarity

$$h_\mu = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{\mu}_j(A) - \hat{\mu}_j(B))^2},$$

$$h_\sigma = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{\sigma}_j(A) - \hat{\sigma}_j(B))^2},$$

$$h_\rho = \sqrt{\frac{1}{(m^2 - m)/2}\sum_{j=1}^{m}\sum_{k=j+1}^{m}(\hat{\rho}_{jk}(A) - \hat{\rho}_{jk}(B))^2}.$$

Large values of h indicate nonstationarity in the corresponding parameter. The problem is to determine how large a value of h is statistically significant. In other words, under the assumption of stationarity, what is the probability that a value as large as h would be observed?

To answer this question, we apply a standard nonparametric testing method. The idea is to randomly reorder the returns, divide the returns into two groups, estimate parameters, and compute measures of nonstationarity. If the returns are stationary, then the new h-measures will not be significantly different than the original h-values. This process is then repeated for many random orderings of the data to determine the empirical distribution of the h-values.

To make this idea more precise, let S be the total number of random orderings generated, indexed by $s = 1, \ldots, S$. For trial s, let π be a random permutation of $(1, 2, \ldots, n)$. The dependence of π on s will be suppressed for convenience. Divide the returns into two groups, where group A consists of

$r_{\pi(1)}, \ldots, r_{\pi(n/2)}$ and group B consists of $r_{\pi(n/2+1)}, \ldots, r_{\pi(n)}$. Denote the sample estimates of means, standard deviations, and correlations by $\hat{\mu}_j^s(A)$, $\hat{\mu}_j^s(B)$, etc. As above, define the measures of nonstationarity h_μ^s, h_σ^s, and h_ρ^s for the reordered data.

Suppose that h_μ, corresponding to the original order of the data, is larger than a fraction p of the h_μ^s-values. If the returns are stationary, then p is an estimate of the probability of observing a value of h_μ or larger. (Note that p is only an estimate, because S is a sample out of the total $n!$ permutations of the original data.) The same procedure is applied to determine the p-values corresponding to the parameters σ and ρ.

The measures of nonstationarity can be combined into an overall measure as follows. Define h_{OV} by

$$h_{OV} = h_\mu + h_\sigma + ch_\rho. \tag{3}$$

The constant c is chosen so that the variability of the h_ρ-values does not dominate the variability in the h_μ-values or the h_σ-values. A reasonable choice for c is 0.20 when using monthly data. Using this overall measure of nonstationarity, an overall p-value can be determined as before.

The p-values computed in this way indicate the probability that the returns are nonstationary. Large p-values are inconsistent with the hypothesis that the returns are stationary. More precisely, under the hypothesis that returns are generated by stationary parameters, p is an estimate of the probability of observing a value of the nonstationarity measure as large or larger than h.

Nonstationarity can occur in many ways. It can occur in any of the parameters and at any time in the data. For this reason, it is impossible to design a single test that is powerful at detecting all kinds of nonstationarity. The test in this paper was designed to be powerful at detecting nonstationarity when the parameters change *halfway* through the data set. Similar tests for nonstationarity can be developed by dividing the data into more groups.

References

[1] M. Best and R. Grauer, On the sensitivity of mean-variance efficient portfolios to changes in asset means: Some analytical and computational results, Rev. Fin. Stud. 4 (1991) 315.

[2] V.K. Chopra, Mean-variance revisited: Near-optimal portfolios and sensitivity to input variations, J. Investing (1993), to appear.

[3] V.K. Chopra, C.R. Hensel and A.L. Turner, Massaging mean-variance inputs: Returns from alternative global investment strategies in the 1980s, Manag. Sci. (1993), to appear.

[4] V.K. Chopra and W.T. Ziemba, The effect of errors in means, variances, and covariances on optimal portfolio choice, J. Portfolio Manag. (Winter 1993) 6.

[5] E.J. Elton and M.J. Gruber, *Modern Portfolio Theory and Investment Analysis*, 4th ed. (Wiley, New York, 1991).

[6] J.D. Jobson, Confidence regions for the mean-variance efficient set: An alternative approach to estimation risk, Rev. Quant. Fin. Accounting 1 (1991) 293.

[7] J.D. Jobson and B. Korkie, Putting Markowitz theory to work, J. Portfolio Manag. (Summer, 1981) 70.
[8] P. Jorion, Bayes–Stein estimation for portfolio analysis, J. Fin. Quant. Anal. 21 (1986) 279.
[9] P. Jorion, Portfolio optimization in practice, Fin. Anal. J. (Jan.–Feb. 1992) 68.
[10] J.G. Kallberg and W.T. Ziemba, Mis-specification in portfolio selection problems, in: *Risk and Capital*, eds. G. Bamberg and A. Spreman, Lecture Notes in Economics and Mathematical Sciences 227 (Springer, New York, 1984).
[11] H.M. Markowitz, *Mean-Variance Analysis in Portfolio Choice and Capital Markets* (Basil Blackwell, New York, 1987).
[12] R. Michaud, The Markowitz optimization enigma: is "optimized" optimal?, Fin. Anal. J. (Jan.–Feb. 1989) 31.
[13] W.F. Sharpe, *AAT: Asset Allocation Tools*, 2nd ed. (Scientific Press, South San Francisco, CA, 1987).