Portfolio Resampling: Review and Critique

Bernd Scherer

A well-understood fact of asset allocation is that the traditional portfolio optimization algorithm is too powerful for the quality of the inputs. Recently, a new concept called "resampled efficiency" has been introduced into the asset management world to deal with estimation error. The objective of this article is to describe this new technology, put it into the context of established procedures, and point to some peculiarities of the approach. Even though portfolio resampling is a thoughtful heuristic, some features make it difficult to interpret by the inexperienced.

*P*ortfolio optimization suffers from error maximization.[1] Because inputs into the efficient frontier algorithm are measured with error, the optimizer tends to pick those assets with the most attractive features (high returns and low risks and/or low correlations) and to short or deselect those with the worst features. These extremes are exactly the cases in which estimation error is likely to be highest; hence, the process maximizes the impact of estimation error on portfolio weights. If, for example, assets have high correlations, they appear to the quadratic programming algorithm to be similar, but an algorithm that takes point estimates as inputs and treats them as if they were known with certainty will react to tiny return differences that are well within measurement error.[2] In other words, the optimization algorithm is too powerful for the quality of the inputs. This problem does not necessarily stem from the mechanism itself; it calls for a refinement of inputs. To deal with the estimation error, a concept called "resampled efficiency" has recently been introduced.[3] This article describes this new technology, puts it into the context of established procedures, and points out some peculiarities of the resampled efficiency approach.

Visualizing Estimation Error

Portfolio sampling allows an analyst to visualize the estimation error in traditional portfolio optimization methods. The estimated parameters used in asset allocation problems (typically point estimates of means, variances, and correlations) are calculated by using only one possible realization of

return history. Even if stationarity (constant mean, non-time-dependent covariances) is assumed, only in very large samples can the point estimates for risk and return inputs equal the true distribution parameters. The effect of the resulting estimation error on optimal portfolios can be captured by the Monte Carlo procedure known as portfolio resampling.[4]

Suppose we estimated both variance–covariance matrix $\hat{\Omega}_0$ and return vector $\hat{\mu}_0$ by using T observations, where Ω is a $k \times k$ covariance matrix of excess returns (asset return minus cash) and μ is a $k \times 1$ vector of average excess returns. The point estimates are random variables (because they are calculated from random returns); that is, another sample of random variables from the same distribution would result in different estimates. This situation is called sampling error.[5] How can we capture the randomness of inputs? One answer is portfolio resampling, which draws repeatedly from the return distribution. We can create a statistically equivalent sample with T observations (original data length), thereby creating a new data set for the estimation of input parameters by either drawing T times without replacement from the empirical distribution (a nonparametric method known as bootstrapping) or sampling from a multivariate normal distribution (a parametric method termed Monte Carlo simulation). Both methods yield virtually the same results, but suppose that to deal with the sampling error, we have chosen the parametric method.[6]

By repeating the sampling procedure n times, we get n new sets of optimization inputs ($\hat{\Omega}_1$, $\hat{\mu}_1$ to $\hat{\Omega}_n$, $\hat{\mu}_n$). For each of these inputs, we can now calculate a new frontier spanning from the minimum-variance portfolio to the maximum-return portfolio. We calculate m portfolios along the frontier and save the corresponding allocation vectors, $\mathbf{w}_{11}, \ldots, \mathbf{w}_{1m}$

Bernd Scherer is the European head of advanced applications at Deutsche Asset Management, Frankfurt am Main.

to $\mathbf{w}_{n1}, \ldots, \mathbf{w}_{nm}$. Evaluating all $n \times m$ portfolios with the original optimization inputs ($\hat{\mathbf{\Omega}}_1, \hat{\mathbf{\mu}}_1$) will force all portfolios to plot below the original efficient frontier. The reason is that no weight vector that is optimal for $\hat{\mathbf{\Omega}}_i$ and $\hat{\mathbf{\mu}}_i$ ($i = 1, \ldots, n$) can be optimal for $\hat{\mathbf{\Omega}}_0, \hat{\mathbf{\mu}}_0$. Hence, because the weights have been derived from data containing estimation error, all portfolio weights lead to portfolios plotting below the efficient frontier. Estimation error in inputs is transformed into uncertainty about the optimal allocation vector.

The mechanics of portfolio resampling are best illustrated through a practical example. Suppose we download 18 years of data and calculate historical means and covariances to obtain the inputs shown in **Table 1**.[7] For these data, running a standard mean–variance optimization (i.e., minimizing portfolio risk subject to a return constraint) in which the returns vary from the return of the minimum-variance portfolio to the return of the maximum-return portfolio results in the asset allocations along the efficient frontier shown in **Figure 1**.[8] For this example, we calculated $m = 25$ portfolios, dividing the return difference between the minimum and maximum return into 25 steps.

As most investors familiar with traditional portfolio optimization would have guessed, the resulting allocations are very concentrated; some assets did not even enter the solution. Also, small changes in risk aversion may lead to widely different portfolios. For example, allocation vectors 20 and 23 are quite different in weightings. Given uncertainty about the degree of an investor's risk aversion, this feature of traditional portfolio optimization is unattractive.

Suppose, instead, we applied the resampling algorithm. In this case, each new weight vector (calculated from resampled inputs) can be interpreted as a set of statistically equivalent weights.

Only the original set of weights, $\mathbf{w}_0$, is optimal, however, for the original set of inputs ($\hat{\mathbf{\Omega}}_0, \hat{\mathbf{\mu}}_0$). All other portfolios must plot below the efficient frontier. Their weight estimates are the direct result of sampling error. **Figure 2** shows the efficient frontier and the resampled portfolios that resulted from using the described resampling technique; the circles represent the resampled portfolios, and the efficient frontier can be thought of as their envelope. The dispersion arises because of the great variation in statistically equivalent weight vectors.

Increasing the number of draws, n, forces the data points closer to the original frontier as the dispersion in inputs becomes smaller. The effect is equivalent to reducing (eliminating) sampling error. But the analysis does not tell us where the new frontier lies—which leads to the next section.

Resampled Efficiency. The method that uses resampled efficiency is intended to deal with the estimation error demonstrated in the previous section. Portfolios along the so-called resampled frontier are defined as ". . . averages of the rank associated mean–variance efficient portfolios" (Michaud 1998, p. 50). Portfolios that carry Rank 1 are the minimum-variance portfolios; portfolios that carry rank m are the maximum-return portfolios. Each other portfolio gets a rank in between that depends on where its expected return ranks. The distance between the minimum-variance and the maximum-return portfolio is equally split.

Averaging maintains an important portfolio characteristic: The weights sum to 1 (even in the case of constraints). This characteristic is probably the main practical justification for this procedure, but keep in mind that there is no economic rationale derived from the optimizing behavior of rational agents that supports this method. Hence, it is a heuristic.

Table 1. Historical Means and Covariances for Portfolio Resampling

Asset	Covariance, $\mathbf{\Omega}_0$								Mean, $\mathbf{\mu}_0$
Equity									
Canadian	30.25	15.85	10.26	9.68	19.17	16.79	2.87	2.83	0.39
French	15.85	49.42	27.11	20.79	22.82	13.30	3.11	2.85	0.88
German	10.26	27.11	38.69	15.33	17.94	9.10	3.38	2.72	0.53
Japanese	9.68	20.79	15.33	49.56	16.92	6.66	1.98	1.76	0.88
U.K.	19.17	22.82	17.94	16.92	36.12	14.47	3.02	2.72	0.79
U.S.	16.79	13.30	9.10	6.66	14.47	18.49	3.11	2.82	0.71
Bonds									
U.S.	2.87	3.11	3.38	1.98	3.02	3.11	4.04	2.88	0.25
European	2.83	2.85	2.72	1.76	2.72	2.82	2.88	2.43	0.27

Source: Michaud (1998, pp. 17, 19).

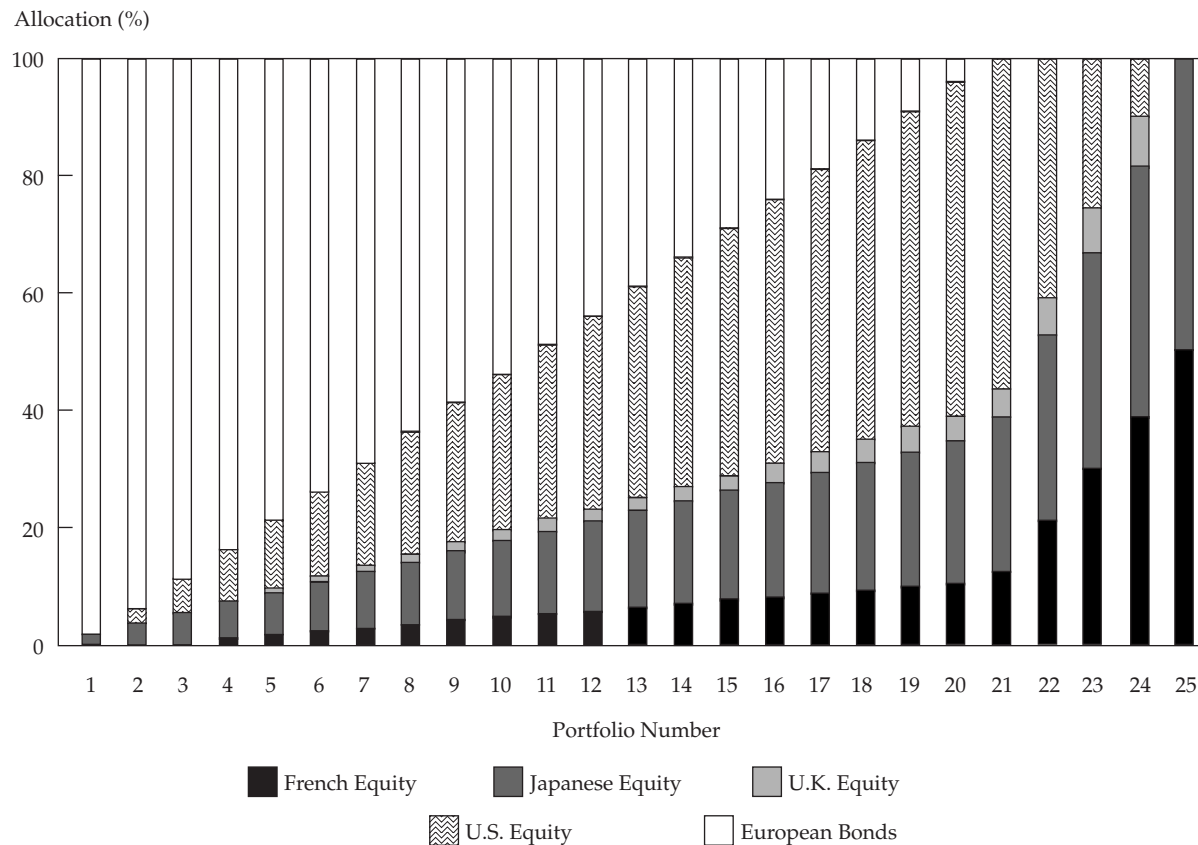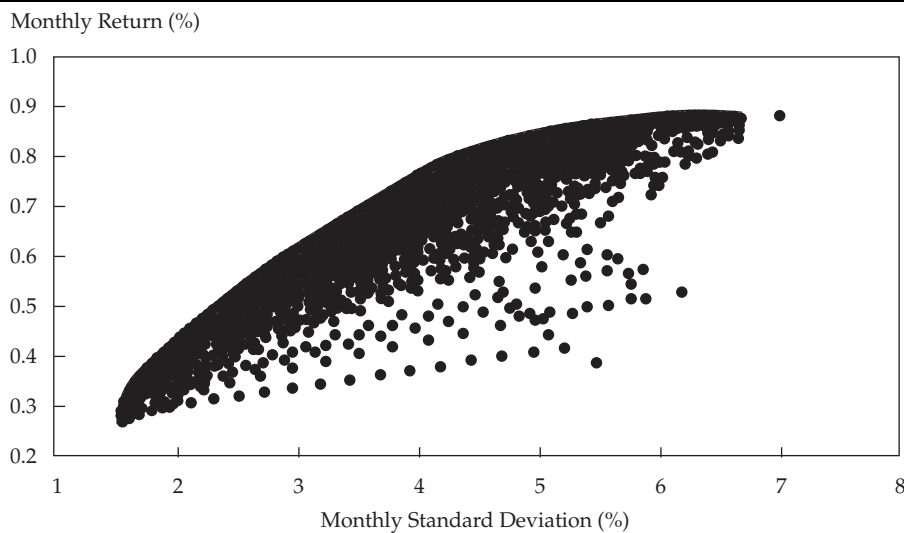**Figure 1. Mean–Variance-Efficient Portfolios: Composition**

Allocation (%)

Portfolio Number

■ French Equity   ▨ Japanese Equity   ▨ U.K. Equity

▨ U.S. Equity   □ European Bonds

Figure 2. Efficient Frontier and Resampled Portfolios

Monthly Return (%)

Monthly Standard Deviation (%)

The resampled weight for a portfolio of rank m (portfolio number m along the frontier) is given by

$$\overline{\mathbf{w}}_m^{resampled} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{w}_{im}, \qquad (1)$$

where $\mathbf{w}_{im}$ denotes the $k \times 1$ vector of the mth portfolio along the frontier for the ith resampling.

Suppose we estimate 100 efficient frontiers (i.e., 1 complete frontier for each set of inputs). Thus, we also have 100 portfolios for each rank. We can now simply calculate the average weight for each asset over all 100 portfolios. Additionally, we can measure the dispersion of portfolio weights to appreciate how the uncertainty in inputs feeds through to the dispersion in outputs.

The procedure can be summarized as follows:

Step 1. Estimate variance–covariance matrix and mean vector of historical inputs. (As an alternative, the inputs could be prespecified.)

Step 2. Resample from inputs (created in Step 1) by taking T draws from input distribution. The number of draws reflects the degree of uncertainty in the inputs. Calculate new variance–covariance matrix from sampled series. Estimation error will result in matrixes that are different from those obtained in Step 1.

Step 3. Calculate efficient frontier for inputs derived in Step 2. Save optimal portfolio weights for m equally distributed return points along the frontier.

Step 4. After repeating Steps 2 and 3 many times, calculate average portfolio weights for each return point. Evaluate frontier of averaged portfolios with variance–covariance matrix from Step 1 to plot the resampled frontier.

Instead of adding up portfolios that share the same rank, we could add up portfolios that show the same risk–return trade-off. It can be easily done by maximizing $U = \mu - 0.5\lambda_m \sigma^2$ for varying risk aversion λ_m and then averaging the λ_m-associated portfolios. Utility-sorted portfolios are theoretically preferable because they indicate risk–return trade-offs an investor with a given risk aversion would actually choose if required to make a choice repeatedly in different risk–return environments.

The resampled portfolios shown in **Figure 3** reflect greater diversification (more assets have entered the solution) than the classic mean–variance-efficient portfolios illustrated in Figure 1. They also exhibit less-sudden shifts (smooth transitions) in allocations as return requirements change. In the eyes of many practitioners, both characteristics are desirable properties.

Because of the apparent overdiversification (relative to return forecasts), the resampled frontier will have different weight allocations from those of the traditional frontier and, as shown in **Figure 4**,

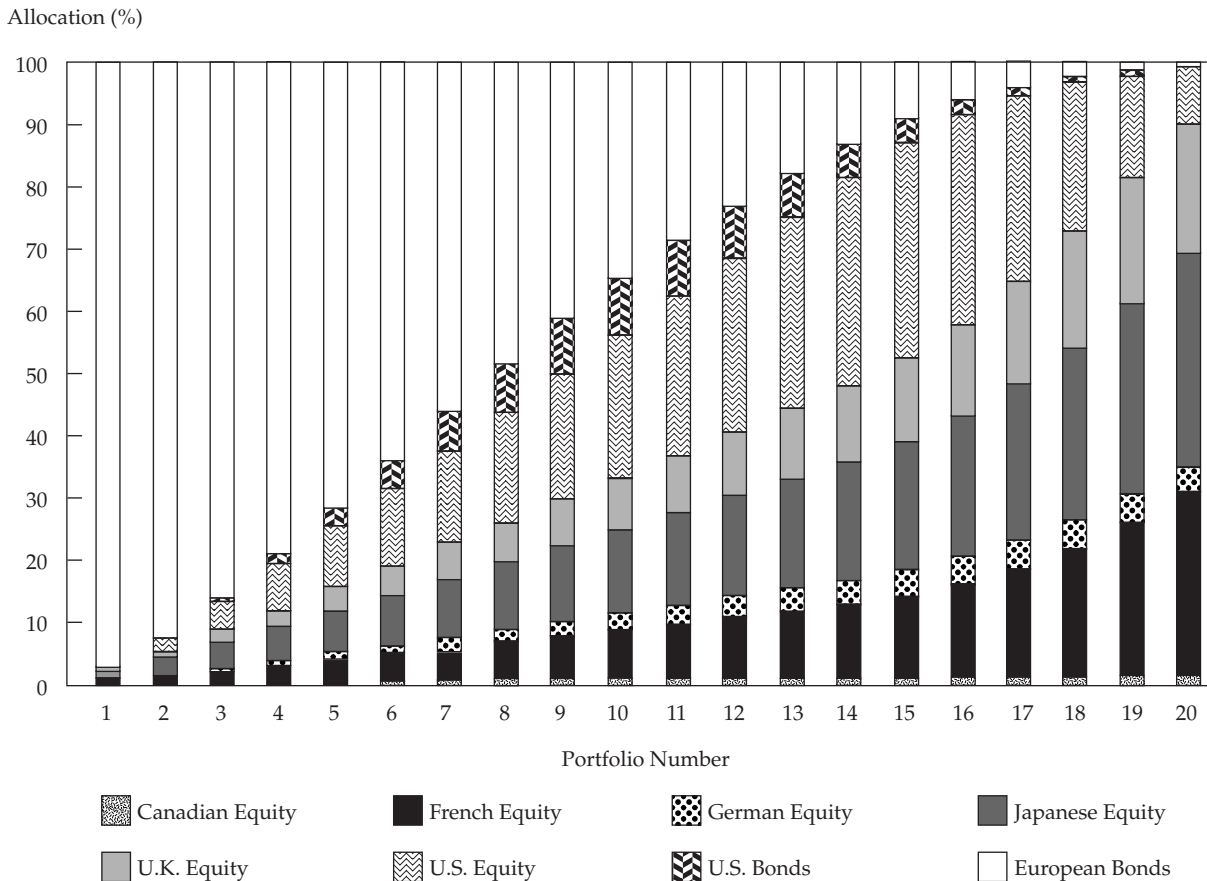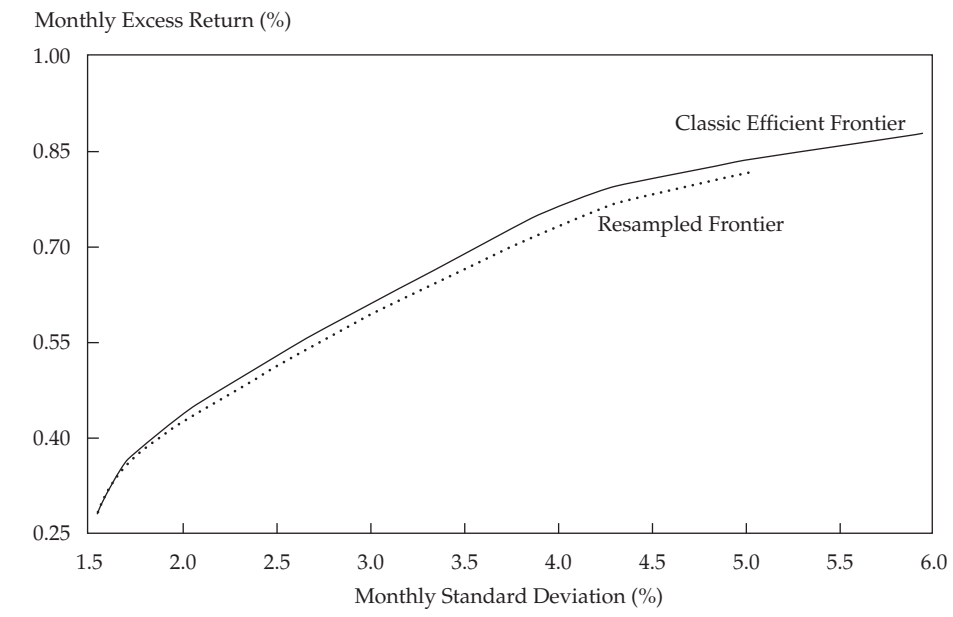**Figure 3. Resampled Portfolios: Composition**

Figure 4. Classic Efficient Frontier versus Resampled Efficient Frontier

Monthly Excess Return (%)

Monthly Standard Deviation (%)

will plot below the traditional frontier. The resampled frontier does not reach the same maximum return as the traditional frontier because of the resampled frontier's greater diversification. Whereas the maximum-return solution in the traditional frontier is made up of 100 percent investment in the highest-return asset, the averaging process prohibits this kind of solution for the resampled frontier. The frontiers are similar in risk–return space but quite different in "weight space," as Figure 3 shows.

One of the problems with using the average criterion can be illustrated by close inspection of the distribution of resampled weights for a particular rank-associated portfolio. Consider Portfolio Rank 12. From Figure 3, we can judge that Portfolio 12 has an average allocation to U.S. equities of about 23 percent. If we look at the distribution of resampled U.S. equity weights for portfolios ranked 12 as shown in **Figure 5**, however, we find that in most of the runs (more than 500 out of 1,000), the actual weight given was 0–5 percent. The average of 23 percent seems to be heavily influenced by a few "lucky" draws (that is, a barbell structure) that led to significant allocation to U.S. equity. Indeed, the 20–25 percent bin is sparsely populated. I will

Figure 5. Distribution of Resampled U.S. Equity Weights for Portfolio Rank 12

Frequency (1,000 runs)

Weight Allocations (%)
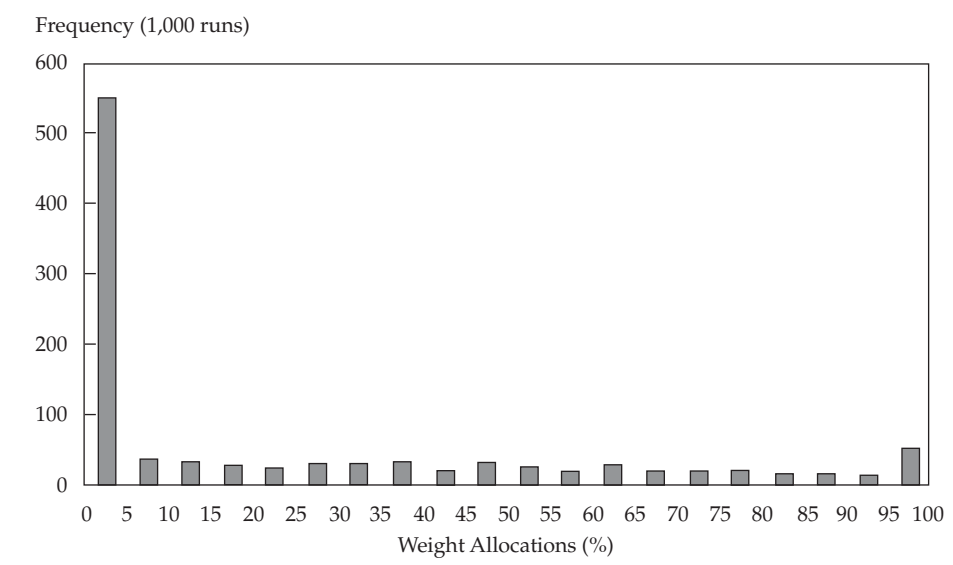
return to this point later; for now, I want to mention two additional issues related to Figure 5. First, averaging over constraint solutions will very likely result in an average allocation that is below the constrained solution because there will always be some draws where the constraint is not binding—because of the randomness of average-return inputs. Second, resampling is likely to include almost all assets in the solution because the likelihood is that at least one favorable draw will allocate to an asset.

Measuring Distance. Effectively, the resampling procedure provides the distribution of portfolio weights. We can now test whether two portfolios are statistically different. This test can be viewed as measuring the distance in a *k*-dimensional vector space. The Euclidean distance measure for the distance of a vector of portfolio weights of portfolio *i* (denoted $\mathbf{w}_i$) with portfolio *p* (denoted $\mathbf{w}_p$) is given by

$$(\mathbf{w}_p - \mathbf{w}_i)'(\mathbf{w}_p - \mathbf{w}_i), \tag{2}$$

where $\mathbf{w}_p - \mathbf{w}_i$ is equivalent to an active weight. Statistical distance, however, is computed as

$$(\mathbf{w}_p - \mathbf{w}_i)'\hat{\mathbf{\Sigma}}^{-1}(\mathbf{w}_p - \mathbf{w}_i), \tag{3}$$

where $\mathbf{\Sigma}$ is the variance–covariance matrix of portfolio weights. This test statistic is distributed as a χ^2 with degrees of freedom equal to the number of assets.[9]

For example, suppose we have two assets, each with 10 percent mean and 20 percent volatility. Suppose further that correlation between the assets

is 0.0 and the risk aversion coefficient is 0.2. Then, the optimal solution without estimation error is

$$
\begin{aligned}
\mathbf{w}^* &= \begin{pmatrix} w_1{}^* \\ w_2{}^* \end{pmatrix} \\[4pt]
&= \lambda \mathbf{\Omega}^{-1} \mathbf{\mu} \\[4pt]
&= 0.2 \begin{bmatrix} \frac{1}{(0.2)^2} & 0 \\ 0 & \frac{1}{(0.2)^2} \end{bmatrix} \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} \\[4pt]
&= \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.
\end{aligned}
$$

Now, we calculate the optimal portfolios without adding up constraints. These portfolios, by definition, do not require holdings to add up to 1, in which case, resampling would make no sense because all resampled weights would plot on a straight line (from 100 percent Weight 1 to 100 percent Weight 2). **Figure 6** shows the results of these calculations. We might be tempted to conclude that these black circles are not portfolios (because the assets do not add up), but cash can be considered a third (filling) asset because cash would leave marginal risks, as well as total risks, unchanged.

Although (as Figure 6 shows) the optimal solution is 50 percent for both assets, the estimated weights are scattered around this solution. Comparing the vector difference with the critical value of the χ^2 yields a measure of how statistically different a portfolio is. The ellipse in Figure 6 shows

Figure 6. Estimation Error and Portfolio Weights
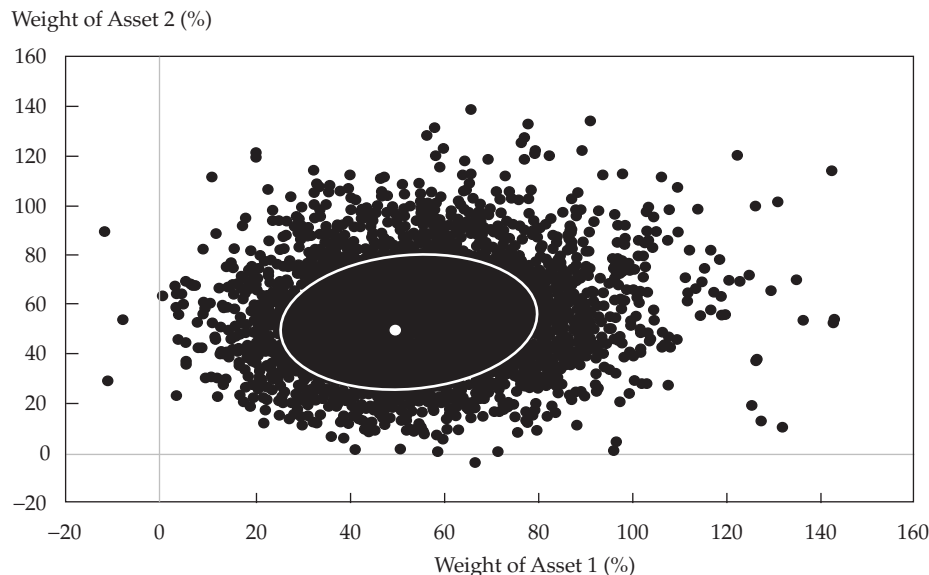
the line of constant density that is consistent with Expression 3 for the vector distance between the optimal portfolio without estimation error and its resamplings. For this two-asset example, lines of constant density can be obtained from

$$p(w_1 - w_1^*, w_2 - w_2^*) = \frac{1}{2\pi \det(\boldsymbol{\Sigma})^{1/2}} e^{-\frac{1}{2} \begin{bmatrix} w_1 - w_1^* \\ w_2 - w_2^* \end{bmatrix} \boldsymbol{\Sigma}^{-1} \begin{bmatrix} w_1 - w_1^* \\ w_2 - w_2^* \end{bmatrix}},$$

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} 27.93 & 0.005 \\ 0.005 & 27.76 \end{bmatrix}$$

where "det" stands for determinant of the matrix in brackets.

Portfolios within this ellipse would be treated as statistically equivalent, whereas portfolios outside this ellipse would be treated as significantly different in portfolio weights. Given this information, we can now visually inspect a new portfolio (new information about markets) to decide whether it is different enough from the current portfolio (after accounting for noise in inputs) to be implemented. In that respect, we could interpret the area within the ellipse as a no-trade zone.

Introducing long-only constraints (truncating weights at zero), however, invalidates the normality assumption for the distribution of portfolio weights. Michaud (1998) used a different distance measure, one that is widely applied in asset management, in recognition that two portfolios with the same risk and return might actually exhibit different allocations. For this approach, the distance between two portfolios is defined as

$$(\mathbf{w}_p - \mathbf{w}_i)' \hat{\boldsymbol{\Omega}}_0 (\mathbf{w}_p - \mathbf{w}_i), \tag{4}$$

which is equivalent to the squared tracking error. The procedure runs as follows.

Step 1. Define portfolio to test difference against. Calculate Expression 4 for all resampled portfolios.

Step 2. Sort portfolios by tracking error in descending order (highest on top).

Step 3. Define TE_α as the critical tracking error for the α percent level (i.e., if 1,000 portfolios are resampled and the critical level is 5 percent, then look at the tracking error of a portfolio that is 50th from the top). Hence, all portfolios for which

$$(\mathbf{w}_p - \mathbf{w}_i)' \hat{\boldsymbol{\Omega}}_0 (\mathbf{w}_p - \mathbf{w}_i) \geq TE_\alpha^2 \tag{5}$$

are labeled statistically different.

Step 4. As a last step, calculate the minimum and maximum allocations for each asset within the confidence region.

For a three-asset example, the uncertainty about the optimal weights can be visualized, but for higher dimensions, such visualization becomes difficult.

Note that similarity is defined with regard to the optimal-weight vector rather than in terms of risk and return. Two portfolios could be very similar in terms of risk and return but very different in allocation, which is well known because risk–return points below the frontier are not necessarily unique. Nevertheless, this test procedure is intuitive. Note also, however, that the dispersion in weights is large, so it will be difficult to reject the hypothesis that the portfolios are statistically equivalent, even if they are not. The power of the suggested test will hence be low.

Resampling and Linear Regression. Without a long-only constraint, optimal portfolios could effectively be found by using a simple regression approach because portfolio optimization in that case is a linear problem. Suppose we have k time series of excess returns—that is, total return, R, minus cash rate, c, with T observations each. Portfolio construction can then be written as a simple textbook regression (which can be estimated using any standard regression software):

$$y = \beta_1 x_1 + \ldots + \beta_k x_k + u$$
$$= 1,$$
$$x_i = R_i - r. \tag{6}$$

Using any econometrics package or Microsoft Excel, simply run a regression of 1's against all asset returns, excluding an intercept.[10] This method will force the regression through the origin in excess-return space and, therefore, maximize the Sharpe ratio.[11] The result can be interpreted as coming closest to a portfolio with zero risk (vector of 1's showing no volatility) and unit return and would thus create an arbitrage opportunity. The regression coefficients would reflect the portfolio weights ($\beta_i = w_i$), which might conveniently be rescaled to sum to 1.

Alternatively, using the same framework, we could run a constrained regression (linear constraint on portfolio weights) to create portfolios meeting particular return requirements. This framework could then be used also to test restrictions on individual regression coefficients (estimated portfolio weights) and restrictions on groups of assets and to test whether the regression coefficients are significantly different from zero.[12]

The regression framework puts a central problem of portfolio construction into a different well-known perspective: Highly correlated asset returns mean highly correlated regressors, with the obvious consequences of multicollinearity—high standard

deviations on portfolio weights (regression coefficients) and identification problems (difficulty in distinguishing between two similar assets). Simply downtesting (stepwise elimination of assets starting with the least-significant asset) and excluding insignificant assets will produce an outcome that is highly dependent on the order of exclusion, with no guidance as to where to start. This problem is familiar to both asset allocators and econometricians.

Portfolio resampling can be interpreted as a simulation approach to arrive at the distribution of weight estimates via Monte Carlo simulation of Equation 6. The center of the distribution is calculated in the same way as in portfolio resampling by averaging over the coefficient estimates for a particular asset. Instead of taking the structural form of the model as given and simulating the error term, resampling simulates a whole new data set, which is equivalent to assuming that regressors are stochastic.[13] By drawing new return data from the variance–covariance matrix and reestimating Equation 6 n times, we can calculate the average weight for asset $j = 1 \ldots k$ via averaging over the estimated regression coefficients ($\hat{\beta}_i = \hat{w}_i$):

$$\bar{\hat{w}}_j = \frac{1}{n} \sum_{i=1}^{n} \hat{w}_{ij}, j = 1 \ldots k. \tag{7}$$

Although such averaging is not necessary for portfolios without long-only constraints (because the distribution of the regressors is known), portfolio resampling is more general than the regression approach. It can also be applied in the case of long-only constraints, where the weight distribution is not known. Essentially, this approach requires bootstrapping the unknown distribution of a *t*-statistic. If an asset is, for example, included in 70 of 1,000 runs for a given rank or utility score, it will get a *p*-value of 7 percent. This approach can also be extended through the use of Bayesian analysis by using standard textbook results. In such analysis, our prior beliefs (priors) are set on the distribution of portfolio weights instead of asset returns.

Pitfalls in Portfolio Resampling

Estimation error will increase portfolio risk. This outcome has been captured in the Bayesian literature on portfolio construction. Consider the simplest case—a two-asset portfolio. In this case, any combination of the two assets will be efficient. All resampled portfolios will still plot on the efficient frontier, and no portfolio will plot below it, although the frontier in that case might be short because, sometimes, the order of assets reverses so the averaged maximum-return portfolio will not contain 100 percent of the higher-returning asset.

For example, suppose we have two uncorrelated assets with estimated volatilities of 10 percent and 15 percent. Suppose we use 60 monthly observations to estimate the frontier. Average returns over cash are 4 percent and 2 percent a year. **Figure 7** plots the resulting efficient frontiers found by a

Figure 7. Traditional versus Resampling versus Bayesian Frontiers
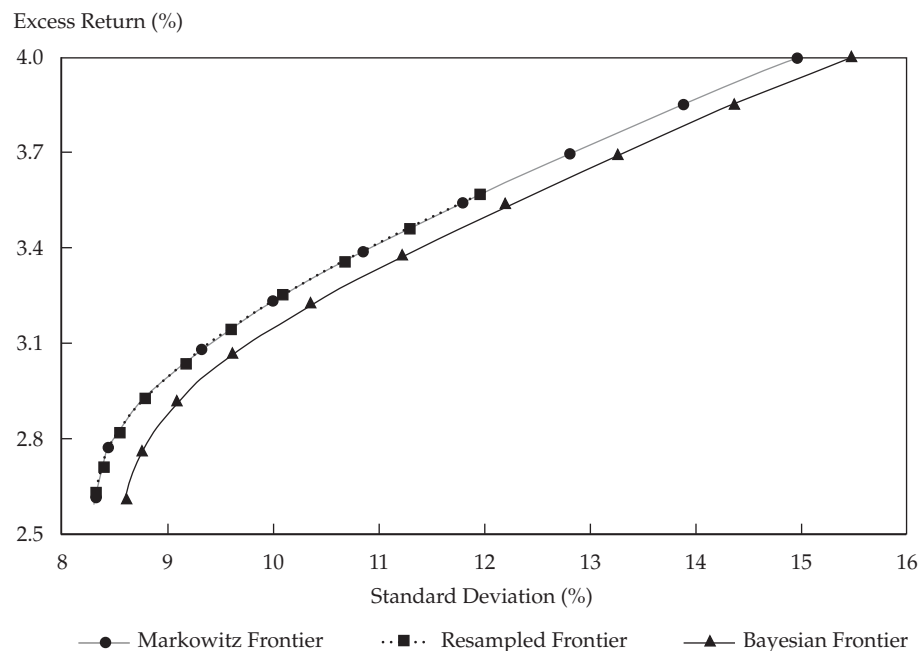
traditional, a resampled, and a Bayesian approach. The increase in risk is captured only by the Bayesian frontier. In the Bayesian method, for the same expected return (expected returns will not change with the introduction of estimation error as an uninformative prior), each portfolio exposes the investor to more risk because Bayesian methods leverage up the variance–covariance matrix but leave the return vector unchanged. In direct contrast, estimation error in the resampled frontier shows up only as a shortening of the frontier, not as an increase in risk for every return level. Instead, uncertainty about the mean causes a reduction in the maximum expected mean return, which is not plausible. Bayesian methods recognize that the exclusive use of sample information will not allow us to tackle the impact of parameter uncertainty on optimal portfolio choice, or as Nobel laureate Harry Markowitz (1987, p. 57) put it, "The rational investor is a Bayesian."

Bayesian methods can use either uninformative or informative priors. Uninformative priors simply express the possible range of parameter estimates. Adding an uninformative prior to a set of return data, therefore, increases the uncertainty about future returns but not the average outcome. Because the priors leave the set of efficient portfolios constant, this method has been little used by practitioners. Informed priors do change the average expected outcome and thus change the set of solutions. An informed prior stating that all average returns are equal will produce the minimum-variance portfolio; an informed prior stating that average returns are close to the implied returns of a benchmark portfolio will move the efficient frontier in the direction of the benchmark portfolio.

Bayesian analysis attempts to combine sample information about asset returns with priors about the return distribution. The more confident researchers are about their forecasts, the more weight they give to it. Together, the priors and the confidence level result in the predictive distribution. Optimal portfolio choice is then based on the predictive distribution. This approach is optimal according to the Neumann–Morgenstern axioms on expected utility.
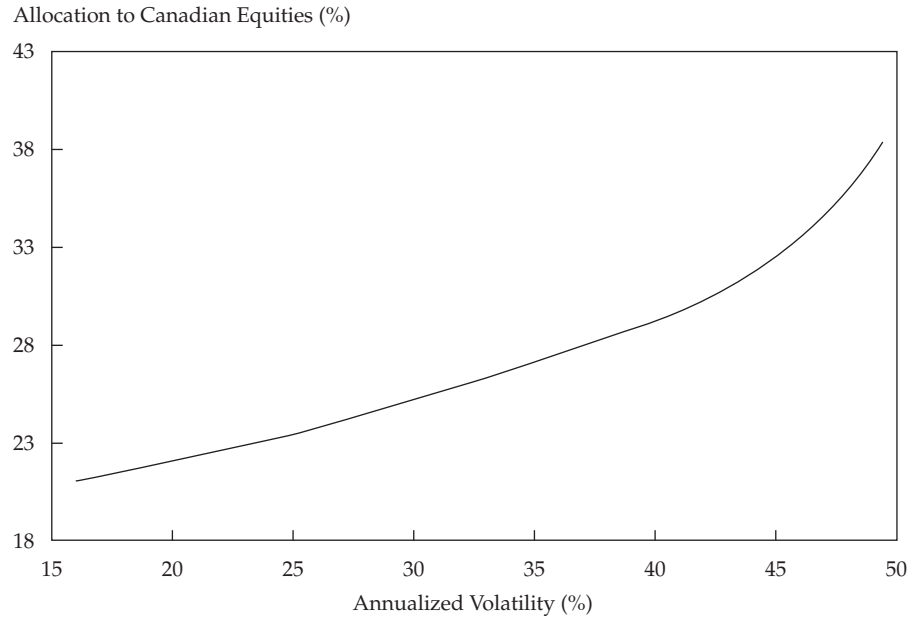
For example, suppose two assets possess the same expected return but one of them has significantly higher volatility. We could think of this example as international fixed-income allocation on a hedged basis and an unhedged basis. In this case, most practitioners (and the mean–variance optimizer) would exclude the higher-volatility asset from the solution (unless it had some desirable correlations). How would resampled efficiency deal with these assets? Repeatedly drawing

from the original distribution would produce draws for the volatile asset with highly negative returns as well as draws with highly positive returns. In the case of the highly positive returns, quadratic programming would heavily invest in this asset; in the case of the highly negative returns, the program would short the asset. Shorting is not allowed for portfolios with long-only constraints, however, so the result would be positive allocation for draws of high positive average returns and zero allocations for draws of high negative average returns.

Unconstrained optimization is different. In the classic approach, large long positions are offset (on average) by large negative positions. Consequently, an increase in volatility would yield an increase in the average allocation; hence, a worsening Sharpe ratio would be accompanied by an increase in weight. This result is not plausible. It arises directly from the averaging rule in combination with a long-only constraint, which creates an optionality for the allocation of the corresponding asset. Assets are either in or out but are never negative.
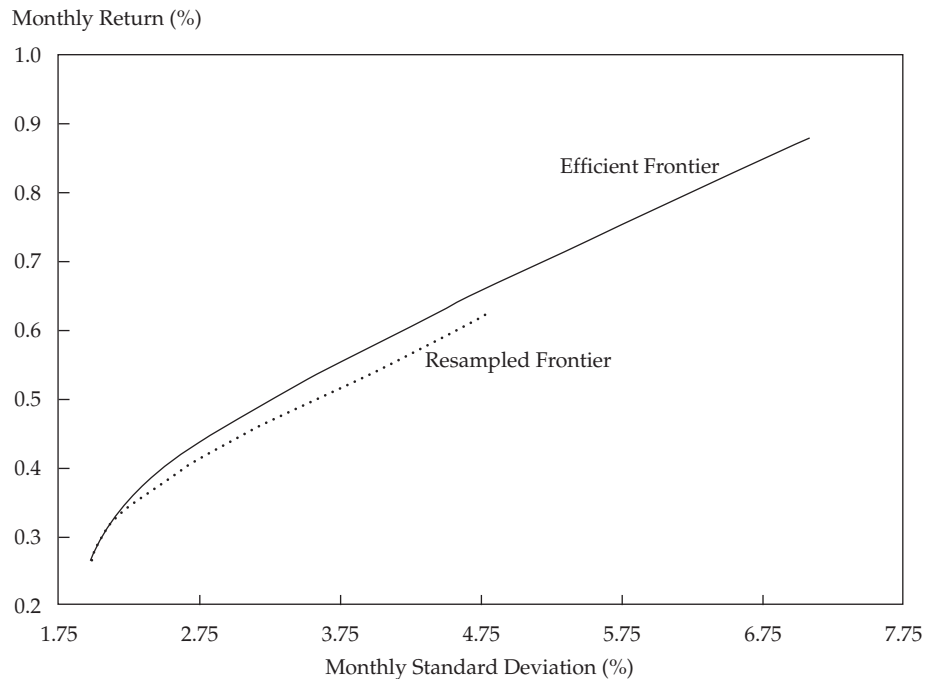
This intuitive line of reasoning can be made explicit with data on the same assets as we used previously. Suppose we are going to put together a portfolio of only equity for Canada, France, and Germany and fixed income for European bonds. Suppose further that we reduce the sample size to 60 monthly observations, which is a realistic time frame for most practical applications. We vary the volatility of only the worst-performing asset, which is Canadian equities (see Table 1). Now, consider this asset's allocation in the maximum-return portfolio. As **Figure 8** shows, even though Canadian equities have the lowest return, their allocation peaks in the maximum-return portfolio. The reason is that in portfolio resampling, as volatility rises (the Sharpe ratio deteriorates), allocations at the high-return end rise. So, a deterioration in the risk–return relationship for Canadian equities is followed by an increased weight. This result does not come from higher volatility leading to higher estimation error; the phenomenon would not arise in long–short portfolios. It results directly from averaging over long-only portfolios. The long-only constraint creates "optionality."

One of the basic properties of efficient-set mathematics is that the efficient frontier does not contain upward-bending parts.[14] An upward-bending part would imply that one could construct portfolios superior to the frontier by linearly combining two frontier portfolios. Could such a forbidden situation arise when using the concept of resampled efficiency?

Figure 8. Allocation and Volatility for Canadian Equities

Allocation to Canadian Equities (%)

(Figure: line chart. X-axis: Annualized Volatility (%) from 15 to 50. Y-axis: 18 to 43.)

For an answer, keep in mind that the difference between the resampled and the traditional efficient frontier arises because resampling provides portfolios that are too diversified. Instances can occur in resampling, however, in which diversification becomes smaller as the maximum-return solution is approached (because all maximum-return solu-

tions tend to be concentrated in the high-return asset anyway). This is exactly what has happened in the resampled frontier shown in **Figure 9**. Certainly, the true test of resampled efficiency is out-of-sample performance in a Monte Carlo study, but convex parts of an efficient frontier are difficult to justify.

Figure 9. Upward-Bending Resampled Frontier

Monthly Return (%)

(Figure: line chart showing Efficient Frontier and Resampled Frontier. X-axis: Monthly Standard Deviation (%) from 1.75 to 7.75. Y-axis: 0.2 to 1.0.)

Efficient Frontier

Resampled Frontier

Moreover, all resamplings are derived from the same vector and covariance matrix ($\hat{\mathbf{\Omega}}_0$, $\hat{\mathbf{\mu}}_0$). The true distribution, however, is unknown. Hence, all resampled portfolios will suffer from the deviation of the parameters $\hat{\mathbf{\Omega}}_0$, $\hat{\mathbf{\mu}}_0$ from $\mathbf{\Omega}_{true}$, $\mathbf{\mu}_{true}$ in much the same way. Averaging will not help greatly in this case because the averaged weights are the result of an input vector, which is itself very uncertain. Hence, it is fair to say that all portfolios inherit the same estimation error. The special importance attached to $\hat{\mathbf{\Omega}}_0$, $\hat{\mathbf{\mu}}_0$ finally limits the analysis.

Conclusion

Portfolio resampling offers an intuitive way to develop tests for the statistical difference between two portfolios (weight vectors). Resampling will thus be the methodology of choice to test for the statistical significance of two portfolios. What is not clear, however, is why averaging over resampled portfolio weights should represent an optimal portfolio construction solution to deal with estimation error. In the case of long–short portfolios, use of averaged resampled portfolios provides no improvement over traditional Markowitz solutions (in fact, the solutions—that is, the frontiers—coincide). In the case of long-only constraints, resampled efficiency leads to more-diversified portfolios, which are well known to beat Markowitz portfolios in out-of-sample tests.[15] Hence, the result presented by Michaud (1998) that resampled efficiency beats simple Markowitz portfolios out-of-sample is hardly surprising.

What is unclear is the extent to which this result can be generalized, because portfolio resampling carries with it some unwanted features. Deteriorating Sharpe ratios (caused by higher volatility) lead to increased allocation of those assets in the high-return portfolios because favorable return draws lead to large allocations whereas unfavorable draws lead to zero allocations at most ("optionality"). Additionally, the efficient frontiers may exhibit turning points (a move from concave to convex). Also interesting is that at least three assets are needed for the resampling methodology to show the increased risk present in the case of estimation error.

Finally, although the ultimate test of any portfolio construction methodology is out-of-sample performance, Markowitz efficiency is not the relevant benchmark for resampled efficiency. Bayesian alternatives, which have a strong foundation in decision theory, are. Therefore, a significant avenue for future research is how resampling compares with Bayesian alternatives.

In short, although why resampled efficiency should be optimal is not clear, resampling remains an interesting heuristic to deal with the important problem of error maximization.

Notes

1. Michaud (1989) and Nawrocki (1996) provide good descriptions of the problem.
2. This problem has been extensively reported and empirically studied. Examples are Best and Grauer (1991), Chopra and Ziemba (1993), and Jobson and Korkie (1983).
3. Michaud (1998) describes his methodology very lucidly in his book *Efficient Asset Management*.
4. Jorion (1992) described portfolio resampling as one way to address sampling error (the true underlying parameters are stable, but there are not enough data to estimate them precisely).
5. Note that resampling deals with sampling error only. In theory, sampling error in means that arises from not having enough data can be cured by lengthening the observation period (in the case of variance, increasing the frequency of observations would help). Because the involved distributions are likely to be nonstationary, however (i.e., the mean and covariance tend to vary over time), enlarging the data set in this way is not always appropriate. Scherer (2002) dealt with this trade-off.
6. All calculations have been done in S-PLUS. Resampling code can be obtained from the author on request.
7. All illustrations in the examples use the original data from Michaud (1998). For these data, $T = 216$ and $k = 8$.
8. For standard mean–variance-efficient frontiers, see Markowitz (1987) or Sharpe (1970).
9. The idea of this test statistic is that looking at weight differences only is obviously not enough. Small weight differences for highly correlated assets might be of higher significance than large weight differences for negatively correlated assets.
10. In Excel, fill one column with 1's and the others with excess returns, and then follow Excel instructions for multiple regressions.
11. See also the exposition in Jobson and Korkie (their Equation 19).
12. Britten-Jones (1999) took this route and showed additionally how this test can be interpreted in light of the well-known Gibbons–Ross–Shanken (1989) procedure.
13. Maddala (2001, p. 600) described how to bootstrap data rather than residuals.
14. For a review of efficient-set mathematics, see Huang and Litzenberger (1988).
15. See Jorion or Chopra, Hensel, and Turner (1993).

References

Best, M., and R. Grauer. 1991. "On the Sensitivity of Mean–Variance Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results." *Review of Financial Studies*, vol. 4, no. 2 (Summer):314–342.

Britten-Jones, M. 1999. "The Sampling Error in Estimates of Mean–Variance Efficient Portfolios." *Journal of Finance*, vol. 54, no. 2 (April):655–671.

Chopra, V., and W. Ziemba. 1993. "The Effects of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice." *Journal of Portfolio Management*, vol. 19, no. 2 (Winter):6–11.

Chopra, V.K., C.R. Hensel, and A.L. Turner. 1993. "Massaging Mean–Variance Inputs: Returns from Alternative Investment Strategies in the 1980s." *Management Science*, vol. 39, no. 7 (July):845–855.

Gibbons, M., S. Ross, and J. Shanken. 1989. "A Test of Efficiency of a Given Portfolio." *Econometrica*, vol. 57, no. 5 (September):1121–52.

Huang, C., and R. Litzenberger. 1988. *Foundations of Financial Economics*. New York: Prentice Hall.

Jobson, J., and B. Korkie. 1983. "Statistical Inference in Two Parameter Portfolio Theory with Multiple Regression Software." *Journal of Financial and Quantitative Analysis*, vol. 18, no. 2 (June):189–197.

Jorion, P. 1992. "Portfolio Optimization in Practice." *Financial Analysts Journal*, vol. 48, no. 1 (January/February):68–74.

Maddala, G. 2001. *Econometrics*. 3rd ed. New York: John Wiley & Sons.

Markowitz, H. 1987. *Mean–Variance Analysis in Portfolio Choice and Capital Markets*. New York: Blackwell.

Michaud, R. 1989. "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" *Financial Analysts Journal*, vol. 45, no. 1 (January/February):31–42.

———. 1998. *Efficient Asset Management*. Boston, MA: Harvard Business School Press.

Nawrocki, D. 1996. "Portfolio Analysis with a Large Universe of Assets." *Applied Economics*, vol. 28, no. 9 (September):1191–98.

Scherer, B. Forthcoming 2002. *Portfolio Construction and Risk Budgeting*. London: Riskwaters.

Sharpe, W. 1970. *Portfolio Theory and Capital Markets*. New York: McGraw-Hill.