Portfolio Optimization in Practice

Philippe Jorion

A major drawback with the classical implementation of mean-variance analysis is that it completely ignores the effect of measurement error on optimal portfolio allocations. A simple simulation approach can provide insight into the distribution of optimal portfolio weights.

As an example, an ex post *optimal portfolio of U.S. and foreign bonds is compared with two benchmarks—a world bond index and a U.S. bond index. Taking sampling variability into account, there is no evidence that the optimal portfolio outperformed the world index over the 1978–88 period. The optimal portfolio did, however, perform significantly better than the U.S. index.*

These results suggest that, over the time period studied, international diversification into foreign bonds has offered some benefits. These benefits are best measured, however, by comparing the performance of a passive world index with that of a U.S. index. An ex post *mean-variance analysis systematically overstates the possible gains from going international.*

The concept of **mean-variance optimization**, developed by Markowitz, is the cornerstone of modern finance theory and a powerful tool for efficiently allocating wealth to different investment alternatives.[1] The technique incorporates investor preferences and expectations of return and risk for all assets considered, as well as diversification effects, which reduce overall portfolio risk.

Given the wide applicability of the mean-variance paradigm, it seems astonishing that investment practitioners do not put it to use more often. One argument often advanced by practitioners is that the optimized portfolios lack investment value in many applications. Instead of implementing nonintuitive decisions dictated by portfolio optimizations, investment managers simply disregard the results, or turn away from the entire approach. Michaud has termed this puzzling situation the "Markowitz optimization enigma."[2]

This article shows that part of the problem lies with the measurement of the necessary inputs. Typically, expected returns, risks and correlations are measured from historical data and fed into an optimizer as if they were known perfectly, when in fact these data are measured with sometimes substantial error. A major drawback of the classical implementation of mean-variance analysis is that it does not recognize the uncertainty inherent in the input parameters, their **estimation risk**.[3] Because of estimation risk, the optimized portfolio can only approximate a true optimal portfolio. In the absence of information about the quality of the approximation, it is not surprising that investment managers often disregard the portfolio optimization process.

This article presents a simple **simulation** method that explicitly measures estimation error. Recognizing that input parameters are in some sense "fuzzy," it reports optimal portfolio weights not only as point estimates, but as point estimates with some measure of dispersion. This approach should give managers a better idea of whether an optimized portfolio makes economic sense. As an illustration, the method is presented in the context of international portfolio choice.

The Classical Approach

Mean-variance analysis assumes that investors prefer portfolios of securities with high expected return in relation to risk. The implementation requires knowledge of the expected returns of all assets under consideration, their standard deviations and all pair-wise correlation coefficients. With this information, a set of efficient portfolios can be calculated. These are defined as the portfolios that minimize risk for various levels of expected returns, and as such, represent the best investment alternatives given the selected assets.

Mean-variance optimization can uniquely integrate portfolio objectives with policy constraints and efficient use of information. For instance, the optimization problem can be formulated with short-sales restrictions, transaction costs, liquidity constraints and turnover constraints. Its ability to incorporate various client constraints makes mean-variance optimization a remarkably flexible tool.

Table I Dollar Returns of World Bond Markets, 1978–1988*

	Government Bonds Denominated in							
	U.S. Doll.	*Can. Doll.*	*Ger. Mark*	*Jap. Yen*	*Brit. Pound*	*Dutch Guilder*	*French Franc*	*World Index*
Avg. Ann.								
Mean Return	9.75	10.03	9.81	15.42	12.57	10.48	10.09	11.31
Std. Dev.	11.21	13.95	16.18	17.31	18.31	15.15	14.13	10.94
				Correlations				
U.S. Doll.	1.000							
Can. Doll.	0.737	1.000						
Ger. Mark	0.358	0.418	1.000					
Jap. Yen	0.292	0.297	0.638	1.000				
Brit. Pound	0.336	0.414	0.543	0.487	1.000			
Dutch Guilder	0.374	0.404	0.963	0.645	0.538	1.000		
French Franc	0.308	0.369	0.880	0.678	0.489	0.891	1.000	

* All data are expressed as per cent per annum. Returns are measured in dollars and include coupon payments. Monthly returns are annualized by multiplying monthly mean returns by 12 and monthly standard deviations by the square root of 12.

Consider, as an illustration, the problem of portfolio choice in the context of a U.S. investor's optimal allocation of U.S. and foreign bonds. International portfolio diversification was advocated by Grubel as early as 1968, and it is traditionally analyzed in a mean-variance framework.[4] Table I reports estimated means, standard deviations and correlations for seven major government bond markets over the period January 1978 to December 1988.[5] Returns are measured in dollars and include coupon payments, price appreciation and exchange rate movements.

The benefits of international diversification are usually expressed in terms of return increments over domestic portfolios with the same level of risk. For example, one portfolio of U.S. government bonds offers 9.75 per cent; an international bond portfolio with the same risk could increase this return to about 12.35 per cent—an increase in performance of 260 basis points. This appears to be a substantial gain.

The basic problem for mean-variance analysis is to identify those combinations of assets that constitute efficient portfolios. Given the difficulty of short selling in foreign markets, the efficient portfolios are constrained to have non-negative weights.[6] These are found by a **quadratic optimization program**. Table II presents the efficient portfolios for a range of expected returns.

The shortcoming of this approach is that it does not recognize, much less quantify, the estimation errors associated with the return outputs. Optimal portfolios, by construction, weight heavily those assets that show the highest returns. As Michaud indicates,

Table II Dollar Returns on Efficient Global Bond Portfolios With No Short Sales, 1978–1988

		Proportion Invested in						
Avg. Ret.	*Stand. Dev.*	*U.S. Doll.*	*Can. Doll.*	*Ger. Mark*	*Jap. Yen*	*Brit. Pound*	*Dutch Guilder*	*French Franc*
10.12	9.95	0.64	0.00	0.00	0.03	0.04	0.00	0.29
10.65	10.03	0.61	0.00	0.00	0.12	0.05	0.00	0.21
11.18	10.25	0.59	0.00	0.00	0.21	0.07	0.00	0.13
11.71	10.61	0.56	0.00	0.00	0.30	0.08	0.00	0.06
12.24	11.10	0.51	0.00	0.00	0.39	0.09	0.00	0.00
12.77	11.77	0.42	0.00	0.00	0.49	0.09	0.00	0.00
13.30	12.62	0.33	0.00	0.00	0.58	0.09	0.00	0.00
13.83	13.63	0.23	0.00	0.00	0.67	0.09	0.00	0.00
14.36	14.75	0.14	0.00	0.00	0.77	0.09	0.00	0.00
14.89	15.96	0.05	0.00	0.00	0.86	0.09	0.00	0.00
15.42	17.31	0.00	0.00	0.00	0.98	0.02	0.00	0.00
Max. Return/Risk:								
11.96	10.82	0.55	0.00	0.00	0.34	0.09	0.00	0.02
World Index								
11.31	10.94	0.46	0.03	0.06	0.14	0.27	0.02	0.02

Glossary

▶***Mean-Variance Optimization:***
A method for determining the amount of funds to commit to each security of a set of specified securities. It assumes that higher portfolio mean returns are preferred to less and that less variance of returns (a measure of risk) is preferred to more. The required inputs are the means, variances and correlations for all specified securities.

▶***Estimation Risk:***
The possibility of errors in the portfolio allocations due to imprecision in the estimated inputs to the portfolio optimization.

▶***Simulation:***
A computer-intensive method used to assess the properties of a decision rule based on statistical data. Repeated samples are obtained from a specified distribution, and the techniques being investigated are then applied to these artificial data sets.

▶***Quadratic Optimization Program:***
A special class of mathematical programming problems having a quadratic objective function and a linear constraint set. In the portfolio selection model, the objective is to minimize some measure of risk (a quadratic function of the portfolio weights) while maximizing return on the total investment.

▶***In-Sample (Out-of-Sample):***
The sample period over which a model is estimated. A decision rule will prove useful if the procedure is validated on a different sample period (out-of-sample), using parameters, such as portfolio weights, previously estimated.

however, these are also the assets most likely to contain positive estimation error.[7] Optimization thus systematically overweights the assets with the highest estimation errors, hence overstates the true efficiency of the optimal portfolio.

Because optimization and performance measurement are conducted over the same period (that is, **in-sample**), the performance of the optimized portfolios must *by construction* improve on the performance of any single asset; at worst, the optimizer would select to remain in a single asset. Moreover, the apparent increase in efficiency offered by optimization will generally grow with the number of assets under consideration. Because of estimation error, optimal weights may be very unstable relative to small changes in expected returns, and portfolio performance may fall sharply in periods outside the sample period used for the optimization.[8]

The next problem is to select one optimal portfolio from among the efficient set. This portfolio can be chosen as a function of the investor's risk-return preferences. For simplicity, assume the investor selects the efficient portfolio that maximizes the overall return-to-risk ratio. This portfolio lies on a line going through the origin and tangent to the efficient frontier. In Table II, the optimal portfolio is invested 55 per cent in U.S. bonds, 34 per cent in Japanese bonds, 9 per cent in British bonds (gilts) and 2 per cent in French bonds. How confident, then, can a portfolio manager be that he has chosen the correct portfolio, given that the input parameters are imprecisely measured? Could it be that the optimal weights are not statistically different from the U.S.-only index? If this is the case, then there is no reliable evidence that an internationally diversified portfolio outperforms a domestic portfolio. The framework presented below answers this question.

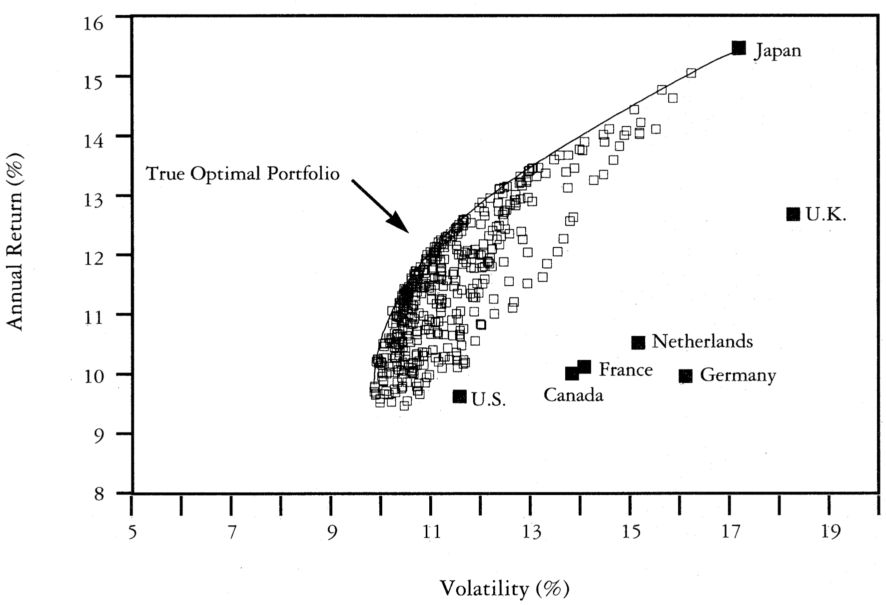
A Simulation Approach

The distribution of an optimal portfolio can be found by simulation analysis. To understand the concept behind the simulation analysis, consider asset returns observed over a given time period. In a traditional statistical framework, the data can be considered to be a random sample drawn from a distribution of returns with unknown means, variances and covariances. The number of observations in the sample is finite, so another set, or random sample, could be drawn from the distribution, and it could have estimated means and a variance-covariance matrix different from that of the first random sample. This new set of input parameters would lead to different optimal portfolios than the first set, even though both were generated by the same underlying distribution.

The simulation proceeds as follows.

1. Compute the means and covariance matrix from the actual sample of historical returns. Define T as the sample size (number of months, say) and N as the number of assets. Perform the optimization, given the stated objective function and investor constraints.
2. Assume that the estimates from Step 1 are true values. From a multivariate standard normal distribution with these parameters, draw one random sample of N joint returns. This represents one month of simulated returns. Sample again until T months are generated.
3. Estimate from these simulated returns a new set of means and a new variance-covariance matrix; perform an optimization using these inputs. The simulated optimal portfolio provides one observation in the distribution of the original optimal portfolio.

Figure A Statistically Equivalent Global Bond Portfolios with No Short Sales, 1978 – 1988



4. Repeat Steps 2 and 3 until the distribution of the optimal portfolio is approximated with enough precision.

For instance, 1000 iterations of Steps 2 to 4 can be performed. A distribution of weights can then be tabulated from the results. Clearly, this distribution is a valuable tool to portfolio managers, because it permits visual inspection of the degree of "fuzziness" of the original optimal investment proportions. In addition, the performance of the 1000 simulated portfolios can be measured in relation to the original data. In terms of the original means and covariance matrix, the simulated portfolios have to be suboptimal. The extent to which their performance falls short of the original optimized performance gives an indication of the error due to estimation risk.

Finally, statistically equivalent portfolios can be generated by selecting a cutoff probability level—say, 5 per cent—and discarding 5 per cent of the portfolios, those with the lowest return-to-risk ratios. The remaining portfolios would then be considered statistically equivalent to the original optimal portfolio.

This procedure is relatively simple to implement, because its major building block consists of the portfolio optimizer, which is already in place. The only additional requirement is a multivariate normal random generator.[9] The procedure is illustrated below.

An Illustration

The simulation described above was used to select a global bond portfolio. Figure A presents the classical efficient set, already described in Table II, as well as the original optimal portfolio that maximizes the return-to-risk ratio. Each additional square represents the performance of a statistically equivalent portfolio (excluding 5 per cent of portfolios with the lowest return-to-risk ratios). Given imprecise measurement of the input parameters, the optimizers could have selected any of these portfolios instead of the original optimal portfolio. The dispersion in the performance of these portfolios suggests that the effect of estimation error is substantial.

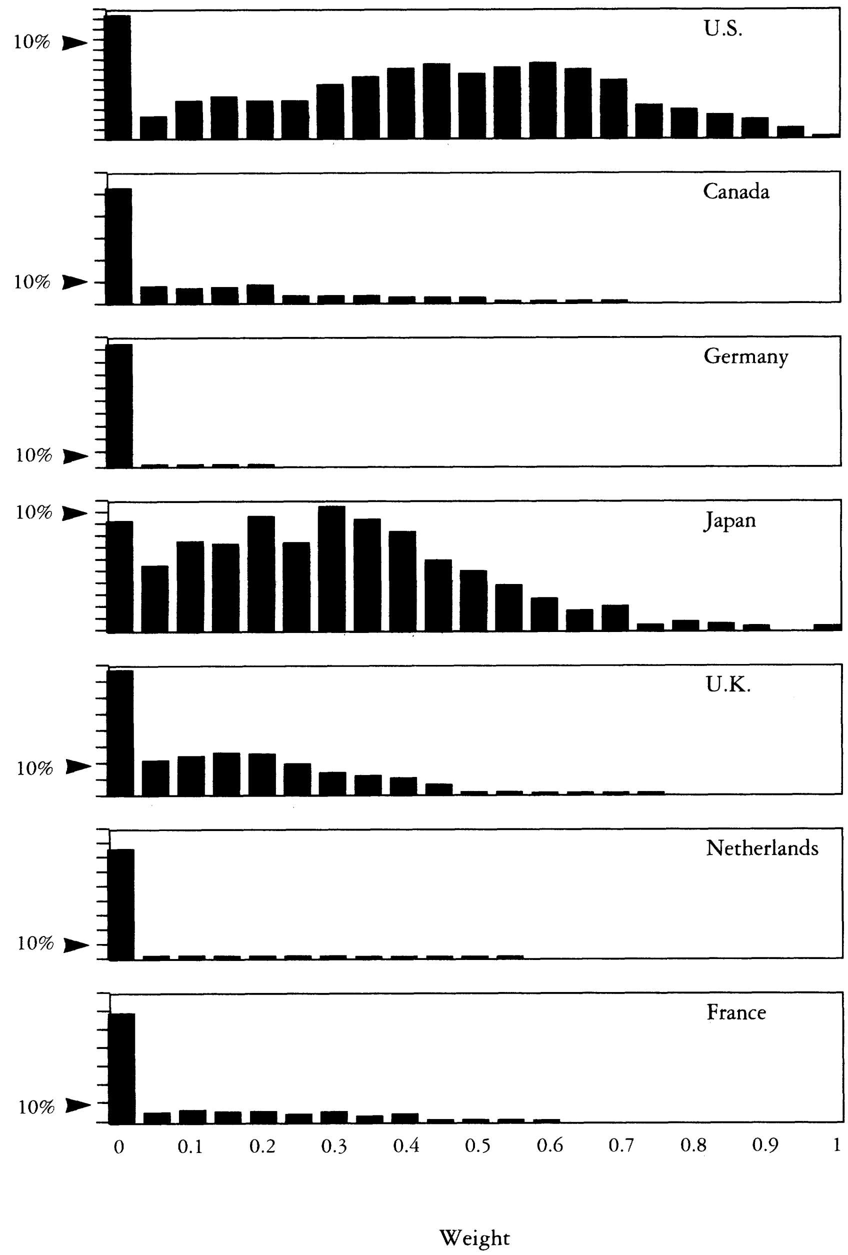
The degree of estimation error changes with the selection criteria used to identify the optimal portfolio. If the objective, for instance, is solely to minimize risk, then the minimum-variance (leftmost) portfolio should be selected. Selection of this portfolio relies only on variances and covariances, not on measured expected returns. Because variances and covariances are estimated with much more precision than expected returns, the portfolios that are statistically equivalent to the true minimum-variance portfolio will be relatively close to it. Dispersion will be much less than that observed in Figure A.

Figure B displays the distributions of the optimal weights. The histograms show the proportion of times a given weight was observed in the simulation. On the horizontal axis, the possible values for the weights are classified into different ranges (0, 0–0.05, 0.05–0.10 and so on up to 0.95–1.00). Each vertical bar indicates the relative number of occurrences within each range.[10]

The average values from the empirical distributions seem to match the point estimates of the previous optimal weights (55 per cent in dollar bonds, 34 per cent in Japanese bonds, 9 per cent in British bonds and 2 per cent in French bonds). In addition, Figure B shows that the portfolio is frequently invested in dollars, in yen and, to some extent, in the pound. Only rarely does the optimal portfolio invest in the remaining currencies. A portfolio that is statistically equivalent to the original optimal portfolio, for example, might be invested 60 per cent in dollar bonds, 30 per cent in Japanese bonds and 10 per cent in British bonds. The reported distribution of weights allows the portfolio manager to assess to what extent the optimal portfolio differs from a reasonable alternative.

More specifically, the portfolio manager can formally test whether the optimal portfolio is statistically better than another given portfolio. Table III reports such tests. The first two rows report the performance of two benchmarks of interest—the world index and the U.S. index.

Figure B Frequency Distribution of Optimal Global Bond Portfolio Weights, 1978 – 1988



On an annual basis, the ratio of return-to-risk for these indexes was 1.034 and 0.870, respectively. As explained before, these numbers must by construction be lower than the ratio of 1.105 obtained for the original optimal portfolio with no short sales.

To test the mean-variance efficiency of a benchmark, a statistic must be constructed that captures the difference between the performance of the optimal portfolio and that of the benchmark. Define θ^* and θ_p as the return-to-risk ratios of the optimal portfolio and the benchmark, respectively. The test statistic used here is:[11]

$$F = k(\theta^{*2} - \theta_p^2)/(1 + \hat{\theta}_p^2).$$

In comparing the observed value of the statistic with its distribution under the null hypothesis, abnormally large values of F imply that the performance of the optimal portfolio far exceeds that of the benchmark; consequently, it is unlikely that the benchmark is efficient.[12]

This particular statistic was chosen because its exact distribution has been derived by Gibbons, Ross and Shanken for the case where negative weights are allowed in the optimal portfolio.[13] With short-sale restrictions, however, the distribution of the statistic must be derived from a simulation under the null hypothesis that the benchmark is truly mean-variance efficient.

Table III reports an F-value of 0.209 for the world index. In the simulation, in 786 instances out of 1000 the value of the statistic was greater than 0.209. There is thus no evidence that the observed value of the statistic is abnormally high, or that the world index is inefficient. Loosely interpreted, the weights of the world index seem to fall within the distribution reported in Figure B.

The value of the test statistic for the U.S. index is 0.656. This corresponds to an empirical marginal significance level of 9.5 per cent. In other words, if the U.S. index were truly efficient, in only 9.5 per cent of the cases would the observed value be exceeded. This suggests that an optimally diversified portfolio of international bonds indeed outperforms an index of U.S. bonds on a risk-return basis.

To investigate the impact of short sales, Figure C displays the portfolios obtained when some of the optimal weights can be negative. Note that the axes in Figure C cover a wider scale than those in Figure B. There thus seems to be much more estimation error in Figure C than in Figure B; the dispersion of the portfolios is much greater than before. Because fewer constraints are imposed on the weights, portfolio positions are more often extreme and more prone to estimation error.

Table III also reports tests of efficiency of the U.S. and world indexes with short sales allowed. With a marginal significance level

Table III Dollar Returns of Portfolios and Tests of Efficiency, 1978–1988*

Portfolio	Mean	Stan. Dev.	Mean/Stan. Dev.	F-Test Statistic (empirical p-value) World	U.S.
World	11.31	10.94	1.034		
U.S.	9.75	11.21	0.870		
Original Optimal with No Short Sales	11.96	10.82	1.105	0.209 (78.6%)	0.656 (9.5%)
With Short Sales	12.35	10.76	1.151	0.353 (88.0%)	0.803 (48.8%)

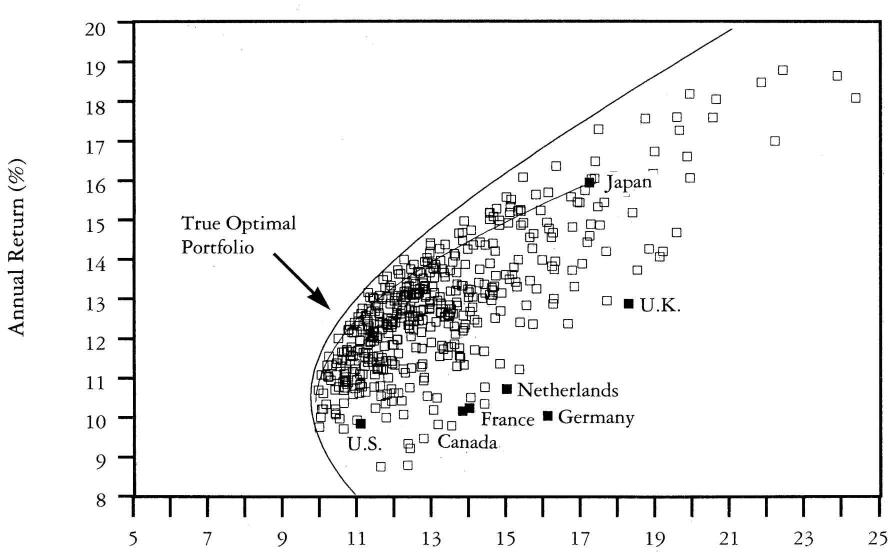
* The optimal portfolios maximize the in-sample return-to-risk ratio. The F statistic measures the difference between the performance of the optimal portfolio and that of the benchmark. The empirical p-values report the proportion of times the value of the test statistic was exceeded from a simulation under the null hypothesis.

of 48.8 per cent, there is no evidence that the U.S. index is inefficient. This contrasts with the results found when short sales are not allowed and suggests that there is more estimation error in an optimal portfolio that allows short sales.

The comparison between the portfolios with and without short sales underlines an important point. Relaxing restrictions or adding assets can only improve the expected performance of efficient portfolios. Because of the interplay between estimation error and optimization, however, adding assets or allowing short sales may actually harm performance, because it can lead to optimal portfolios that are more imprecisely measured.

Consider, for example, the case of two highly correlated assets with truly identical expected returns—say, 10 per cent. In a finite sample, the average returns on these assets will in general differ slightly from each other. Assume assets A and B have average returns of 10.1 and 9.9 per cent, respectively. With short-sale restrictions, the optimizer will select the asset with the highest return; given that the two assets are very similar, this choice cannot be far from the correct choice. In the absence of short-sale restrictions, however, and without transaction costs, the optimizer will typically go long asset A and short asset B in about the same amount; because the data indicate a quasi-arbitrage opportunity of 20 basis points with little risk, this spread can involve very large positions, such as 100 times the initial investment. Such positions are a result of estimation error combined with an unrestricted optimization.

Figure C Statistically Equivalent Global Bond Portfolios with Short Sales, 1978 – 1988



Finally, a few limitations of this framework should be mentioned. The mean-variance approach implicitly assumes that, for each time period under consideration, there is an unknown but stable set of parameters. As the length of the sample period increases, these parameters are estimated more accurately. In practice, unfortunately, very long horizons are required for these tests to be meaningful. Given that average returns are much smaller than volatilities, the tests generally require many years of monthly data. With long horizons, however, expected returns may not be stationary. Hence there is a tradeoff between precision in the optimal weights and validity outside the sample period. Nevertheless, the approach presented here is useful, if only to indicate how imprecise the optimal portfolio weights can be.

Conclusions

The evidence presented here has important implications for the interpretation of mean-variance optimization results. Levy and Lerman, for instance, found that an internationally diversified portfolio containing both stocks and bonds outperforms a portfolio of stocks only.[14] Clearly, with more assets to choose from, a more widely diversified portfolio cannot do worse, in-sample, than a portfolio of stocks only. Along similar lines, Madura and Reiff report a dramatic increase in portfolio performance when currency hedging is allowed.[15] Because currency hedging is equivalent to allowing short positions in Eurocurrency deposits, the performance improvement is a direct result of lifting restrictions on the menu of assets. In these two studies, the interesting question—one that remains unanswered—is whether the performance improvement is statistically significant.[16] The simulation approach presented here provides a framework that can be used to answer this question.[17]

Footnotes

1. *H. M. Markowitz,* Portfolio Selection: Efficient Diversification of Investments *(New York: John Wiley, 1959).*
2. *R. O. Michaud, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?"* Financial Analysts Journal, *January/February 1989.*
3. *Markowitz himself, however, recognized that security analysts should participate in the process of estimating reasonable probability beliefs as inputs to the portfolio selection problem.*
4. *M. Grubel, "Internationally Diversified Portfolios: Welfare Gains and Capital Flows,"* American Economic Review, *1968.*
5. *The data are computed from the Salomon Brothers international bond indexes, which are value-weighted indexes of major government bond markets. The indexes include all outstanding fixed-coupon bonds with at least five years to maturity.*
6. *Short-selling is not allowed in most foreign bond markets. This restriction can be to some extent avoided with the recently created futures contracts on foreign government bonds. However, this may not be a feasible option for many institutional investors.*
7. *Michaud, "The Markowitz Optimization Enigma,"* op. cit.
8. *P. Jorion ("International Portfolio Diversification with Estimation Risk,"* Journal of Business *58 (1985) pp. 259–78), for instance, observed that the out-of-sample performance of optimal portfolios systematically falls short of what was expected. Statistical techniques to alleviate the impact of estimation error are presented in P. Jorion, "Bayes-Stein Estimation for Portfolio Analysis,"* Journal of Financial and Quantitative Analysis, *September 1986.*
9. *Given the ever falling cost of computing power, the simulation can be performed with relative ease. For instance, performing 1000 experiments, which is often more than necessary, with seven assets and a sample size of 132 months takes less than one hour of CPU time on an AT-class personal computer.*
10. *Some information is lost in Figure B because it reports only the marginal distributions, obtained for each individual weight by ignoring the other weights. It should be noted, for instance, that for each sample all weights must add up to unity; this constraint, however, does not appear in the figure. Although the full multivariate distribution is more informative, it must be summarized to be reported on a two-dimensional surface.*
11. *The parameter k is defined as* $k = T(T - N - 1)/N(T - 2)$*, where T is the number of observations and N the total number of assets.*
12. *The test can be interpreted as a transformed measure of the distance between the set of portfolio allocations, since it implicitly involves the weights of the optimal portfolio and of the benchmark.*
13. *The test is derived in Gibbons, Ross and Shanken, "A Test of the Efficiency of a Given Portfolio,"* Econometrica, *September 1989. This statistical test is based on the maximized value of the likelihood function and has the desirable property of being the most powerful test available against a defined alternative. The asymptotic distribution of the optimal weights allowing short sales is described by Jobson and Korkie, "Estimation for Markowitz Efficient Portfolios,"* Journal of the American Statistical Association, *September 1980.*
14. *H. Levy and Z. Lerman, "The Benefits of International Diversification in Bonds,"* Financial Analysts Journal, *September/October 1988.*
15. *Madura and Reiff, "A Hedge Strategy for International Portfolios,"* Journal of Portfolio Management, *Fall 1985.*
16. *The benefits of currency hedging are analyzed in a statistical framework in P. Jorion, "Currency Hedging for International Portfolios" (Columbia University).*
17. *I thank Michael Adler and Richard Michaud for their helpful comments.*

Cornell *footnotes concluded from page 67.*

lowing the collapse of Drexel, but there is no reliable volume database.

19. *Annual data on new issues and on the outstanding stock of low-grade bonds are presented in Altman, "Measuring Corporate Bond Mortality,"* op. cit. *and Asquith, Mullins and Wolff, "Original Issue High-Yield Bonds,"* op. cit.
20. *I thank the Investment Company Institute and Dimensional Fund Advisers for making data available and Michael Brennan, Richard Roll and Walter Torous for their helpful comments.*

Wong et al. *footnotes concluded from page 52.*

8. *See Wong, "Neural Forecaster,"* op. cit.
9. *See, for example, S. Romaniuk and L. Hall, "Parallel Connectionist Expert Systems,"* Proceedings of the International Symposium on Expert Systems, Theory and Applications, *Zurich, Switzerland, June 1989; I. B. Turkesen and L. Guo, "A Fuzzy-Inference System with a Neural Network,"* Proceedings of the North America Fuzzy Information Processing Society (NAFIPS'90), *Toronto, Canada, June 1990; and M. Cohen and D. Hudson, "An Expert System Based on Neural Network Techniques,"* NAFIPS'90, *Toronto, Canada, June 1990.*
10. *The theory of the FN gate can be explained by the Falling Shadow Representation Theory. See P-Z Wang, "Random sets in fuzzy set theory," in M. G. Singh, ed.,* Systems & Control Encyclopedia *(Elmsford, NY: Pergamon Press, 1988).*