Components of Investment Performance

Author(s): Eugene F. Fama
Source: *The Journal of Finance*, Vol. 27, No. 3 (Jun., 1972), pp. 551-567
Published by: Wiley for the American Finance Association
Stable URL: http://www.jstor.org/stable/2978261
Accessed: 08-05-2017 14:02 UTC

The Journal of FINANCE

VOL. XXVII | JUNE 1972 | NO. 3

COMPONENTS OF INVESTMENT PERFORMANCE*

EUGENE F. FAMA‡

I. INTRODUCTION

THIS PAPER SUGGESTS methods for evaluating investment performance. The topic is not new. Important work has been done by Sharpe [21, 22], Treynor [23], and Jensen [13, 14]. This past work has been concerned with measuring performance in two dimensions, return and risk. That is, how do the returns on the portfolios examined compare with the returns on other "naively selected" portfolios with similar levels of risk?

This paper suggests somewhat finer breakdowns of performance. For example, methods are presented for distinguishing the part of an observed return that is due to ability to pick the best securities of a given level of risk ("selectivity") from the part that is due to predictions of general market price movements ("timing"). The paper also suggests methods for measuring the effects of foregone diversification when an investment manager decides to concentrate his holdings in what he thinks are a few "winners."

Finally, most of the available work concentrates on single period evaluation schemes. Since almost all of the relevant theoretical material can be presented in this context, much of the analysis here is likewise concerned with the one-period case. Eventually, however, a multiperiod model that allows evaluations both on a period-by-period and on a cumulative basis is presented.

II. FOUNDATIONS

The basic notion underlying the methods of performance evaluation to be presented here is that the returns on managed portfolios can be judged relative to those of "naively selected" portfolios with similar levels of risk. For purposes of exposition, the definitions of a "naively selected" portfolio and of "risk" are obtained from the two-parameter market equilibrium model of Sharpe [20], Lintner [15, 16], Mossin [18] and Fama [10, 11]. But it is well to note that the two-parameter model just provides a convenient and somewhat familiar set of naively selected or "benchmark" portfolios against which

* Research on this paper was supported by a grant from the National Science Foundation.

‡ Graduate School of Business, University of Chicago.

the investment performance of managed portfolios can be evaluated. As indicated later, other risk-return models could be used to obtain benchmark portfolios consistent with the same general methods of performance evaluation.

In the simplest one-period version of the two-parameter model, the capital market is assumed to be perfect—that is, there are no transactions costs or taxes, and all available information is freely available to everybody—and investors are assumed to be risk averse expected utility maximizers who believe that return distributions for all portfolios are normal. Risk aversion and normally distributed portfolio returns imply that the expected utility maximizing portfolio for any given investor is mean-standard deviation efficient.[1] In addition, investors are assumed to have the same views about distributions of one-period returns on all portfolios (an assumption usually called "homogeneous expectations"), and there is assumed to be a riskless asset f, with both borrowing and lending available to all investors at a riskless rate of interest R_f.

It is then possible to show that in a market equilibrium all efficient portfolios are just combinations of the riskless asset f and one portfolio of risky assets m, where m, called the "market portfolio," contains every asset in the market, each weighted by the ratio of its total market value to the total market value of all assets. That is, if $\tilde{R}_m$, $E(\tilde{R}_m)$ and $\sigma(\tilde{R}_m)$ are the one-period return, expected return, and standard deviation of return for the market portfolio m, and if x is the proportion of investment funds put into the riskless asset f, then all efficient portfolios are formed according to[2]

$$\tilde{R}_x = xR_f + (1 - x)\tilde{R}_m \qquad x \leqslant 1, \tag{1}$$

so that

$$E(\tilde{R}_x) = xR_f + (1 - x)E(\tilde{R}_m) \tag{2}$$

$$\sigma(\tilde{R}_x) = (1 - x)\sigma(\tilde{R}_m). \tag{3}$$

Geometrically, the situation is somewhat as shown in Figure 1. The curve b m d represents the boundary of the set of portfolios that only include risky assets. But efficient portfolios are along the line from R_f through m. Points below m (that is, $x \geqslant 0$) involve lending some funds at the riskless rate R_f and putting the remainder in m, while points above m (that is, $x < 0$) involve borrowing at the riskless rate with both the borrowed funds and the initial investment funds put into m.

In this model the equilibrium relationship between expected return and risk for any security j is

$$E(\tilde{R}_j) = R_f + \left[\frac{E(\tilde{R}_m) - R_f}{\sigma(\tilde{R}_m)}\right] \frac{\text{cov}(\tilde{R}_j, \tilde{R}_m)}{\sigma(\tilde{R}_m)} \qquad (\textit{Ex ante} \text{ market line}). \tag{4}$$

Here cov $(\tilde{R}_j,\tilde{R}_m)$ is the covariance between the return on asset j and the return

1. By definition, a mean-standard deviation efficient portfolio must have the following property: No portfolio with the same or higher expected one-period return has lower standard deviation of return.

2. Tildes (~) are used throughout to denote random variables. When we refer to realized values of these variables, the tildes are dropped.

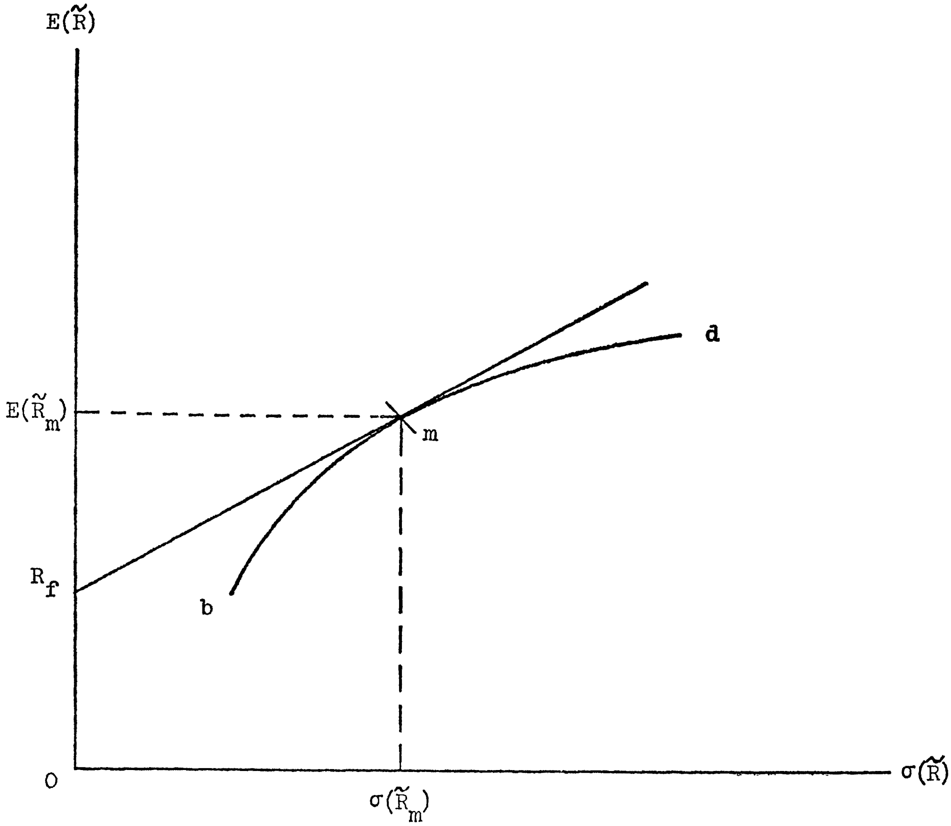

FIGURE 1
The Efficient Set with Riskless Borrowing and Lending

on the market portfolio m. In the two-parameter model $\sigma(\tilde{R}_m)$ is a measure of the total risk in the return on the market portfolio m. Since the only risky assets held by an investor are "shares" of m, it would seem that, from a portfolio viewpoint, the risk of an asset should be measured by its contribution to $\sigma(\tilde{R}_m)$. In fact this contribution is just $\text{cov}(\tilde{R}_j,\tilde{R}_m)/\sigma(\tilde{R}_m)$. Specifically, if x_{jm} is the proportion of asset j, $j = 1, \ldots, N$, in the market portfolio m

$$\sigma(\tilde{R}_m) = \sum_{j=1}^{N} x_{jm} \frac{\text{cov}(\tilde{R}_j, \tilde{R}_m)}{\sigma(\tilde{R}_m)}. \tag{5}$$

In this light (4) is a relationship between expected return and risk which says that the expected return on asset j is the riskless rate of interest R_f plus a risk premium that is $[E(\tilde{R}_m) - R_f]/\sigma(\tilde{R}_m)$, called the market price per unit of risk, times the risk of asset j, $\text{cov}(\tilde{R}_j,\tilde{R}_m)/\sigma(\tilde{R}_m)$.

Equation (4) provides the relationship between expected return and risk for portfolios as well as for individual assets. That is, if x_{jp} is the proportion of asset j in the portfolio p (so that $\sum_{j=1}^{N} x_{jp} = 1$), then multiplying both sides of (4) by x_{jp} and summing over j, we get

$$E(\tilde{R}_p) = R_f + \left[\frac{E(\tilde{R}_m) - R_f}{\sigma(\tilde{R}_m)}\right] \frac{\text{cov}(\tilde{R}_p, \tilde{R}_m)}{\sigma(\tilde{R}_m)} \tag{6}$$

where, of course,

$$\tilde{R}_p = \sum_{j=1}^{N} x_{jp}\tilde{R}_j.$$

But (4) and (6) are expected return-risk relations derived under the assumption that investors all have free access to available information and all have the same views of distributions of returns on all portfolios. In short, the market setting envisaged is a rather extreme version of the "efficient markets" model in which prices at any time "fully reflect" available information. (See, for example [7].) But in the real world a portfolio manager may feel that he has access to special information or he may disagree with the evaluations of available information that are implicit in market prices. In this case the "homogeneous expectations" model underlying (4) provides "benchmarks" for judging the manager's ability to make better evaluations than the market.

The benchmark or naively selected portfolios are just the combinations of the riskless asset f and the market portfolio m obtained with different values of x in (1). Given the *ex post* or realized return R_m for the market portfolio, for the naively selected portfolios, *ex post* return is just

$$\tilde{R}_x = xR_f + (1 - x)\tilde{R}_m, \tag{7}$$

that is, (1) without the tildes. Moreover,[3]

$$\beta_x = \frac{\text{cov}(\tilde{R}_x, \tilde{R}_m)}{\sigma(\tilde{R}_m)} = \frac{\text{cov}([1 - x]\tilde{R}_m, \tilde{R}_m)}{\sigma(\tilde{R}_m)} = (1 - x)\sigma(\tilde{R}_m) = \sigma(\tilde{R}_x). \tag{8}$$

That is, for the benchmark portfolios risk and standard deviation of return are equal. And the result is quite intuitive: In the homogeneous expectations model these portfolios comprise the efficient set, and for efficient portfolios risk and return dispersion are equivalent.

For the naively selected portfolios, (7) and (8) imply the following relationship between risk β_x and *ex post* return R_x:

$$R_x = R_f + \left(\frac{R_m - R_f}{\sigma(\tilde{R}_m)}\right)\beta_x \qquad (\textit{ex post} \text{ market line}). \tag{9}$$

That is, for the naively selected portfolios there is a linear relationship between risk and return that is of precisely the same form as (4) except that the expected returns that appear in (4) are replaced by realized returns in (9).

In the performance evaluation models to be presented, (9) provides the benchmarks against which the returns on "managed" portfolios are judged. These "benchmarks" are used in a sequence of successively more complex suggested performance evaluation settings. First we are concerned with one-period models in which a portfolio is chosen by an investor at the beginning of the

3. Henceforth the risk $\text{cov}(\tilde{R}_j, \tilde{R}_m)/\sigma(R_m)$ of an asset or portfolio j will be denoted as β_j.

period, its performance is evaluated at the end of the period, and there are no intermediate cash flows or portfolio decisions. Then we consider multiperiod evaluation models that also allow for fund flows and portfolio decisions between evaluation dates. We find, though, that almost all of the important theoretical concepts in performance evaluation can be treated in a one-period context.

III. The Benchmark Portfolios: Some Empirical Issues

Before introducing the evaluation models, however, it is well to discuss some of the empirical issues concerning the so-called "market lines" (4) and (9). Since this paper is primarily theoretical, and since empirical problems are best solved in the context of actual applications, the discussion of empirical issues will be brief.

First of all, to use (9) as a benchmark for evaluating *ex post* portfolio returns requires estimates of the risk, β_p, and dispersion, $\sigma(\tilde{R}_p)$, of the managed portfolios as well as an estimate of $\sigma(\tilde{R}_m)$, the dispersion of the return on the market portfolio. If performance evaluation is to be objective, it must be possible to obtain reliable estimates of these parameters from historical data. Fortunately, Blume's evidence [3, 4, 5] suggests that at least for portfolios of ten or more securities, β_p and $\sigma(\tilde{R}_p)$ seem to be fairly stationary over long periods of time (e.g., ten years), and likewise for $\sigma(\tilde{R}_m)$.

But other empirical evidence is less supportive. Thus throughout the analysis here normal return distributions are assumed, though the data of Fama [6], Blume [3], Roll [19] and others suggest that actual return distributions conform more closely to non-normal two-parameter stable distributions. It would conceptually be a simple matter to allow for such distributions in the evaluation models (cf. Fama [11]). But since the goal here is just to suggest some new approaches to performance evaluation, for simplicity attention will be restricted to the normal model.

Finally, the available empirical evidence (e.g., Friend and Blume [12], Miller and Scholes [17], and Black, Jensen and Scholes [2]) indicates that the average returns over time on securities and portfolios deviate systematically from the predictions of (4). Though the observed average return-risk relationships seem to be linear, the tradeoff of risk for return (the price of risk) is in general less than would be predicted from (4) or (9). In short, the evidence suggests that (4) and (9) do not provide the best benchmarks for the average return-risk tradeoffs available in the market from naively selected portfolios.

Even these results do little damage to the performance evaluation models. They indicate that other benchmark portfolios than those that lead to (9) might be more appropriate, but given such alternative "naively selected" portfolios, the analysis could proceed in exactly the manner to be suggested. For example, Black, Jensen and Scholes [2] compute the risks (β's) for each security on the New York Stock Exchange, rank these, and then form ten portfolios, the first comprising the .1N securities with the highest risks and the last comprising the .1N securities with the lowest risks, where N is the total number of securities. They find that over various subperiods from 1931-65 the average monthly returns among these portfolios are highly correlated, and when plotted against risk the average returns on these portfolios lie along a straight line with slope somewhat less than would be implied by the "price of risk" in (4) or (9). As benchmarks for performance evaluation models, their empirical risk-return lines seem to be natural alternatives to (9). And with these alternative benchmarks, performance evaluation could proceed precisely as suggested here. But again, for simplicity, we continue on with the more familiar benchmarks given by (9).

It would be misleading, however, to leave the impression that all important empirical problems relevant in the application of performance evaluation models have been solved. To a large extent the practical value of such models depends on the empirical validity of the model of market equilibrium—that is, the expected return-risk relationship—from which the benchmark or "naively selected" portfolios are derived. And though much interesting work is in progress, it would be rash to claim that all empirical issues concerning models of market equilibrium have been settled.

For example, an important (and unsolved) empirical issue in models of market equilibrium is the time interval or "market horizon period" over which the hypothetical expected return-risk relationship is presumed to hold. Does the model hold continuously (instant by instant), or is the market horizon period some discrete time interval? This is an important issue from the viewpoint of performance evaluation since if the market horizon period is discrete, evaluation periods should be chosen to coincide with horizon periods.

The evidence of Friend and Blume [12] and that of Black, Jensen, and Scholes [2] suggests that meaningful relationships between average returns and risk can be obtained from monthly data, while the evidence of Miller and Scholes [17] indicates that this is not true for annual periods. Within these broad bounds, however, the sensitivity of risk-return relations to the time interval chosen remains an open issue.

But unsolved empirical questions are hardly a cause for disheartenment. It is reasonable to expect that some of the empirical issues will be solved in the process of applying the theory. And in any case, application of a theory invariably involves some empirical approximations. The available evidence on performance evaluation, especially Jensen's [13, 14], suggests that the required approximations need not prevent even more complicated evaluation models from yielding useful results.

IV. Performance Evaluation in a One-Period Model When There Are No Intraperiod Fund Flows

Let $V_{a,t}$ and $V_{a,t+1}$ be the total market values at t and t + 1 of the actual (a = actual) portfolio chosen by an investment manager at t. With all portfolio activity occurring at t and t + 1, that is, assuming that there are no intraperiod fund flows, the one-period percentage return on the portfolio is

$$R_a = \frac{V_{a,t+1} - V_{a,t}}{V_{a,t}}.$$

One benchmark against which the return R_a on the chosen portfolio can be compared is provided by $R_x(\beta_a)$, which by definition is the return on the combination of the riskless asset f and the market portfolio m that has risk β_x equal to β_a, the risk of the chosen portfolio a. One measure of the performance of the chosen portfolio a is then

$$\text{Selectivity} = R_a - R_x(\beta_a). \tag{10}$$

That is, *Selectivity* measures how well the chosen portfolio did relative to a naively selected portfolio with the same level of risk.

Selectivity, or some slight variant thereof, is the sole measure of performance in the work of Sharpe [21, 22], Treynor [23] and Jensen [13, 14]. But more detailed breakdowns of performance are possible. Thus consider

$$\overbrace{[R_a - R_f]}^{\substack{\text{Overall}\\ \text{Performance}}} = \overbrace{[R_a - R_x(\beta_a)]}^{\text{Selectivity}} + \overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}}. \tag{11}$$

That is, the *Overall Performance* of the portfolio decision is the difference between the return on the chosen portfolio and the return on the riskless asset. The Overall Performance is in turn split into two parts, *Selectivity* (as above) and *Risk*. The latter measures the return from the decision to take on positive amounts of risk.[4] It will be determined by the level of risk chosen (the value of β_a) and, from (9), by the difference between the return on the market portfolio, R_m, and the return on the riskless asset, R_f.

These performance measures are illustrated in Figure 2. The curly bracket along the vertical axis shows *Overall Performance* which in this case is positive. The breakdown of performance given by (11) can be found along the vertical line from β_a. In this example, *Selectivity* is positive: A portfolio was chosen that produced a higher return than the corresponding "naively selected" portfolio along the market line with the same level of risk. *Risk* is also positive, as it is whenever a positive amount of risk is taken and the return on the market portfolio turns out to be higher than the riskless rate.

A. *Selectivity: A Closer Look*

If the portfolio chosen represents the investor's total assets, in the mean-variance model the risk of the portfolio to him is measured by $\sigma(\tilde{R}_a)$, the standard deviation of its return. And the risk of the portfolio to the investor, $\sigma(\tilde{R}_a)$, will be greater than what might now be called its "market risk," β_a, as long as the portfolio's return is less than perfectly correlated with the return on the market portfolio. To see this, note that the correlation coefficient k_{am} between R_a and R_m is

4. For greater descriptive accuracy, we should, of course, say "return from risk" or even "return from bearing risk," rather than just *Risk*. Likewise, "return from selectivity," would be more descriptive than *Selectivity*. But (hopefully) the shorter names save space without much loss of clarity.

$$k_{am} = \frac{cov(\tilde{R}_a, \tilde{R}_m)}{\sigma(\tilde{R}_a)\sigma(\tilde{R}_m)}.$$

It follows that

$$\beta_a = \frac{cov(\tilde{R}_a, \tilde{R}_m)}{\sigma(\tilde{R}_m)} = k_{am}\sigma(\tilde{R}_a)$$

so that $\beta_a \lesseqgtr \sigma(\tilde{R}_a)$ depending on whether $k_{am} \lesseqgtr 1$.[5]

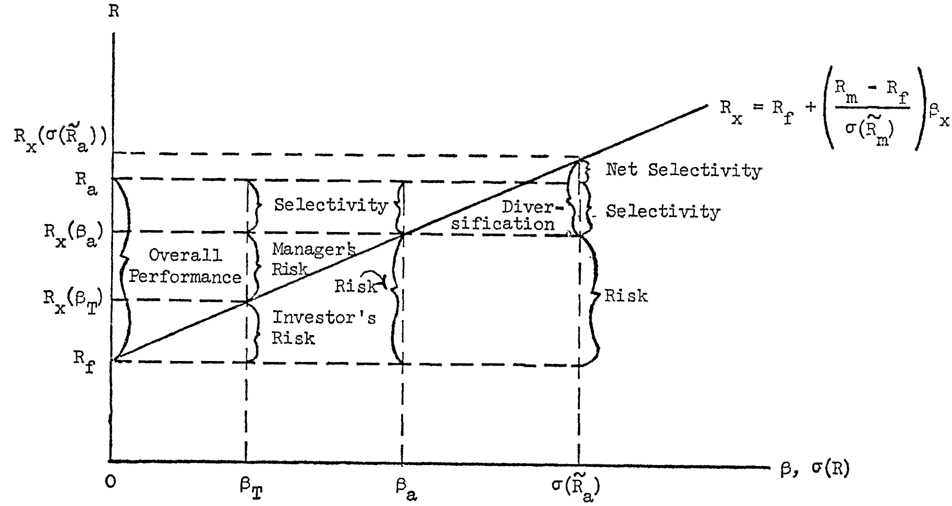

FIGURE 2
An Illustration of the Performance Measures of Equations (11), (12), and (13).

Intuitively, to some extent the portfolio decision may have involved putting more eggs into one or a few baskets than would be desirable to attain portfolio efficiency—that is, the manager places his bets on a few securities that he thinks are winners. In other words, to the extent that $\sigma(\tilde{R}_a) > \beta_a$, the portfolio manager decided to take on some portfolio dispersion that could have been diversified away because he thought he had some securities in which it would pay to concentrate resources. The results of such a decision can be evaluated in terms of the following breakdown of *Selectivity*:

$$\overbrace{[R_a - R_x(\beta_a)]}^{\text{Selectivity}} = \text{Net Selectivity} + \overbrace{[R_x(\sigma(\tilde{R}_a)) - R_x(\beta_a)]}^{\text{Diversification}}; \qquad (12a)$$

or

$$\text{Net Selectivity} = \overbrace{[R_a - R_x(\beta_a)]}^{\text{Selectivity}} - \overbrace{[R_x(\sigma(\tilde{R}_a)) - R_x(\beta_a)]}^{\text{Diversification}}. \qquad (12b)$$

By definition, $R_x(\sigma(\tilde{R}_a))$ is the return on the combination of the riskless asset f and the market portfolio m that has return dispersion equivalent to

5. In fact the naively selected portfolios are the only ones whose returns are literally perfectly correlated with those of the market portfolio (cf. equation (8)). But the theoretical work of Fama [9] and the empirical work of Black, Jensen and Scholes [2] suggests that the return on any well-diversified portfolio will be very highly correlated with R_m.

that of the actual portfolio chosen. Thus *Diversification* measures the extra portfolio return that the manager's winners have to produce in order to make concentration of resources in them worthwhile. If *Net Selectivity* is not positive, the manager has taken on diversifiable risk that his winners have not compensated for in terms of extra return.

Note that, as defined in (12), *Diversification* is always non-negative, so that *Net Selectivity* is equal to or less than *Selectivity*. When $R_m > R_f$, *Diversification* measures the additional return that would just compensate the investor for the diversifiable dispersion (that is, $\sigma(\tilde{R}_a) - \beta_a$) taken on by the manager. When $R_m < R_f$ (so that the market line is downward sloping), *Diversification* measures the lost return from taking on diversifiable dispersion rather than choosing the naively selected portfolio with market risk *and* standard deviation both equal to β_a, the market risk of the portfolio actually chosen.

The performance measures of (12) are illustrated in Figure 2 along the dashed vertical line from $\sigma(\tilde{R}_a)$. In the example shown, *Selectivity* is positive but *Net Selectivity* is negative. Though the manager chose a portfolio that outperformed the naively selected portfolio with the same level of market risk, his *Selectivity* was not sufficient to make up for the avoidable risk taken, so that *Net Selectivity* was negative.

The breakdown of *Selectivity* given by (12) is the only one that is considered here. The rest of Section IV is concerned with successively closer examinations of the other ingredient of *Overall Performance, Risk*. Before moving on, though, we should note that (12) itself is *only* relevant when diversification is a goal of the investor. And this is the case only when the portfolio being evaluated constitutes the investor's entire holdings, and the investor is risk averse. For example, an investor might allocate his funds to many managers, encouraging each only to try to pick winners, with the investor himself carrying out whatever diversification he desires on personal account. In this case *Selectivity* is the relevant measure of the managers' performance, and the breakdown of *Selectivity* of (12) is of no concern.

B. *Risk: A Closer Look*

If the investor has a target risk level β_T for his portfolio, the part of *Overall Performance* due to *Risk* can be allocated to the investor and to the portfolio manager as follows:

$$\overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}} = \overbrace{[R_x(\beta_a) - R_x(\beta_T)]}^{\text{Manager's Risk}} + \overbrace{[R_x(\beta_T) - R_f]}^{\text{Investor's Risk}} \qquad (13)$$

$R_x(\beta_T)$ is the return on the naively selected portfolio with the target level of market risk. Thus *Manager's Risk* is that part of *Overall Performance* and of *Risk* that is due to the manager's decision to take on a level of risk β_a different from the investor's target level β_T, while *Investor's Risk* is that part of *Overall Performance* that results from the fact that the investor's target level of risk is positive. These performance measures are illustrated in Figure 2 along the dashed vertical line from β_T.

Manager's Risk might in part result from a timing decision. That is, in part at least the manager might have chosen a portfolio with a level of risk higher or lower than the target level because he felt risky portfolios in general would do abnormally well or abnormally poorly during the period under consideration. But if an estimate of $E(\tilde{R}_m)$ is available, a more precise measure of the results of such a timing decision can be obtained.[6] Specifically, making use of the *ex ante* market line of (4)[7] we can subdivide *Risk* as follows:

$$\overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}} = \overbrace{\{\underbrace{[R_x(\beta_a) - E(\tilde{R}_x(\beta_a))]}_{\text{Total Timing}} - \underbrace{[R_x(\beta_T) - E(\tilde{R}_x(\beta_T))]}_{\text{Market Conditions}}\}}^{\text{Manager's Timing}}$$
$$+ \underbrace{[E(\tilde{R}_x(\beta_a)) - E(\tilde{R}_x(\beta_T))]}_{\text{Manager's Exp. Risk}} + \underbrace{[R_x(\beta_T) - R_f]}_{\text{Investor's Risk}}. \quad (14)$$

The first three terms here sum to the *Manager's Risk* of (13). *Manager's Expected Risk* is the incremental expected return from the manager's decision to take on a nontarget level of risk. *Market Conditions* is the difference between the return on the naively selected portfolio with the target level of risk and the expected return of this portfolio. It answers the question: By how much did the market deviate from expectations at the target level of risk? *Total Timing* is the difference between the *ex post* return on the naively selected portfolio with risk β_a and the *ex ante* expected return. It is positive when $R_m > E(\tilde{R}_m)$ (and then more positive the larger the value of β_a), and it is negative when $R_m < E(\tilde{R}_m)$ (and then more negative the larger the value of β_a). The difference between *Total Timing* and *Market Conditions* is *Manager's Timing*: it measures the excess of *Total Timing* over timing performance that could have been generated by choosing the naively selected portfolio with the target level of risk. *Manager's Timing* is only positive when the sign of the difference between β_a and β_T is the same as the sign of the difference between R_m and $E(\tilde{R}_m)$, that is, when the chosen level of market risk is above

6. $E(\tilde{R}_m)$ might be estimated from past average returns on the market portfolio m. Alternatively, past data might be used to estimate the average difference between R_m and R_f. In any case, it should become clear that the expected values used must be naive or mechanical estimates (or at least somehow external to those being evaluated), otherwise the value of the timing measures is destroyed.

Admittedly, given the current status of empirical work on the behavior through time of average returns on risky assets, we can at most sepeculate about the best way to estimate $E(\tilde{R}_m)$. Hopefully empirical work now in progress will give more meaningful guidelines. And perhaps the development of theoretical methods of performance evaluation will itself stimulate better empirical work on estimation procedures. In any case, the discussion in the text should help to emphasize that one cannot obtain precise measures of returns from timing decisions without mechanical or naive estimates of equilibrium expected returns.

7. That is,

$$E(\tilde{R}_x(\beta_a)) = R_f + \left[\frac{E(\tilde{R}_m) - R_f}{\sigma(\tilde{R}_m)}\right]\beta_a$$

and similarly for $E(\tilde{R}_x(\beta_T))$.

(below) the target level and R_m is above (below) $E(\tilde{R}_m)$. It is thus somewhat more sensitive than *Total Timing* as a measure of the results of a timing decision.

A target level of risk will not always be relevant in evaluating a manager's performance. For example, an investor may allocate his funds to many managers, with the intention that each concentrates on selectivity and/or timing, with the investor using borrowing or lending on personal account to attain his desired level of market risk.

If a target level of risk is not relevant but the expected value or *ex ante* market line is still available, a breakdown of *Risk* similar to (14) can be obtained by treating the market portfolio (or the appropriate proxy)[8] as the target portfolio. That is,

$$\overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}} = \overbrace{\{\underbrace{[R_x(\beta_a) - E(\tilde{R}_x(\beta_a))]}_{\text{Total Timing}} - \underbrace{[R_m - E(\tilde{R}_m)]}_{\text{Market Conditions}}\}}^{\text{Manager's Timing}} + \underbrace{[E(\tilde{R}_x(\beta_a)) - E(\tilde{R}_m)]}_{\substack{\text{Expected Deviation} \\ \text{from Market}}} + \underbrace{[R_m - R_f]}_{\text{Market Risk}}. \quad (15)$$

The idea here is that even in the absence of a target level of risk, the measure of *Manager's Timing* must be standardized for the deviation of the market return from the expected market return, that is, for the "average" spread between the *ex post* and *ex ante* market lines.

Finally, the goal of this paper is mainly to suggest some ways in which available theoretical and empirical results on portfolio and asset pricing models can provide the basis of useful procedures for performance evaluation. But the various breakdowns of performance suggested above are hardly unique. Indeed any breakdown chosen should be tailored to the situation at hand. For example, if a target level of risk is relevant but the subdivision of Risk given by (14) is regarded as too complicated, then the approximate effects of the timing decision might still be separated out as follows:

$$\overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}} = \overbrace{[R_x(\beta_a) - E(\tilde{R}_x(\beta_a))]}^{\text{Total Timing}} + \overbrace{[E(\tilde{R}_x(\beta_a)) - E(\tilde{R}_x(\beta_T))]}^{\text{Manager's Expected Risk}} + \overbrace{[E(\tilde{R}_x(\beta_T)) - R_f]}^{\text{Investor's Expected Risk}}. \quad (16)$$

The one new term here is *Investor's Expected Risk,* which measures the expected contribution to *Overall Performance* of the investor's decision to have a positive target level of risk. Alternatively if a target level of risk is not

8. For example, if one were faced with portfolio evaluation in a multiperiod context, one might use the average of past levels of market risk chosen by the manager as a proxy for the target risk level when the latter is not explicitly available.

relevant for the situation at hand, but an expected value line is available, *Risk* can nevertheless be subdivided as follows,

$$\overbrace{[R_x(\beta_a) - R_f]}^{\text{Risk}} = \overbrace{[R_x(\beta_a) - E(\tilde{R}_x(\beta_a))]}^{\text{Total Timing}} + \overbrace{[E(\tilde{R}_x(\beta_a)) - R_f]}^{\text{Total Expected Risk}}. \tag{17}$$

And these few suggestions hardly exhaust the possibilities.

V. Components of Performance: Multiperiod Models with Intraperiod Fund Flows

In the one-period evaluation model presented above, (i) the time at which performance is evaluated is assumed to correspond to the portfolio horizon date, that is, the time when portfolio funds are withdrawn for consumption; and (ii) there are assumed to be no portfolio transactions or inflows and outflows of funds between the initial investment and withdrawal dates, so that there is no reinvestment problem. If in a multiperiod context we are likewise willing to assume that: (i) though there are many of them, evaluation dates nevertheless correspond to the dates when some funds are withdrawn for consumption, and (ii) all reinvestment decisions and other portfolio transactions are also made at these same points in time, then generalization of the one-period model to the multiperiod case is straightforward.[9] Indeed the basic procedure could be period-by-period application of the performance measures presented in the one-period model. The major embellishments would not be in the nature of new theory, but rather would arise from the fact that multiperiod performance histories allow statistically more reliable estimates of the various one-period performance measures.

But this pure case is unlikely to be met in any real world application. Often performance evaluation would be carried out by someone with little or no knowledge of the dates when funds are needed for consumption by the owner of the portfolio, and often (e.g., in the case of a mutual fund or a pension fund) the portfolio is owned by many different investors with different consumption dates. As a result evaluation dates, withdrawal dates, and reinvestment dates do not usually coincide.

The rest of this paper is concerned with how the concepts of the one-period model must be adjusted to deal with such intraevaluation period (or more simply, intraperiod) fund flows. The procedure is to first present detailed definitions of variables of interest in models involving intraperiod fund flows, and then to talk about actual measures of performance. And it is well to keep in mind that though the analysis is carried out in a multiperiod context, the problems to be dealt with arise from intraperiod fund flows. With such fund flows, the same problems would arise in a one-period evaluation model.

A. *Definitions*

Suppose the investment performance of a portfolio is to be evaluated at discrete points in time, but that there can be cash flows between evaluation

9. For the development of the underlying models of consumer and market equilibrium for this case see [8].

dates. That is, there can be intraperiod inflows in the form of either cash receipts (dividends, interest) on existing portfolio holdings or net new contributions of capital by new or existing owners. And there can be intraperiod outflows in the form of dividend payments to the portfolio's owner(s) (e.g., a mutual fund declares dividends) or withdrawals of capital (e.g., by a mutual fund's shareholders).

In simplest terms, the major problem with intraperiod cash flows is obtaining a measure of the return on the beginning of period market value of a portfolio that abstracts from the effects of intraperiod new contributions and withdrawals on the end of period value of the portfolio. One approach is what might be called the mutual fund method. Specifically, when performance evaluation is first contemplated, the market value of the portfolio is subdivided into "shares." Subsequently, whenever there are contributions of new capital or withdrawals of capital from the portfolio, the current market value of a share is computed and the number of shares outstanding is adjusted to reflect the effects of the cash flow.[10]

Thus let evaluation dates correspond to integer values of t and define

$V'_{a,t}$ = actual market value of the portfolio at time t. It thus includes the effects of investment of new capital or reinvestment of any cash income received on securities held in the portfolio, and it is net of any dividends paid out to owners or other withdrawals of funds prior to t.

$V_{a,t}$ = market value the portfolio would have had at t if no dividends were paid out to owners since the previous evaluation date. In computing $V_{a,t}$ it is simply assumed that dividends paid to the portfolio's owners were instead reinvested in the entire portfolio. At the beginning of each evaluation period, however, $V_{a,t}$ is set equal to $V'_{a,t}$.

n_t = number of shares outstanding in the portfolio at t. As indicated above, this is adjusted when new capital comes into the portfolio and when capital is withdrawn, but it is unaffected by reinvestment of cash income received on securities held or by dividends paid to the portfolio's owners.

$p'_{a,t} = V'_{a,t}/n_t$ = actual market value at t of a share in the portfolio.

$p_{a,t} = V_{a,t}/n_t$ = value of a share at t under the assumption that dividends paid to owners of the portfolio were instead reinvested in the entire portfolio.

$R_{a,t} = (p_{a,t} - p'_{a,t-1})/p'_{a,t-1}$. Assuming t corresponds to an evaluation date, this is the one-period return on a share with reinvestment of all dividends paid on a share since the last evaluation date.

$R_{a,t}$ is an unambiguous measure of the return from $t - 1$ to t on a dollar invested in the portfolio at $t - 1$. This is not to say, however, that it is unaffected by intraperiod fund flows. Such fund flows are usually associated with redistributions of portfolio holdings across securities and these affect the

10. This is in fact the method of accounting used by open end mutual funds. It is also closely related to the "time-weighted rate of return" approach developed by Professor Lawrence Fisher. On this point see [1, Appendix I and p. 218].

return on a share. Moreover, $R_{a,t}$ as defined above is not the only unambiguous measure of the return from $t-1$ to t on funds invested in the portfolio at $t-1$. For example, one could define $R_{a,t} = (p'_{a,t} + d_t - p'_{a,t-1})/p'_{a,t-1}$, where d_t is the dividend per share paid during the evaluation period to the portfolio's owners. The more complicated definition, that is, with dividends assumed to be reinvested, is "purer" (especially for the purpose of inter-portfolio comparisons of performance) in the sense that funds invested at the beginning of a period remain invested for the entire period, but it is less pure in the sense that it assumes a reinvestment policy not actually followed in the portfolio.

The next step is to define prices per share for the benchmark or naively selected portfolios that also take account of intraperiod fund flows.

$p_{xt}(\beta_T)$ = price at t per share of the naively selected portfolio with the target risk level. To avoid double-counting of past performance, at the beginning of any evaluation period (for example, just after an evaluation takes place at $t-1$) this price is set equal to the price per share of the actual portfolio. Then this amount is invested in the naively selected portfolio with the target risk level, and the behavior of the market value of this portfolio during the evaluation period determines the end-of-period price per share, $p_{xt}(\beta_T)$. Any intraperiod cash income generated by the securities of this naively selected portfolio is assumed to be reinvested in this portfolio.

These conventions for the treatment of beginning-of-period values and intraperiod cash income will be taken to apply in the definitions of all the benchmark portfolios. Thus

$p_t(R_f)$ = price at t per share of the naively selected portfolio obtained by investing all funds available at $t-1$ in the riskless asset.

The benchmarks provided by $p_{xt}(\beta_T)$ and $p_t(R_f)$ are unaffected by intraperiod fund flows in the actual portfolio. This is not true of the following two benchmarks.

$p_{xt}(\beta_a)$ = price at t per share of the naively selected portfolio with market risk equal to that of the actual portfolio. At the beginning of any evaluation period and after any transaction in the actual portfolio during an evaluation period (that is, after any cash flow or exchange of shares in the actual portfolio) the market risk of the actual portfolio is measured, and the current price per share of this benchmark is shifted into the naively selected portfolio with that level of market risk. Thus the value of β_a could be shifting more or less continuously through time as a result of inflows and outflows of funds and decisions to shift the holdings in the portfolio.[11]

11. Indeed even if there are no transactions taking place, the value of β_a shifts continuously through time as a result of shifts in the relative market values of individual securities in the portfolio. Aside from adjusting the value of β_a at the beginning of each evaluation period, we have chosen to ignore the effects of such "non-discretionary" shifts here.

$p_{xt}(\sigma(\tilde{R}_a))$ = price at t per share of the naively selected portfolio with return dispersion equal to that of the actual portfolio. The definition of $p_{xt}(\sigma(\tilde{R}_a))$ is obtained by substituting $\sigma(\tilde{R}_a)$ for β_a in the definition of $p_{xt}(\beta_a)$ above.

Thus $p_{xt}(\beta_a)$ and $p_{xt}(\sigma(\tilde{R}_a))$ take account of changes in β_a and $\sigma(\tilde{R}_a)$ that result from intraperiod fund flows and portfolio shifts. Computationally, keeping tract of β_a and $\sigma(\tilde{R}_a)$ in the way required for these benchmarks is not a difficult problem. At any point in time the market risk β_a of the chosen portfolio is just the weighted average of the market risks of the individual assets in the portfolio, where the weights are the proportions of total portfolio market value represented by each asset. Thus if one has estimates of the market risks of the assets from which portfolios are chosen, the value of β_a is updated by combining these with current measures of the weights of individual assets in the chosen portfolio. And a similar procedure can be followed with respect to updating values of $\sigma(\tilde{R}_a)$.[12]

B. *Multiperiod Measures of Performance*

Given the beginning and end-of-period prices per share for these benchmark portfolios, their one-period returns are obtained in the usual way. Then the performance history of a portfolio can be built up (for example) through period-by-period application of the breakdowns given by (11)-(13). Alternatively, one can define performance measures in terms of profit per share rather than return. Thus, in line with (13) and using end of evaluation period prices, define

$$\overbrace{[p_{a,t} - p_t(R_f)]}^{\substack{\text{Overall}\\ \text{Performance}}} = \overbrace{[p_{a,t} - p_{xt}(\beta_a)]}^{\text{Selectivity}} + \overbrace{[p_{xt}(\beta_a) - p_{xt}(\beta_T)]}^{\text{Manager's Risk}} + \overbrace{[p_{xt}(\beta_T) - p_t(R_f)]}^{\text{Investor's Risk}}. \quad (18)$$

This type of breakdown can of course be computed both period-by-period and cumulatively. And from such multiperiod histories one can get more reliable measures of a portfolio manager's true abilities than can be obtained from a one-period analysis. For example, one can determine whether his *Selectivity* is systematically positive or simply randomly positive in some periods.

For some purposes one may wish to compare the multiperiod performance histories of different portfolios. For example, an investment company may be interested in the relative abilities of its different security analysts and portfolio managers. Or an investor who has allocated his funds to more than one manager may be interested in comparing their performances. On a period-by-period basis such performance comparisons can be carried out in terms of percentage returns. Alternatively, if the prices of shares in different portfolios

12. Keeping track of $\sigma(\tilde{R}_a)$ is especially simple if one assumes that returns are generated by the so-called "market model." On this, and for additional computational suggestions, see Blume [3, 4, 5].

are set equal at the beginning of comparison periods, profit-based performance measures such as (18) could be computed both on a period-by-period basis and cumulatively.

One must not get the impression, however, that all the problems caused by intraperiod fund flows have been solved. Though the performance of a "share" during any given evaluation period (or across many periods) gives an unambiguous picture of the investment history of funds invested in a given portfolio at a given point in time, comparisons of the performances of shares in different portfolios are not completely unambiguous. This is due to the fact that even when things are done on a per share basis, intraperiod fund flows necessitate portfolio decisions that usually have some effect on the performance of a share. And when such fund flows occur at different times (and thus during different market conditions) in different portfolios, the observed performances of shares in the portfolios may differ, even if the portfolios are managed by the same person trying to follow the same policies in all of his portfolio decisions. But though such ambiguities seem unavoidable and to some extent unsolvable, their effects on performance comparisons should be minor except in cases where portfolios experience large cash flows (relative to their total market values) in short periods of time and/or when evaluation periods are long.

Finally, if an *ex ante* market line is available to compute expected values through time for the three benchmarks, $p_{xt}(\beta_T)$, $p_{xt}(\beta_a)$ and $p_{xt}(\sigma(\tilde{R}_a))$, then the one-period performance breakdowns of (14)-(17) can be carried out either in terms of returns or market values, and these can be used as the basis of even more detailed multiperiod performance histories.

But we terminate the discussion at this point. We do this not because of a lack of additional interesting problems, but because in the absence of actual applications, suggested solutions become increasingly speculative and thus of less likely usefulness.

VI. Summary

Some rather detailed methods for evaluating portfolio performance have been suggested, and some of the more important problems that would arise in implementing these methods have also been discussed. In general terms, we have suggested that the return on a portfolio can be subdivided into two parts: the return from security selection (*Selectivity*) and the return from bearing risk (*Risk*). Various finer subdivisions of both *Selectivity* and *Risk* have also been presented.

To a large extent the suggested models can be viewed as attempts to combine concepts from modern theories of portfolio selection and capital market equilibrium with more traditional concepts of what constitutes good portfolio management.

For example, the return from *Selectivity* is defined as the difference between the return on the managed portfolio and the return on a naively selected portfolio with the same level of market risk. Both the measure of risk and the definition of a naively selected portfolio are obtained from modern capital market theory, but the goal of the performance measure itself is just to test how good the portfolio manager is at security analysis. That is, does he show any ability to uncover information about individual securities that is not already implicit in their prices?

Likewise, traditional discussions of portfolio management distinguish between security analysis and market analysis, the latter being prediction of general market price movements rather than just prediction of the special factors in the returns on individual securities. The various timing measures suggested in this paper provide estimates of the returns obtained from such attempts to predict the market. And modern capital market theory again plays a critical role in defining these estimates.

REFERENCES

1. Bank Administration Institute. *Measuring the Investment Performance of Pension Funds.* Park Ridge, Illinois: B.A.I., 1968.
2. Fisher Black, Michael Jensen, and Myron Scholes. "The Capital Asset Pricing Model: Some Empirical Tests." To appear in *Studies in the Theory of Capital Markets,* edited by Michael Jensen and published by Praeger.
3. Marshall Blume. "The Assessment of Portfolio Performance," unpublished Ph.D. dissertation, University of Chicago, 1968.
4. ———. "Portfolio Theory: A Step Toward Its Practical Application." *Journal of Business* XLIII (April, 1970), 152-173.
5. ———. "On the Assessment of Risk," *Journal of Finance* XXVI (March, 1971), 1-10.
6. Eugene F. Fama. "The Behavior of Stock Market Prices," *Journal of Business* XXXVIII (January, 1965), 34-105.
7. ———. "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* XXV (May, 1970), 383-417.
8. ———. "Multiperiod Consumption-Investment Decisions," *American Economic Review* XL (March, 1970), 163-174.
9. ———. "Portfolio Analysis in a Stable Paretian Market," *Management Science* XII (January, 1965), 404-419.
10. ———. "Risk, Return and Equilibrium: Some Clarifying Comments," *Journal of Finance* XXIII (March, 1968), 29-40.
11. ———. "Risk, Return, and Equilibrium," *Journal of Political Economy* LXXIX (January-February, 1971), 30-55.
12. Irwin Friend, and Marshall Blume. "Measurement of Portfolio Performance under Uncertainty," *American Economic Review* XL (September, 1970), 561-575.
13. Michael Jensen. "The Performance of Mutual Funds in the Period 1945-64," *Journal of Finance* XXIII (May, 1968), 389-416.
14. ———. "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios," *Journal of Business* XLII (April, 1969), 167-247.
15. John Lintner. "Security Prices, Risk, and Maximal Gains from Diversification," *Journal of Finance* XX (December, 1965), 587-615.
16. ———. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics* XLVII (February, 1965), 13-37.
17. Merton Miller, and Myron Scholes. "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings." To appear in *Studies in the Theory of Capital Markets,* edited by Michael Jensen, and published by Praeger.
18. Jan Mossin. "Equilibrium in a Capital Asset Market," *Econometrica* XXXIV (October, 1966), 768-783.
19. Richard Roll. *The Behavior of Interest Rates: Application of the Efficient Market Model to U.S. Treasury Bills.* New York: Basic Books, Inc., 1970.
20. William F. Sharpe. "Capital Assets Prices: A Theory of Market Equilibrium under Conditions of Risks," *Journal of Finance* XIX (September, 1964), 425-442.
21. ———. "Mutual Fund Performance," *Journal of Business* XXXIX (Special Supplement, January, 1966), 119-138.
22. ———. "Risk Aversion in the Stock Market," *Journal of Finance* XX (September, 1965), 416-422.
23. Jack L. Treynor. "How to Rate Management of Investment Funds," *Harvard Business Review* XLIII (January-February, 1965), 63 75.