Measurement of Portfolio Performance Under Uncertainty

By IRWIN FRIEND AND MARSHALL BLUME*

Harry Markowitz's pioneering work in portfolio theory and James Tobin's subsequent extension forced a complete reevaluation of the pricing of capital assets under uncertainty. A by-product of this reevaluation was a theory of equilibrium in the capital markets, which was independently discovered by William Sharpe (1964), John Lintner (1965a), and Jack Treynor (1965). The theory, which in this paper will be called the "market-line" theory, led to several different, although related, one-parameter measures of the investment performance of an asset or a portfolio.

The entire rationale of one-parameter measures of investment performance is to replace two-parameter measures of performance—rate of return and risk—with a single measure which uses market data to combine the two different dimensions of performance into a single measure which adjusts for differences in risk. A single risk-adjusted measure of performance is not only simpler than a combination of risk and return measures, but permits, at least theoretically, a definitive comparison of performance of investments with different returns and risks. It is not surprising, therefore, that one-parameter measures of performance are receiving more and more attention, but it is surprising that there has been virtually no statistical analysis of the extent to which the risk-adjusted rates of return successfully abstract from risk.

The usefulness of these one-parameter measures depends, of course, upon the validity of the asumptions underlying the market-line theory. Section I of this paper briefly reviews this theory, including the assumptions on which it is based, and discusses the different one-parameter measures of performance that have been derived from the theory. Section II examines the adequacy of the one-parameter measures of performance by measuring empirically the relationship between these measures and the risk from which they are supposed to abstract. It then attempts to explain the apparent discrepancies between the market-line theory and the empirical findings in terms of specific deficiencies in the underlying assumptions.

The importance of these one-parameter measures of performance, and the associated theory of equilibrium in the capital markets, lies not only in their usefulness for analyzing investment management and market efficiency—areas of investigation to which they have already been applied—but also in their relevance and potential utility for cost of capital problems.[1] It is, therefore, essential that a careful investigation be made of the validity of these

* The authors are professor of economics and finance and associate professor of finance, respectively at the University of Pennsylvania. They wish to thank Deno Papageorge, now of the Bank of New York, for his statistical assistance and the Oppenheimer Foundation for its financial support.

[1] Since it is typically easier to estimate the cost of capital for the market as a whole than for an individual corporation, the reasonableness of the estimated cost of capital for a corporation can be tested against the equilibrium relation between the rate of return on an individual security and the rate of return for the market. See equation (1) in Section I where $E(\tilde{R}_i)$ is a measure of the "cost" of equity on an individual security and $E(\tilde{R}_m)$ the cost of equity for stocks as a whole.

measures—a goal towards which this paper is a small first step.

I. *One-Parameter Performance Measures and Their Rationale*

The key assumptions underlying the market-line theory are 1) Every investor is a one-period expected utility maximizer and exhibits diminishing marginal utility of terminal wealth.[2] 2) All investors have the same one-period time horizon. 3) Every investor feels that he can evaluate a portfolio solely in terms of the mean and dispersion or variance of one-period returns.[3] 4) It is assumed that there are no transaction or information costs, that the borrowing and lending rates are equal and the same for all individuals, and that investors will only select portfolios with optimal combinations of risk and return. 5) All investors hold identical or homogeneous expectations about the distributions of future returns. 6) The capital market is in equilibrium.[4]

Under these assumptions, Sharpe and Lintner[5] have shown that the expected return for asset or portfolio i, $E(\tilde{R}_i)$, is related to the expected return on the market portfolio,[6] $E(\tilde{R}_m)$, by the equation

$$(1) \qquad E(\tilde{R}_i) - R_f = \beta_i[E(\tilde{R}_m) - R_f]$$

where R_f is the risk-free rate for borrowing or lending and β_i is defined as the $\text{Cov}(\tilde{R}_i, \tilde{R}_m)$ divided by $\text{Var}(\tilde{R}_m)$. The variable β_i is a measure of systematic or nondiversifiable risk. The tilde superscript indicates a random variable.

If the equilibrium relationship contained in equation (1) held for all assets on an *ex ante* basis, there would be no opportunity for abnormal profit: All assets would be correctly priced. Only if some of the above assumptions did not strictly hold for all investors and for all securities could there be incorrectly priced securities. In reality, these assumptions are not likely to hold completely, but equation (1) may, nonetheless, be an adequate approximation of reality for most securities. Yet, to explicitly recognize that not all securities are in equilibrium, equation (1) can be rewritten as

$$(2) \qquad E(\tilde{R}_i) - R_f = \eta_i + \beta_i[E(\tilde{R}_m) - R_f]$$

where η_i is a measure of disequilibrium. If η_i equals zero, the portfolio or asset is in equilibrium. If η_i is greater than zero, the expected return is larger than one would anticipate on the basis of the equilibrium relationship. In the market terminology, this would represent an undervalued security. If η_i is less than zero, the security is overvalued.

[2] Eugene Fama (1970) has shown under perfect capital markets and very general assumptions about consumer behavior that an individual who actually faces a multiperiod decision would act as if he followed assumption 1).

[3] The adequacy of this assumption hinges upon how closely the individual's subjective distributions of returns can be described by two-parameters. The development in this paper assumes that the variance of these distributions is defined, which is the traditional assumption. Michael Jensen (1969) shows that his one-parameter measure can be developed with the variance undefined.

The empirical evidence of this approximation to a two-parameter distribution is conflicting: Fred Arditti finds for individual securities on the New York Stock Exchange (*N.Y.S.E.*) that at least three parameters are necessary to describe these distributions, but the empirical evidence of Blume suggests that the distribution of returns for well-diversified portfolios of *N.Y.S.E.* stocks can be very accurately approximated by two-parameter stable distributions. This latter finding is, of course, more relevant to an investor who would typically hold a well-diversified portfolio.

[4] There is a substantial body of literature pertaining to the question of equilibrium. For a recent bibliography, the reader is referred to Jensen (1969).

[5] Fama (1968) has recently shown that the two developments are for all essential purposes identical.

[6] The market portfolio is defined theoretically as the portfolio consisting of all wealth whose return is uncertain. Wealth here is construed quite broadly. Yet, in every known empirical use of the model, the return on the market portfolio has been measured by some index of common stocks listed on the *N.Y.S.E.* This practice will be followed in this paper in the absence of feasible alternatives.

The one-parameter measures of investment performance of Sharpe (1966), Treynor (unpublished), and Jensen (1968) follow from equation (2). The simplest measure is Jensen's, which is merely η_i. Since this measure is in the same units as $E(R_i)$ or R_f, which are rates of returns, it can be interpreted as the rate of return above and beyond that justified by the equilibrium relationship (1), an easily interpretable measure of investment performance.

Treynor's measure follows from (2) by dividing both sides of the equation by β_i to obtain

$$(3) \quad \frac{E(\tilde{R}_i) - R_f}{\beta_i} = \frac{\eta_i}{\beta_i} + [E(\tilde{R}_m) - R_f]$$

The ratio on the left of the above equation is Treynor's measure. If η_i equals zero, Treynor's measure is equal to $[E(\tilde{R}_m) - R_f]$, a term which is independent of the level of the systematic risk β_i. Further insight into Treynor's measure follows from the rewriting of (3) as

$$(4) \quad \frac{\eta_i}{\beta_i} = \frac{E(\tilde{R}_i) - R_f}{\beta_i} - [E(\tilde{R}_m) - R_f]$$

Since $E(R_m) - R_f$ is a constant, Treynor's measure is merely a translation of Jensen's measure divided by the systematic risk β_i. The dimension of Treynor's measure is therefore return per unit of systematic risk.

The derivation of Sharpe's measure from equation (2) proceeds by replacing β_i by its definition to obtain

$$(5) \quad E(\tilde{R}_i) - R_f = \eta_i + \frac{\mathrm{Cov}\,(\tilde{R}_i, \tilde{R}_m)}{\sigma^2(\tilde{R}_m)} [E(\tilde{R}_m) - R_f]$$

or noting that $\mathrm{Cov}(\tilde{R}_i, \tilde{R}_m) = \rho(\tilde{R}_i, \tilde{R}_m) \cdot \sigma(\tilde{R}_i)\sigma(\tilde{R}_m)$,

$$(6) \quad E(\tilde{R}_i) - R_f = \eta_i + \frac{\rho(\tilde{R}_i, \tilde{R}_m)\sigma(R_i)}{\sigma(\tilde{R}_m)} [E(\tilde{R}_m) - R_f]$$

If portfolios i are only taken to be efficient portfolios, Sharpe (1964) has shown that the market-line theory implies that $\rho(\tilde{R}_i, \tilde{R}_m)$ equals one. Noting this property and dividing by $\sigma(\tilde{R}_i)$, one has

$$(7) \quad \frac{E(\tilde{R}_i) - R_f}{\sigma(\tilde{R}_i)} = \frac{\eta_i}{\sigma(\tilde{R}_i)} + \frac{E(\tilde{R}_m) - R_f}{\sigma(\tilde{R}_m)}$$

The ratio on the left of (7) is Sharpe's measure. Similarly to Treynor's measure, Sharpe's is a translation of Jensen's measure divided by the standard deviation of return. An important difference between Sharpe's measure and the previous two is that whereas Jensen's and Treynor's measures can be used for any portfolio as well as individual securities. Sharpe's measure should only be applied to portfolios which are purported to be efficient.

These three one-parameter measures of investment performance were developed in terms of *ex ante* values, which at first glance might present a difficult problem in using the measures to evaluate performance. Jensen, however, has shown that at least for his measure it is possible to obtain unbiased estimates of η_i, providing β_i and R_f are constant over time.[7] If these two parameters are constant, (2) can be rewritten in *ex post* or historical data as

$$(8) \quad R_{it} - R_f = \eta_i + \beta_i[R_{mt} - R_f] + \epsilon_{it}$$

where R_{it} is the return for portfolio or asset i in period t, R_{mt} is the market return in period t, and ϵ_{it} is a disturbance term whose expectation is zero and which is independent of R_{mt}.[8] The constant η_i

[7] Jensen (1968) examines the biases in the estimates of η_i for a specific type of non-stationarity in the risk measure.

[8] Because the market return includes asset i, the disturbances cannot be independent of R_{mt}. For a

can be estimated by standard regression techniques.

Although (8) follows from (2) providing the risk free rate is constant, Jensen (1968) in some of his empirical work allows the risk free rate to change over time, so that R_f in (8) would be subscripted by t.

Assuming again that the risk free rate is constant, one can use (8) to derive a consistent estimate of Treynor's measure using *ex post* data. If T observations are used in estimating the parameters and if $\hat{\eta}_i$, $\hat{\beta}_i$, and $\hat{\epsilon}_{it}$ are the corresponding least-squares estimates, equation (8) can be summed over T and averaged to obtain

$$(9) \qquad \overline{R_i} - R_f = \hat{\eta}_i + \hat{\beta}_i[\overline{R_m} - R_f]$$

where the bar indicates an average. Upon dividing through by $\hat{\beta}_i$, one has on the left a consistent estimate of Treynor's index.

Sharpe's measure also follows from (9) if, in addition, one assumes that $\sigma(R_i)$ is constant over time. The mathematical development is virtually the same as that which was used in deriving (7) from (2).

In the empirical part of this paper, the three performance measures will be estimated using monthly data. In estimating Sharpe's and Treynor's measures, the risk-free rate will be estimated as the average risk-free rate over the sample period. In estimating Jensen's, the risk-free rate will be allowed to vary over time following his procedures. Rate of return will be measured in two ways: the monthly investment relative, and the continuously compounded rate of return. The monthly investment relative can be defined as the wealth at the end of the month resulting from a one dollar investment at the beginning of the month with dividends reinvested. The continuously compounded rate of return is just the natural logarithm of the investment relative. Jensen argues that this is the appropriate measure of return if the market has an infinitesimal time horizon.

II. *Theoretical vs. Empirical Relationship of One-Parameter Performance Measures*

Theoretically, it would be expected that for random portfolios the Sharpe, Treynor, and Jensen one-parameter measures of performance would be independent of the corresponding measures of risk unless one or more of the following four conditions holds:

1) The assumptions underlying the market-line theory are invalid, i.e., are not realistic approximations of the real world, even for the *ex ante* magnitudes to which the theory applies.
2) The *ex post* distributions of return and values of risk differ substantially from their *ex ante* magnitudes.
3) Measurement errors, especially in the risk variables, result in biased estimates of the relationship between performance and risk.
4) There are in fact real systematic differences among the risk-adjusted performances of portfolios characterized by different degrees of risk.

The last possibility does not seem meaningful unless there are appreciable differences between *ex post* and *ex ante* magnitudes. If random portfolios do exhibit over very long periods of time significant dependencies between the one-parameter measures of performance and risk, the first and third conditions are more likely than the second to explain this result since *ex post* magnitudes would not be expected to deviate from their *ex ante* values indefinitely though such deviations are possible for extended periods.

The empirical analysis to follow suggests

mathematical discussion, the reader is referred to Jensen (1969) who indicates that the resulting bias is extremely small.

that the invalidity of a key assumption in the market-line theory does bias systematically the one-parameter measures of performance in all periods, while discrepancies between *ex post* and *ex ante* values (and perhaps also the invalidity of other assumptions) affect these measures of performance in different ways depending on the underlying market conditions. Measurement errors apparently do not substantially bias the estimates of the relationship between performance and risk.

To examine the relationship of one-parameter performance measures to risk, both performance and risk measures were derived for 200 random portfolios. These portfolios were selected from the universe of 788 common stocks listed on the New York Stock Exchange (*N.Y.S.E.*) throughout the period January 1960 through June 1968.[9] These 200 random portfolios consist of 50 individual portfolios of 25 securities and a like number for portfolios of 50, 75, and 100 stocks. An equal investment is assumed in each stock. A stratified random sampling procedure was used to insure that the entire spectrum of risk would be covered.[10]

Three basic performance measures computed for these 200 random portfolios were regressed against each of two measures of portfolio risk. The performance measures used were those developed by Sharpe, Treynor, and Jensen described in Section I. The two measures of risk were *Beta* coefficients, the covariance of portfolio and market return divided by the variance of market return, and the standard deviation of portfolio return. The first risk measure is implicit in the measures of Treynor and Jensen, and the second in Sharpe's. These regressions were derived using the performance and risk measures calculated with both the monthly investment relatives and the continuously compounded returns or the natural logarithms of the monthly relatives. Table 1 presents the resulting 12 regressions for the entire period January 1960 through June 1968. Scatter diagrams were also plotted for each of these 12 regressions, but only one is presented in view of space considerations (Figure 1).[11]

The results are striking. In all cases, risk-adjusted performance is dependent upon risk: The relationship is inverse and highly significant.[12] While rate of return is normally found to be positively related to risk, the adjustment of the rate of return for risk which would be expected to eliminate this relationship actually reverses it. For cross-sectional data, the correlations are quite high. The correlations for the *log relatives* are higher than the other correlations, with the differences particularly large for the Jensen regressions. The highest correlation is associated with the regression of Jensen's performance

[9] The monthly data for individual securities were obtained from University of Chicago updated tapes. These monthly measures of return reflect all capital gains as well as dividends and are adjusted for all capital changes (i.e., stock dividends, splits, etc.) as described in Lawrence Fisher and James Lorie. The market return refers to Fisher's combination link relatives, as updated for this study. These relatives assume an equal investment in each of the *N.Y.S.E.* stocks. The risk-free rate, also required to estimate the one-parameter performance measures, is the three-month yield on Treasury bills adjusted to a one-month basis.

[10] The *N.Y.S.E.* securities were ranked in ascending order according to the values of the covariances of the monthly security investment relatives and Fisher's link relatives over the entire sample period. A portfolio of 25 securities, one of 50, one of 75, and one of 100, were drawn at random from the first 200 securities. Four more portfolios of the four different sizes were drawn from the ranked 13th through 212th securities. This process was repeated again and again, each time increasing the bounds of the stratum by 12 securities until 200 portfolios were obtained.

[11] In view of the non-linearity observed in the scatter diagrams and the possible dependencies among portfolios, the regressions and t-values in Table 1 should be regarded as rough approximations.

[12] Section I contains formulae for the expected values of the constant terms in the theoretical relationships. The expected value is zero for the Jensen relationships and greater than zero for the other two.

TABLE 1—REGRESSIONS OF ONE-PARAMETER PERFORMANCE MEASURES ON RISK
Random Portfolios, January 1960–June 1968

Performance Measure	=	a	+	b Risk Measure	$\overline{R}^2$	Standard Error
1) S	=	0.2677 (45.50)	−	0.0557 X_1 (−9.17)	0.2944	0.0215
2) S	=	0.2724 (42.00)	−	1.3614 X_2 (−9.01)	0.2871	0.0216
3) T	=	0.0134 (45.10)	−	0.0039 X_1 (−12.82)	0.4510	0.0011
4) T	=	0.0136 (39.82)	−	0.0921 X_2 (−11.59)	0.4012	0.0011
5) J	=	0.0028 (11.34)	−	0.0018 X_1 (−7.13)	0.2004	0.0009
6) J	=	0.0029 (10.42)	−	0.0429 X_2 (−6.61)	0.1768	0.0009
7) S'	=	0.2648 (44.57)	−	0.0741 X_1' (−12.04)	0.4199	0.0214
8) S'	=	0.2714 (41.49)	−	1.8356 X_2' (−11.91)	0.4146	0.0215
9) T'	=	0.0130 (44.98)	−	0.0046 X_1' (−15.29)	0.5391	0.0010
10) T'	=	0.0133 (39.92)	−	0.1097 X_2' (−13.99)	0.4946	0.0011
11) J'	=	0.0345 (31.04)	−	0.0311 X_1' (−27.03)	0.7857	0.0040
12) J'	=	0.0372 (29.80)	−	0.7698 X_2' (−26.16)	0.7745	0.0041

Note: S represents the Sharpe measure of performance; T represents the Treynor measure of performance; J represents the Jensen measure of performance; X_1 represents the *Beta* coefficient of a random portfolio; X_2 represents the standard deviation of portfolio return; $\overline{R}^2$ represents the coefficient of determination adjusted for degrees of freedom.

The figures in parentheses are t-values. The unprimed variables are calculated using monthly relatives; the primed variables using the logarithm of the monthly relatives.

measure using *log relatives*, the variant preferred by Jensen, on the corresponding *Beta* coefficient. The large values of the coefficients of determination adjusted for degrees of freedom would suggest if these performance measures are valid that, at least for the period covered, the best way of ensuring good performance was to select a non-risky portfolio: In other words, risk is a good inverse proxy for performance. Not only is the proportion of variance in performance explained by risk very high in these relationships, but the implied magnitude of the impact of variations in risk on performance is sizable.

It has not proved feasible to extend our random portfolios back before 1960, but the performance measures for 115 mutual funds over the period 1945–64 presented by Jensen in the May 1968 issue of the *Journal of Finance* provide some useful insights into the earlier period. There have been many analyses indicating that, at least until recent years, the average performance of mutual funds in their stock investment has not deviated in any important way from that of random portfolios of *N.Y.S.E.* securities, although the funds have not confined themselves to investments in *N.Y.S.E.* issues.

The regression of the 1945–64 performance measures for these 115 funds on

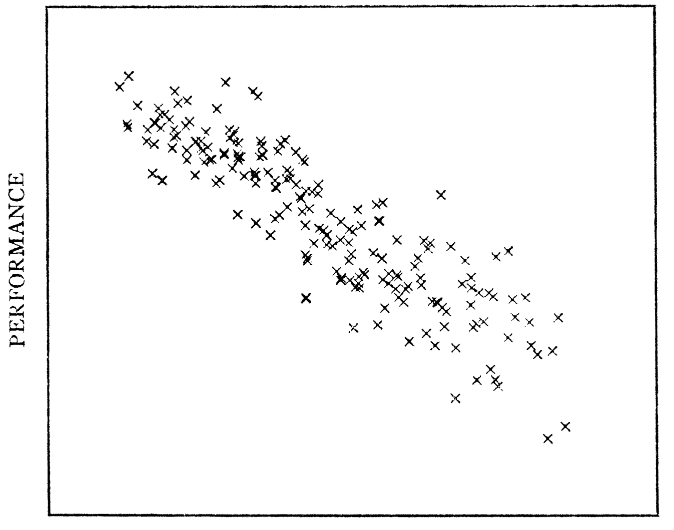

FIGURE 1. SCATTER DIAGRAM OF JENSEN'S PERFORMANCE MEASURE* ON RISK.
January 1960—June 1968
* Using *log relatives*.

the corresponding *Beta* coefficients[13] is

$$J' = .0399 - .0606X' \qquad \bar{R}^2 = .39$$
$$\qquad (6.48) \quad (-8.57)$$

where J' is Jensen's performance measure using *log relatives* rather than the relatives themselves, X' is the associated *Beta* coefficient, the variables are estimated from annual rather than monthly data, the terms in parentheses are t-values, and $\bar{R}^2$ is the coefficient of determination adjusted for degrees of freedom.

Again the data over a long period of time indicate that performance and risk are strongly inversely correlated, with the riskiest portfolios performing very much worse than the less risky portfolios. Jensen (1969) shows that the performance funds have the highest *Beta* coefficients and the balanced funds the lowest coefficients, suggesting that these coefficients may serve as a reasonable proxy for risk. Neither Jensen, nor Donald Farrar who commented on Jensen's paper, made any reference to the apparent bias against the riskier funds over this period in the Jensen one-parameter measure of performance.

The first question that must be answered, in attempting to explain this apparent bias of one-parameter measures of performance against the riskier funds, is whether the bias is purely statistical in origin reflecting random errors in measurement of the risk variables which appear on both sides of the regression.[14] It can be shown, however, that if errors involved in measuring the risk variables are random (either with constant variance or with variance proportional to the square of risk), the estimate of the regression coefficient of the risk variable in the Jensen regression is biased downward, but the

[13] These coefficients were obtained from Jensen (1968, 1969).

[14] The risk variables appear in the denominator of the dependent variable and simultaneously as the independent variable in the Sharpe and Treynor regressions, and with opposite signs on each side of the Jensen regression.

magnitude of the bias is so small that it can be ignored.[15] Similarly, the estimate of the slope coefficient in the Sharpe and Treynor measures is asymptotically downward biased,[16] but again the likely magnitude of the bias is such that it could not account for the results in Table 1. Therefore, the assumptions underlying the market theory will be examined for further insights into the apparent bias in these performance measures.

[15] If the expected value of the coefficient of the risk variable in the Jensen regressions is b' and the estimated value is b, then it can be shown that the expected value of b

$$E(b) = \frac{b' - k\dfrac{\sigma_\epsilon^2}{\sigma_x^2}}{1 + \dfrac{\sigma_\epsilon^2}{\sigma^2}}$$

where k is the difference between the rate of return on the market (R_m) and the riskless rate of return (R_f), ϵ is the measurement error in x, and x is the measure of risk. For the random portfolios using investment relatives, the variance of the measurement error, σ_ϵ^2, is estimated to be 0.0009883; σ_x^2 is 0.06198, and k is roughly .008. Therefore, if $b'=0$, the bias in its estimate is approximately -0.000125, which in absolute value is considerably less than the actual estimate -0.0557. For Jensen's sample of mutual funds, estimates of σ_ϵ^2 and σ_x^2 are, respectively, 0.0092 and 0.0413 and k is at most 0.16, so that if $b'=0$, the bias is approximately -0.02915 which is compared to the actual estimate of -0.0606. Thus, the bias for the random portfolios is trivial, and for Jensen's mutual funds can only explain about half of the actual estimate.

[16] For example, in the Treynor regressions, if $y_i = R_i - R_f$ and $x_i = \beta_i$, the risk coefficient, the correct relationship between x and y (dropping the subscripts) will be

$$\frac{y}{x} = a' + b'x + v$$

where a' and b' would be the regression coefficients without measurement error in x, and v is a random disturbance independent of x and with zero expectation. If, however, x is measured with random error ϵ, independent of x and with zero expectation, the estimated regression will be

$$\frac{y}{x+\epsilon} = a + b(x+\epsilon) + \nu$$

where a and b are the estimated coefficients and ν a random disturbance independent of x and ϵ and with zero expectation.

Then, if n is the number of observations, the estimated coefficient b will be

$$\frac{b'[\Sigma(x-\bar{x})^2] + n\bar{x}\left(\overline{\dfrac{y}{x}} - \overline{\dfrac{y}{x+\epsilon}}\right) - n\bar{\epsilon}\,\overline{\dfrac{y}{x+\epsilon}}}{\Sigma(x-\bar{x})^2 + \Sigma(\epsilon-\bar{\epsilon})^2 + 2\Sigma(x-x)(\epsilon-\bar{\epsilon})}$$

In the limit, the last term in the numerator will be zero because the limit of a product of sequences is the product of the limits. The last term in the demoninator will be zero because x and ϵ are assumed independent.

The middle term in the numerator will probably be negative under the observed distribution for x and ϵ. (The estimated risk measures, $x+\epsilon$, were always positive and it seems reasonable to assume that the values of x are always positive.) If it is, assuming that the measurement error ϵ is symmetrically distributed about zero and in absolute value is less than the minimum value of x, which seems appropriate for the random portfolios, the following inequality holds for any x:

$$\frac{1}{x(x+\epsilon)} < \frac{1}{x(x-\epsilon)}, \qquad \epsilon > 0$$

Since (ϵ) and (-ϵ) are equally likely,

$$\mathrm{E}\left(\frac{\epsilon}{x(x+\epsilon)}\,\Big|\, x\right)$$

will be negative for all x, so that

$$E\left(\frac{\epsilon}{x(x+\epsilon)}\,\Big|\, x\right)$$

$=h$ will be negative. Under the market-line theory, $E(y) = E(R_m) - R_f = k$ is greater than zero, which implies in the limit that

$$\bar{x}\left(\overline{\frac{y}{x}} - \overline{\frac{y}{x+\epsilon}}\right) = khE(x) < 0$$

or that in the limit

$$b = \frac{b' + \dfrac{khE(x)}{\sigma_x^2}}{1 + \dfrac{\sigma_\epsilon^2}{\sigma_x^2}}, \qquad h < 0$$

which yields a downward biased estimate of b'.

An estimate of the bias for the regression of Treynor's measure calculated with the investment relatives on the *Beta* coefficient (Table 1, equation 3) was derived. The expected value of x, $E(x)$, was estimated as 0.934. The estimates of σ_ϵ^2 and σ_x^2 are the same as those given in the preceding footnote. Assuming x and ϵ are normally distributed with means of 0.93 and 0.0 and with variances as given above, a simulation using 10,000 drawings yielded -0.00114 as an estimate of h. Thus, the bias is estimated as -0.000136, which should be compared to -0.0039, so that the bias is negligible. In view of the similarity between Treynor's and Sharpe's measure as well as the empirical results in Table 1, the relative bias in the Sharpe regressions is likely also to be trivial.

Of the key assumptions underlying the market theory leading to one-parameter measures of performance, the one which most clearly introduces a bias against risky portfolios is the assumption that the borrowing and lending rates are equal and the same for all investors. Since the borrowing rate for an investor is typically higher than the lending rate, the assumption of equality might be expected to bias the one-parameter measures of performance against risky portfolios because, for such portfolios, investors do not have the same option of increasing their return for given risk by moving from an all stock portfolio to an investment with additional stock financed with borrowings at the lending rate.

An examination of the scatter diagrams of the one-parameter performance measures against the risk measures confirms the strong inverse correlation between "risk-adjusted performance" and risk indicated by the regression in Table 1. The scatters show a fairly steady decline in these performance measures throughout the observed range of *Beta* coefficients with some evidence of a tapering in the rate of decline for the coefficients in excess of one (e.g., see Figure 1). The absolute values of the performance measures are in excess of market expectations for funds with *Beta* coefficients below one and below expectations for higher coefficients. These findings suggest that an "optimal" portfolio consisting of positive investments in both a risky portfolio and a riskless asset does not contain the market portfolio as its risky component, contrary to the usual assumption.[17] The risky portfolio involved in such optimal combinations appears to be considerably less risky than the market portfolio. If the difference between the borrowing and lending rates is sufficiently large, only risky assets would be held in portfolios with *Beta* coefficients beyond some low value. However, with more moderate differences between the borrowing and lending rates, both risky and riskless assets may be held, even in portfolios with high Beta coefficients.

Of the other departures from the perfect market assumptions of the market-line theory, none seems likely to introduce a substantial bias against risky portfolios. Differential taxation of capital gains and dividend income would tend to make a unit of before-tax return on risky portfolios larger on an after-tax basis than an equivalent before-tax return on less risky portfolios. The difference in after-tax return of the random portfolios for given before-tax return is likely to be quite small, however, since they are confined to *N.Y.S.E.* stocks. At a maximum, this difference between random portfolios with relatively small and those with relatively large *Beta* coefficients is likely to be not much over .3 of 1 percent annually,[18] and the difference would be negligible for portfolios with only moderate variations in these coefficients. Differential transaction and information costs might also affect moderately the comparison of performance of risky vs. less risky portfolios, but as discussed subsequently any bias which may exist from this source favors risky portfolios rather than the reverse.

[17] If borrowing and lending rates were equal, then it would be expected that the market line would be tangent to the efficient set at the point representing the market portfolio; but if the borrowing and lending rates are not equal, the market line (from the lending rate to the efficient set of risky securities) which represents optimal combinations of lending at the risk free rate and investing in a risky portfolio may touch the efficient set at a point representing higher or lower risk than that of the market portfolio.

[18] This estimate assumes that the average dividend yield on the risky portfolio is as much as 2 percent lower than on other portfolios which is offset by 2 percent more price appreciation; the former is assumed to be subject to a 15 percent higher tax rate than the latter. It should be pointed out that there is a significant though small group of corporate investors for whom the capital gains tax is higher than the tax on dividend income.

Still another limitation in the market-line theory which might help to explain at least part of the observed results is the assumption that it is possible to stipulate a holding period or planning horizon over which it is planned to hold all assets. If this is not possible, then risk-free assets may be desired for liquidity purposes[19] and their overall rate of return to the investor may be understated by the market rate. Such an understatement would tend to bias the one-parameter measures of performance downward for high risk portfolios (*Beta* coefficients higher than one) and upward for low risk portfolios. However, for realistic values of the understatement of the effective risk-free rate, the resulting biases would typically explain only a small part of the deviations of the one-parameter measures from their expected values.[20]

The only other explanation which comes to mind for the apparent bias against risky portfolios over the 1945–64 and 1960–68 periods is a difference between *ex ante* and *ex post* magnitudes, such that the *ex post* returns for risky portfolios in both periods were lower and the *ex post* risk higher than the respective *ex ante* values. However, as pointed out in Section I, Jensen has shown that if the risk measure *Beta* and the risk-free rate are constant over time, differences between *ex ante* expectations of returns and *ex post* realizations are irrelevant. The values of *Beta* for the random portfolios are remarkably constant over time,[21] and as shown by Jensen (1969), these same measures are reasonably stationary over time for his mutual funds. In addition, it seems intuitively highly unlikely that investors underestimated the risk of risky portfolios at the beginning of either of these periods since the whole evolution of the investment climate in the following years was a gradual realization that cyclical risks were no longer as great as they had been in our earlier history. Finally, the observed variations in the risk-free rate are likely to introduce only trivial biases.

Even though the risk measure *Beta*, which is really a measure of covariation with respect to general market movements, can be assumed to be stationary over time at least for the random portfolios, the performance measures may yet be biased because Jensen in his proof assumes that the *ex post* return on an individual security can be explained by two orthogonal factors: a market factor common to all securities and a unique factor. This is of course a simplistic view of the determination of returns. There are certainly industry factors, i.e., factors affecting a subset of securities, and possibly factors affecting stocks with different *Beta* coefficients. If the *ex post* values of these factors were such that the *ex post* return for portfolios with high *Beta* coefficients were lower than their *ex ante* values, the observed bias would result. However, the relationships between the average return on risky stocks and on other stocks in the 1945–64 and 1960–68 periods were somewhat more favorable to the risky stocks than in the preceding periods, 1926–45 and 1945–60, respectively.[22]

At this stage of our analysis, therefore, it appears that the unreality of the assumption of equal borrowing and lending rates is the most important factor explaining the bias of existing one-parameter measures of

[19] See Reuben Kessel.

[20] Thus, the bias in the Jensen measures resulting from an understatement of ΔR_f in the risk-free rate is $(1-\beta_i)\Delta R_f$ where β_i is the *Beta* coefficient for the i^{th} portfolio. At most, the value of ΔR_f might be on the order of 0.1 percent (per month). For $\Delta R_f = 0.1$ percent, the bias for a portfolio with a low β_i say of 0.48 would be 0.00052, which might be compared with corresponding Jensen performance measure of 0.0022.

[21] The correlation between the *Betas* of the same portfolio for 1960–64 and 1964–68 was .96.

[22] These comparisons are based on data discussed in Friend and Paul Taubman. The data in that paper for the years 1926–60 have since been updated.

performance against risky portfolios over the extended periods of time analyzed.

Further insight into the relative importance of differences between borrowing and lending rates and of differences between *ex ante* and *ex post* magnitudes can be obtained by breaking down our longer period into shorter spans of time, segregating in particular those recent years when the market seemed to favor speculative issues to an unusual degree. We have broken down the period of January 1960 through June 1968 into two equal intervals: January 1960 through March 1964, and April 1964 through June 1968, the latter corresponding to a period of speculative fervor and booming prices of risky issues.

Tables 2 and 3 present regressions for each of the two sub-periods, January 1960 through March 1964, and April 1964 through June 1968, corresponding to those presented in Table 1 for the period as a whole. Figures 2 and 3 present the corresponding scatter diagrams. For the first of these two sub-periods, most of the same tendencies characterizing the period as a whole are observed but in even stronger form (Table 2 and Figure 2). The negative correlations between the performance and risk measures are extremely high, and the effects of the impact of variations in risk on performance, implied by the slope coefficients, are very substantial.[23] For the most recent period, however, the situation is reversed (Table 3 and Figure 3). The correlations between performance and risk measures are significantly positive and the slope coefficients fairly sizable though all eight slope coefficients and six of the eight correlations[24] are much lower than in the earlier period.

The most plausible explanation of the new results for April 1964 through June 1968 seems to be that in this interval the *ex post* returns for risky portfolios were higher than the respective *ex ante* values conditional on the general market factor[25] or that the *ex post* risks were lower than their *ex ante* values, and that these differences between *ex post* and *ex ante* magnitudes more than offset the normal bias operating in the opposite direction as a result of differences between borrowing and lending rates. There is a substantial body of evidence pointing to unanticipated high returns on risky issues in recent years as reflected in the much greater upsurge of price-earnings ratios on such securities than on the less risky issues. This upsurge in price-earnings ratios and hence in total return probably reflected a changed risk valuation of the riskier issues fully as much as higher than expected earnings. This changed risk valuation may have partially reflected a growing recognition both of the reduction in cyclical risks and of the potential for reducing risk through diversification. Whether such higher returns are likely to be retained in more normal periods might be questioned for reasons which we are developing in a forthcoming study, but this is not relevant to the subject of the present paper.[26]

[23] The correlations for the *log relatives* are no longer higher than the other correlations, and in the Jensen regressions the former sizable differences are reversed.

[24] The exceptions are the two Jensen regressions using the *log relatives*.

[25] Even though the values of *Beta* show little variation over time, *ex post* and *ex ante* values conditional on the market may very well differ if, as seems quite plausible, there are factors besides market and unique factors determining the returns of individual securities. This argument was developed above.

[26] Two possible additional reasons for the changed results from the earlier years to the more recent period should be mentioned: First, instead of a single market line there may exist segmented market lines for investors with greatly different tastes in risk, and a much higher proportion of investable funds may have been flowing in recent years to investors (e.g., the "performance" funds) with relatively little risk aversion. This explanation, which is perhaps less plausible than that previously presented, would be inconsistent with the assumption of identical or homogeneous expectations implicit in one-parameter performance measures and would again cast doubt on their validity.

Second, transactions costs might be expected to bias

TABLE 2—REGRESSIONS OF ONE-PARAMETER PERFORMANCE MEASURES ON RISK
Random Portfolios, January 1960–March 1964

Performance Measure	=	a	+	b Risk Measure	$\bar{R}^2$	Standard Error
1) S	=	0.4008	−	0.2494 X_1	0.7534	0.0331
		(40.22)		(−24.68)		
2) S	=	0.4102	−	5.7307 X_2	0.7365	0.0342
		(37.98)		(−23.60)		
3) T	=	0.0188	−	0.0119 X_1	0.7478	0.0016
		(38.92)		(−24.31)		
4) T	=	0.0192	−	0.2722 X_2	0.7225	0.0017
		(36.10)		(−22.78)		
5) J	=	0.0105	−	0.0088 X_1	0.7145	0.0013
		(27.12)		(−22.34)		
6) J	=	0.0109	−	0.2020 X_2	0.7029	0.0013
		(26.22)		(−21.72)		
7) S'	=	0.4016	−	0.2727 X_1'	0.7847	0.0328
		(40.35)		(−26.95)		
8) S'	=	0.4123	−	6.3085 X_2'	0.7695	0.0339
		(38.19)		(−25.80)		
9) T'	=	0.0187	−	0.0128 X_1'	0.7787	0.0016
		(39.18)		(−26.48)		
10) T'	=	0.0191	−	0.2956 X_2'	0.7562	0.0016
		(36.45)		(−24.87)		
11) J'	=	0.0686	−	0.0389 X_1'	0.5397	0.0082
		(27.45)		(−15.31)		
12) J'	=	0.0702	−	0.9021 X_2'	0.5330	0.0083
		(26.63)		(−15.10)		

Note: See Table 1.

TABLE 3—REGRESSIONS OF ONE-PARAMETER PERFORMANCE MEASURES ON RISK
Random Portfolios, April 1964–June 1968

Performance Measure	=	a	+	b Risk Measure	$\bar{R}^2$	Standard Error
1) S	=	0.1515	+	0.1328 X_1	0.4632	0.0384
		(15.75)		(13.14)		
2) S	=	0.1378	+	3.3747 X_2	0.4381	0.0393
		(12.36)		(12.50)		
3) T	=	0.0088	+	0.0036 X_1	0.2257	0.0018
		(19.73)		(7.68)		
4) T	=	0.0081	+	0.0981 X_2	0.2477	0.0018
		(16.36)		(8.16)		
5) J	=	−0.0031	+	0.0037 X_1	0.2693	0.0016
		(−7.53)		(8.62)		
6) J	=	−0.0038	+	0.1038 X_2	0.3057	0.0016
		(−8.45)		(9.41)		
7) S'	=	0.1434	+	0.1214 X_1'	0.4133	0.0381
		(14.75)		(11.88)		
8) S'	=	0.1317	+	3.1075 X_2'	0.3851	0.0390
		(11.68)		(11.21)		
9) T'	=	0.0081	+	0.0032 X_1'	0.1956	0.0017
		(18.64)		(7.03)		
10) T'	=	0.0076	+	0.0883 X_2'	0.2097	0.0017
		(15.48)		(7.34)		
11) J'	=	−0.0048	+	0.0154 X_1'	0.8615	0.0016
		(−11.50)		(35.19)		
12) J'	=	−0.0068	+	0.4086 X_2'	0.8626	0.0016
		(−14.52)		(35.36)		

Note: See Table 1.

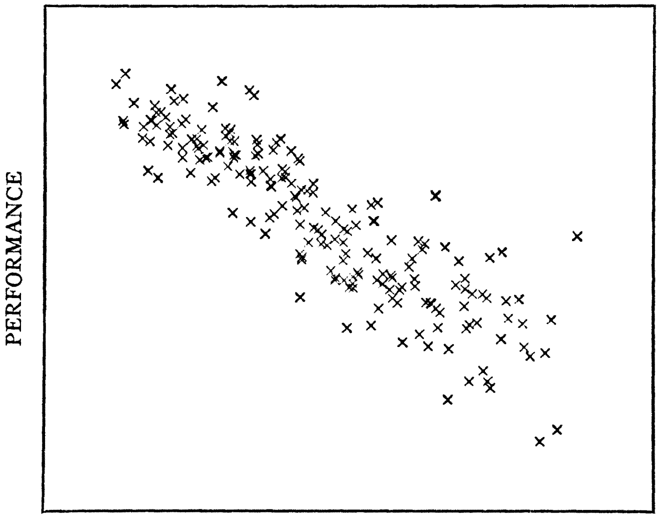

FIGURE 2. SCATTER DIAGRAM OF JENSEN'S PERFORMANCE MEASURE* ON RISK.
January 1960—March 1964

* Using *log relatives*.

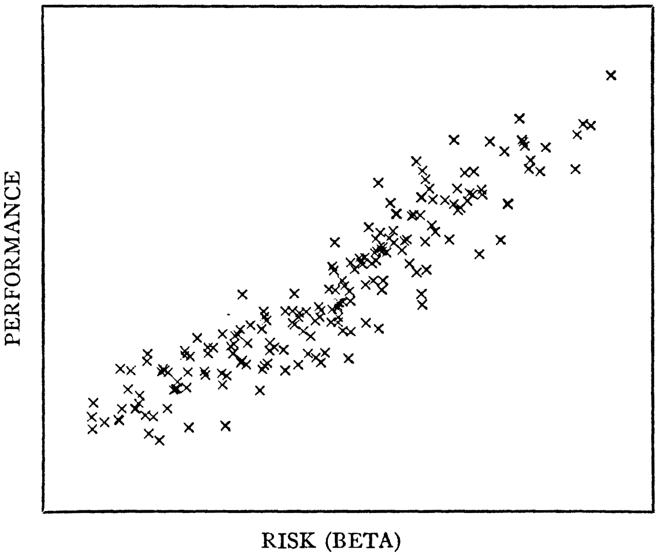

FIGURE 3. SCATTER DIAGRAM OF JENSEN'S PERFORMANCE MEASURE* ON RISK.
April 1964—June 1968

* Using *log relatives*.

III. *Some Concluding Comments*

While the market-line theory has made a substantial contribution to our understanding of the theoretical mechanism of capital asset pricing under uncertainty, our analysis raises some questions about the usefulness of the theory in its present form to explain market behavior. The Sharpe, Treynor, and Jensen one-parameter measures of portfolio performance based on this theory seem to yield seriously biased estimates of performance, with the magnitudes of the bias related to portfolio risk. Thus, the numerous studies of mutual fund performance based on these one-parameter measures are suspect (e.g., Sharpe (1966), Jensen (1968), and Lintner (1965b)) especially when they attempt to appraise individual portfolios, or when the average risk of these portfolios differs from that of the market as a whole.

Somewhat improved measures of portfolio performance for any period could be obtained by adjusting the Sharpe, Treynor or Jensen measures of performance for the portfolio in question by the relationship between the corresponding measures of performance and risk of random portfolios in that same period, with the precise adjustment factor depending on the degree of risk in the non-random portfolio. However, in view of the remaining uncertainty as to the reasons for the observed biases, it is probably preferable at the present stage of knowledge to use the traditional two parameters—rate of return and risk—to measure portfolio performance, in preference to the more elegant but also more dangerous one-parameter measures, since in the former it is not necessary to stipulate an explicit functional relationship between risk and return.

the one-parameter performance measures somewhat in favor of the riskier portfolios. In estimating these measures on a monthly basis it is assumed that there is a risk-free investment (essentially a Treasury bill maturing in exactly one month) with a known return and no capital risk. For investors to take advantage of such instruments would require a turnover of this part of the portfolio every month, which would typically involve much higher turnover and probably somewhat higher relative costs than for the rest of the portfolio. The risk-free measure actually used in our empirical relationships is the three-month yield on Treasury bills adjusted to a one-month basis which in a period of markedly rising interest rates would probably introduce a small positive discrepancy on the average between actual and expected bill rates. This unanticipated increase in bill rates would bias the one-parameter measures of performance in favor of the riskier portfolios in the same manner as transaction costs, but the effect should be small.

REFERENCES

F. Arditti, "Risk and the Required Return on Equity," *J. Finance,* Mar. 1967, *22,* 19–36.

M. Blume, "Portfolio Theory: A Step Towards Its Practical Application," *J. Business,* Apr. 1970, *43,* 152-74.

E. Fama, "Risk, Return, and Equilibrium: Some Clarifying Comments," *J. Finance,* Mar. 1968, *23,* 29–40.

———, "Multiperiod Consumption-Investment Decisions," *Amer. Econ. Rev.,* Mar. 1970, *60,* 163–74.

D. Farrar, "Discussion: The Performance of Mutual Funds in the Period 1945–1964 by Michael C. Jensen," *J. Finance,* May 1968, *23,* 417–19.

L. Fisher, "Some New Stock-Market Indexes," *J. Business,* Jan. 1966, *39,* 191–225.

——— **and J. Lorie,** "Rates of Return on Investments in Common Stocks," *J. Business,* Jan. 1964, *37,* 1–21.

I. Friend and P. Taubman, "Risk and Stock Market Performance," *Proc. Center for Research on Security Prices,* Univ. Chicago November 1966.

M. Jensen, "The Performance of Mutual Funds in the Period 1945–1964," *J. Finance,* May 1968, *23,* 389–416.

———, "Risk, Capital Assets, and the Evaluation of Investment Portfolio," *J. Business,* Apr. 1969, *42,* 167–247.

R. Kessel, *The Cyclical Behavior of the Term Structure of Interest Rates,* New York 1965.

J. Lintner, (1965a) "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Rev. Econ. Statist.*, Feb. 1965, *47*, 13–37.

———, (1965b) "Security Prices, Risk, and Maximal Gains from Diversification," *J. Finance*, Dec. 1965, *20*, 587–616.

H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, New York 1959.

W. Sharpe, "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *J. Finance*, Sept. 1964, *19*, 425–42.

———, "Mutual Fund Performance," *J. Business*, Jan. 1966, *39*, 119–38.

J. Tobin, "Liquidity Preference as Behavior Towards Risk," *Rev. Econ. Stud.*, Feb. 1958, *25*, 65–85.

J. Treynor, "How to Rate Management of Investment Funds," *Harvard Bus. Rev.*, Jan.-Feb. 1965, *43*, 63–75.

———, "Toward a Theory of Market Value of Risky Assets," unpublished manuscript, undated.