MUTUAL FUND PERFORMANCE*

WILLIAM F. SHARPE†

I. INTRODUCTION

WITHIN the last few years considerable progress has been made in three closely related areas—the theory of portfolio selection,[1] the theory of the pricing of capital assets under conditions of risk,[2] and the general behavior of stock-market prices.[3] Results obtained in all three areas are relevant for evaluating mutual fund performance. Unfortunately, few of the studies of mutual funds have taken advantage of the substantial backlog of theoretical and empirical material made available by recent studies in these related areas. However, one paper pointing the direction for future studies of mutual fund performance has appeared. Drawing on results obtained in the field of portfolio analysis, Jack L. Treynor has suggested a new predictor of mutual fund performance[4]—one that differs from virtually all those used previously by incorporating the volatility of a fund's return in a simple yet meaningful manner.

This paper attempts to extend Treynor's work by subjecting his proposed measure to empirical test in order to evaluate its predictive ability. But we will also attempt to do something more —to make explicit the relationships between recent developments in capital theory and alternative models of mutual fund performance and to subject these alternative models to empirical test.

II. IMPLICATIONS OF RECENT DEVELOPMENTS IN CAPITAL THEORY

A. PORTFOLIO ANALYSIS THEORY[5]

The theory of portfolio analysis is essentially normative; it describes efficient techniques for selecting portfolios on the basis of predictions about the performance of individual securities. The key element in the portfolio analyst's view of the world is his emphasis on both expected return and risk. The selection of a preferred combination of risk and expected return must, in the final analysis, depend on the preferences of the investor and cannot be made solely by the tech-

* I am grateful to Norman H. Jones, of the RAND Corporation, and Eugene F. Fama, of the University of Chicago, for helpful comments and suggestions.

† Associate professor of economics and operations research, University of Washington, and consultant, the RAND Corporation. Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of the RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors.

[1] The original work in the field was that of H. Markowitz; see his "Portfolio Selection," *Journal of Finance*, XII (March, 1952), 71–91, or the subsequent expanded version, *Portfolio Selection, Efficient Diversification of Investments* (New York: John Wiley & Sons, 1959). For extensions see my "A Simplified Model for Portfolio Analysis," *Management Science*, IX (January, 1963), 277–93, and Eugene F. Fama, "Portfolio Analysis in a Stable Paretian Market," *Management Science*, XI (January, 1965), 404–19.

[2] See, e.g., my "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, XIX (September, 1964), 425–42.

[3] For a summary of this work see Eugene F. Fama, "The Behavior of Stock-Market Prices," *Journal of Business*, XXXVIII (January, 1965), 34–105.

[4] "How To Rate Management of Investment Funds," *Harvard Business Review*, XLIII (January–February, 1965), 63–75.

[5] The material in this section is based on the references given in n. 1.

nician. However, the technician can (and should) attempt to find *efficient* portfolios—those promising the greatest expected return for any given degree of risk. The *portfolio analyst's* tasks are thus (1) translating predictions about security performance into predictions of portfolio performance, and (2) selecting from among the large number of possible portfolios those that are efficient. The *security analyst's* task is to provide the required predictions of security performance (including the interrelationships among the performances of securities). The *investor's* task is to select from among the efficient portfolios the one that he considers most desirable, based on his particular feelings regarding risk and expected return.

What tasks are implied for the mutual fund by this view of the investment process? Certainly those of security analysis and portfolio analysis. The emphasis in mutual fund advertising on diversification and the search for incorrectly priced securities reflects the importance accorded these aspects of the process. Portfolio analysis theory, unfortunately, does not make clear the manner in which the third function should be performed. A mutual fund cannot practically determine the preference patterns of its investors directly. Even if it could, there might be substantial differences among them. The process must work in the other direction, with the mutual fund management selecting an attitude toward risk and expected return and then inviting investors with similar preferences to purchase shares in the fund. At one extreme, the fund might attempt to describe an entire pattern of relative preference for expected return vis-à-vis risk (i.e., a pattern of indifference curves). A much more likely method, and one that seems to be followed in practice, involves merely a description of the general degree of risk planned for the fund's portfolio; the fund then simply attempts to select the efficient portfolio for that degree of risk (i.e., the one with the greatest expected return).

Portfolio analysis theory per se makes no assumptions about the pattern of security prices or the skill of investment managers. Thus few implications can be drawn concerning the results obtained by different mutual funds. Performance ex post might vary in two respects. First, different funds could exhibit different degrees of variability in return, due either to conscious selection of different degrees of risk or to erroneous predictions of the risk inherent in particular portfolios. Second, funds holding portfolios with similar variability in return might exhibit major differences in average return, due to the inability of some managers to select incorrectly priced securities and/or to diversify their holdings properly. In short, if sound mutual fund management requires the selection of incorrectly priced securities, effective diversification, and the selection of a portfolio in the chosen risk class, there is ample room for major and persisting differences in the performance of different funds.

B. THE BEHAVIOR OF STOCK-MARKET PRICES[6]

Recent work on the general behavior of stock-market prices has raised serious questions concerning the importance of one of the functions of mutual fund management. The theory of random walks asserts that the past behavior of a security's price is of no value in predicting its future price. The impressive evidence supporting this theory suggest that it

[6] The material in this section is based on Fama's "The Behavior of Stock-Market Prices," *op. cit.*

may be very difficult (and very expensive) to detect securities that are incorrectly priced. If so, it is not because security analysts are not doing their job properly, but because they are doing it very well. However, if this is the case, it may not pay the manager of a particular fund to devote extensive resources to the search for incorrectly priced securities; and the fund that does so may provide its investors a poorer net performance (after costs) than one that does not.

Under these conditions, what are the tasks of the mutual fund? Broadly defined, they still include security analysis, portfolio analysis, and the selection of a portfolio in the desired risk class. But the emphasis is changed. Security analysis is directed more toward evaluating the interrelationships among securities—the extent to which returns are correlated. And portfolio analysis is concerned primarily with diversification and the selection of a portfolio of the desired risk. In a perfect capital market, any properly diversified portfolio will be efficient; the mutual fund manager must select from among alternative diversified portfolios the one with the appropriate degree of risk.

Strictly speaking, the implications of this view of the world for mutual fund performance do not differ from those of the theory of portfolio analysis. Ex post, funds can be expected to exhibit differences in variability of return, due to intentional or unintentional selection of different risk classes. And the portfolios of some funds may be more efficient than others (i.e., give greater average return at the same level of variability) if managers differ in their ability to diversify effectively. However, the likelihood that persistent differences in efficiency will occur is greatly reduced. Recent work has shown that the task of diversification may be much simpler than formerly supposed, requiring only the spreading of holdings among standard industrial classes.[7] If so, most funds are likely to hold portfolios that are efficient ex ante. Any differences in efficiency ex post are thus probably transitory. The only basis for persistently inferior performance would be the continued expenditure of large amounts of a fund's assets on the relatively fruitless search for incorrectly valued securities.

C. THE THEORY OF CAPITAL-ASSET PRICES UNDER CONDITIONS OF RISK

Empirical work on the behavior of stock-market prices supports the view that the market responds very rapidly to new information affecting the value of securities. A natural reaction to these results is the construction of a model of a perfectly informed market in which each participant used his information in the manner suggested by portfolio analysis theory. Such an approach has been described elsewhere;[8] only the major features will be given here.

The predicted performance of a portfolio is described with two measures: the expected rate of return (E_i) and the predicted variability or risk, expressed as the standard deviation of return (σ_i). All investors are assumed to be able to invest funds at a common risk-free interest rate and to borrow funds at the same rate (at least to the desired extent). At any point of time, all investors share the same predictions concerning the future performance of securities (and thus portfolios). Under these conditions all ef-

[7] See Benjamin F. King, "Market and Industry Factors in Stock Price Behavior," *Journal of Business*, XXXIX, No. 1, Part II (Supplement, January, 1966).

[8] In my "Capital Asset Prices . . . ," *op. cit.*

ficient portfolios will fall along a straight line of the form[9]

$$E_i = p + b\sigma_i ,$$

where p is the pure (riskless) interest rate and b is the risk premium. Since investors are assumed to be risk-averse, b will be positive.

These results follow immediately from a relationship first described by James Tobin.[10] If an investor can borrow or lend at some riskless interest rate p and/or invest in a portfolio with predicted performance (E_i, σ_i), then by allocating his funds between the portfolio and borrowing or lending he can attain any point on the line

$$E = p + \left(\frac{E_i - p}{\sigma_i}\right) \sigma .$$

Any portfolio will thus give rise to a complete (linear) boundary of E,σ combinations. The best portfolio will be the one giving the best boundary; clearly it is the one for which $(E_i - p)/\sigma_i$ is the greatest. If more than one portfolio is to be efficient, all must lie along a common line and give identical values of this ratio.

The capital-market model described here deals with predictions of future performance. Since the predictions cannot be obtained in any satisfactory manner, the model cannot be tested directly. Instead, ex post values must be used—the average rate of return of a portfolio must be substituted for its expected rate of return, and the actual standard deviation of its rate of return for its predicted risk. We denote these measures by A_i and V_i (the latter for variability).

The capital-market model implies that ex post values for A_i and V_i for efficient portfolios should lie along a straight line, with higher values of V_i associated with higher values of A_i. Because there is risk in the stock market, the points will not lie precisely along such a line, even if the model is completely correct. But the relationship should be present, visible, and statistically significant.

The implications of this model for mutual fund performance are relatively straightforward. If all funds hold properly diversified portfolios and spend the appropriate amount for analysis and administration, they should provide rates of return giving A_i, V_i values lying generally along a straight line. Points that diverge from the underlying relationship should reflect only transitory effects and not persistent differences in performance. On the other hand, if some funds fail to diversify properly, or spend too much on research and/or administration, they will persistently give rates of return yielding inferior A_i, V_i values. Their performance will be poorer and can be expected to remain so.

III. PERFORMANCE OF 34 OPEN-END MUTUAL FUNDS, 1954–63

To test some of the implications drawn in the previous section, the annual rates of return for thirty-four open-end mutual funds[11] during the period 1954–63 were analyzed in the manner described above. The annual rate of return for a fund is based on the sum of dividend payments, capital gains distributions, and changes in net asset value; it is thus a measure of *net* performance—gross yield less the expenses of management and administration. The latter range from 0.25 per

[9] By definition, for inefficient portfolios: $E_i < p + b\sigma_i$.

[10] In his "Liquidity Preference as Behavior towards Risk," *Review of Economic Studies*, XXV (February, 1958), 65–86.

[11] The funds used for this and all subsequent analyses were those for which annual rates of return were given by Weisenberger for at least the last 20 years. All data are from Arthur Weisenberger & Co., *Investment Companies* (1953, 1962, 1964 eds.).

cent of net assets per year to 1.5 per cent per year. Most funds also charge an initial fee of approximately 8.5 per cent for selling costs when shares are purchased; annual rates of return are *not* net of such costs.

The average annual rate of return (A_i) and the standard deviation of annual rate of return (V_i) for each fund are shown in Figure 1; the values are listed in Table 1. The relationship predicted by the theory of capital asset prices is clearly present—funds with large average returns typically exhibit greater variability than those with small average returns. Moreover, the relationship is approximately linear and significant.[12] However, there are differences in efficiency; a number of funds are actually dominated (i.e., some other fund provided both a greater value of A_i and a smaller value of V_i). To analyze the differences, we need a single measure of performance; once such a measure is specified, any persistent differences can be investigated by testing alternative measures for predicting performance.

An intuitively appealing and theoretically meaningful measure of performance is easily derived from the Tobin effect. With substitution of the ex post measures (A and V) for the ex ante measures (E and σ), the formula described in Section II becomes

$$A = p + \left[\frac{A_i - p}{V_i}\right] V .$$

By investing in fund i and borrowing or lending at the riskless rate p, an investor could have attained any point along the line given by this formula. In 1953 it was possible to purchase a ten-year U.S. government bond at a price that would have guaranteed a return of slightly less than 3 per cent if held to maturity. Using 3 per cent as the estimate of p for the period, Boston Fund, shown by point Y in Figure 1, plus borrowing or lending could have provided any combination of average return and variability lying along line PYZ. Incorporated Investors, shown by point Q, could have provided any combination lying along line PQ. Clearly the former is better than the latter, since for *any* level of risk it offered a greater average return. Indeed, the steepness[13] of the line associated with a fund provides a useful measure of performance—one that incorporates both risk and average return. We define this as the reward-to-variability ratio: For Boston Fund the ratio is equal to the distance XP on Figure 1 divided by the distance XY. The larger the ratio, the better the performance.

An alternative interpretation of the ratio gives rise to the name—reward-to-variability ratio (R/V). The numerator shows the difference between the fund's average annual return and the pure interest rate; it is thus the *reward* provided the investor for bearing risk. The denominator measures the standard deviation of the annual rate of return; it shows the amount of risk actually borne. The ratio is thus the reward per unit of variability.

The final column of Table 1 shows the values of the R/V ratio for the thirty-four funds. They vary considerably—from almost 0.78 (the Boston Fund) to slightly over 0.43 (Incorporated Investors). Those who view the market as nearly perfect and managers as good diversifiers would argue that the differ-

[12] The results of statistical tests on these data are reported in my "Risk Aversion in the Stock Market: Some Empirical Evidence," *Journal of Finance*, September, 1965, pp. 416–22.

[13] The contangent of the angle made by the line with the horizontal axis is equal to the R/V ratio. Putting it another way, the reciprocal of the slope of the line (dA/dV) equals the R/V ratio.

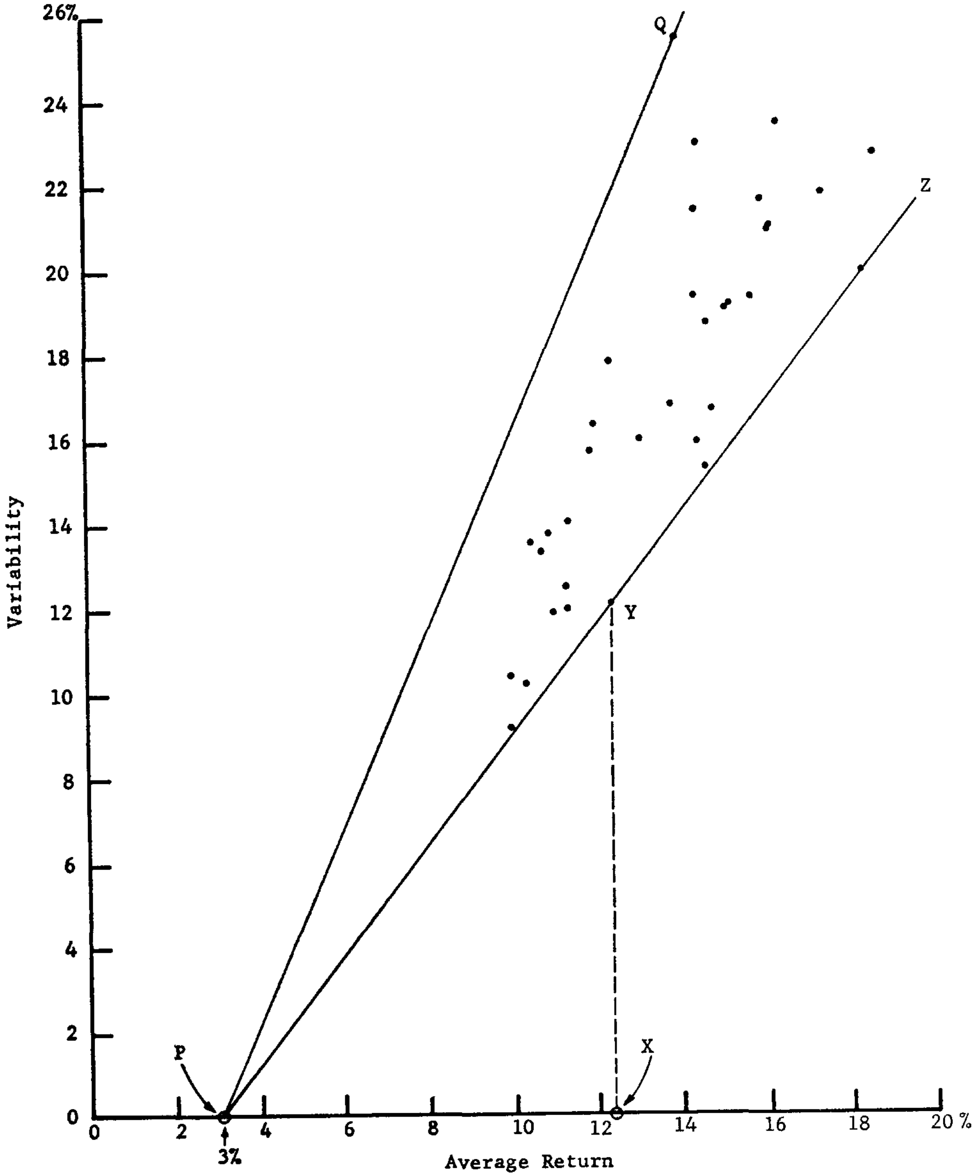


FIG. 1.—Average return and variability, 34 open-end mutual funds, 1954–63

ences are either transitory or due to excessive expenditures by some funds. Others would argue that the differences are persistent and can be attributed (at least partially) to differences in management skill. The remainder of this paper attempts to test these alternative explanations, using pre-1954 data to predict performance from 1954 to 1963, and in the tradition of empirical studies of mutual funds, provides comparisons with the performance of the securities used to compute the Dow-Jones Industrial Average.

IV. THE PERSISTENCE OF DIFFERENCES IN PERFORMANCE

To determine the extent to which differences in performance continue through time, the returns from the thirty-four funds during the period 1944–53 were used to compute R/V ratios for the decade preceding the one previously investigated.[14] The funds were ranked in each

[14] Since the long-term interest rate during this period was somewhat lower than that prevailing in the latter period, a pure interest rate of 2.5 per cent was used in the calculations (this was approximately the yield on a ten-year U.S. government bond in

TABLE 1

PERFORMANCE OF 34 MUTUAL FUNDS, 1954–63

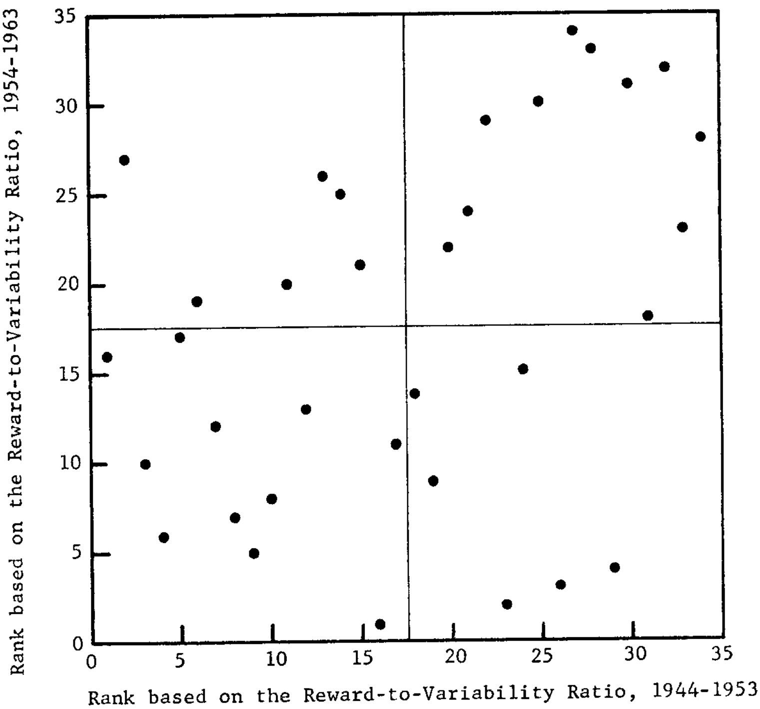
Mutual Fund	Average Annual Return (Per Cent)	Variability of Annual Return (Per Cent)	Reward-to-Variability Ratio (R/V)*
Affiliated Fund	14.6	15.3	0.75896
American Business Shares	10.0	9.2	.75876
Axe-Houghton, Fund A	10.5	13.5	.55551
Axe-Houghton, Fund B	12.0	16.3	.55183
Axe-Houghton, Stock Fund	11.9	15.6	.56991
Boston Fund	12.4	12.1	.77842
Broad Street Investing	14.8	16.8	.70329
Bullock Fund	15.7	19.3	.65845
Commonwealth Investment Company	10.9	13.7	.57841
Delaware Fund	14.4	21.4	.53253
Dividend Shares	14.4	15 9	.71807
Eaton and Howard, Balanced Fund	11.0	11.9	.67399
Eaton and Howard, Stock Fund	15.2	19.2	.63486
Equity Fund	14.6	18.7	.61902
Fidelity Fund	16.4	23.5	.57020
Financial Industrial Fund	14.5	23.0	.49971
Fundamental Investors	16.0	21.7	.59894
Group Securities, Common Stock Fund	15.1	19.1	.63316
Group Securities, Fully Administered Fund	11.4	14.1	.59490
Incorporated Investors	14.0	25.5	.43116
Investment Company of America	17.4	21.8	.66169
Investors Mutual	11.3	12.5	.66451
Loomis-Sales Mutual Fund	10.0	10.4	.67358
Massachusetts Investors Trust	16.2	20 8	.63398
Massachusetts Investors—Growth Stock	18.6	22.7	.68687
National Investors Corporation	18 3	19.9	.76798
National Securities—Income Series	12.4	17.8	.52950
New England Fund	10.4	10.2	.72703
Putnam Fund of Boston	13.1	16 0	.63222
Scudder, Stevens & Clark Balanced Fund	10.7	13.3	.57893
Selected American Shares	14.4	19.4	.58788
United Funds—Income Fund	16.1	20.9	.62698
Wellington Fund	11.3	12 0	.69057
Wisconsin Fund	13.8	16.9	0.64091

* R/V ratio = (average return — 3.0 per cent)/variability. The ratios shown were computed from original data and thus differ slightly from the ratios obtained from the rounded data shown in the table.

period, from 1—the fund with the highest (best) R/V ratio, to 34—the fund with the lowest (worst) R/V ratio. Figure 2 plots the rankings in the two periods. Although the relationship is far from perfect, there is a general upward trend, suggesting that funds ranking low in the early period tend to rank low in the later period, while those ranking high in the early period tend to rank high in the later period. The value of Spearman's rank correlation coefficient (+.360) bears this out[15] as does the count of points in the four quadrants (shown in the lower

1943). Although different assumptions regarding the pure interest rates in the two periods would substantially affect the rankings of individual funds, the relative predictive ability of the measures considered in this paper is not significantly affected by altering the assumed rates over a considerable range.

[15] The standard error for the sample size used in this and all subsequent calculations is 0.174.



Rank correlation coefficient = .360

	best	worst
worst	6	11
best	11	6

Fig. 2.—Predictions based on the reward-to-variability ratio

portion of the figure). Put rather crudely, the latter shows that an investor selecting one of the seventeen best funds in the first period would have an 11:6 chance of holding one of the seventeen best in the second period. Conversely, if he had selected one of the seventeen worst in the first period he would have an 11:6 chance of holding one of the seventeen worst in the second period. Simple regression analysis using the actual values of the R/V ratios gives similar results: the correlation coefficient is +.3157 and the t-value for the slope coefficient +1.88.

These results show that differences in performance can be predicted, although imperfectly. However, they do not indicate the sources of the differences. Equally important, there is no assurance that past performance is the best predictor of future performance. We consider next the alternative measure proposed by Treynor.

V. THE TREYNOR INDEX

In a perfect capital market no securities would be incorrectly priced. Thus one function of the mutual fund (finding such securities) would be eliminated, leaving only the tasks of diversification and selecting the appropriate risk class. Moreover, under such conditions it has been shown[16] that all truly diversified portfolios will move with the over-all market, giving high returns when the market in general provides high returns and low returns when the market provides low returns. The data bear out this hypothesis. During the period 1954–63, almost 90 per cent of the variance of the return on the typical fund in our sample was due to its comovement with the return on the thirty securities used to compute the Dow-Jones Industrial Average; moreover, as shown in Figure 3, the percentage was quite similar for most of the thirty-four funds. Treynor has taken advantage of this relationship by using the *volatility* of a fund as a measure of its risk instead of the total *variability* used in the R/V ratio. Since the returns on all diversified portfolios move with the market, the extent to which changes in the market are reflected in changes in a fund's rate of return can stand as a good measure of the total variability of the fund's return over time. By observing this relationship over some past period, a reasonably good estimate of volatility—the change in the rate of return on a fund associated with a 1 per cent change in the rate of return on, say, the Dow-Jones portfolio—can be obtained. We will use B_i to represent this value for the ith fund.[17]

The measure that we will term the Treynor Index (TI) can be obtained by simply substituting volatility for variability in the formula for the R/V ratio:

$$TI = \frac{A_i - p}{B_i}.$$

Stated in this manner, the relationship between the two measures is clear. And the extent of the contribution of volatility to over-all variability makes the ranking of funds on the basis of the Treynor Index very close to that based on the R/V ratio. Figure 4 shows the rankings using the two measures for the period 1954–63. Since the mutual funds in our sample all hold highly diversified portfolios, the similarity of the rankings is not surprising. And the cost of using the Treynor Index as a measure of past

[16] In my "Capital Asset Prices . . . ," *op. cit.*

[17] To be consistent with the notation in my "A Simplified Model for Portfolio Analysis," *op. cit.*

performance is relatively slight[18] since it mirrors the R/V ratio quite well. However, if some relatively undiversified funds (or more likely, privately held portfolios) had been included, the results could have been significantly different, since the Treynor Index cannot capture the portion of variability that is due to lack of diversification. For this reason it is an inferior measure of *past* performance. But for this reason it may be a superior measure for predicting *future* performance.

If mutual funds hold well-diversified portfolios, any major discrepancies between the variability of their returns and that portion due to movements in the market are likely to be due to transitory effects. By concentrating on the systematic part of a fund's variability—that is, its volatility—we can avoid paying attention to these transitory effects and

[18] Treynor apparently intended that his index be used both for measuring a fund's performance after the fact and for predicting its performance in the future: "When one talks about the historical performance pattern of a fund, he is looking at the past; but when he considers the preferences of individual investors and their choices among funds, he is talking about their appraisal of the future. We shall continue to talk about the performance of funds in terms of historical performance patterns, even though actual investor choices among funds are necessarily based on expectations regarding future performance patterns. The implication is that a good historical performance pattern is one which, if continued into the future, would cause investors to prefer it to others" (Treynor, *op. cit.*, p. 67).

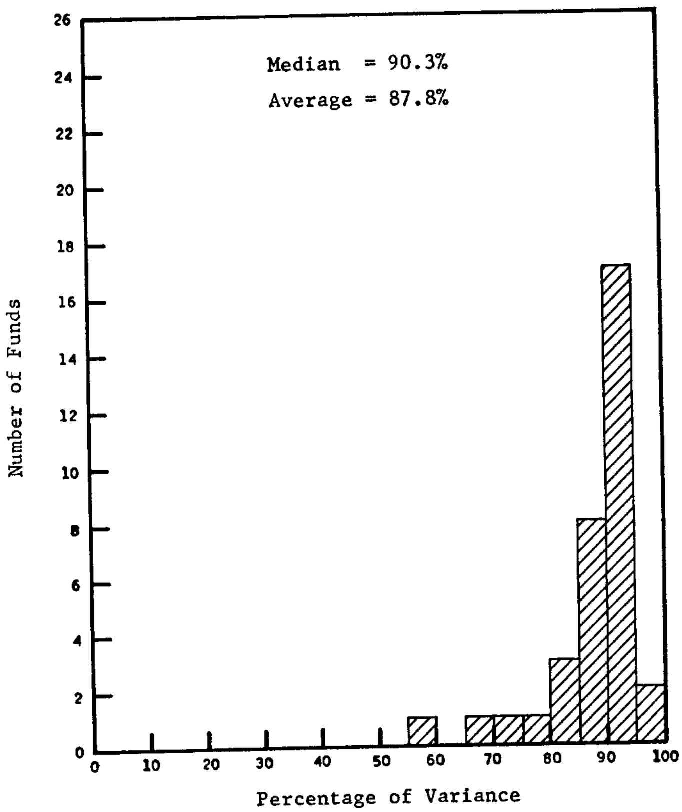
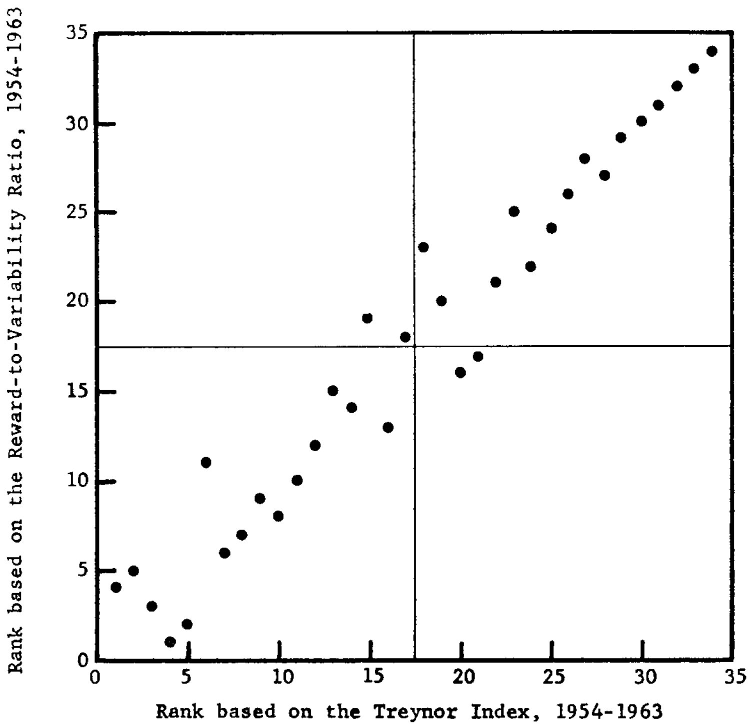


FIG. 3.—Percentage of variance due to comovement with the Dow-Jones Industrial Average, 34 open-end mutual funds, 1954–63.

concern ourselves with the more permanent relationships. Thus, given some reasonable assurance that a fund will perform its diversification function well, the Treynor Index may provide better predictions of future performance than the R/V ratio.

As Figure 5 shows, the data bear out this suspicion. Using the rankings of funds based on the Treynor Index computed from 1944–53 data to predict rankings based on the performance (measured by the R/V ratio) in the subsequent ten years gives somewhat better results than those obtained before. The odds of remaining in the selected half are now 12 to 5 (instead of 11 to 6) and the rank correlation coefficient is substantially higher—.454 instead of .360. Simple regression using actual values gives similar results—the correlation coefficient is +.4008 and the t-value for the slope co-



Rank correlation coefficient = .974

	best	worst
worst	2	15
best	15	2

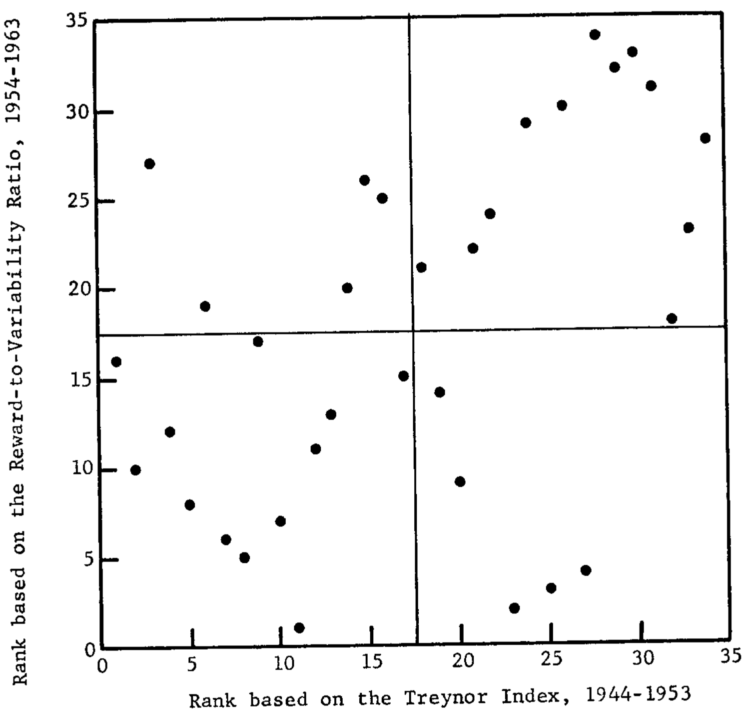
FIG. 4.—Correlation between ranks, reward-to-variability ratio versus Treynor Index

efficient +2.47. While the differences between these results and those based on the R/V ratio are far from overwhelming, the Treynor Index does appear to be the better predictor.

Before turning to other methods for predicting performance, the relationship between the measure we have termed the Treynor Index and that suggested by Treynor needs to be clarified. Our measure (TI) is similar in form to the R/V ratio. One measure proposed by Treynor[19] was simply the negative of ours:

$$\text{Slope angle} = -TI.$$

This form has a major advantage, since it can be transformed into a rather different quantity—the rate of return on a market portfolio (e.g., the Dow-Jones portfolio) that would cause the fund in



Rank correlation coefficient = .454

	best	worst
worst	5	12
best	12	5

FIG. 5.—Predictions based on the Treynor Index

[19] Described on p. 69 in Treynor, *op. cit.*:

$$\text{tangent } \alpha = \frac{(\mu - \mu^*)}{\sigma}.$$

question to provide a rate of return equal to the pure interest rate.[20] This substitute measure has considerable intuitive appeal. Moreover, if predictions are to be made subjectively (using some judgment) rather than objectively (solely on the basis of historical data), the advantages of this alternative index are substantial. For the purposes of this paper, however, such considerations are relatively unimportant, since we are dealing only with objective methods of prediction. Accordingly we will continue to state the Treynor Index in the form directly comparable to the R/V ratio.

VI. EXPENSE RATIOS AND SIZE

Past performance appears to provide a basis for predicting future performance, especially when measured with the Treynor Index. But this does not necessarily imply that differences in performance are due to differences in management skill. The high correlation among mutual fund rates of return suggests that most accomplish the task of diversification rather well. Differences in performance are thus likely to be due to either differences in the ability of management to find incorrectly priced securities or to differences in expense ratios. If the market is very efficient, the funds spending the least should show the best (net) performance. If it is not, funds devoting more resources to management may gain enough to more than offset the increased expenditure and thus show better net performance. Intimately related with such considerations is the impact of size. A fund with substantial assets can obtain a given level of security analysis by spending a smaller percentage of its income than can a smaller fund; alternatively, by spending the same percentage it can obtain more (and/or better) analysis. On the other hand, more analysis may be required for a large fund than for a small one. In any event, both influences should be considered.

Figure 6 shows the relationship between the expenses incurred by the thirty-four funds and their performance. Since we are concerned with prediction, the funds were ranked on the basis of the ratio of expenses to net assets during 1953 (needless to say, expense ratios

[20] The original index ("SA" for slope angle) is

$$SA_i = -TI_i = -\left(\frac{A_i - p}{B_i}\right) = \frac{p - A_i}{B_i}.$$

If the relationship between the rate of return on the fund (r_i) and that on the Dow-Jones portfolio (D) is

$$r_i = a_i + B_i D ,$$

then the average return on the fund is related to that on the Dow-Jones portfolio (D) by

$$A_i = a_i + B_i D .$$

Thus

$$SA_i = \frac{p - A_i}{B_i} = \frac{p - (a_i + B_i D)}{B_i}$$

$$= \left(\frac{p - a_i}{B_i}\right) - D .$$

D_i^*—the return on the Dow-Jones portfolio required to make the return on the fund equal to the pure interest rate (p) is

$$p = r_i = a_i + B_i D_i^* , \qquad D_i^* = \frac{p - a_i}{B_i} .$$

Therefore,

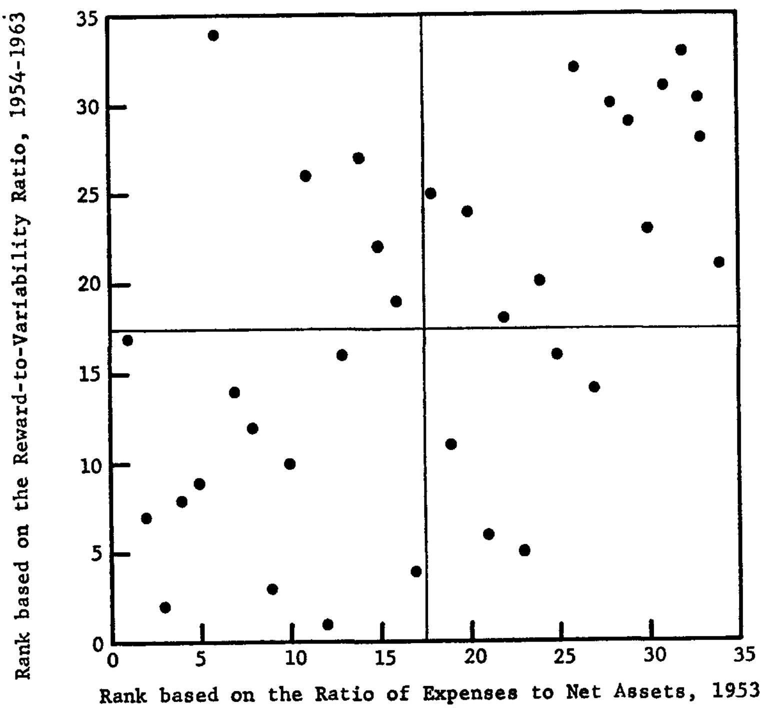
$$SA_i = D_i^* - \bar{D} .$$

Clearly, if the slope angle for one fund is higher than that of another, the return on the Dow-Jones required to make it yield the pure interest rate will be also. Since D_i^* is simply the slope angle plus a constant ($\bar{D}$), both measures will give the same ranking. The advantage of using D_i^* lies in its intuitive appeal and the fact that if future values are being estimated directly, no prediction about market performance is required. The insensitivity of Treynor's rankings (using either of the two measures) to predictions about market behavior follows directly from his assumption that variability is due *only* to the response of the fund's return to changes in the market. If there is another source of variability, the rankings of funds (using the R/V ratio) will depend to some extent on the behavior of the market.

changed somewhat during the subsequent ten years). Expenses ranged from 0.27 per cent of net assets (rank 1) to 1.49 per cent of net assets (rank 34).

The results tend to support the cynics: good performance is associated with low expense ratios. One of our summary measures (the rank correlation coefficient)[21] suggests that expense ratios provide somewhat better predictions than the Treynor Index; another suggests that the two are equally good (the odds of remaining in the selected half are 12 to 5 in both cases). The third suggests that the expense ratio is a slightly poorer predictor than the Treynor Index: Simple regression using actual values gave a

[21] Some of the funds had equal expense ratios (to the accuracy of two decimal places, as computed by Weisenberger). In such cases ranks were assigned alphabetically. For this reason the rank correlation coefficient is not as trustworthy as those computed for the other comparisons.



Rank correlation coefficient = .505

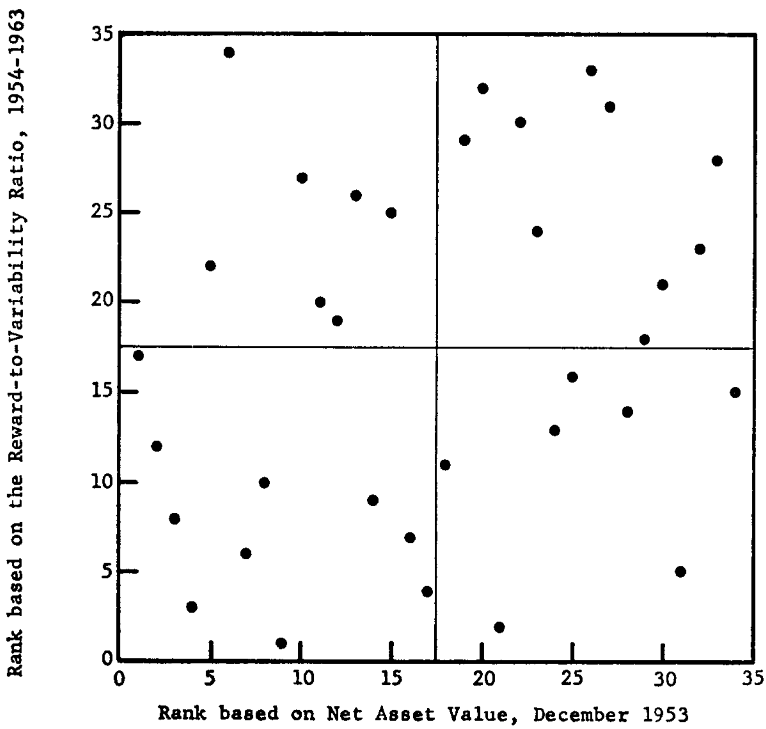
worst	5	12
best	12	5
	lowest	highest

Fig. 6.—Predictions based on the ratio of expenses to net assets

correlation coefficient of −.3746 (the Treynor Index gave +.4008) with a t-value of −.229 (instead of +2.47). Selecting a fund with a low ratio of expense to net assets may not be as foolish as some have suggested.

Figure 7 provides information concerning the predictive ability of the amount of a fund's assets. Size, measured by net asset value at the end of 1953, ranged from $522 million (rank 1) to $5.26 million (rank 34). Although the data show some correlation—with the larger funds exhibiting somewhat better performance—the relationship is marginal at best. This is borne out by the correlation coefficient based on actual values; it is +.1523, with a t-value of +0.87.

Multiple correlations, using actual values of the variables, gave the results shown in Table 2. The extremely small t-values for the slope coefficients relating



Rank correlation coefficient = .234

	largest	smallest
worst	7	10
best	10	7

FIG. 7.—Predictions based on size

R/V ratios to size, plus the fact that the signs differ from case to case, support the assertion that size per se is an unimportant factor in predicting future performance. The t-values for the Treynor Index and the exprense ratio support the assertion that both are useful for such predictions (although neither is highly significant).

Unfortunately, the results do not provide strong confirmation for either of the views of mutual fund management described in the earlier sections of this

TABLE 2

DEPENDENT VARIABLE: R/V RATIO, 1953–63

Multiple Correlation Coefficients	t-Values for Slope Coefficients		
	Treynor Index, 1944–53	Expense/ NAV 1953	Size (NAV) Dec. 1953
.4633........	+1.67	−1.43	−0.46
.4572........	+1.64	−1.38	......
.4015........	+2.26	......	+ .15
.3767........	......	−2.07	−0.24

paper. Expense ratios account for a substantial portion of the differences in performance, but so does another measure (the Treynor Index). Thus differences in management skill may be important. However, one reservation is in order. Expense ratios as reported do not include all expenses; brokers' fees are omitted. Thus the expense ratio does not capture all the differences in expenses among funds. It is entirely possible that funds with performance superior to that predicted by the traditional expense ratio engage in little trading, thereby minimizing brokerage expense. It was not feasible to attempt to measure total expense ratios for this study; had such ratios been used, a larger portion of the difference in performance might have been explained in this manner, and the apparent differences in management skill might have been smaller. Clearly, more work is needed before the traditional view of the importance of the search for incorrectly priced securities can be accepted.

VII. THE RELATIVE RISK OF MUTUAL FUNDS

All the calculations presented thus far deal with a measure of performance (the R/V ratio) that disregards risk per se, concentrating instead on the relationship between the reward obtained from the fund and the risk actually experienced. As shown earlier, the investor should be most concerned with this relationship, if he can arrange his other commitments (either by borrowing additional funds or by investing in some riskless security) to complement the risk inherent in a particular fund in any manner he desires. However, to do this he must have some idea of the variability the fund will actually experience. If mutual fund managers do not perform the second of the three tasks we have outlined (staying in a selected risk class), investors will find it difficult to arrange their over-all holdings in the most desirable manner. Holders of mutual fund shares presumably expect that funds will show reasonable consistency over time with regard to the variability of returns.

Figure 8 provides some evidence on this point. The funds were ranked in each of the two periods studied (1944–53 and 1954–63) on the basis of variability —rank 1 indicating the smallest amount and rank 34 the largest. A reasonable amount of consistency between periods is evident, but a number of major shifts appear.[22] In some cases this may have been due to announced changes in man-

[22] Simple regression analysis using actual values gave a correlation coefficient of +.4518 and a t-value of +2.86.

agement philosophy; in others it was probably inadvertent. Whatever the cause, the prevalence of such shifts in our sample is likely to disappoint some investors. On the other hand, there is no well-defined standard against which the results can be compared; given the difficulties involved,[23] one might reasonably argue that the data show that mutual fund managers fulfil remarkably well the obligation to stay within their selected risk classes.

[23] In "The Variation of Certain Speculative Prices," *Journal of Business*, XXXVI (October, 1963), 394–419, B. Mandelbrot has shown that predicting the variability of the changes in security prices is very difficult indeed. Presumably the law of large numbers cannot be relied upon to eliminate enough of the difficulty to make predictions of variability of the return on portfolios relatively simple.

VIII. MUTUAL FUNDS VERSUS THE DOW-JONES INDUSTRIAL AVERAGE

We have dealt at length with comparisons among mutual funds but have not

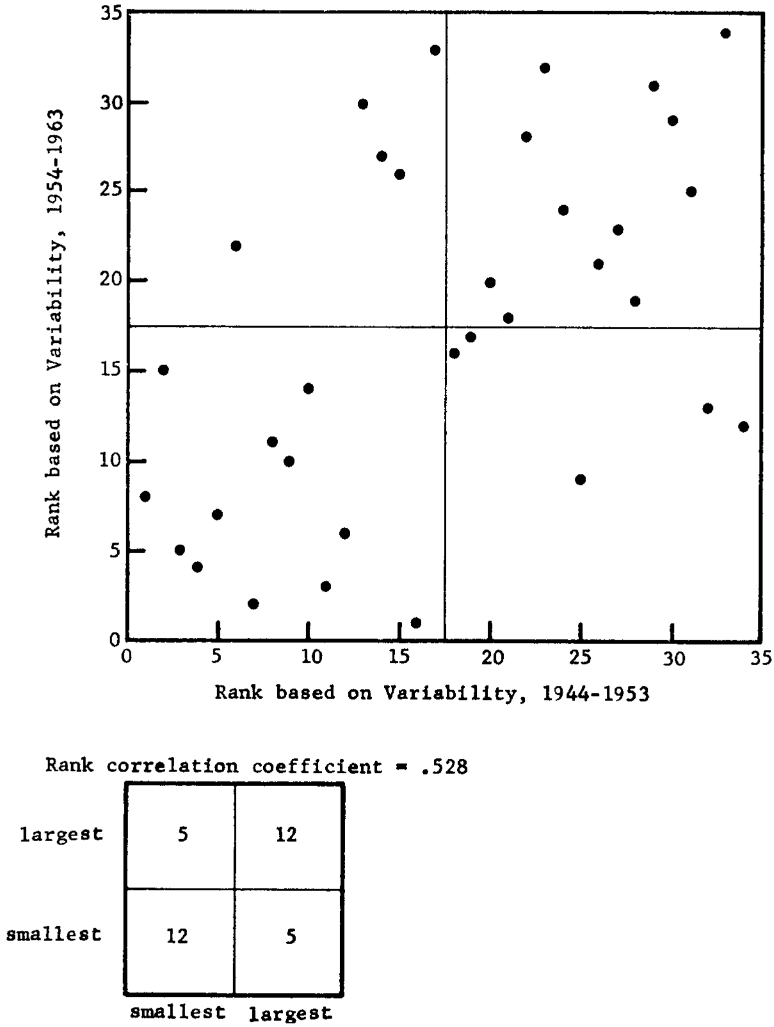


FIG. 8.—Consistency of risk class

considered an alternative strategy—investing directly in a reasonably diversified group of securities. To investigate such an alternative we must specify the portfolio to be held; following tradition, the thirty securities used to compute the Dow-Jones Industrial Average will be used.

When calculating the returns from the Dow Jones portfolio, no costs (brokerage, management, or administrative) are deducted. To some extent this overstates the performance available from such a direct investment. On the other hand, the initial selling (load) charge is not deducted when determining the returns from mutual funds; thus the results from both types of investments are overstated. The magnitudes of the differences between the measures we use and those relevant for a particular investor depend on a number of factors,[24] but for most investors the comparison made here should give results similar to those obtained if all the relevant costs had been considered.

Figure 9 shows the distribution of the *R/V* ratios for the thirty-four funds (based on their performance during the

[24] For example, the amount to be invested, the number of years the portfolio is to be held, and the extent to which dividends are to be reinvested.

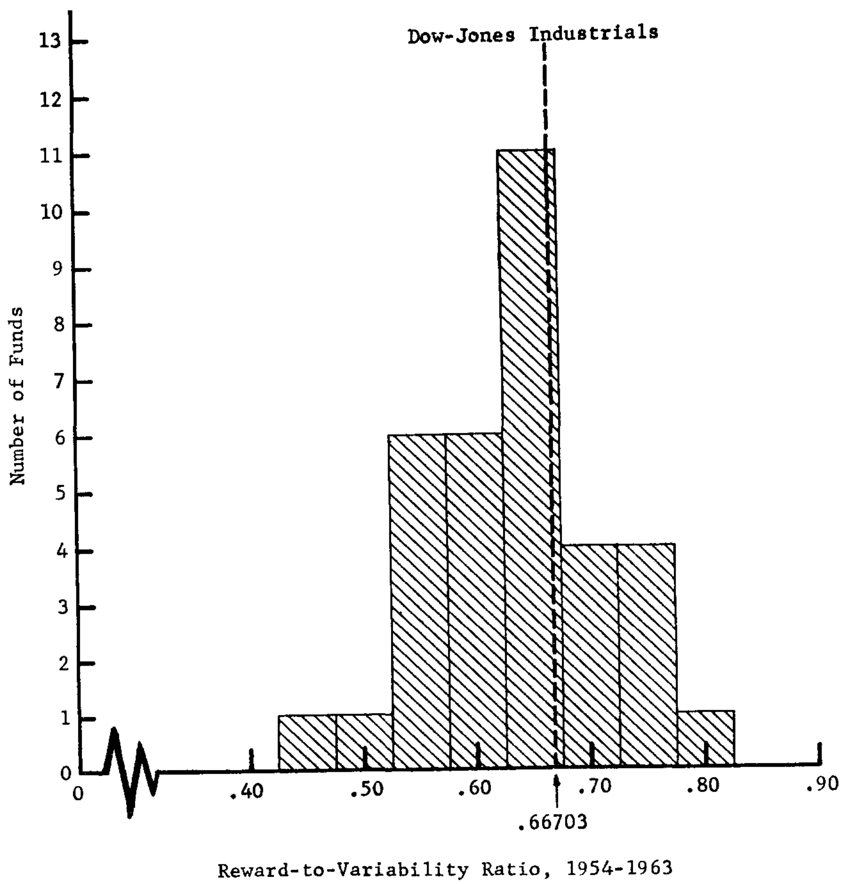


FIG. 9.—Mutual fund performance versus Dow-Jones Industrials, 1954–63

period 1954–63). The vertical line represents the R/V ratio for the Dow-Jones portfolio—its return average 16.3 per cent during the period with a variability of 19.94 per cent, giving an R/V ratio of 0.667. The average R/V ratio for the funds in our sample was 0.633—considerably smaller than that of the Dow-Jones Average. Although another group of mutual funds would give different results, the odds are greater than 100 to 1 against the possibility that the average mutual fund did as well as the Dow-Jones portfolio from 1954 to 1963.[25] In this group, only eleven funds did better than the Dow-Jones portfolio, while twenty-three did worse.

From the standpoint of the investor the comparison shown in Figure 9 is the most relevant. But to account for the relatively poor performance of most mutual funds it is instructive to compare *gross* performance (i.e., before deducting expenses) with that of the Dow-Jones portfolio. Such a comparison[26] shows that nineteen funds did better than the Dow-Jones portfolio, and only fifteen did worse. Although another group of funds would give different results, it is unlikely that the gross performance of the average mutual fund was worse than that of the Dow-Jones portfolio from 1954 to 1963.[27]

While it may be dangerous to generalize from the results found during one ten-year period, it appears that the average mutual fund manager selects a portfolio at least as good as the Dow-Jones Industrials, but that the results actually obtained by the holder of mutual fund shares (after the costs associated with the operation of the fund have been deducted) fall somewhat short of those from the Dow-Jones portfolio. This is consistent with our previous conclusion that, all other things being equal, the smaller a fund's expense ratio, the better the results obtained by its stockholders.

IX. CONCLUSIONS

This paper represents an attempt to bring to bear on the measurement and prediction of mutual fund performance some of the results of recent work in capital theory and the behavior of stock-market prices. We have shown that performance can be evaluated with a simple yet theoretically meaningful measure that considers both average return and risk. This measure precludes the "discovery" of differences in performance due solely to differences in objectives (e.g., the high average returns typically obtained by funds who consciously hold risky portfolios). However, even when performance is measured in this manner there are differences among funds; and such differences do not appear to be en-

[25] The standard deviation of the R/V values for the 34 funds was 0.08067. If the population of mutual funds had a mean of 0.667 and a standard deviation of 0.08067, the distribution of sample means for groups of 34 would have a standard deviation of 0.01383 ($=0.08067/\sqrt{34}$) and be roughly normally distributed. The observed mean of 0.633 is 2.46 standard deviations below the assumed mean of 0.667; the odds are 144 to 1 that under the hypothesized conditions a sample of 34 funds would have an average R/V value as low as 0.633.

[26] The comparison was made by assuming that each fund maintained its 1953 ratio of expenses to net assets throughout the subsequent ten years. Under these conditions the only change required to compute the R/V ratios for gross performance was to add each fund's expense ratio to its average return before the R/V ratio was computed.

[27] The standard deviation of the gross R/V values for the 34 funds was 0.08304. If the populatoin of mutual funds had a mean of 0.667 and a standard deviation of 0.08304, the distriubtion of sample means for groups of 34 would have a standard deviation of 0.01424 ($=0.08304/\sqrt{34}$) and be roughly normally distributed. The observed mean of 0.677 is 0.74 standard deviations above the assumed mean of 0.667; the odds are 3.36 to 1 that under the hypothesized conditions a sample of 34 funds would have an average gross R/V value as high as 0.677.

tirely transitory. To a major extent they can be explained by differences in expense ratios, lending support to the view that the capital market is highly efficient and that good managers concentrate on evaluating risk and providing diversification, spending little effort (and money) on the search for incorrectly priced securities. However, past performance per se also explains some of the differences. Further work is required before the significance of this result can be properly evaluated. But the burden of proof may reasonably be placed on those who argue the traditional view—that the search for securities whose prices diverge from their intrinsic values is worth the expense required, even for a mutual fund operating under severe constraints on the proportion of funds invested in any single security.[28] Fortunately many who hold this view have both the means and the data required to perform extensive analyses; we will all look forward to their results.

[28] By law, no more than 5 per cent of the assets of a fund may be invested in any given security, and no more than 10 per cent of the assets of a given firm may be held by the fund. Moreover, the very size of many mutual funds makes it impossible to invest even the legal maximum in a security without driving its price up to a point substantially above the original (bargain) amount.