Eric C. Chang
University of Iowa

Wilbur G. Lewellen
Purdue University

Market Timing and Mutual Fund Investment Performance*

The parametric statistical procedure recently developed by Henriksson and Merton to test jointly for the presence of either superior market timing or security selection ability in managed portfolios is employed to evaluate empirically the investment performance of a sample of mutual funds. This procedure and the associated findings are compared with those of prior investment performance evaluations. The new technique produces a more favorable judgment about mutual fund security selection performance in the aggregate, and it alters the performance evaluations of various individual funds. Nonetheless, few fund managers appear to have displayed much market-timing skill, and the general conclusion that they have been unable collectively to outperform a passive investment strategy still seems valid.

The investment performance of mutual fund portfolios has been the subject of extensive examination in the literature of finance. Evidence about the *collective* performance of such funds is relevant to the efficient market hypothesis and thereby to an understanding of the process of security price determination, because of its potential implications about differential investment information availability in the marketplace. Evidence about the *relative* performance of individual mutual funds is, of course, of obvious interest to entities with investment funds to allocate.

Performance evaluations of this sort have typically employed a one-parameter risk/return benchmark like that developed by Jensen (1968, 1969) and refined by Black, Jensen, and Scholes (1972) and Blume and Friend (1973). Such investigations have effectively focused on fund managers' security selection skills (or lack thereof), since the examined portfolios' risk levels have been assumed to be stationary through time.

* The investigation reported herein consists of a portion of the National Bureau of Economic Research program of research in Financial Markets and Monetary Economics. We gratefully acknowledge financial support for the study from the National Science Foundation under grant no. SOC-7825789. Any opinions, findings, and conclusions or recommendations expressed are ours, however, and do not necessarily reflect the views either of the National Science Foundation or of the National Bureau of Economic Research. We are indebted to Gary G. Schlarbaum for helpful comments on both the exposition and the analysis.

(*Journal of Business*, 1984, vol. 57, no. 1, pt. 1)

0021-9398/84/5701-0007$01.50

Portfolio managers might, however, achieve differential return performance by engaging in successful "macro" market-timing activities as well as careful "micro" security selection efforts. That is, they can shift the overall risk composition of their portfolios in anticipation of broad market price movements. Fama (1972) and Jensen (1972) addressed this issue and pointed out the empirical measurement problems involved in evaluating properly the constituents of investment performance when portfolio risk levels are nonstationary.

A variety of studies have since picked up on the point, including those by Kon and Jen (1978, 1979), Fabozzi and Francis (1979), Alexander and Stover (1980), and Miller and Gressis (1980). They find at least some evidence that mutual fund portfolios do not in fact maintain a constant risk posture over time and conclude that attempts at market timing may well be a dimension of fund managers' decision processes. The conceptual framework in which decisions to change risk levels are made, however, has not until recently had a clear theoretical foundation. Our objective here is to build on such a theory to provide some additional pertinent evidence.

I. Methodology

The vast majority of empirical investigations of managed-portfolio investment performance have been conducted within the framework of the single-factor market model using a regression equation specification for portfolio p of the form

$$Z_p(t) - R(t) = \alpha_p + \beta_p[Z_m(t) - R(t)] + \epsilon_p(t), \tag{1}$$

where $Z_p(t)$ is the observed rate of return on the portfolio during the period, $R(t)$ is the contemporaneous rate of return on a riskless asset, $Z_m(t)$ is the return on the fully diversified "market" portfolio of all risky assets during t, β_p is an index of the systematic risk level of portfolio p, α_p is the average residual or "abnormal" component of that portfolio's return, and $\epsilon_p(t)$ is a random error term with $E[\epsilon_p(t)] = 0$. In this specification, β_p is assumed to be constant over time, and a statistically significant positive (negative) value for the estimated intercept α_p is taken to be an indication of superior (inferior) return performance attributable to the security-selection efforts of the portfolio manager. The possibility that the latter may engage in macro-price-forecasting designed consciously to alter β_p over time is ignored.

Merton (1981), however, addresses that possibility with a model that characterizes portfolio managers who attempt market timing as doing so by forecasting time periods in which equities in the aggregate will outperform bonds $[Z_m(t) > R(t)]$ or bonds will outperform stocks $[Z_m(t) \leq R(t)]$, not as predicting the *magnitude* of the differentials. He

shows that the sum of the conditional probabilities, p_1 and p_2, of a correct prediction of a coming down- or up-market period, where

$$p_1 = \text{prob}[\text{forecast } Z_m(t) \leq R(t) | Z_m(t) \leq R(t)]$$

$$p_2 = \text{prob}[\text{forecast } Z_m(t) > R(t) | Z_m(t) > R(t)],$$

will be the appropriate statistic for the evaluation of forecasting ability, and a sufficient condition for a manager's predictions to have positive value is that $p_1 + p_2 > 1$. In a subsequent paper, Henriksson and Merton (1981) portray the market-timing portfolio manager as having an asset allocation policy involving investments divided among the market portfolio of equities and riskless bonds according to the following rule: At the beginning of time period t, if the manager's forecast is that bonds will outperform stocks during the period, then $100\eta_1\%$ of the managed assets will be invested in the market portfolio and $100(1 - \eta_1)\%$ in bonds; conversely, if the forecast is that stocks will outperform bonds, the allocation will be $100\eta_2\%$ to the market portfolio and $100(1 - \eta_2)\%$ to bonds. Accordingly, the systematic risk level of the portfolio, $\beta(t)$, is a decision variable, η_1 and η_2 are its respective predicted down-market and up-market values, and we would expect that $\eta_1 < \eta_2$ for a rational forecaster.[1]

Merton demonstrated that, up to an additive noise term, the periodic returns per dollar invested in a managed portfolio using the indicated market-timing rule are the same as those that would be generated by a strategy of investing $[p_2\eta_2 + (1 - p_2)\eta_1]$ dollars in the market portfolio; acquiring $[p_1 + p_2 - 1][\eta_2 - \eta_1]$ free put options on the market portfolio having an exercise price (per dollar of the market)[2] equal to $R(t)$; and investing the remainder in riskless securities. On that basis, Henriksson and Merton arrived at a least squares regression equation to estimate empirically the separate contributions of selectivity and timing to a managed portfolio's return, of the form

$$Z_p(t) - R(t) = \alpha + \beta_1 X(t) + \beta_2 Y(t) + \epsilon_p(t), \tag{2}$$

where α is the expected excess rate of return on the portfolio due to the manager's security selection ability, $X(t) \equiv Z_m(t) - R(t)$, and $Y(t) \equiv \max[0, -X(t)]$. They showed that plim $\hat{\beta}_2 = [p_1 + p_2 - 1][\eta_2 - \eta_1]$, the large-sample least squares estimate of β_2, is an appropriate measure of the portfolio manager's market-timing skill. The true β_2 will equal zero if the forecaster either has no timing ability (i.e., $p_1 + p_2 = 1$) or

1. Implicit in this portrayal is the notion that changes in the managed portfolio's β are either solely or predominantly attributable to deliberate attempts at market timing by the manager. There may, of course, be other causes as well; see Kon and Jen (1979) for a review.

2. For convenience, Merton assumed that one "unit" (or share) of the market portfolio is of a size such that \$1 is the beginning-of-period investment needed to acquire one share. Hence, stock "splits" in the ratio $1 + Z_m(t)$ occur at the end of every period.

does not act on his or her forecasts (i.e., $\eta_1 = \eta_2$). Hence, the timing null hypothesis is $H_0: \beta_2 = 0$. Similarly, the large-sample least squares estimate of α, which was shown to be plim $\hat{\alpha} = \bar{Z}_p - \bar{R} -$ plim $\hat{\beta}_1 \bar{X} -$ plim $\hat{\beta}_2 \bar{Y}$, permits a test of the selectivity null hypothesis $H_0: \alpha = 0$.

The regression specification in equation (2) is more general than that of equation (1), the latter being the special case for which $\beta_2 = 0$ and timing activity is not a component of the portfolio decision process. Moreover, should timing be present,[3] it turns out that the assessment of the manager's security selection ability obtained by estimating the α_p in equation (1) will be biased—in a direction and to an extent that can be identified. Note that, in the absence of knowledge about the portfolio manager's particular market forecasts, the systematic risk level $\beta(t)$ of a "timed" portfolio will appear as a random variable. Following Henriksson and Merton, let b denote the unconditional (on the forecast) expected value of $\beta(t)$. Thus,

$$b = q[p_1\eta_1 + (1 - p_1)\eta_2] + (1 - q)[p_2\eta_2 + (1 - p_2)\eta_1],$$

where q is equal to the unconditional (on the forecast) probability that $Z_m(t) \leq R(t)$. If we define the random variable $\theta(t)$ as equal to $[\beta(t) - b]$, $\theta(t)$ will be the unanticipated component of portfolio systematic risk. Henriksson and Merton demonstrate that, conditional on $X(t)$, the expected value of θ will be

$$\begin{aligned} E[\theta|X] &= \theta_1 = (1 - q)(p_1 + p_2 - 1)[\eta_1 - \eta_2] \quad \text{for } X(t) \leq 0 \\ E[\theta|X] &= \theta_2 = q(p_1 + p_2 - 1)[\eta_2 - \eta_1] \quad \text{for } X(t) > 0. \end{aligned} \tag{3}$$

If market timing is done well, therefore, a managed portfolio will display a positive θ_2 and a negative θ_1.

Using equation (1), the least squares estimate of β_p will be

$$\text{plim } \hat{\beta}_p = \frac{\text{cov}[Z_p(t), X(t)]}{\text{var}[X(t)]}. \tag{4}$$

If we define the down-market and up-market return differentials between stocks and bonds to be, respectively, $X_1(t) \equiv \min[0, X(t)]$ and $X_2(t) \equiv \max[0, X(t)]$, and denote

$$\begin{aligned} \text{var}[X_1(t)] &= \sigma_1^2 \qquad \text{var}[X_2(t)] = \sigma_2^2 \\ E[X_1(t)] &= \bar{X}_1 \qquad E[X_x(t)] = \bar{X}_2, \end{aligned}$$

then we have $E[X(t)] \equiv \bar{X} = \bar{X}_1 + \bar{X}_2$, and the expression in equation (4) can be rewritten (see App. A) as

$$\text{plim } \hat{\beta}_p = b + \theta_2 - \left\{\frac{\theta_2}{q[1 + (\sigma_2^2 - \bar{X}_1\bar{X}_2)/(\sigma_1^2 - \bar{X}_1\bar{X}_2)]}\right\}. \tag{5}$$

3. Or, perhaps more accurately, should it occur even inadvertently that there are systematic movements in a portfolio's β with general market conditions over time.

The least squares estimate of α_p from equation (1) therefore will be

$$\text{plim } \hat{\alpha}_p = E[Z_p(t)] - E[R(t) - [\bar{X}][\text{plim } \hat{\beta}_p] \tag{6}$$

$$\begin{aligned} \text{plim } \hat{\alpha}_p &= \bar{Z}_p - \bar{R} - (\bar{X}_1 + \bar{X}_2)(b + \theta_2) \\ &+ (\bar{X}_1 + \bar{X}_2)\left\{\frac{\theta_2}{q[1 + (\sigma_2^2 - \bar{X}_1\bar{X}_2)/(\sigma_1^2 - \bar{X}_1\bar{X}_2)]}\right\}. \end{aligned} \tag{7}$$

Conversely (also, see App. A), if equation (2) is used for estimation,

$$\text{plim } \hat{\alpha} = \bar{Z}_p - \bar{R} - \bar{X}_1(b + \theta_1) - \bar{X}_2(b + \theta_2) \tag{8}$$

and the difference between the two is

$$\begin{aligned} \text{plim } \hat{\alpha}_p - \text{plim } \hat{\alpha} = \bar{X}&\left\{\frac{\theta_2}{q[1 + (\sigma_2^2 - \bar{X}_1\bar{X}_2)/(\sigma_1^2 - \bar{X}_1\bar{X}_2)]}\right\} \\ &- \bar{X}_1(\theta_2 - \theta_1). \end{aligned} \tag{9}$$

In a risk-averse investment environment, of course, $\bar{X}$ should be positive; by construction, $\bar{X}_1$ is negative and $\bar{X}_2$ is positive; and, for a portfolio manager who is able correctly to time the market, $\theta_2 > 0$ while $\theta_1 < 0$, from above. Accordingly, the difference plim $\hat{\alpha}_p$ − plim $\hat{\alpha}$ will be positive and the specification in equation (1) will produce an estimate of α_p that will overstate the associated security-selection skills of the successful market-timing manager. Similarly, for a systematically unsuccessful such manager, $\theta_2 < 0$ and $\theta_1 > 0$, which will cause plim $\hat{\alpha}_p$ − plim $\hat{\alpha}$ to be negative and introduce the opposite bias; his or her talent for security selection will be underestimated by fitting equation (1) to observed return data. The potential for the misinterpretation of empirical evidence on investment performance is obvious.

Henriksson and Merton went on to develop an alternative version of equation (2) for which the coefficients to be estimated have a particularly appealing intuitive meaning. Via a linear transformation, that equation was shown to be equivalent to a regression specification of the form

$$Z_p(t) - R(t) = \alpha^* + \beta_1^* X_1(t) + \beta_2^* X_2(t) + \epsilon_p^*(t). \tag{10}$$

Because $X_1(t) = 0$ and $X_2(t) = X(t)$ when $X(t) > 0$, β_2^* can be characterized as the "up-market beta" of a managed portfolio; symmetrically, β_1^* can be labeled its "down-market beta." As Henriksson and Merton demonstrated, the test for market timing ability—on β_2—embedded in estimates of equation (2) is equivalent to a test from equation (10) of whether β_1^* is significantly different from β_2^* ($H_0: \beta_1^* = \beta_2^*$), and the economic interpretation of the estimated coefficients is rather more transparent in the latter formulation.[4] Moreover, the large-

4. In a companion paper in this issue of the *Journal*, Henriksson (1984) employs instead the format of our equation (2) in his parametric analysis, testing for the presence

sample properties of α^* in equation (10) are the same as those of α in equation (2), and it thereby retains its role as a measure of the incremental contribution of security-selection ability to observed portfolio returns. Accordingly, equation (10) is the version we shall estimate here, for comparison with the estimates of performance obtained from equation (1).

There are some points to be made about that process. Since equation (10) contains three variables, the least squares estimates of their coefficients form, in principle, a plane in three-dimensional space (Johnston 1972). However, because of the special character of variables $X_1(t)$ and $X_2(t)$, the least squares estimates obtained will describe only two straight lines in that space, given that the value of one of these variables will inevitably be zero when the other is nonzero.

The market-timing and security-selection test methodology assayed by estimating equation (10) thereby involves (*a*) separating the portfolio return data for the sample of mutual funds studied into two subsets, based on the sign of $X(t)$ for each time interval examined—that is, partitioning the data into up-market and down-market conditions; (*b*) estimating the least squares lines in each of the two market conditions for every mutual fund, pursuant to the requirement that the lines share a common intercept for each fund; and (*c*) testing whether the slope-coefficient estimates for the two lines, $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$, differ significantly.[5] Appendix B provides the details and the relevant proofs. This procedure, of course, is simply a special case of covariance analysis or dummy-variable regression, which techniques are designed to identify differences in the parameters of hypothesized relationships in different data regimes. Prior investigations by Alexander and Stover (1980) and Veit and Cheney (1982) are in that spirit, although without the theoretical underpinning provided by the Merton (1981) and Merton and Henriksson (1981) analyses to surface the key interdependence between the return contributions of security selection and market timing.

II. The Sample and the Data

The sample of professionally managed portfolios examined consists of 67 mutual funds encompassing a wide range of stated investment objectives (see App. C), for which complete monthly rate of return data were available covering the period January 1971 through December

of a nonzero β_2. The relevant transformation is that, in equation (10), β_1^* is equal to the difference between the β_1 and β_2 of equation (2) and $\beta_2^* = \beta_1$. Substituting these values in equation (10) yields equation (2). Hence, β_2 is equal to our $\beta_2^* - \beta_1^*$.

5. As Henriksson and Merton (1981, p. 530) note, there may be a problem with heteroscedasticity in the error term $\epsilon_p^*(t)$ in this specification, which would compromise the efficiency of ordinary least squares estimates of the parameters. Henriksson (1984) finds in his analysis, however, that the difference between the results using weighted and ordinary least squares are de facto minimal, and his overall conclusions match ours.

1979.[6] The value-weighted stock index of the Center for Research in Security Prices (CRSP) was employed as a representation of the market portfolio of equities. The monthly returns calculated both for that portfolio and the mutual funds included dividends as well as capital gains and losses. The yields on U.S. Treasury Bills having approximately 1 month to maturity, as of the beginning of each month, furnished the risk-free rate series, $R(t)$. The number of up-market $[Z_m(t) > R(t)]$ and down-market $[Z_m(t) \leq R(t)]$ observations contained within the period studied were 52 and 56, respectively.

There is, obviously, no special virtue inherent in a 1-month interval as the time unit of account for attempting to identify the influence of macro-price-forecasting efforts on portfolio managers' investment strategies—although any true underlying tendency for higher (lower) portfolio systematic-risk levels to coincide with upward (downward) general equities price movements should be revealed by a sequence of repeated observations on the two variables, whatever the observation interval chosen. Still, there is the issue of the transactions costs of portfolio revision and perhaps therefore a question as to the duration of a predicted broad price movement that might induce a deliberate response by the portfolio manager. For this reason, we also transformed the indicated mutual fund and market portfolio monthly return series into quarterly figures,[7] used Treasury Bills having 3 months to maturity to measure the contemporaneous risk-free rates, and reestimated the equations on that basis. Both sets of results are presented below.

III. The Empirical Evidence

Table 1 contains a summary of the regression statistics produced by estimating equations (1) and (10) on the 1971–79 monthly return data described, for the 67 mutual funds of interest. Several phenomena are apparent from those figures. The most obvious is that, in the aggregate, there is precious little evidence of any market-timing ability on the part of the fund managers. Indeed, if anything, the suggestion is that they tend to have the reverse "skill": the average estimated down-market beta of the 67 portfolios ($\hat{\beta}_1^* = 0.993$) is actually slightly higher than its up-market counterpart ($\hat{\beta}_2^* = 0.955$). Consistent with that finding, the average of the estimated intercepts of the Henriksson-Merton regression specifications ($\hat{\alpha}^* = 0.116\%$ per month) is, as predicted by our analysis above, greater than the average estimate ($\hat{\alpha}_p = 0.049\%$) obtained from the standard regression specification, which assumes port-

6. The particular time period selected was influenced by the data needs and data availability associated with a parallel research project. Nonetheless, we feel it is lengthy enough and recent enough to allow meaningful empirical conclusions to be drawn.

7. By the formula $1 + Z(t, t + 2) = [1 + Z(t)][1 + Z(t + 1)][1 + Z(t + 2)]$, where $Z(t, t + 2)$ denotes the quarterly rate of return spanning months t, $t + 1$, and $t + 2$.

TABLE 1 Summary Mutual Fund Performance Statistics: 1971–79 Monthly Data

Parameter	Mean Value	Standard Deviation	Extreme Values	
			Maximum	Minimum
A. From equation (1)—market timing ignored:				
$\hat{\alpha}_p$ (%)	.049	.526	2.878	− .719
$\hat{\beta}_p$	.972	.264	1.808	.225
R^2	.747	.159	.944	.234
B. From equation (10)—marketing timing considered:				
$\hat{\alpha}^*$ (%)	.116	.443	1.298	−1.061
$\hat{\beta}_1^*$ (Down)	.993	.267	1.709	.263
$\hat{\beta}_2^*$ (Up)	.955	.327	2.656	.192
adj. R^2	.748	.159	.943	.235
$\hat{\beta}_2^* - \hat{\beta}_1^*$	−.038	.288	1.833	−.502

NOTE.—Rate of return figures in percentages per month.

folio betas to be constant over time. That is, if mutual fund managers are generally poor market timers, the standard approach will tend to underestimate the concurrent contribution of their security-selection activities to observed overall investment performance. The value of the richer model employed here lies in its potential to distinguish properly between the two constituents, if either is present, and there is at least a hint that both may be.

It must also be said, however, that this hint is an extremely mild one. Although of the 67 mutual funds studied 42 yielded estimates of down-market betas that exceeded the corresponding up-market estimates for the period, in only four instances were the differences statistically significant at the 5% confidence level or better. Approximately this number might be expected to emerge by chance alone. Of the four, three had higher down-market betas. Apparently, then, there was really not much macro-price-forecasting going on within the sample or, if there was, it was either not being acted on or was overwhelmed by other portfolio design factors.

A similar conclusion applies to the evaluation of the fund managers' security-selection abilities. Even using the improved timing-sensitive Henriksson-Merton model, although for 41 of the 67 fund portfolios the estimates of the "selectivity" excess-return intercept $\hat{\alpha}^*$ were positive, in only five cases out of the 67 was the estimate statistically significant and, of these, three were negative values—that is, an abnormal, inferior security selection performance was indicated. Again, chance alone could produce virtually the same finding. Interestingly, the only two mutual funds for which $\hat{\alpha}^*$ was significantly positive were also two

TABLE 2 Summary Mutual Fund Performance Statistics: 1971–79 Quarterly Data

Parameter	Mean Value	Standard Deviation	Extreme Values	
			Maximum	Minimum
A. From equation (1)—market timing ignored:				
$\hat{\alpha}_p$ (%)	.015	1.438	8.418	−2.104
$\hat{\beta}_p$	.961	.255	1.694	.295
R^2	.799	.134	.962	.333
B. From equation (10)—marketing timing considered:				
$\hat{\alpha}^*$ (%)	.021	1.527	6.906	−3.106
$\hat{\beta}_1^*$ (Down)	.971	.289	1.661	.367
$\hat{\beta}_2^*$ (Up)	.951	.387	3.063	.216
adj. R^2	.802	.132	.964	.336
$\hat{\beta}_2^* - \hat{\beta}_1^*$	−.020	.446	2.593	−.976

NOTE.—Rate of return figures in percentages per quarter.

of the three for which the difference $\hat{\beta}_2^* - \hat{\beta}_1^*$ was significantly negative (they were perverse market timers).

The results obtained when quarterly rate of return data (and up/down market partitioning) were used to estimate the regression equations are shown in table 2. As is evident, the message is essentially a duplicate of that implied by table 1 from an analysis of the monthly data. The sampled funds' average estimated up-market beta is somewhat below the down-market figure, and the mean Henriksson-Merton selectivity intercept estimate marginally exceeds that produced by using equation (1) instead. In this instance, seven of the 67 funds displayed statistically significant differences between $\hat{\beta}_2^*$ and $\hat{\beta}_1^*$; once more, however, the majority (five) of these were in a negative direction, suggesting poor market timing; and there were three apparently superior ($\hat{\alpha}^* > 0$, significantly) and one apparently inferior ($\hat{\alpha}^* < 0$) security-selection performances among the group over the period. None of these results can be described as overwhelming evidence of collective portfolio management skill, either in micro or macro price forecasting.

Both the similarities in, and the contrasts between, the empirical messages conveyed by estimating the parameters of the two models at issue merit some further attention. It is evident from tables 1 and 2 that the overall "fit" of equation (1) to the mutual fund return data examined is little different from that of equation (10). The R^2 values for either the monthly or quarterly measurement intervals, averaged over the 67 funds, are quite high for both equations and within less than a percentage point of each other in both instances. If we take the 67

estimates of α_p and α^* as observations on the funds' security-selection abilities, we are unable to reject the null hypothesis that $\hat{\alpha}_p = \hat{\alpha}^*$; the same is true of the 67 observations on $\hat{\beta}_p$, $\hat{\beta}_1^*$, and $\hat{\beta}_2^*$. In the aggregate, therefore, it is tempting to conclude that not much has been gained by the use of the more sophisticated model.

We are more sanguine, however. The major reason for the similar overall performance of the two models is the fact that indeed there does seem not to have been much systematic market-timing activity undertaken by the funds' portfolio managers during the 1970s and, to the extent that there was, it was about as often in the wrong as in the right direction. We feel this is information worth having, and which estimates of equation (10) provide—even though the individual fund impacts tend to wash out in the composite findings when equation (1) is estimated.

Moreover, at the individual fund level, the distinctions are more substantial and the subtleties more apparent. To take the quarterly measurement-interval results as an illustration, the estimates of equation (1), which ignores market-timing phenomena, for each of the 67 funds yielded statistically significant excess-return intercepts in seven cases—all of which, in the usual interpretation, would be attributed to security selection decisions by the portfolio managers involved. Only one of these cases, however, coincided with those for which the Henriksson-Merton regression specification gave rise to a significant such intercept. The fund in question was adjudged by both models to display superior security selection performance, the explanation for that congruence being that its up-market and down-market betas were not significantly different by the Henriksson-Merton criterion. Hence, the two models should produce effectively identical estimates in a situation of this sort.

Of the other six funds, four had negative and two had positive estimated intercepts according to equation (1)—but in all instances, the equation (10) estimates suggested that these performance differentials could be imputed to market-timing behavior ($\hat{\beta}_2^* - \hat{\beta}_1^*$ negative for the first four, positive for the last two) rather than to security selection activities. Conversely, there were several additional funds for which the equation (10) estimates indicated a combination of significant timing and selectivity phenomena to be present—in opposite directions—that the corresponding equation (1) estimates were unable, of course, to detect at all. Thus, Henriksson and Merton's approach has the clear potential to provide a much richer insight into the nature and sources of managed portfolio performance differentials, when and if they exist. Based on the data examined here, however, we are forced to conclude that such differentials do exist only to a very minor extent in the recent investment track record of U.S. mutual funds, and that many of those that can be identified are of a character which provides little cause for celebration.

IV. Summary

A parametric statistical procedure that allows a joint test for the presence of either superior market-timing or security-selection skills in managed portfolios has been applied to examine empirically the investment performance of a sample of mutual funds during the decade of the 1970s. The test permits a more complete appraisal of the constituents of that performance than did prior methodologies and can eliminate certain biases in the estimates provided by the latter. We have discussed these in detail, and the empirical results are consistent with the new model's predictions. Nonetheless, those same results suggest that neither skillful market timing nor clever security selection abilities are evident in abundance in observed mutual fund return data, and the general conclusion of the prior literature that mutual funds have been unable collectively to outperform a passive investment strategy still seems valid.

Appendix A

Estimates of the Regression Parameters

According to the variable definitions $X(t) \equiv Z_m(t) - R(t)$, $Y(t) \equiv \max[0, -X(t)]$, $X_1(t) \equiv \min[0, X(t)]$, and $X_2(t) \equiv \max[0, X(t)]$, we have

$$E[X(t)] \equiv \bar{X} = E[X_1(t)] + E[X_2(t)] = \bar{X}_1 + \bar{X}_2$$

$$E[Y(t)] \equiv \bar{Y} = -\bar{X}_1$$

$$\mathrm{var}[Y(t)] \equiv \sigma_y^2 = \mathrm{var}[X_1(t)] \equiv \sigma_1^2$$

$$\mathrm{var}[X(t)] \equiv \sigma_x^2 = \sigma_1^2 + \sigma_2^2 - 2\bar{X}_1\bar{X}_2$$

$$\mathrm{cov}[X(t), Y(t)] \equiv \sigma_{xy} = \bar{X}_1\bar{X}_2 - \sigma_1^2$$

$$\mathrm{cov}[Z_p(t), X(t)] \equiv \sigma_{px} = (b + \theta_1)(\sigma_1^2 - \bar{X}_1\bar{X}_2) + (b + \theta_2)(\sigma_1^2 - \bar{X}_1\bar{X}_2)$$

$$\mathrm{cov}[Z_p(t), Y(t)] \equiv \sigma_{py} = (b + \theta_2)(\bar{X}_1\bar{X}_2) - (b + \theta_1)\sigma_1^2.$$

From this, and equation (2) in the text, it follows that the large-sample least squares estimates of β_1 and β_2 can be written as

$$\mathrm{plim}\ \hat{\beta}_1 = \frac{\sigma_{px}\sigma_y^2 - \sigma_{py}\sigma_{xy}}{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2} = b + \theta_2 = p_2\eta_2 + (1 - p_2)\eta_1$$

$$\mathrm{plim}\ \hat{\beta}_2 = \frac{\sigma_{py}\sigma_x^2 - \sigma_{px}\sigma_{xy}}{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2} = \theta_2 - \theta_1 = [p_1 + p_2 - 1][\eta_2 - \eta_1]$$

and, thus,

$$\mathrm{plim}\ \hat{\alpha} = E[Z_p(t)] - E[\bar{R}] - \mathrm{plim}\ \hat{\beta}_1\bar{X} - \mathrm{plim}\ \hat{\beta}_2\bar{Y}$$

$$\mathrm{plim}\ \hat{\alpha} = \bar{Z}_p - \bar{R} - \bar{X}_1(b + \theta_1) - \bar{X}_2(b + \theta_2).$$

If instead a regression specification like that of equation (1) in the text is used, the large-sample least squares estimate of β_p will be

$$\text{plim } \hat{\beta}_p = \frac{\text{cov}[Z_p(t), X(t)]}{\text{var}[X(t)]} = \frac{\sigma_{px}}{\sigma_x^2}$$

$$\text{plim } \hat{\beta}_p = \frac{(b + \theta_1)(\sigma_1^2 - \bar{X}_1\bar{X}_2) + (b + \theta_2)(\sigma_2^2 - \bar{X}_1\bar{X}_2)}{\sigma_1^2 + \sigma_2^2 - 2\bar{X}_1\bar{X}_2}$$

$$\text{plim } \hat{\beta}_p = b + \frac{\theta_1(\sigma_1^2 - \bar{X}_1\bar{X}_2) + \theta_2(\sigma_2^2 - \bar{X}_1\bar{X}_2)}{\sigma_1^2 + \sigma_2^2 - 2\bar{X}_1\bar{X}_2}.$$

Since $\theta_1 = \theta_2[1 - (1/q)]$, we therefore have

$$\text{plim } \hat{\beta}_p = b + \theta_2 - \theta_2 \Big/ \left[\frac{q(\sigma_1^2 - \bar{X}_1\bar{X}_2) + q(\sigma_2^2 - \bar{X}_1\bar{X}_2)}{\sigma_1^2 - \bar{X}_1\bar{X}_2}\right]$$

$$\text{plim } \hat{\beta}_p = b + \theta_2 - \left\{\frac{\theta_2}{q[1 + (\sigma_2^2 - \bar{X}_1\bar{X}_2)/(\sigma_1^2 - \bar{X}_1\bar{X}_2)]}\right\}.$$

Appendix B

Nature of the Up-Market/Down-Market Parameter Estimation Methodology

Equation (10) in the text can be written for estimation as

$$Z(t) = a + bX_1(t) + cX_2(t) + \epsilon_p(t)$$

where $Z(t) \equiv Z_p(t) - R(t)$. The corresponding sum of squared residuals about the plane to be estimated in ZX_1X_2 space is, for n observations,

$$\sum_{t=1}^{n} [\epsilon_p(t)]^2 = \sum_{t=1}^{n} [Z(t) - a - bX_1(t) - cX_2(t)]^2,$$

and the least squares estimates of coefficients a, b, and c can be expressed in matrix form (Johnston 1972) as

$$\hat{B} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} n & \sum^{d} X(t) & \sum^{u} X(t) \\ \sum^{d} X(t) & \sum^{d} [X(t)]^2 & 0 \\ \sum^{u} X(t) & 0 & \sum^{u} [X(t)]^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum^{n} Z(t) \\ \sum^{d} X(t)Z(t) \\ \sum^{u} X(t)Z(t) \end{bmatrix},$$

where

$\sum^{n} Z(t)$ = the sum of all excess returns on the portfolio p;

$\sum^{u} X(t)$ = the sum of all up-market excess returns on the market portfolio;

$\sum^{u}[X(t)]^2$ = the sum of squares of all up-market excess returns on the market portfolio;

$\sum^{d}X(t)$ = the sum of all down-market excess returns on the market portfolio;

$\sum^{d}[X(t)]^2$ = the sum of squares of all down-market excess returns on the market portfolio;

$\sum^{u}X(t)Z(t)$ = the sum of the products of $X(t)$ and $Z(t)$ for all up-market observations; and

$\sum^{d}X(t)X(t)$ = the sum of the products of $X(t)$ and $Z(t)$ for all down-market observations.

It can be shown that these least squares estimates are the same as the estimates obtained from minimizing the sum of the sum of the squares of the deviations about two straight lines in planes ZX_1 and ZX_2, respectively, subject to the constraint that they have a common intercept. This is because, when the data points are plotted in ZX_1X_2 space, all up-market observations lie on the ZX_2 plane and all down-market ones on the ZX_1 plane. The sum of the squared residuals about the two lines is

$$\xi = \sum^{d}[Z(t) - a - bX(t)]^2 + \sum^{u}[Z(t) - a - cX(t)]^2.$$

Taking the partial derivatives with respect to a, b, and c and equating to zero, we obtain

$$\frac{\partial\xi}{\partial a} = \sum^{n}Z(t) - an - b\sum^{d}X(t) - c\sum^{u}X(t) = 0$$

$$\frac{\partial\xi}{\partial b} = \sum^{d}X(t)Z(t) - a\sum^{d}X(t) - b\sum^{d}[X(t)]^2 = 0$$

$$\frac{\partial\xi}{\partial c} = \sum^{u}X(t)Z(t) - a\sum^{u}X(t) - c\sum^{u}[X(t)]^2 = 0.$$

Hence, the normal equations for a, b, and c are

$$\begin{bmatrix} n & \sum^{d}X(t) & \sum^{u}X(t) \\ \sum^{d}X(t) & \sum^{d}[X(t)]^2 & 0 \\ \sum^{u}X(t) & 0 & \sum^{u}[X(t)]^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum^{n}Z(t) \\ \sum^{d}X(t)Z(t) \\ \sum^{u}X(t)Z(t) \end{bmatrix},$$

which are exactly the same as above for $\hat{B}$.

Appendix C

The Mutual Fund Sample

Fund	Stated Objectives	Fund	Stated Objectives
Affiliated	G,I	Keystone Custodian (K2)	G
Anchor Growth	G	Keystone Custodian (S1)	G,I
Boston	G,I,S	Keystone Custodian (S3)	G
Century Shares Trust	G	Keystone Custodian (S4)	MCG
Chase Fund of Boston	MCG	Life Insurance Investors	G
Chase Special	MCG	Loomis Sayles Mutual Fund	G,I,S,
Chemical	G	Manhattan	G
Colonial	I,G	Massachusetts	S,I,G
Colonial Income	I	Massachusetts Investors Growth	G,I
Delaware	G,I,S	Massachusetts Investors Trust	G
Dividend Shares	I,G	Mathers	MCG
Dreyfus	G	National Investors Corporation	G
Dreyfus Leverage	MCG	National Securities Stock Series	I,G

Eaton and Howard Balanced	I,G,S	National Securities Income Series	I
Eaton and Howard Special	MCG	Omega	MCG
Eaton and Howard Stock	G,I	One William Street	G,I
Eberstadt	G	Oppenheimer	MCG
Explorer	MCG	Pioneer Fund	G,I
Commonwealth Investment Company	G,I,S	Puritan	I
Fairfield	MCG	Putnam Duofund Capital Shares	MCG
Fidelity Capital	G	Putnam Growth	G
Fidelity	G,I	Putnam Investors	G
Fidelity Trend	G	Scudder Special	MCG
Financial Dynamics	MCG	Security Equity	MCG
Fletcher Capital	MCG	Selected Special Shares	MCG
Foursquare	G	Sigma Trust Shares	S,I,G
Fundamental Investors	G,I	Source Capital	I
George Putnam Fund of Boston	S,I,G	State Street Investment Company	G,I
Industries Trend	G	Stein Roe & Farnham Balanced	G,I,S
Investment Company of America	G,I	Technology	G
Istel Fund, Inc.	G,S,I	Value Line Special Situation	MCG
Ivest	MCG	Vance Sanders Special	MCG
Keystone Custodian (B4)	I	Windsor	G
Keystone Custodian (K1)	I		

* G = growth; I = income; S = stability of capital; and MCG = maximum capital gain.

References

Alexander, Gordon J., and Stover, Roger D. 1980. Consistency of mutual fund performance during varying market conditions. *Journal of Economics and Business* 32 (Spring): 219–25.

Black, Fisher; Jensen, Michael C.; and Scholes, Myron. 1972. The capital asset pricing model: Some empirical tests. In Michael C. Jensen (ed.), *Studies in the Theory of Capital Markets*. New York: Praeger.

Blume, Marshall E., and Friend, Irwin. 1973. A new look at the capital asset pricing model. *Journal of Finance* 28 (March): 19–33.

Fabozzi, Frank J., and Francis, Jack Clark. 1979. Mutual fund systematic risk for bull and bear markets. *Journal of Finance* 34 (December): 1243–50.

Fama, Eugene F. 1972. Components of investment performance. *Journal of Finance* 27 (June): 551–67.

Henriksson, Roy D. 1984. Market timing and mutual fund performance: An empirical investigation. In this issue.

Henriksson, Roy D., and Merton, Robert C. 1981. On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *Journal of Business* 54 (October): 513–33.

Jensen, Michael C. 1968. The performance of mutual funds in the period 1945–1964. *Journal of Finance* 23 (May): 389–416.

Jensen, Michael C. 1969. Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business* 42 (April): 167–247.

Jensen, Michael C. 1972. Optimal utilization of market forecasts and the evaluation of investment performance. In G. P. Szego and K. Shell (eds.), *Mathematical Methods in Investment and Finance*. Amsterdam: North-Holland.

Johnston, J. 1972. *Econometric Methods*. 2d ed. New York: McGraw-Hill.

Kon, Stanley, and Jen, Frank C. 1978. Estimation of time-varying systematic risk and performance for mutual fund portfolios: An application of switching regression. *Journal of Finance* 33 (May): 457–76.

Kon, Stanley, and Jen, Frank C. 1979. The investment performance of mutual funds: An empirical investigation of timing, selectivity, and market efficiency. *Journal of Business* 52 (April): 263–89.

Merton, Robert C. 1981. On market timing and investment performance. I. An equilibrium theory of value for market forecasts. *Journal of Business* 54 (July): 363–406.

Miller, Tom W., and Gressis, Nicholas. 1980. Nonstationarity and evaluation of mutual fund performance. *Journal of Financial and Quantitative Analysis* 15 (September): 639–54.

Veit, E. Theodore, and Cheney, John M. 1982. Are mutual funds market timers? *Journal of Portfolio Management* 8 (Winter): 35–42.