Ravi Jagannathan
Robert A. Korajczyk
Northwestern University

Assessing the Market Timing Performance of Managed Portfolios*

A number of techniques have been proposed to measure portfolio performance and to distinguish between performance due to forecasting security-specific returns and performance due to forecasting market-wide events. We show theoretically and empirically that it is possible to construct portfolios that show artificial timing ability when no true timing ability exists. In particular, investing in options or levered securities will show spurious market timing. These types of securities will also induce the negative correlation between measured selectivity and timing ability found by others. We suggest specification tests to help distinguish between spurious and true timing ability. In addition, the tests can be used to distinguish between different models of the manager's reaction function.

I. Introduction

Assessments of managers' investment performance are likely to influence the manner in which investors allocate their wealth across various professionally managed portfolios and directly or indirectly to influence the compensation of managers (see Korschot 1978; and Smith 1978). In addition, tests of abnormal performance can be viewed as evidence regarding the validity of the efficient market hypothesis. Hence accurate measurement of portfolio managers' investment performance is an important topic in finance.

It is common practice to divide portfolio performance into two main components, security selection and market timing.[1] A number of techniques are available to estimate the separate components of portfolio performance.[2] We will limit our attention to those parametric techniques that only assume knowledge of the managed portfolios' returns and hence do not

* The authors are grateful to Gregory Connor, Mark Grinblatt, Bruce Grundy, Lars P. Hansen, Allan Kleidon, Thomas Lys, Peter E. Rossi, Daniel Siegel, Sheridan Titman, and the referee for helpful comments.

1. Finer divisions of performance are possible (e.g., see Fama 1972).

2. For example, see Treynor and Mazuy 1966; Fama 1972; Jensen 1972; Kon and Jen 1979; Henriksson and Mert on 1981; Kon 1983; Modest 1983; Pfleiderer and Bhattacharya 1983; Chang and Lewellen 1984; and Henriksson 1984.

(*Journal of Business*, 1986, vol. 59, no. 2, pt. 1)

0021-9398/86/5902-0002$01.50

require direct observation of the managers' market forecasts or portfolio composition.

The empirical evidence on timing ability (e.g., Kon 1983; Chang and Lewellen 1984; Henriksson 1984; and Lockwood and Kadiyala 1985) seems to indicate that significant timing ability is rare. However, the most puzzling aspect of the empirical evidence is that the average timing measures across mutual funds are negative and that those funds that do exhibit significant timing performance show negative performance more often than positive performance. Also, Kon (1983) and Henriksson (1984) find that there is negative correlation (cross-sectionally) between the measures of security selection and market timing.[3] After performing a particularly careful set of diagnostics on the market-timing tests, Henriksson (1984, p. 93) concludes that the "specification used in the parametric tests must be questioned because of the persistence of a negative correlation between $\hat{\alpha}$ and $\hat{\beta}_2$," where $\hat{\alpha}$ and $\hat{\beta}_2$ are the selectivity and timing performance measures, respectively. He suggests a number of potential explanations for these results, including errors-in-variables bias, misspecification of the market portfolio, and use of a single-factor rather than a multifactor asset-pricing model.[4] We suggest an explanation for the empirical results that relies on the nonlinear payoff structure of options and "option-like" securities.[5]

In Section II we show that the portfolio strategy of buying call options (in this case, calls on the market) will exhibit positive timing performance and negative security selection even though no market forecasting or security-specific forecasting is being done. The insight of the analysis of Henriksson and Merton (1981) and Merton (1981) is that market-timing ability, of the type specified in their papers, can be viewed as a free option. In our case, in which one invests in call options, the same payoff structure is obtained (hence positive market timing), but our return is reduced by the premium paid for the option (leading to negative security selection). Thus we predict the negative cross-sectional correlation between measures of timing and security selection if managers are purchasing options or option-like securities such as the common stock of highly levered firms.

This analysis of artificial timing is incomplete since it seemingly requires funds to sell call options or buy put options in order to explain the generally negative market-timing performance found in the empir-

3. The negative correlation may be partially due to sampling error. However, the results in Henriksson (1984) and below indicate that this is unlikely to be the sole cause of the negative correlation.

4. Negative timing performance can be rational for investors with increasing absolute risk aversion, such as investors with quadratic utility (see Grinblatt and Titman 1985).

5. The option-like characteristics of common stock need not be due solely to leverage effects. Any fixed costs that induce differential skewness in the return distributions across firms will lead to the same results.

ical literature. Note, however, that the market proxy (the value-weighted NYSE stock index) is a portfolio of stocks that are, to a greater or lesser extent, options (due to their varying levels of risky debt). In this case the sign of the "artificial" market-timing performance of a given mutual fund will depend on whether the "average" stock held by the fund has more or less of an option effect than the "average" stock in the index.[6] Thus funds that tend to invest in stocks with little or no risky debt (blue chips?) will show negative-timing performance, while funds that invest in small, highly levered stocks will show positive-timing performance. The average negative timing performance found in Kon (1983), Chang and Lewellen (1984), and Henriksson (1984) may be due to the fact that the mutual funds in the sample tend to invest in larger, more established, and lower-levered firms than the average firm on the NYSE. In Section III we show empirically that one can obtain statistically significant artificial timing using naive portfolio formation rules.

In Section IV we propose a general specification test that may be used to determine whether "significant" measured timing performance is due to some violation of the theoretical underpinnings of the model (such as the stochastic properties of asset returns). In addition, the test may be useful in choosing among different specifications of the managers' reaction functions.

The specification tests are applied to the techniques of Henriksson and Merton (1981) and Pfleiderer and Bhattacharya (1983). For the former model the specification tests perform well in the sense that they reject the specification of the model when statistically significant artificial timing is found. For the latter model the tests seem to have lower power, probably because the model is a closer approximation to the return structure of option-like securities. The empirical results highlight the usefulness and limitations of our proposed specification tests. A summary of our conclusions and proposals for extensions of this work are presented in Section V.

II. Creation of Artificial Timing Performance

In this section we show that the proposed parametric techniques for measuring timing and selectivity performance can show spurious timing and selectivity performance (of opposite signs) when applied to option-like securities. We use the parametric tests proposed by Henriksson and Merton (1981, pp. 525–31) as our performance measurement technique. Their method is chosen because it is a widely known and tested technique.

6. Here "average" means a weighted average where the weights are the *portfolio* weights.

A. *The Henriksson and Merton Parametric Tests*

We briefly describe the Henriksson and Merton technique (see Henriksson and Merton 1981). Assume a statistical model for asset returns given by

$$R_i(t) = \alpha + \beta_i x(t) + \epsilon_i(t), \quad i = 1, \ldots, N, \tag{1}$$

where $R_i(t)$ is the "excess return" on security i, that is, the return on security i in excess of the risk-free rate; $x(t)$ is the excess return on the market portfolio; and $\epsilon(t)$ is a random error satisfying

$$E[\epsilon(t)] = 0,$$

$$E[\epsilon(t)|x(t)] = 0,$$

$$E[\epsilon(t)|\epsilon(t-\tau)] = 0, \quad \tau = 1, 2, 3, \ldots$$

Also assume that equilibrium asset expected returns are determined according to the Capital Asset Pricing Model (CAPM).[7] Each period we assume that the manager attempts to forecast whether the market will have positive or negative excess returns (i.e., $x[t] > 0$ or $x[t] < 0$). A manager who believes a nonpositive value of $x(t)$ is more likely will choose a level of systematic risk (β) equal to η_1; otherwise a level of risk equal to η_2 is chosen ($\eta_2 > \eta_1$). The HM test for market timing is a test of the hypothesis $\beta_2 = 0$ in

$$R_p(t) = \alpha + \beta_1 x(t) + \beta_2 y(t) + \epsilon(t), \tag{2}$$

where $y(t) = \max[0, -x(t)]$. If we define the long-run average beta of the portfolio as b, then Henriksson and Merton (1981) show that (under the assumptions of the model)

$$\text{plim } \hat{\beta}_1 = b + \bar{\theta}_2, \tag{3}$$

$$\text{plim } \hat{\beta}_2 = \bar{\theta}_2 - \bar{\theta}_1, \tag{4}$$

where

$$\bar{\theta}_1 = E[\beta(t) - b|x(t) \leq 0],$$

$$\bar{\theta}_2 = E[\beta(t) - b|x(t) > 0],$$

$$b = E[\beta(t)],$$

and $\hat{\beta}_1$ and $\hat{\beta}_2$ are the OLS estimates of β_1 and β_2. If the manager has no timing ability or does not act on market forecasts ($\eta_1 = \eta_2$), then $\bar{\theta}_1 = \bar{\theta}_2 = 0$. In addition, $\hat{\alpha}$ (the OLS estimate of α) is a consistent estimate of security selection ability. A significantly positive value of $\hat{\beta}_2$ indicates superior timing ability, while significantly negative values of $\hat{\beta}_2$ indicate perverse timing activity (i.e., $\eta_2 < \eta_1$).

7. The Henriksson and Merton (1981) analysis can accommodate multifactor intertemporal asset-pricing models in addition to the CAPM.

B. *Options and Timing*

Consider the portfolio policy of buying one-period call options on the market each period. Clearly, there is neither market timing nor security selection taking place. We show that this portfolio policy will lead to positive timing performance and negative selection performance in (2). Let $V(t)$ denote the value of the market index (for simplicity we will assume no dividends are paid); K is the exercise price of the call option; and C is the current $(t - 1)$ value of the call. The excess return on our portfolio is merely the excess return on the call option:

$$\begin{aligned}
R_p(t) &= \max\left[\frac{V(t) - K}{C}, 0\right] - (1 + R_F) \\
&= \frac{V(t-1)}{C} \max\left[\frac{V(t) - V(t-1)}{V(t-1)} - R_F \right. \\
&\quad \left. + \frac{V(t-1)(1 + R_F) - K}{V(t-1)}, 0\right] - (1 + R_F) \\
&= \frac{V(t-1)}{C} \max\left[x(t) + \frac{V(t-1)(1 + R_F) - K}{V(t-1)}, 0\right] \\
&\quad - (1 + R_F).
\end{aligned}$$

Assume that the economy is stationary in the sense that R_F and the moments of $x(t)$ and $y(t)$ are constant through time. This implies that $V(t - 1)/C$ is a constant that we will denote by λ. For simplicity, choose $K = V(t - 1)(1 + R_F)$, then

$$R_p(t) = \lambda \max[x(t), 0] - (1 + R_F).$$

Using the fact that $x(t) = \max[0, x(t)] + \min[0, x(t)]$ and $y(t) = -\min[0, x(t)]$, one can show that OLS estimates of the parameters in (2) for this portfolio are (see App.):

$$\hat{\beta}_1 = \beta_1 = \lambda > 0 \tag{5a}$$

$$\hat{\beta}_2 = \beta_2 = \lambda > 0 \tag{5b}$$

$$\hat{\alpha} = \alpha = -(1 + R_F) < 0. \tag{5c}$$

That is, this portfolio yields an exact fit to the regression equation (2). Thus we have a positive measure of market timing (β_2) and a negative measure of security selection (α) for this portfolio even though there is no timing or selection involved in the portfolio construction process.

These values for α, β_1, and β_2 are dependent on the particular choice of the exercise price (in this case $K = V[t - 1][1 + R_F]$). This dependence can be seen by noting that, if one were to choose K equal to zero, then $\beta_1 = 1$, $\beta_2 = 0$, and $\alpha = 0$. Thus a portfolio manager can choose to show varying degrees of market timing by choosing options with different exercise prices (or stocks of firms with lower or higher proportions of risky debt).[8] However, there is no free lunch since higher levels of measured timing performance are associated with lower levels of measured security selection performance. Also, artificial timing created through intraperiod trading by the portfolio manager (see Pfleiderer and Bhattacharya 1983) also induces a negative relation between measured timing and selectivity. A natural measure of total performance would be the value of the manager's timing ability plus the manager's selection ability. Merton (1981) has shown that the total value of the manager's timing and selectivity in this model is $\beta_2 p + \alpha/(1 + R_F)$, where p is the value of a one-period put on the market index, with current value normalized to one and a strike price of $1 + R_F$. We know from the put-call parity relationship for European options that $c = 1 + p - \text{PV}(1 + R_F)$, where c is the value of a one-period call on the index with exercise price $= 1 + R_F$ when the current index value $= 1$, and PV denotes present value. Hence $c = p$. We also know that the value of a call on the index with current value V and strike price $(1 + R_F)V$ is $V \cdot c$. Therefore

$$\beta_2 p + \frac{\alpha}{1 + R_F} = \frac{V}{C} p + \frac{-(1 + R_F)}{(1 + R_F)} = \frac{V}{V \cdot c} \cdot p - 1 = 0.$$

Unfortunately, this one-for-one tradeoff between artificial timing and selectivity does not generalize to different exercise prices and option maturities longer than one period. This measure of total performance also requires the valuation of a one-period put option on the market index. Thus more conservative bounds on total performance may be useful. An upper bound on total performance will be $[\alpha + \beta_2 E(x_2)]/(1 + R_F)$ since $E(x_2)/(1 + R_F) \geq c = p$, where $x_2 = \max(0, x)$. Conversely, a lower bound on total performance will be $[\alpha + \beta_2 E(y)]/(1 + R_F)$ since $E(y)/(1 + R_F) \leq p$. Hence one method of identifying artificial performance will be to check whether $\hat{\alpha} + \hat{\beta}_2 \bar{y} > 0$ versus $\hat{\alpha} + \bar{\beta}_2 \bar{y} \leq 0$. An advantage of these bounds is that they do not require explicit valuation of the put. Disadvantages of the bounds are that they are strictly valid only when total artificial performance is zero and that they may have low power.

Table 1 presents simulation results of the OLS estimates of (2) when the portfolio consists of a call option written on an asset that is highly correlated with x.[9] Note that the values of $\bar{\hat{\beta}}_2$ ($\bar{\hat{\alpha}}$) increase (decrease)

8. Dybvig and Ross (1985, p. 395) also note that market-timing measures are not robust to allowing portfolios of options or common stock of highly levered firms.

9. Rather than use an option with 1 month to maturity we calculate the monthly returns from holding options with 12 months to maturity, assuming the Black-Scholes (1973) option-pricing model holds. We simulate a monthly index whose correlation with the market is .99, for a 14-year time period (168 observations) with $E(x) = .0047$ and $\sigma(x) = .05295$. One hundred replications of the regression are performed.

TABLE 1 **Simulation Results for Henriksson-Merton Market Timing Test Where Portfolio Consists of a Call Option ($R_p = \alpha + \beta_1 x + \beta_2 y - \epsilon$)**

Exercise Price ÷ Market Price	$\bar{\hat{\alpha}}$	$\bar{\hat{\beta}}_1$	$\bar{\hat{\beta}}_2$	% Rejection of $\alpha = \beta_2 = 0$ (5% Level)	Average $\hat{\alpha} + \hat{\beta}_2\bar{y}$	% Rejection White Test (5% Level)	% Rejection Exclusion Restriction (5% Level)
.6	−.002	2.636	.027	14	−.001	0	8
	(.002)	(.040)	(.075)				
.7	−.003	3.400	.057	14	−.002	0	15
	(.002)	(.052)	(.096)				
.8	−.010	4.652	.396	84	−.003	6	19
	(.003)	(.073)	(.130)				
.9	−.027	6.505	1.261	100	−.005	42	61
	(.004)	(.101)	(.182)				
1.0	−.063	9.115	3.133	100	−.011	82	95
	(.006)	(.175)	(.294)				
1.1	−.116	12.348	5.889	100	−.014	95	100
	(.009)	(.293)	(.462)				
1.2	−.188	16.097	9.664	100	−.028	93	100
	(.014)	(.524)	(.763)				
1.3	−.278	20.277	14.269	100	−.032	95	100
	(.022)	(.887)	(1.259)				
1.4	−.375	24.593	19.303	100	−.044	98	100
	(.031)	(1.300)	(1.829)				
1.5	−.486	29.241	25.094	100	−.063	94	100
	(.046)	(1.975)	(2.765)				

NOTE.—Figures reported are averages of OLS parameter estimates across 100 replications. Standard errors in parentheses are the average standard errors of OLS estimates using White's correction for heteroscedasticity. Monthly returns with $E(x) = .0047$ per month; $\sigma(x) = .0530$; number of observations = 168. Options are assumed to be written on an asset whose returns have correlation with x of .99, 12 months to maturity. Option prices are simulated using the Black-Scholes (1973) option-pricing model.

monotonically as the ratio of the exercise price to the value of the market index increases. In fact, the sample correlation (across exercise prices) between $\bar{\hat{\alpha}}$ and $\bar{\hat{\beta}}_2$ is -0.9994. Also note that when the hypothesis that $\alpha = \beta_2 = 0$ is tested using a χ^2 statistic (at the 5% level) we get rejection rates of 14% for exercise-price/current-price ratios of 0.6 and 0.7 and rejection rates of 84%–100% for ratios of 0.8 or greater. The standard errors used in these tests are not the usual OLS standard errors. Rather, they are the heteroscedasticity-consistent standard errors proposed by White (1980*a*), Hansen (1982), and Hsieh (1983).[10] This is done because Henriksson and Merton (1981) show that the errors in (2) are heteroscedastic when the manager has timing ability. Breen, Jagannathan, and Ofer (1985) show that the usual (uncorrected) standard errors of the OLS parameter estimates are likely to have substantial downward bias when a manager is a market timer. An alternative to our procedure, suggested in Henriksson (1984), is to use weighted least squares (WLS). The difference between OLS and WLS is discussed in Section IV.

The average value of $\hat{\alpha} + \hat{\beta}_2(1 + R_F)p$ increases monotonically with the ratio of exercise price to market price from -0.001 to 0.040. Also, the values of $\bar{\hat{\alpha}} + \bar{\hat{\beta}}_2\bar{y}$ are negative and monotonically decreasing as the ratio of exercise price to market price increases. The values range from -0.001 to -0.063. This is consistent with the fact that there is no real ability used to construct these portfolios and that $\hat{\alpha} + \hat{\beta}_2\bar{y}$ is a rough estimate of the lower bound for total ability.

The simulation results are only meant to be suggestive since the typical portfolio manager would not hold such a portfolio. In the next section we demonstrate the existence of artificial timing using actual portfolios.

10. The usual estimate of the covariance matrix for the OLS parameter estimates ($\text{var}\{T^{1/2}[\hat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta}]\}$) in the regression

$$Y_t = \mathbf{Z}_t\boldsymbol{\beta} + \epsilon_t, \quad t = 1, \ldots, T,$$

is given by $s^2(Z'Z/T)^{-1}$ where $Z = T \times k$ matrix of independent variables;

$$s^2 = \sum_{t=1}^{T} \frac{(Y_t - \mathbf{Z}_t\hat{\boldsymbol{\beta}}_{OLS})^2}{T}.$$

The heteroscedasticity-corrected covariance matrix of the OLS parameter estimates is given by

$$\left(\frac{Z'Z}{T}\right)^{-1}\left[T^{-1}\sum_{t=1}^{T}(Y_t - \mathbf{Z}_t\hat{\boldsymbol{\beta}}_{OLS})^2\, Z_t'Z_t\right]\left(\frac{Z'Z}{T}\right)^{-1}.$$

For our regressions the corrected standard errors are consistently larger than the usual standard errors. For the actual portfolio regressions reported in Sec. III it is not uncommon for the corrected standard errors to be two or three times the size of the uncorrected standard errors.

III. Artificial Timing: Empirical Results

In Section II we suggest an explanation of the tendency for mutual funds to show negative measures of market timing, on average. If the funds being analyzed tend to hold assets that are less option-like than the average asset in the market proxy, then one would expect to see negative timing and positive selectivity measures. On the other hand, funds holding assets that are more option-like than the assets in the market proxy should show positive measures of market timing and negative measures of security selectivity.

To demonstrate the possibility of artificial timing two naive portfolio strategies are used. Portfolio VW will be a market-value-weighted portfolio of NYSE stocks (the value-weighted index constructed by the Center for Research in Security Prices [CRSP]), and portfolio EW will be an equally weighted portfolio of NYSE stocks (the equal-weighted index from CRSP). Clearly, these two portfolio "strategies" involve no timing or security selection ability. However, if our conjecture is correct, they will show significant measured timing performance.

Assume that one chooses the equally weighted portfolio as the proxy for the market portfolio. Portfolio VW, according to our conjecture, should show negative timing and positive selectivity because it is composed of the common stock of firms that are less levered than the proxy, while EW will show no timing or selection ability. Conversely, if one uses the value-weighted index as the proxy for the market, then portfolio VW will show no timing, while portfolio EW should show positive timing and negative selectivity.

Table 2 contains the results of the regressions

$$x_{\mathrm{VW}}(t) = \alpha_{\mathrm{VW}} + \beta_{1\mathrm{VW}} x_{\mathrm{EW}}(t) + \beta_{2\mathrm{VW}} y_{\mathrm{EW}}(t) + \delta_{\mathrm{VW}} d(t) + \epsilon_{\mathrm{VW}}(t),$$

$$x_{\mathrm{EW}}(t) = \alpha_{\mathrm{EW}} + \beta_{1\mathrm{EW}} x_{\mathrm{VW}}(t) + \beta_{2\mathrm{EW}} y_{\mathrm{VW}}(t) + \delta_{\mathrm{EW}} d(t) + \epsilon_{\mathrm{EW}}(t),$$

where x is the excess returns for the portfolio, $y = \max(0, -x)$, and $d(t)$ is a dummy variable that is equal to unity during January and zero otherwise. Excess returns are portfolio returns in excess of the short-term Treasury bill rate from Ibbotson and Sinquefield (1982).

There is substantial evidence that smaller (larger) firms earn higher (lower) returns relative to the CAPM and that this differential occurs primarily in January (see Keim 1983). Because the equally weighted portfolio has much higher concentration in small firms than the value-weighted portfolio, failure to include the dummy would lead to estimates of α that combine artificial timing and the small-firm January seasonal effect. Since we are interested in isolating the effects of option-like securities on the value of α and β_2, the January dummy variable $d(t)$ is included.

The results are reported for the 56-year period January 1926–December 1981. In addition, we report results for four 14-year sub-

TABLE 2 **Henriksson-Merton Market Timing Model Equal- and Value-Weighted NYSE Portfolios ($R_p[t] = \alpha\ \beta_1 x[t] + \beta_2 y[t] + \delta d[t] + \epsilon[t]$)**

Period	$\hat{\alpha}$	$\hat{\delta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	p-Value: $\alpha = 0, \beta_2 = 0$	p-Value: White Test	p-Value: Exclusion Test
				Value Weighted vs. Equal Weighted			
1926–81	.008	−.022	.574	−.295	<.001	.002	.026
	(.002)	(.003)	(.047)	(.064)			
1926–39	.016	−.028	.534	−.370	<.001	.005	.145
	(.003)	(.006)	(.058)	(.080)			
1940–53	.007	−.021	.602	−.266	<.001	<.001	.126
	(.002)	(.005)	(.046)	(.064)			
1954–67	.003	−.020	.797	−.130	.073	.111	.195
	(.001)	(.004)	(.044)	(.072)			
1968–81	.002	−.027	.665	−.111	.487	.437	.062
	(.002)	(.007)	(.066)	(.095)			
				Equal Weighted vs. Value Weighted			
1926–81	−.008	.034	1.453	.347	.002	.004	.034
	(.002)	(.005)	(.096)	(.123)			
1926–39	−.017	.046	1.606	.542	<.001	.045	.762
	(.004)	(.011)	(.127)	(.157)			
1940–53	−.002	.031	1.262	.054	.645	.904	.602
	(.002)	(.008)	(.075)	(.099)			
1954–67	.002	.027	.962	−.143	.435	.686	.279
	(.002)	(.005)	(.073)	(.112)			
1968–81	.002	.042	1.142	−.102	.888	.595	.519
	(.004)	(.011)	(.154)	(.225)			

NOTE.—Test for $\alpha = \beta_2 = 0$ is a Wald test using heteroscedasticity-corrected covariance matrix. Exclusion test is a t-test for $\gamma = 0$ in $R_p = \alpha + \beta_1 x + \beta_2 y + \delta d + \gamma x^2 + \epsilon$. Portfolios are CRSP value-weighted and equal-weighted indexes. Risk-free rate from Ibbotson and Sinquefield (1982).

periods. Note that the last subperiod (January 1968–December 1981) corresponds closely to the period studied in Henriksson (1984).

In the first half of table 2 we treat the value-weighted index as our portfolio and the equal-weighted index as the proxy for the market portfolio. Since the value-weighted portfolio has lower concentration in option-like securities than the equal-weighted index, our conjecture predicts that $\alpha > 0$ and $\beta_2 < 0$ for these regressions. The parameter estimates have the predicted signs for the overall period and every subperiod. Also, the hypothesis that there is no timing or selection ability (i.e., $\alpha = \beta_2 = 0$) can be rejected at any reasonable level of significance for the overall period and the first two subperiods and can be rejected at the 10% level for the third subperiod (see col. 6).

In the second half of table 2 we treat the equal-weighted portfolio as our fund and the value-weighted index as the proxy for the market. Our conjecture implies that $\alpha < 0$ and $\beta_2 > 0$ for these regressions. The parameter estimates have the predicted signs for the overall period and the first two subperiods, while the estimates have the opposite signs for the last two subperiods. For these last two periods, however, we cannot reject $\alpha = \beta_2 = 0$. For the periods that have parameter estimates whose signs correspond to our predictions, we can reject $\alpha = \beta_2 = 0$ (at the 1% level) for the overall period and the first subperiod.

The results in table 2, in general, support our conjecture. Over the entire 56-year period the regressions in both halves of table 2 have the predicted signs and are statistically significant (at the 1% level). For the eight sets of subperiod regressions, six have the predicted signs (three significant at 1%, one significant at 10%, and two not significant), while two have the opposite signs (neither significant). Also, note that across subperiods there is a strong negative correlation between $\hat{\alpha}$ and $\hat{\beta}_2$. In addition, the specification test of White (1980*b*), discussed in more detail below, rejects linearity of the model in precisely those cases for which we reject $\alpha = \beta_2 = 0$. Thus we see significant artificial timing in those periods in which we can reject the hypothesized model's specification. A comparison of the White specification test and an alternative exclusion test (discussed below) indicates that the White test seems to be more powerful in this case.[11]

An alternative to the model of market-timing behavior in Henriksson and Merton (1981) is given in Pfleiderer and Bhattacharya (1983). In the former model the manager's choice between two discrete levels of systematic risk depends on whether positive or negative excess market returns are considered more likely. In the latter model the manager chooses among a continuum of possible values of risk, depending on

11. Estimates of the model using portfolios of NYSE and AMEX stocks for the 1968–81 subperiod provide similar results. Therefore they are not reported here. A complete series of AMEX data is not available for the other three subperiods.

the point estimate of the market forecast and the precision of the forecast where beta is a linear function of the manager's market forecast. This implies a quadratic relation between R_p and x.[12] Table 3 contains estimates of the Pfleiderer-Bhattacharya market-timing model for the same portfolios used in table 2. The results are essentially identical to those of table 2. The signs of the parameter estimates and the "p-values" of the test for the absence of timing or selectivity performance correspond closely across tables 2 and 3. The main difference between the two models seems to be the power of the specification tests discussed below.

This evidence, using the naive portfolio strategies of holding either an equally weighted or a value-weighted portfolio, indicates the potential for artificially creating measured-timing ability at the expense of measured-security-selection ability. It also shows that specification tests may be useful in detecting spurious timing.

IV. Specification Tests for Market Timing Models

Artificial timing created by investing in options or stocks of highly levered firms exists because the actual return-generating process deviates from the assumed return-generating process (the market model). Option-like securities have returns that are nonlinearly related to the return on the "market portfolio" used here.

We propose two methods of testing the specification of market-timing models that rely on the fact that in most reasonable cases the misspecification causes a nonlinear relation between portfolio returns and the independent variables in the timing models. The first specification test is suggested in White (1980*b*), while the second involves testing restrictions on the coefficients of additional independent variables in the regression.

A. Specification Test Based on OLS and WLS Parameter Estimates

White (1980*b*) suggests a method of testing linearity that involves looking at the difference between OLS and WLS parameter estimates. Assume that the true model is

$$y_t = g(\mathbf{Z}_t) + \epsilon_t, \quad t = 1, \ldots, T, \tag{6}$$

where g is an unknown function and $(\mathbf{Z}_t, \epsilon_t)$ are random $1 \times (p + 1)$ vectors with $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma_\epsilon^2 < \infty$, $E(\mathbf{Z}_t'\epsilon_t) = 0$, and $E[g(\mathbf{Z}_t)^2] < \infty$.

12. The estimation of market-timing ability by a regression of R_p on x and x^2 was first considered in Treynor and Mazuy (1966) and extended by Jensen (1972) and Pfleiderer and Bhattacharya (1983). This model is a special case of the method of performance measurement considered in Admati and Ross (1985). A multifactor application of this quadratic-regression technique is provided in Lehmann and Modest (1985).

TABLE 3 **Pfleiderer-Bhattacharya Market Timing Model Equal- and Value-Weighted NYSE Portfolios ($R_p[t] = \alpha + b_1 x[t] + b_2 x[t]^2 + \delta d[t] + v[t]$)**

Period	$\hat{\alpha}$	$\hat{\delta}$	$\hat{b}_1$	$\hat{b}_2$	p-Value: $\alpha = 0,\ b_2 = 0$	p-Value: White Test	p-Value: Exclusion Test
				Value Weighted vs. Equal Weighted			
1926–81	.003	−.024	.750	−.399	<.001	.129	.619
	(.001)	(.003)	(.014)	(.094)			
1926–39	.006	−.031	.754	−.413	<.001	.241	.455
	(.002)	(.007)	(.021)	(.098)			
1940–53	.004	−.020	.715	−.755	.001	.002	<.001
	(.001)	(.005)	(.023)	(.236)			
1954–67	.002	−.020	.864	−.694	.030	.533	.025
	(.001)	(.004)	(.024)	(.302)			
1968–81	.001	−.026	.727	−.413	.179	.300	.011
	(.001)	(.007)	(.028)	(.225)			
				Equal Weighted vs. Value Weighted			
1926–81	−.003	.035	1.262	.827	.002	.087	.844
	(.001)	(.005)	(.034)	(.217)			
1926–39	−.006	.044	1.311	.876	.006	.508	.247
	(.003)	(.011)	(.052)	(.290)			
1940–53	−.001	.031	1.241	.215	.491	.923	.516
	(.001)	(.008)	(.037)	(.245)			
1954–67	.000	.026	1.031	−.529	.773	.089	.016
	(.001)	(.005)	(.032)	(.785)			
1968–81	.001	.042	1.195	−.617	.842	.241	.374
	(.002)	(.011)	(.059)	(1.127)			

NOTE.—Test for $\alpha = b_2 = 0$ is a Wald test using heteroscedasticity-corrected covariance matrix. Exclusion test is a t-test for $\gamma = 0$ in $R_p = \alpha + b_1 x + b_2 x^2 + \delta d + \gamma x^3 + v$.

White (1980*b*) considers the linear approximation

$$Y_t = \mathbf{Z}_t^*\boldsymbol{\beta} + u_t, \tag{7}$$

where $\mathbf{Z}_t^*$ is a l*xk* vector whose elements are functions of elements of $\mathbf{Z}_t$. White (1980*b*, theorem 2) proves, under mild assumptions, that $\hat{\boldsymbol{\beta}}_{\mathbf{OLS}} \rightarrow \boldsymbol{\beta}^*$, where $\boldsymbol{\beta}^*$ is the parameter vector that solves

$$\min_{\boldsymbol{\beta}} \sigma^2(\boldsymbol{\beta}) = \int [g(\mathbf{Z}) - \mathbf{Z}^*\boldsymbol{\beta}]^2 dF(\mathbf{Z}) + \sigma_\epsilon^2.$$

Note that $\boldsymbol{\beta}^*$ will, in general, depend on the density function $dF(\mathbf{Z})$ when $g(\mathbf{Z}) \neq \mathbf{Z}^*\boldsymbol{\beta}_0$. If $g(\mathbf{Z}) = \mathbf{Z}^*\boldsymbol{\beta}_0$, then $\boldsymbol{\beta}^* = \boldsymbol{\beta}_0$. Let $W(\mathbf{Z}_t^*)$ denote the weights used in the WLS regression. If the model is linear, the parameter estimates from OLS and WLS ($\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}$ and $\hat{\boldsymbol{\beta}}_{\mathbf{WLS}}$) are both consistent estimates of $\boldsymbol{\beta}^*$ for any nonperverse choice of weights, $W(\mathbf{Z}_t^*)$. Thus $\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}$ and $\hat{\boldsymbol{\beta}}_{\mathbf{WLS}}$ should not be "too far" from each other if $g(\mathbf{Z}) = \mathbf{Z}^*\boldsymbol{\beta}_0$. Weighted least squares alters the density function $dF(\mathbf{Z})$ and, hence, will alter $\boldsymbol{\beta}^*$ if $g(\mathbf{Z})$ is not linear. The specification test is based on the difference $\hat{\boldsymbol{\beta}}_{\mathbf{OLS}} - \hat{\boldsymbol{\beta}}_{\mathbf{WLS}}$. Theorem 4 of White (1980*b*) shows that if $g(\mathbf{Z}) = \mathbf{Z}^*\boldsymbol{\beta}_0$ then

$$T(\hat{\boldsymbol{\beta}}_{\mathbf{OLS}} - \hat{\boldsymbol{\beta}}_{\mathbf{WLS}})' \, \hat{\phi}^{-1}(\hat{\boldsymbol{\beta}}_{\mathbf{OLS}} - \hat{\boldsymbol{\beta}}_{\mathbf{WLS}}) \tag{8}$$

has an asymptotic distribution that is chi-square with k degrees of freedom (k is the number of independent variables). The matrix $\hat{\phi}$ in (8) is given by

$$\begin{aligned}\hat{\phi} = {} & \left(\frac{Z^{*\prime}Z^*}{T}\right)^{-1} \hat{V}_{\mathrm{OLS}}\left(\frac{Z^{*\prime}Z^*}{T}\right)^{-1} \\ & + \left(\frac{Z^{*\prime}\Omega^{-1}Z^*}{T}\right)^{-1} \hat{V}_{\mathrm{WLS}}\left(\frac{Z^{*\prime}\Omega^{-1}Z^*}{T}\right)^{-1} \\ & - \left(\frac{Z^{*\prime}Z^*}{T}\right)^{-1} \hat{U}\left(\frac{Z^{*\prime}\Omega^{-1}Z^*}{T}\right)^{-1} \\ & - \left(\frac{Z^{*\prime}\Omega^{-1}Z^*}{T}\right)^{-1} \hat{U}\left(\frac{Z^{*\prime}Z^*}{T}\right)^{-1},\end{aligned} \tag{9}$$

where

$$\hat{V}_{\mathrm{OLS}} = \left(\frac{1}{T}\right) \sum_{t=1}^{T} (Y_t - \mathbf{Z}_t^*\hat{\boldsymbol{\beta}}_{\mathbf{OLS}})^2 \mathbf{Z}_t^{*\prime}\mathbf{Z}_t^*,$$

$$\hat{V}_{\mathrm{WLS}} = \left(\frac{1}{T}\right) \sum_{t=1}^{T} (Y_t - \mathbf{Z}_t^*\hat{\boldsymbol{\beta}}_{\mathbf{WLS}})^2 \, W(\mathbf{Z}_t^*)^2 \mathbf{Z}_t^{*\prime}\mathbf{Z}_t^*,$$

$$\hat{U} = \left(\frac{1}{T}\right) \sum_{t=1}^{T} (Y_t - \mathbf{Z}_t^*\hat{\boldsymbol{\beta}}_{\mathbf{OLS}})(Y_t - \mathbf{Z}_t^*\hat{\boldsymbol{\beta}}_{\mathbf{WLS}}) W(\mathbf{Z}_t^*)\mathbf{Z}^{*\prime}_t\mathbf{Z}_t^*,$$

$$\Omega^{-1} = \mathrm{diag}[W(\mathbf{Z}_t^*)^2].$$

Values of (8) larger than the critical value of a χ^2_k statistic (at a specified level of significance) will lead one to reject the null hypothesis that the model is well specified (H_0: $g[\mathbf{Z}] = \mathbf{Z}^*\boldsymbol{\beta}_0$).

Column 6 of table 1 indicates the frequency with which the White test rejects linearity (at the 5% level) in the call-option simulations. To determine Ω we use the WLS procedure suggested in Henriksson (1984, eq. [13]). The absolute value of the OLS residual is regressed on $x_1 = \min(0, x)$ and $x_2 = \max(0, x)$:

$$|\xi(t)| = \phi_0 + \phi_1 x_1(t) + \phi_2 x_2(t) + \xi(t), \tag{10}$$

and the weighting function is defined by

$$W(t) = [\hat{\phi}_0 + \hat{\phi}_1 x_1(t) + \hat{\phi}_2 x_2(t)]^{-1}. \tag{11}$$

For low exercise prices (in the money options) the test does not seem to have very much power. However, there is a dramatic increase in power as the exercise price/value ratio increases from 0.8 to 0.9. For at-the-money and out-of-the-money options the frequency of rejection ranges from 82% to 98%.

In tables 2 and 3 the specification test is applied to the Henriksson-Merton and the Pfleiderer-Bhattacharya models, respectively. The results in table 2 lend support to our conjecture that spurious timing is related to misspecification of the model. In precisely those cases in which we find significant (say, at the 5% level) spurious selection and timing ability we also reject the hypothesis that the model is well specified, using White's test. In table 3, however, there is not such a close correspondence between significant performance and rejection of the model's specification.

B. *Specification Test Based on Exclusion Restrictions*

A second method of testing misspecification resulting in nonlinearity of the market-timing model is to include additional variables (specifically, nonlinear transformations of some of the independent variables) in the regression function. If the model is linear, then the regression coefficients on the additional variables should be close to zero. Thus to test H_0 ($g[\mathbf{Z}] = \mathbf{Z}^*\boldsymbol{\beta}_0$) we might estimate

$$Y_t = \mathbf{Z}_t^*\boldsymbol{\beta} + \mathbf{h}(\mathbf{Z}_t^*)\boldsymbol{\gamma} + u_t, \tag{12}$$

where $\mathbf{h}(\mathbf{Z}_t^*)$ is some ($1xq$) vector transformed values of $\mathbf{Z}_t^*$. Testing $g(\mathbf{Z}) = \mathbf{Z}^*\boldsymbol{\beta}_0$ is equivalent to testing $\boldsymbol{\gamma} = \mathbf{0}$ in (12). Standard "t-tests" and "F-tests" can be used to test the restrictions. However, because of potential heteroscedasticity, it is advisable to use the heteroscedasticity-corrected covariance matrix of the parameter estimates (see n. 10 above) when constructing test statistics.

For example, a possible test of the specification of (2) would be to test $\gamma = 0$ in the regression[13]

$$R_p = \alpha + \beta_1 x + \beta_2 y + \gamma x^2 + \eta. \tag{13}$$

A t-test (corrected for heteroscedasticity) can be used here. One advantage of this approach is that it is easier to implement than the White specification test. There is no need to calculate both OLS and WLS estimates; OLS is sufficient.

Ideally one would like to compare the two specification tests on the basis of their respective power functions. However, the power depends on the form chosen for $\mathbf{h}(\mathbf{Z}_t^*)$. Column 8 of table 1 indicates the frequency of rejection of $\gamma = 0$ in (13) for the call-option simulations. For this particular case the second technique uniformly dominates the first in terms of power (assuming the sizes of the tests are equivalent in small samples). However, in the tests reported in tables 2 and 3 neither technique exhibits uniformly higher power.

C. *Choosing between Alternative Models of Manager's Reaction Functions*

Nonlinearity in market-timing models, such as (2), need not be due solely to violation of the assumed return-generating process. Nonlinearity may also result when the manager reacts to market forecasts in a manner that is different from that assumed in the model. As an example, consider the models of Henriksson and Merton (1981) and Pfleiderer and Bhattacharya (1983). If the manager's "true" reaction function is such that the relation between R_p and x is quadratic as in Pfleiderer and Bhattacharya (1983), then one should be able to reject $\gamma = 0$ in (13), given a sufficient number of observations. The specification tests can be used to discriminate, in some sense, between potential models of the manager's reaction function. For example, if one were to apply the tests to the Henriksson-Merton and Pfleiderer-Bhattacharya models and reject the specification of one while not rejecting the specification of the other, then one could argue that the model whose specification is not rejected is more plausible. However, testing the linearity of each model is not the same as testing one model against the other. In the hypothetical tests above it is possible to accept both models or reject both models. Alternatively, one may wish to use the tests proposed in MacKinnon, White, and Davidson (1983) or a Bayesian approach (see Zellner 1971, chap. 10).

There is no reason to expect that all portfolio managers have the same type of reaction function for market forecasts. Therefore a reasonable approach when analyzing many funds may be to choose the

13. The appropriate-choice form of $\mathbf{h}(\mathbf{Z}_t^*)$ is not, a priori, obvious. We choose to include x^2 in (13) because (a) it plays an important role in the model of Pfleiderer and Bhattacharya (1983) and (b) one might view higher orders of the independent variables as polynomial approximations to an unknown function.

form of the timing model for each fund using the specification tests and then estimate fund-specific timing models.

V. Conclusion

In this paper we demonstrate that it is possible to create artificial market timing, as measured by commonly used parametric models of timing, by investing in option-like securities. This artificial timing ability is obtained at the cost of poorer measured security selectivity. When the proxy for the market portfolio contains option-like securities, portfolios with greater (lower) concentration in option-like securities will show positive (negative) timing performance and negative (positive) selectivity.

This provides a possible explanation of previous empirical findings that indicate that mutual funds have negative timing ability, on average, and that selectivity and timing performance are negatively correlated (cross-sectionally). If mutual funds tend to invest in "higher quality" (i.e., less option-like) securities, then one would expect to find average measured timing performance to be negative. Also, one would expect to see negative correlation between measured selectivity and timing performance.

We suggest two types of specification tests to check for artificial timing. The tests perform adequately when applied to the Henrikkson-Merton parametric timing test. That is, the tests reject linearity when spurious timing is statistically significant. However, the tests are less able to detect misspecification due to spurious timing when the Pfleiderer-Bhattacharya timing model is used.[14] However, we refrain from using the specification tests to choose between the models for the portfolios used here since both models exhibit significant measured-timing ability when no true timing ability exists.

A useful extension of this work may be to measure performance of mutual funds grouped by type of fund (balanced, growth, dual purpose, etc.), which may show important differences among groups (some of which may be due to artificial timing). Determining the robustness of nonparametric tests (such as those suggested in Henriksson and Merton 1981) to artificial timing relative to their loss of power when true timing ability exists will help one choose the appropriate testing procedure.

Appendix

Let $x_1 = \min(0, x)$ and $x_2 = \max(0, x)$. Note that $x_1 = -y$ and $x = x_1 + x_2$. For notational simplicity the time index t will be eliminated when there is no risk of

14. This may be due to the specific portfolios that we use. Different portfolio formation rules may lead to changes in the relative powers of the tests.

confusion. The probability limits of the OLS estimates of the parameters in (2) are given by

$$\text{plim } \hat{\beta}_1 = \frac{\sigma_{px}\sigma_y^2 - \sigma_{py}\sigma_{xy}}{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2}, \tag{A1}$$

$$\text{plim } \hat{\beta}_2 = \frac{\sigma_{py}\sigma_x^2 - \sigma_{px}\sigma_{xy}}{\sigma_x^2\sigma_y^2 - \sigma_{xy}^2}, \tag{A2}$$

$$\text{plim } \hat{\alpha} = E(R_p) - \text{plim } \hat{\beta}_1 E(x) - \text{plim } \hat{\beta}_2 E(y), \tag{A3}$$

where σ_{px} is the covariance of R_p and x, σ_{py} is the covariance of R_p and y etcetera. Since $R_p = \lambda x_2 - (1 + R_F)$ and $x_1 x_2 = 0$, the relevant variances and covariances are given by

$$\begin{aligned}
\sigma_{px} &= \lambda[\sigma_2^2 - E(x_1)E(x_2)] \\
\sigma_{py} &= \lambda E(x_1)E(x_2) \\
\sigma_{xy} &= E(x_1)E(x_2) - \sigma_1^2 \\
\sigma_p^2 &= \lambda^2\sigma_2^2 \\
\sigma_y^2 &= \sigma_1^2 \\
\sigma_x^2 &= \sigma_1^2 + \sigma_2^2 - 2E(x_1)E(x_2).
\end{aligned}$$

By using these values in (A1)–(A3), the probability limits of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\alpha}$ are given by

$$\begin{aligned}
\text{plim } \hat{\beta}_1 &= \lambda, \\
\text{plim } \hat{\beta}_2 &= \lambda, \\
\text{plim } \hat{\alpha} &= -(1 + R_F).
\end{aligned}$$

Using these parameter values to calculate residuals yields

$$\begin{aligned}
\epsilon &= R_p + (1 + R_F) - \lambda(x_1 + x_2) + \lambda x_1 \\
&= R_p - [\lambda x_2 - (1 + R_F)] = 0.
\end{aligned}$$

Thus the regression errors are identically zero as was claimed in Section II*B*.

References

Admati, Anat R., and Ross, Stephen A. 1985. Measuring investment performance in a rational expectations equilibrium model. *Journal of Business* 58 (January): 1–26.

Black, Fischer, and Scholes, Myron. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (May/June): 637–54.

Breen, William; Jagannathan, Ravi; and Ofer, Aharon R. 1985. Market timing and investment performance—a comment. Working Paper no. 111. Evanston, Ill.: Northwestern University, Banking Research Center.

Chang, Eric C., and Lewellen, Wilbur G. 1984. Market timing and mutual fund investment performance. *Journal of Business* 57 (January): 57–72.

Dybvig, Philip H., and Ross, Stephen A. 1985. Differential information and performance measurement using a security market line. *Journal of Finance* 40 (June): 383–99.

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

Grinblatt, Mark, and Titman, Sheridan. 1985. Portfolio performance evaluation: Old issues and new insights. Working Paper no. 16–84. Los Angeles: University of California, Los Angeles, Graduate School of Management.

Hansen, Lars Peter. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50 (July): 1029–54.

Henriksson, Roy D. 1984. Market timing and mutual fund performance: An empirical investigation. *Journal of Business* 57 (January): 73–96.

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.

Hsieh, David A. 1983. A heteroscedasticity-consistent covariance matrix estimator for time series regressions. *Journal of Econometrics* 22 (August): 281–90.

Ibbotson, Roger G., and Sinquefield, Rex A. 1982. *Stocks, Bonds, Bills and Inflation: The Past and the Future*. Charlottesville, Va.: Financial Analysts Research Foundation.

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.

Keim, Donald B. 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 12 (June): 13–32.

Kon, Stanley J. 1983. The market-timing performance of mutual fund managers. *Journal of Business* 56 (July): 323–47.

Kon, Stanley J., 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.

Korschot, Benjamin C. 1978. Quantitative evaluation of investment research analysts. *Financial Analysts Journal* (July/August): 41–46.

Lehmann, Bruce N., and Modest, David M. 1985. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. Typescript. New York: Columbia University, Graduate School of Business.

Lockwood, Larry J., and Kadiyala, K. Rao. 1985. Utilization of market forecasts in portfolio management: An empirical analysis. Typescript. Arlington: University of Texas at Arlington, College of Business Administration.

MacKinnon, James G.; White, Halbert; and Davidson, Russell. 1983. Tests for model specification in the presence of alternative hypotheses. *Journal of Econometrics* 21 (January): 53–70.

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.

Modest, David M. 1983. Technical indicators, macroeconomic variables and market timing: A test of the ability to predict stock returns. Typescript. New York: Columbia University, Graduate School of Business.

Pfleiderer, Paul, and Bhattacharya, Sudipto. 1983. A note on performance evaluation. Technical Report no. 714. Stanford, Calif.: Stanford University, Graduate School of Business.

Smith, Keith V. 1978. Is fund growth related to fund performance? *Journal of Portfolio Management* (Spring): 49–54.

Treynor, Jack L., and Mazuy, Kay K. 1966. Can mutual funds outguess the market? *Harvard Business Review* 44 (July/August): 131–36.

White, Halbert. 1980*a*. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

White, Halbert. 1980*b*. Using least squares to approximate unknown regression functions. *International Economic Review* (February): 149–70.

Zellner, Arnold. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.