TESTS OF MARKET TIMING

AND MUTUAL FUND PERFORMANCE

Roy Henriksson .

WP 1136-80   May 1980

TESTS OF MARKET TIMING

AND MUTUAL FUND PERFORMANCE

Roy Henriksson   .

WP 1136-80   May 1980

INTRODUCTION

Can investment managers beat the market? Or more specifically, can they determine when a portfolio replicating the stock market as a whole will provide a greater return than riskless government securities? Using a model developed by Merton (1979), we will attempt to answer this important question.

There have been numerous studies, both theoretical and empirical, dealing with the evaluation of the performance of investment managers. This is justifiable considering the magnitudes of the invested assets and management fees involved. Proper evaluation would allow investors to analyze the trade-off between expected fund performance and the size of the management fees, while allowing the best managers to charge the highest fees.

A second important reason for study of the subject is the inconsistency of superior forecasting ability and stronger forms of the Efficient Markets Hypothesis.[1] If managers actually do possess superior forecasting ability, it can not be true that market prices reflect all available information, assuming investors prefer more to less.

Forecasting skills are usually considered to be of two forms: "micro" forecasting and "macro" forecasting. Micro forecasting focuses on individual securities, seeking to determine securities that are under- or overvalued in light of the forecaster's expectation of future returns. In terms of the Capital Asset Pricing Model (CAPM), such securities would lie above or below the Security Market Line.[2]

Macro forecasting or market-timing strategies attempt to determine when the stock market as a whole will outperform riskless securities [i.e., $Z_m(t) > R(t)$] and when riskless securities will outperform the stock market [i.e., $Z_m(t) < R(t)$] where $Z_m(t)$ is the one-period per-dollar return on the stock market and $R(t)$ is the one-period per-dollar return on riskless securities.

Treynor and Black (1973) show that through the use of an "active" portfolio, investment managers can effectively separate actions related to security analysis from those related to market timing. The investment manager's optimal portfolio can be thought of as a combination of three portfolios: the active portfolio, a portfolio replicating the market and a portfolio comprised of riskless securities. The active portfolio[3] reflects the manager's micro forecasts and will be comprised of long positions in securities that are thought to be undervalued and short positions in securities that are thought to be overvalued. Knowing the composition of the active portfolio, the manager will achieve the desired level of market-related risk, which will take into account any market timing forecasts, by adjusting their positions in the market portfolio and riskless securities.

In this paper, we will focus on macro forecasting or market timing. Fama (1972) and Jenson (1972a) develop a theoretical structure for evaluation of the performance of investment managers using both micro and macro forecasting, based on the ex post performance of the manager's fund and the market. Jenson also shows that there are difficulties

in empirically distinguishing among the individual contributions of
the two types of forecasts. Without knowing the actual forecasts or
the beta (i.e., level of market risk) of the fund at each point in
time, regression analysis of portfolio returns cannot be used to
distinguish excess returns that are the result of micro forecasting
from those that are the result of market timing.

Jenson assumes that the market timer forecasts what the return
on the market portfolio will be and adjusts his portfolio accordingly.
The magnitude of the adjustment will be a function of the level of
his forecast and the risk preferences of the fund's investors. Their
risk preferences will in part reflect the volatility of market returns.
The higher the estimate of the market return, the greater the amount
of market-related risk they will be willing to bear. Jensen shows
that a market timer's forecast can be measured by ρ, the correlation
between the market timer's forecast and the actual performance of the
market.[4] The expected return of the fund will increase as ρ
increases. This is because one would expect a greater proportion than
average of the fund's assets to be invested in the market portfolio
when the realized return on the market was greater than the average
and a greater proportion than average of the fund's assets to be
invested in riskless securities when the realized return on the market
was less than the average.

Grant (1977) explains how market timing actions will effect
the results of empirical tests that focus only on micro forecasting
skills, when the true forecasts are not known. He shows that market

timing ability will cause a downward bias in estimates of the excess returns resulting from micro forecasting ability. All of these studies use a parametric structure assuming a CAPM framework.

Treynor and Mazuy (1966) use a CAPM framework with a quadratic term to test for market timing ability. In the standard regression equation, a portfolio's return is a linear function of the return on the market portfolio. However, if the portfolio manager can forecast market returns, he will hold a greater proportion of the market portfolio when the return on the market is high and a smaller proportion when the market return was small. Thus, the portfolio return would be a non-linear function of the market return. Using annual returns for 57 open-end mutual funds, they find that only one of the funds can reject the hypothesis of no market-timing ability at a 95% confidence level.

THE MODEL

Merton develops a model that does not depend on CAPM. It takes the simple form that the investment manager either forecasts that stocks will outperform riskless securities or that riskless securities will outperform stocks. Based on his forecast, the investment manager will adjust the relative proportions of the market portfolio and riskless securities that he will hold in his fund. A perfect forecaster (who cannot borrow or sell short) will always invest all of his assets in the market portfolio if $Z_M(t) > R(t)$ and all of his assets in riskless securities if $Z_M(t) < R(t)$.

Merton (1979) shows that the return from successful market timing will be virtually indistinguishable from certain option investment strategies (either a strategy of holding call options and treasury bills or a strategy of holding protective put options and the stock). The value of market timing ability can be determined using methods for valuing options.

The key factors in determining the value are the sum of the conditional probabilities $[p_1(t) + p_2(t)]$ and the magnitude of the portfolio change in response to the market-timing forecast. The probabilities are conditional on what the market does. Given that $Z_M(t) \leq R(t)$, $p_1(t)$ is the conditional probability of a correct forecast. Given that $Z_M(t) > R(t)$, $p_2(t)$ is the conditional probability of a correct forecast.

The conditional probabilities are segmented around $R(t)$ as this is the optimal switching point between stocks and bonds. Since superior market forecasting ability is inconsistent with an efficient

markets equilibrium for security prices, for such an equilibrium to exist, it is necessary that an investment manager's forecast have no value.

Merton (1979) proves that a necessary and sufficient condition for market-timing forecasts to have value is $p_1(t) + p_2(t) > 1$. With rational investors, it should never be true that $p_1 + p_2 < 1$ as the opposite of the forecasts would result in $p_1 + p_2 > 1$ and would thus have value. Thus, a simple test of a forecaster's market timing ability is to estimate $p_1 + p_2$ and determine if one can reject the null hypothesis that $p_1 + p_2 = 1$. This is a test of independence between the market timer's forecast and whether or not the return on the market portfolio is greater than the return on riskless securities.[5]

An unconditional probability of a successful forecast $p > 0.5$ is not a necessary or sufficient condition for a forecaster's market timing ability to have value, as is shown in Merton (1979). There is one exception to this. One can use the unconditional probability of a successful forecast if there is good reason to believe that $p_1 = p_2$, i.e., the forecaster has the same ability at predicting up-markets as down-markets.

The use of this nonparametric test eliminates the empirical problem that extreme outcomes can swamp the results in estimating a regression equation, without eliminatimg them from the sample. Differences between our non-parametric test and estimations of Jenson's ρ are caused by differences in the weightings of actual market

outcomes. Jenson's test gives a larger weighting to extreme outcomes, while our test weighs all outcomes equally.

Since the volatility of stock market returns is large relative to the expected return over short time periods, it would be quite difficult to attach much accuracy to point estimate forecasts. For quarterly data from 1926:1 to 1978:4, the average three-month rate of return was 2.7% with a standard deviation of 12.6%.

Our nonparametric formulation characterizes a more subjective approach to market timing; one where the forecaster predicts only if stocks will provide a greater return than bonds, but not by how much. It is this forecasting ability that we will examine.

ESTIMATION

To test for market timing, we will examine the null hypothesis that the timer has no forecasting ability, i.e., H_0: $p_1 + p_2 = 1$, where p_1 and p_2 are not known. We want to determine the probability, P, that a given outcome from our sample came from a population that satisfies H_0. Define the following variables:

$N_1 \equiv$ Number of observations where $Z_M < R$.

$N_2 \equiv$ Number of observations where $Z_M > R$.

$N \equiv N_1 + N_2$.

$n_1 \equiv$ Number of successful predictions, given $Z_M < R$.

$n_2 \equiv$ Number of unsuccessful predictions, given $Z_M > R$.

$n \equiv n_1 + n_2 =$ number of times forecasted $Z_M < R$.

From our null hypothesis, we have that $E(n_1/N_1) = p_1$ and $E(n_2/N_2) = 1 - p_2 \underset{H_0}{=} p_1$ where E is the expected value operator. It then follows that $E(n/N) = p_1 \equiv p$. Both n_1/N_1 and n_2/N_2 have the same expected value yet they are drawn from independent subsamples.

Both n_1 and n_2 are random variables from binomial distributions. Because of this, the probability that $n_i = x$ from a subsample of N_i drawings is:

$$P(x; N_i, p) = \binom{N_i}{x} p^x (1 - p)^{N_i - x} \quad .$$

Given H_0, we can use Bayes' Theorem to determine the probability that $n_1 = x$ given N_1, N_2, and n, i.e., $P(n_1 = x | N_1, N_2, n)$. Assume that B: $n_1 = x$ and $n_2 = m - x$; and A: $n = m$. By Bayes' Theorem:

$$P(B|A) = \frac{P(B + A)}{P(A)} = \frac{P(B)}{P(A)}$$

$$= \frac{\binom{N_1}{x}\binom{N_2}{m-x} p^x (1 - p)^{N_1-x} p^{m-x} (1 - p)^{N_2-m+x}}{\binom{N}{m} p^m (1 - p)^{N-m}}$$

$$= \frac{\binom{N_1}{x}\binom{N_2}{m-x}}{\binom{N}{m}}$$

which follows a hypergeometric distribution and is independent of p.

Given N_1, N_2, and n, one can determine the probability distribution of n_1, realizing that the feasible range for n_1 is:

$$Max[0, n - N_2] \le n_1 \le Min[N_1, n] \quad .$$

Thus we can use the distribution of $P(n_1 = x | N_1, N_2, n)$ over the feasible range to establish confidence intervals for the null hypothesis. This is straightforward for small samples as the calculations can be done using factorials (i.e., gamma functions).

However, for large samples, the calculations may become quite cumbersome. Fortunately, for large samples, the hypergeometric distribution can be approximated by the normal distribution.[6] From the

hypergeometric distribution, $E(n_1) = n \cdot \dfrac{N_1}{N}$ and the variance of n_1, $\sigma^2(n_1) = n \cdot \dfrac{N_1}{N} \cdot \dfrac{N - N_1}{N} \cdot \dfrac{N - n}{N - 1}$.

Our test of the null hypothesis is a one-tailed test. This is because of our rationality assumption that $p_1 + p_2 \geq 1$. Thus, our null hypothesis can only be rejected in the right-hand tail of the hypergeometric distribution. By considering only one tail, we are making it easier to reject the null hypothesis, either at the 95% or 99% confidence level.

If we wanted to test for market timing using the unconditional probability of a correct forecast,[7] then our null hypothesis would be H_0: $p = 0.5$. The distribution of outcomes drawn from a population that satisfied H_0 would be the binomial distribution:

$$P(k,N,p) = \binom{N}{k} p^k \cdot (1 - p)^{N-k}$$

$$= \binom{N}{k} (0.5)^N$$

where k the number of successful predictions, N is the total number of observations, and p is the unconditional probability of a successful forecast.

Tables 1-3 give values of n_1 that reject the null hypothesis at the 99% confidence level for different values of N_1, N_2, and m. As would be expected, the required estimated value of $p_1(t) + p_2(t)$ decreases as the size of the total sample increases. Tables 1-3 also show that the normal distribution is an excellent approximation for

determining the confidence intervals for the hypergeometric distribution, even for observation samples as small as 50.

Figures 1 and 2 show the relationship between the required number of correct forecasts using conditional probabilities and unconditional probabilities for sample sizes of 30 and 200 for different values of N_1 and n.

-13-

Table 1: REQUIRED OUTCOMES TO REJECT THE NULL HYPOTHESIS
WITH 99 PERCENT CONFIDENCE

N	N_1	n	True Distribution			Normal Approximation		
			Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + 1 - \frac{n_2}{N_2}$	Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + 1 - \frac{n_2}{N_2}$
50	25	25	18	36	1.44	18	36	1.44
50	35	25	22	34	1.43	22	34	1.43
50	15	25	12	34	1.43	12	34	1.43
50	25	15	12	34	1.36	12	34	1.36
50	25	35	22	34	1.36	22	34	1.36
50	15	15	9	38	1.43	9	38	1.43
50	15	35	15	30	1.43	15	30	1.43
50	35	15	15	30	1.43	15	30	1.43
50	35	35	29	38	1.43	29	38	1.43

Table 2: REQUIRED OUTCOMES TO REJECT THE NULL HYPOTHESIS WITH 99 PERCENT CONFIDENCE

N	N_1	n		True Distribution			Normal Approximation		
				Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + \frac{1-n_2}{N_2}$	Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + \frac{1-n_2}{N_2}$
100	50	50		32	64	1.28	32	64	1.28
100	75	50		43	61	1.29	44	63	1.35
100	25	50		18	61	1.29	19	63	1.35
100	50	75		43	61	1.22	44	63	1.26
100	50	25		18	61	1.22	19	63	1.26
100	25	25		12	74	1.31	12	74	1.31
100	25	75		24	48	1.28	24	48	1.28
100	75	25		24	48	1.28	24	48	1.28
100	75	75		62	74	1.31	62	74	1.31

Table 3: REQUIRED OUTCOMES TO REJECT THE NULL HYPOTHESIS WITH 99 PERCENT CONFIDENCE

N	N_1	n	True Distribution			Normal Approximation		
			Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + 1 - \frac{n_2}{N_2}$	Required Value of n_1	Total Correct Forecasts	$\frac{n_1}{N_1} + 1 - \frac{n_2}{N_2}$
200	100	100	59	118	1.18	60	120	1.20
200	50	100	33	116	1.21	34	118	1.24
200	150	100	83	116	1.21	84	118	1.24
200	100	150	83	116	1.16	84	118	1.18
200	100	50	33	116	1.16	34	118	1.18
200	50	50	20	140	1.20	20	140	1.20
200	50	150	44	88	1.17	45	90	1.20
200	150	50	44	88	1.17	45	90	1.20
200	150	150	120	140	1.20	120	140	1.20

Figure 1 : Number of Correct Forecasts Required To Reject H₀ with 99% Probability Conditional vs. Unconditional Tests Sample Size = 30
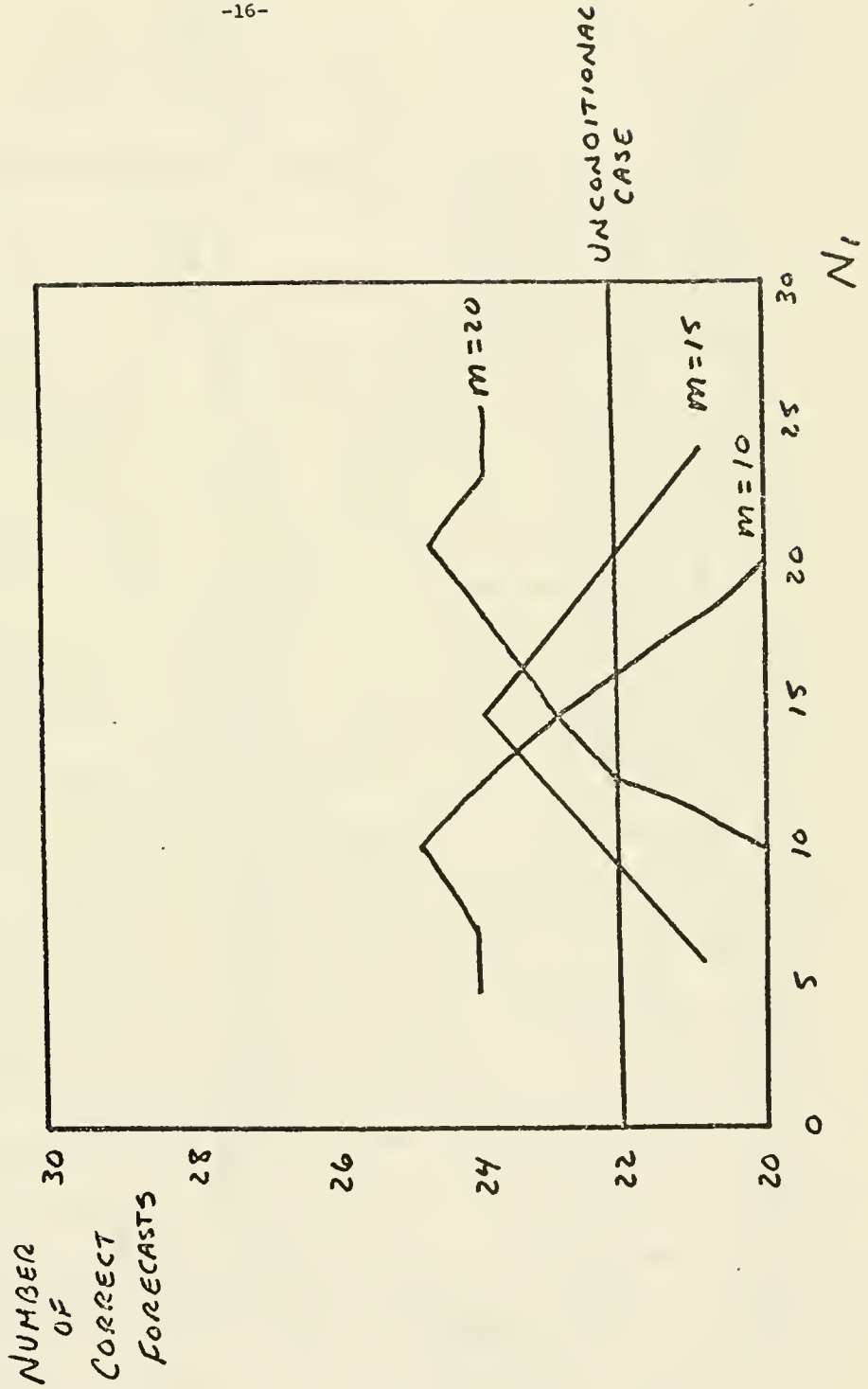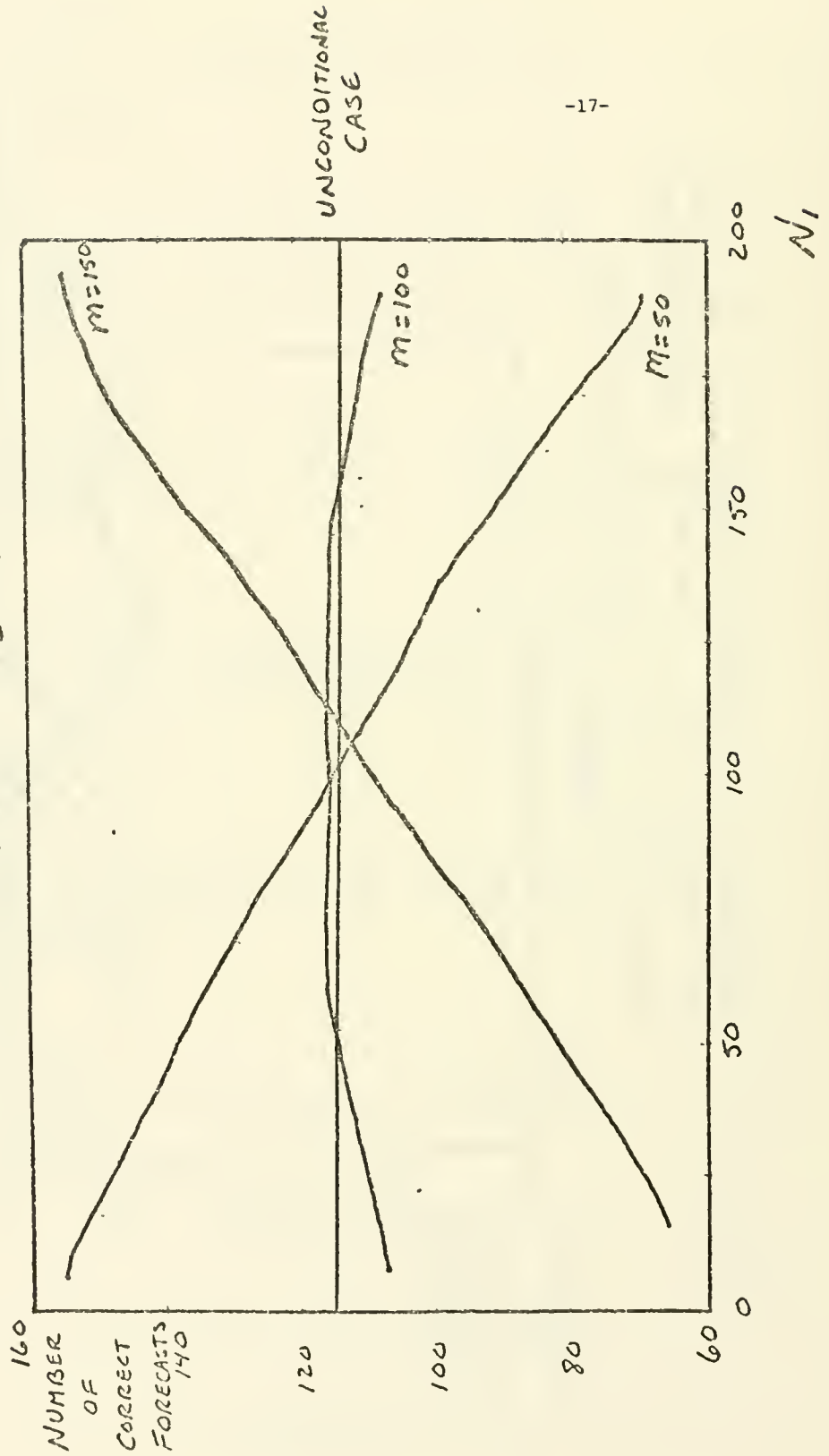
FIGURE 2: NUMBER OF CORRECT FORECASTS REQUIRED TO REJECT H_0 WITH 99% PROBABILITY CONDITIONAL VS. UNCONDITIONAL TESTS SAMPLE SIZE = 200

THE TESTS

As most funds do not explicitly claim to undertake market timing,[8] it is necessary to have the data itself signal the investment manager's forecast of how the stock market will perform. In a Capital Asset Pricing Model type world, a measure of this would be changes in the β of the fund. However, as many of the funds hold more than one hundred different securities, this would be difficult to determine from the raw data, especially when one considers that the β of the individual securities may be changing over time.

Instead, changes in the relative proportions of the funds in stocks and bonds[9] are used as the proxy for the market timing fore-cast. When one considers the frequency with which investors move in and out of mutual funds, it is likely that open-ended mutual funds are constantly receiving new cash or selling off securities in order to redeem shares in the mutual fund being cashed in. This gives the managers of mutual funds the opportunity to change the level of market-related risk in their funds, at least to a limited amount, at a very low marginal cost as the transactions costs involved would have been incurred regardless of the market timing action.

Any time a manager is changing the level of market-related risk in his fund, he is implicitly undertaking a market timing strategy as he is changing the risk characteristics of the fund. Treynor and Black (1972) have shown that managers can incorporate micro forecasts into their portfolios in a manner which does not effect the level of market-related risk in the portfolio. Black (1976) has also shown

that it is inefficient to undertake a market timing strategy if one has no ability to forecast the relative performance of the stock market. Such a strategy would cause the forecaster's portfolio to have greater variance for a given expected return than a passive (with respect to market timing) strategy. Black refers to this as the loss of time diversification.

Thus, changes in the level of market-related risk in a mutual fund should reflect some expectation of how the market will perform. Merton (1979) has shown that for an individual with no market timing ability, $p_1(t) + p_2(t) = 1$. From this, even if we include in our sample funds which are not attempting to forecast market performance, for whom changes in the proportions of their fund held in stocks and bonds are independent of market expectations, this will not bias our results. And as almost all the funds in the sample have a number of quarterly changes of greater than two percent, it can be argued that most of the funds are undertaking some implicit strategy of market timing.

Five different tests of market timing were run using quarterly data spanning the period from March 31, 1973-March 31, 1980. The tests were run on each of 186 mutual funds.[10] Over this period, the return on the market exceeded the riskless rate in 15 of the 28 quarters. The return from investing $1 in the market portfolio[11] for the 28 quarters was $1.23. For U.S. Treasury bills[12] the return was $1.62. A perfect market timing strategy on a quarterly basis would have returned $4.31.

The tests were also run using six-month time intervals, running from September 30, 1973-March 31, 1980. Over this interval the return on the market exceeded the return from the riskless asset during 5 of the 13 six-month periods. The return from investing $1 in the market portfolio for the 13 periods was $1.23. The return on U.S. Treasury bills was $1.56. A perfect market timing strategy on a semi-annual basis would have returned $3.30.

To define the tests that were run, let $\gamma(t) \equiv$ the market timer's forecast where $\gamma(t) = 1$ if the forecast is $Z_M(t) > R(t)$ and $\gamma(t) = 0$ if the forecast is that $Z_M(t) < R(t)$. The tests run were as follows:

Test #1: $\gamma(t) = 1$ if the proportion of the fund in stocks increases;

$\gamma(t) = 0$ if the proportion of the fund in bonds increases.

Test #2: $\gamma(t) = 1$ if the proportion of the fund in stocks increases by one percent or more of the total value of the fund;

$\gamma(t) = 0$ if the proportion of the fund in bonds increases by one percent or more of the total value of the fund.

There is no forecast if neither porportion increases by one percent or more.

Test #3: $\gamma(t) = 1$ if the proportion of the fund in stocks increases by two percent or more of the total value of the fund;

$\gamma(t) = 0$ if the proportion of the fund in bonds increases by two percent or more of the total value of the fund.

There is no forecast if neither proportion increases by two percent or more.

Test #4: $\gamma(t) = 1$ if the proportion of the fund in stocks increases by one percent or more of the total value of the fund;

$\gamma(t) = 0$ if the proportion of the fund in bonds increases by one percent or more of the total value of the fund;

$\gamma(t) = (t - 1)$ if neither proportion increases by one percent or more.

Test #5: $\gamma(t) = 1$ if the proportion of the fund in stocks increases by two percent or more of the total value of the fund;

$\gamma(t) = 0$ if the proportion of the fund in bonds increases by two percent or more of the total value of the fund;

$$\gamma(t) = \gamma(t - 1) \qquad \text{if neither proportion increases by two}$$

percent or more.

Test #6 is the same as Test #1 for six-month forecast intervals.

Test #7 is the same as Test #2 for six-month forecast intervals.

Test #8 is the same as Test #3 for six-month forecast intervals.

There are two major reasons for using quarterly data. First, it is the shortest time interval for which data could be obtained. Second, it is sufficiently long enough to allow managers of the funds to react to fluctuations in market values. Changes in the proportions of stocks and bonds in the funds caused by changes in the market values of the individual securities, and the resulting change in the funds' market-related risk, could be offset by the managers of the funds if they desired.

The use of a filter band in some of the tests allows for the possibility that changes in market values occurred just previous to the observation dates, or were small enough that it would not be worth incurring the transactions costs to make the necessary adjustments. Since the size of the bands is small, a manager who thought he could predict the performance of the market would surely adjust his portfolio weights by more than one or two percent.

Tests #4 and #5 allow for the possibility that the range over which a manager can vary the proportions of stocks and bonds in a fund are limited. This could be the result of restrictions imposed as a matter of fundamental policy or as a requirement for maintaining tax-

exemption for distributed income. An example of a fund where restrictions would be relevant would be Merton's hypothetical fund which was either 100 percent in stocks or 100 percent in bonds. If the forecast stayed the same as in the previous period, the position of the fund would not change.

Three of the tests were run using six-month forecast intervals as well as three-month forecast intervals to consider the possibility that fund managers do have market timing ability, but they either have a longer forecast horizon than three months or quarterly intervals do not give them sufficient time to adjust their portfolios.

TEST RESULTS

The results of the tests are shown in Table 4. It is clear that we cannot reject the null hypothesis, H_0: $p_1(t) + p_2(t) = 1$, for any of the eight tests. Only one fund in only one of the tests had an estimated $p_1(t) + p_2(t)$ that differed from H_0 with 99 percent probability. We would expect to observe more outliers than this from random sampling from a population that satisfies H_0. The same is true for the 95 percent confidence level. At the most, five of the more than 180 funds could reject the null hypothesis with 95 percent probability in any of the tests.

Our tests also show no evidence that investment managers have a longer forecast horizon than three months. When we lengthen the forecast period to six months, in all three cases our estimation mean falls below one, none of the funds can reject the null hypothesis with 99 percent probability, and at most four of the funds can reject the null hypothesis with 95 percent probability.

One problem, which occurs especially in Tests #3 and #8, is that the use of a filter makes the number of usable observations become quite small for some of the funds in the sample. For a few of the funds it actually becomes impossible to reject H_0, even with 95 percent confidence. However, the estimation mean and distribution of outcomes for all eight tests seem to be consistent with H_0. The number of outcomes in the two tails of the distributions are approximately the same, as is shown in Figures 3-10.

Table 4: TEST RESULTS

	Sample Size	Estimation Mean	Estimation Std. Dev.	Estimation Median	Number of Funds That Reject H_0	
					With 95% Probability	With 99% Probability
Test #1	186	0.998	.191	1.00	2	0
#2	186	1.001	.227	1.00	4	0
#3	183	1.008	.251	1.00	0	0
#4	186	1.001	.201	1.01	4	1
#5	185	1.010	.197	1.00	5	0
#6	186	0.956	.301	0.96	4	0
#7	186	0.944	.320	0.93	2	0
#8	185	0.927	.362	0.93	0	0

FIGURE 3: TEST # 1 DISTRIBUTION
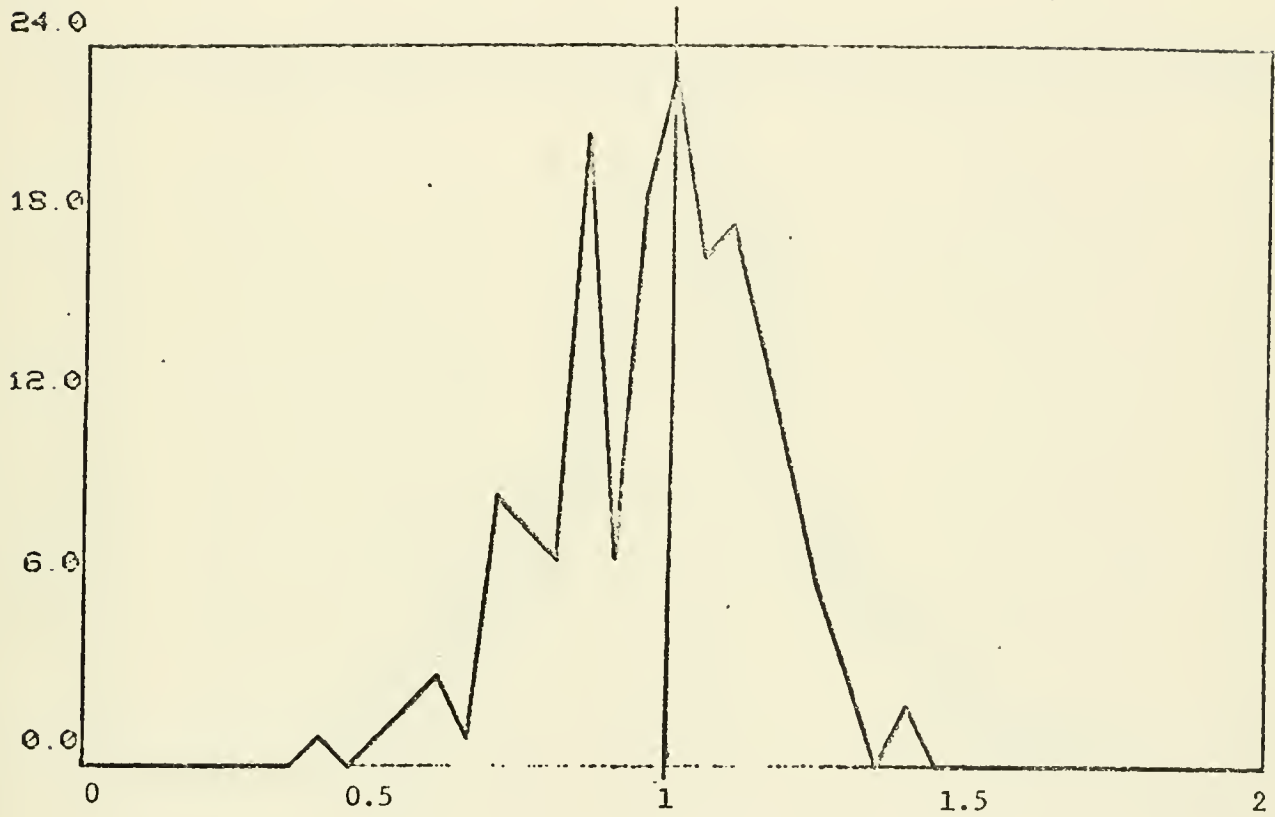
FIGURE 4:   TEST # 2   DISTRIBUTION
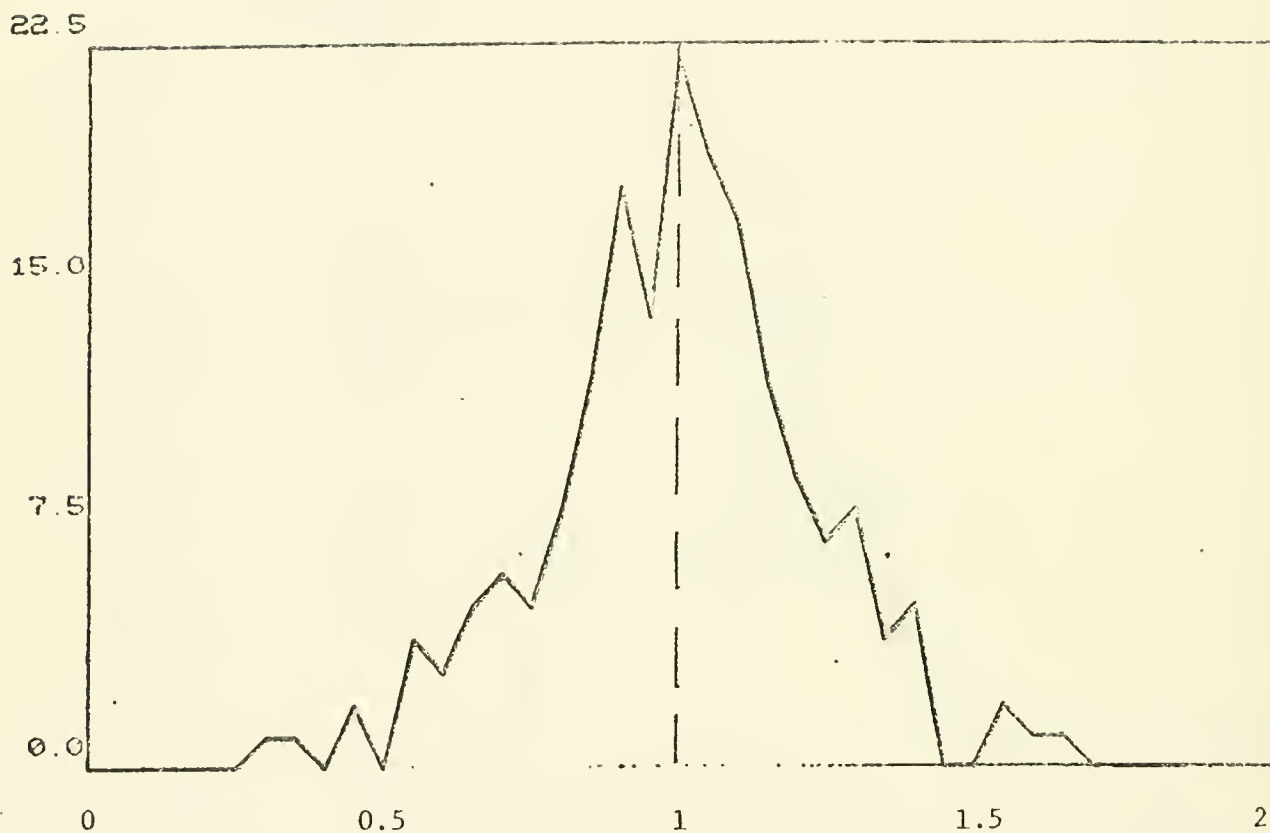
FIGURE 5:   TEST # 3 DISTRIBUTION

FIGURE 6:   TEST # 4 DISTRIBUTION

FIGURE 7:   TEST # 5 DISTRIBUTION
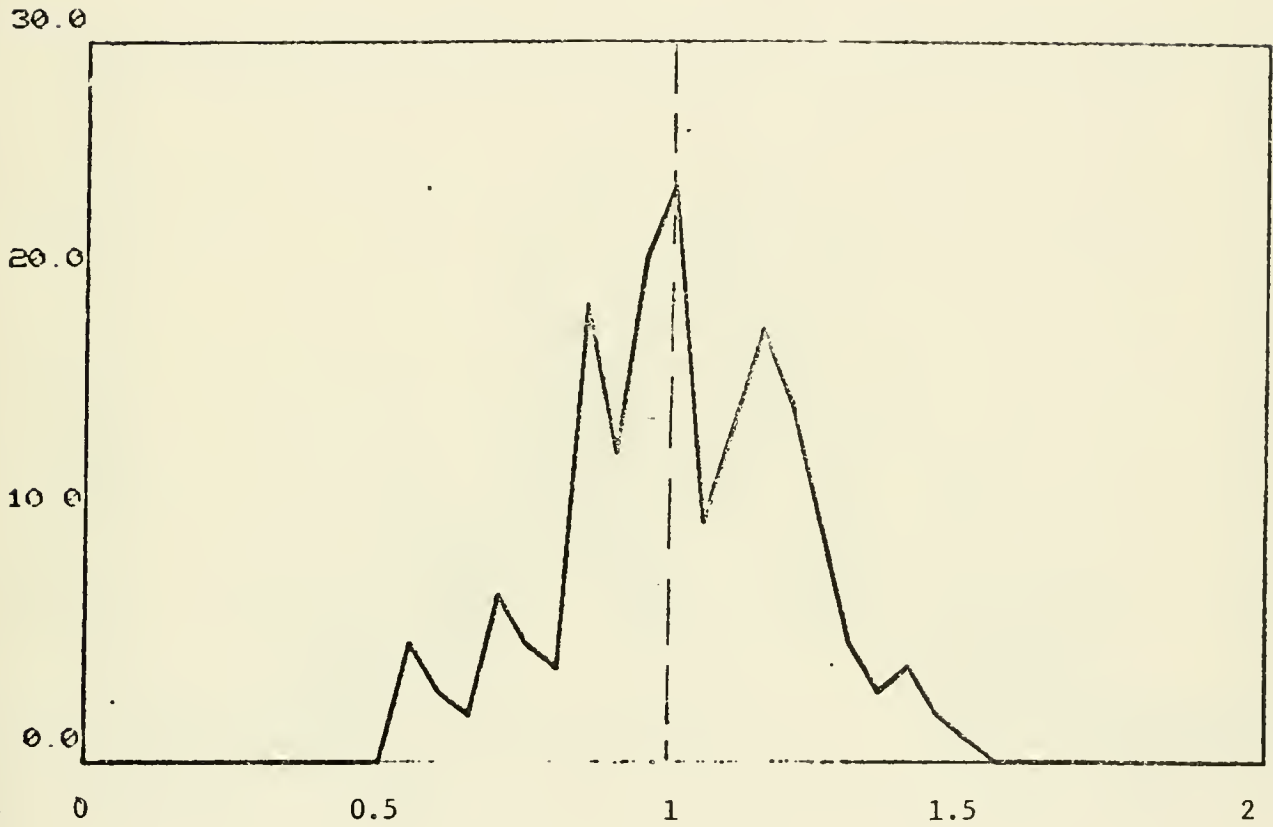
FIGURE 8:   TEST # 6 DISTRIBUTION

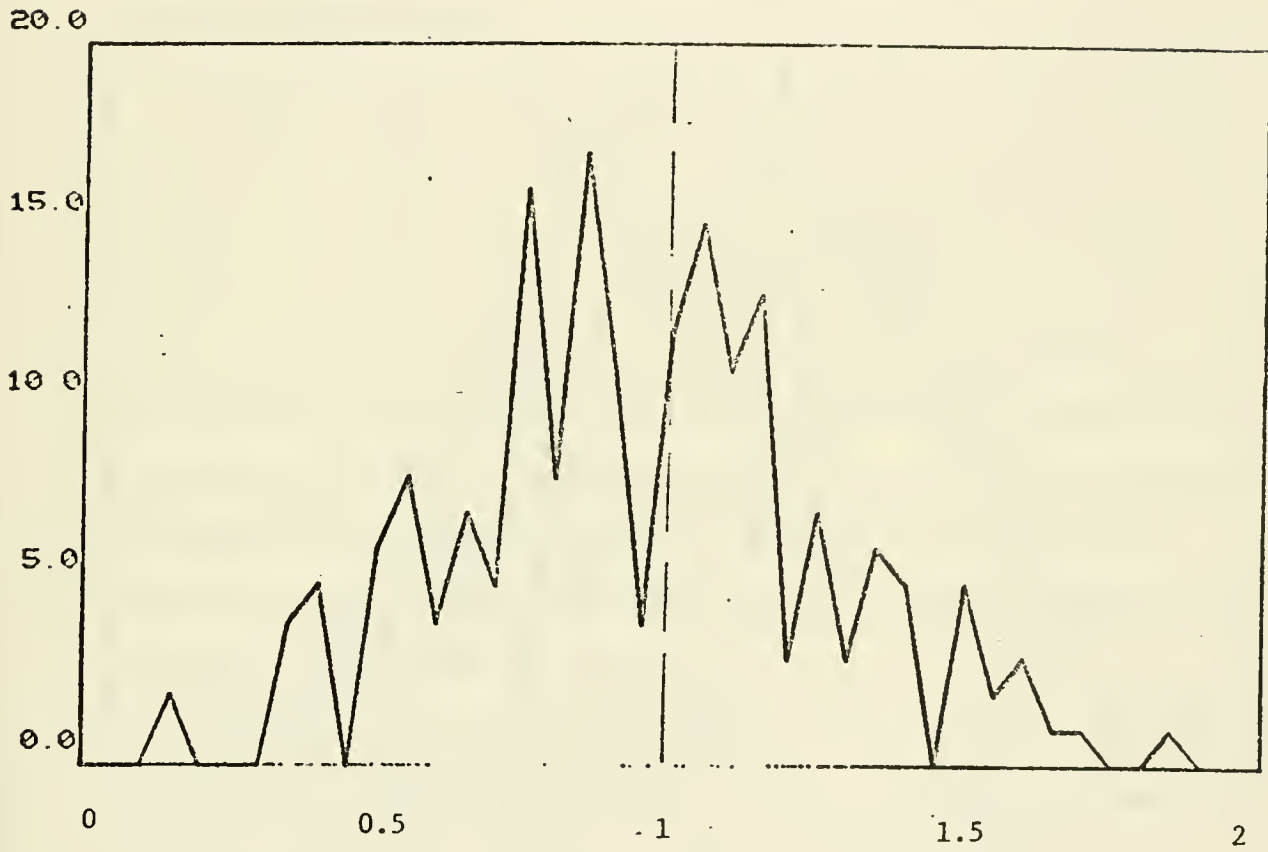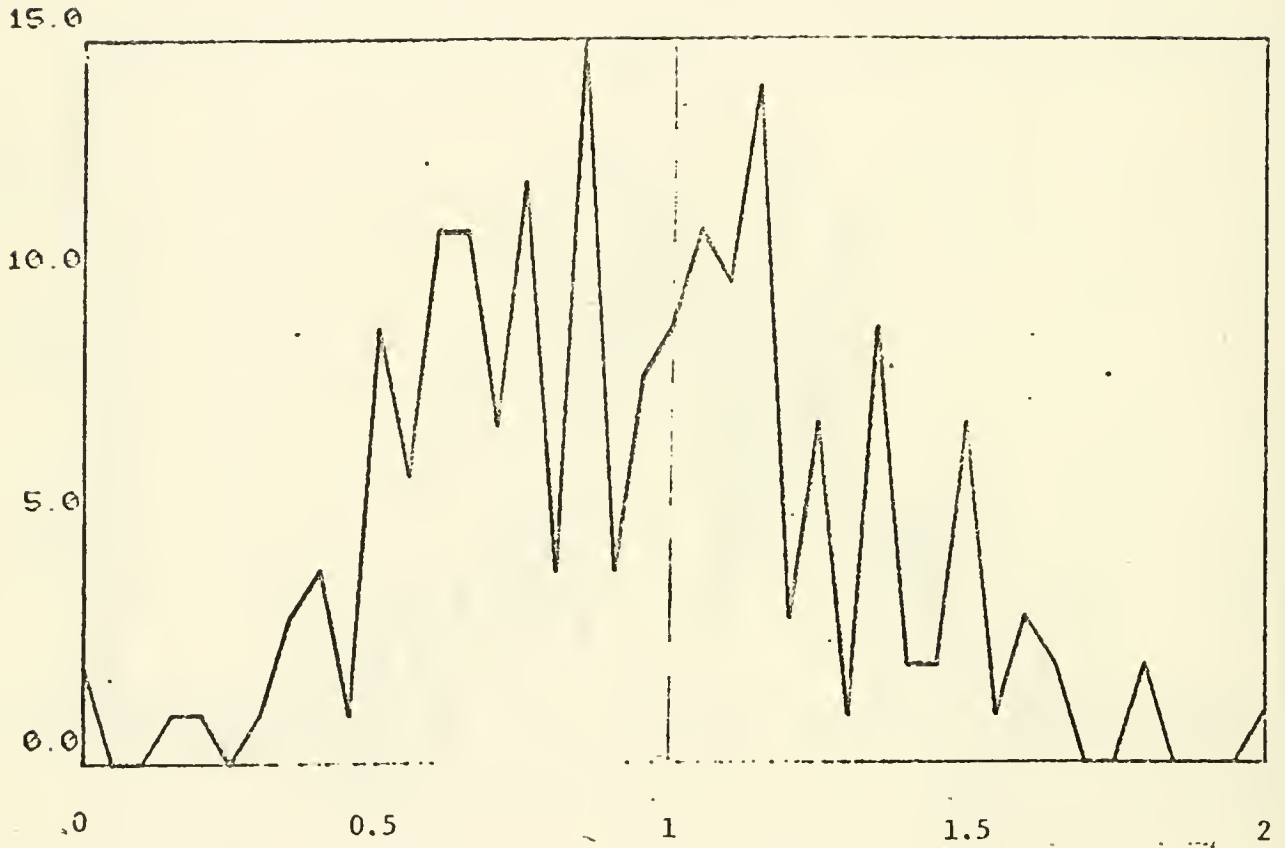FIGURE 9:   TEST # 7 DISTRIBUTION

-33-

FIGURE 10: TEST # 8 DISTRIBUTION

Another problem that is more difficult to deal with is potential errors in identifying the correct market forecast. One potential error is due to the discrete quarterly timing of the data used to determine the forecasts. The adjustment process may differ from the data intervals used.

An example of this problem would be as follows. The proportion of the fund in stocks increased from 0% to 50% from March 31 to June 30, 1980, thus indicating a forecast that the return on the market will exceed the return from bonds. However, the proportion of the fund in stocks may have been 100% on May 31. This would imply that the managers of the fund had actually been reducing the proportion of the fund in stocks over the last month of the quarter, thus indicating a forecast that bonds would outperform stocks. Thus, it would be useful to examine data on the composition of the funds more often than quarterly, if such data was available.

It may also be true that the market timer's forecasting horizon is shorter than three months. Then it would be necessary to use shorter observation intervals, such as monthly data, for actual market returns in testing for forecasting ability. Even if the data required to determine the forecasts is not available for less than three-month intervals, the quarterly forecast estimates can be tested against the actual returns for shorter intervals. If the market timing performance improves, it is likely that the forecast horizon is shorter than three months.

An even more difficult problem to deal with is the possibility

that investment managers have some desirable average proportion of stocks. In the absence of market forecasting, one would expect some pattern of adjustment back to this average in response to changes in market values. If such an adjustment process exists, there may be periods when it acts in the opposite direction of the adjustment due to the market timer's forecast. Depending on the magnitudes of the two adjustments, it may be impossible to determine the market timer's forecast from just the net change in the fund's proportions of stocks and bonds.

To deal with this problem, it is necessary to determine the desired average proportions and the pattern of the reversion adjustment. We could then examine adjustments independent of the reversion adjustment to determine the forecast. Simple tests along this line might use the average proportions for the sample and some sort of a moving average correction.

None of the identification problems would appear to introduce a systematic bias to our forecast estimates as they are just as likely to bias our forecast decision in either direction. However, it will make it more difficult to actually recognize market timing ability for a manager who actually can consistently beat the market. This is the result of the noise introduced into our estimation by measurement error.

One possible way of dealing with the measurement error problem is to aggregate the forecasts, under the assumption that the errors of the individual funds are uncorrelated with each other and thus will cancel out in the aggregate. Unfortunately, it is unlikely that the

measurement errors are uncorrelated, especially with respect to the
reversion adjustment as changes in market values will tend to effect
the funds in the same direction.

Nevertheless, we aggregated the forecasts of all of the funds
in our sample for each of the eight tests. As Table 5 shows, for
each of the five tests using quarterly data, the aggregate forecasts
result in estimations of $p_1(t) + p_2(t) > 1$, although in none of the
five tests could we reject H_0.

This seems surprising in light of the previously reported results.
However, over the seven years of the data set, the performance of the
stock market was streaky; it was common for the relative performance
of stocks, with respect to riskless securities, to stay the same for a
number of quarters. The mutual funds as a group did quite poorly at
predicting turning points. In fact, as Table 5 also shows, the
aggregate forecasts follow closely the relative performance of the
stock market in the previous quarter or six-month period.

This may show that investment managers follow a strategy of
relative strength selection.[13] Or more likely, it may just mean that
many of the funds in the sample do not fully adjust their portfolios
in response to changes in market values. In either case, the results
do not support the hypothesis that investment managers possess superior
market timing ability.

Table 5: AGGREGATE FORECASTS AND PROBABILITIES FOR THE EIGHT TESTS

	$\dfrac{n_1}{N_1}$	$\dfrac{1-n_2}{N_2}$	$\dfrac{n_1}{N_1} + \dfrac{1-n_2}{N_2}$	Aggregate Unconditional Probability of a Correct Forecast	Comparison of Aggregate Forecasts With Last Period's Market Performance $Z_M > R$	$Z_M < R$
Test #1	9/15	8/13	1.22	50.3%	13/15	12/13
#2	8/15	7/12	1.12	49.9	12/14	12/13
#3	9/15	6/13	1.06	49.7	14/15	11/13
#4	9/15	8/13	1.22	50.2	13/15	12/13
#5	9/15	6/13	1.06	50.9	14/15	11/13
#6	1/5	4/8	0.70	48.0	5/5	8/8
#7	1/5	4/8	0.70	46.9	5/5	8/8
#8	1/5	4/8	0.70	46.1	5/5	8/8

CONCLUSION

This paper has developed the statistical framework to test for
market timing ability using a model where forecasters predict only
if stocks will provide a greater return than bonds or if bonds will
provide a greater return than stocks. With this model, a necessary
and sufficient condition for a market timer to have superior forecasting
ability is that $p_1(t) + p_2(t) > 1$ where $p_1(t)$ and $p_2(t)$ are the
probabilities of a successful forecast, conditional on the actual
return on the market. Using changes in the relative proportions of
stocks and bonds held by a mutual fund as a proxy for their market
timing forecast, we found no evidence of superior forecasting ability.

This model may seem simplistic in comparison with the Jenson
(1972a) formulation where forecasters predict the actual outcome of
the market and performance is evaluated by ρ, the correlation coeffi-
cient between the market timer's forecast and the actual returns on
the market portfolio. However, our nonparametric formulation characterizes
a more subjective approach to market timing and only requires that investors
prefer more to less.

The use of our nonparametric tests allow us to avoid most of the
estimation problems involved with Jenson's model. However, because
we must estimate the forecasts, we are faced with possible errors in
our estimates. This is the result of having only quarterly data for

the estimation of the forecasts and the possibility that the funds
follow some sort of a mean reversion process in reaction to changes in
market values that will offset any action in response to market timing.
Because of these possible problems, the test results should be considered
preliminary with further tests required to examine the extent of these
potential errors.

FOOTNOTES

1. For a description of the different forms of the Efficient Markets
    Hypothesis, see Fama (1970) or Brealey and Myers (1981).

2. The Capital Asset Pricing Model (CAPM) provides a framework for
    the pricing of individual securities based on homogeneous
    expectations and a mean-variance criterion function. Jenson
    (1972b) provides an excellent review of the theory and empirical
    tests of the model.

3. Treynor and Black called it the active portfolio because positions
    in the portfolio were based on special information which typically
    propagates rapidly as the market becomes aware of the information.
    Thus, this portfolio may experience a large volume of trading.

4. This result was also obtained by Treynor and Black (1973).

5. This test is actually a special case of a 2 x 2 contingency
    table test. I thank Fischer Black for pointing this out.

6. See Lehmann (1975) theorem 19 for a general proof that is appli-
    cable to the hypergeometric distribution. I thank Arnold Barnett
    for pointing out this proof.

7. Remember, this is only a correct test if one has some strong
    reason to believe that $p_1(t) = p_2(t)$.

8. For a sample of some of the stated objectives, see the individual
    mutual fund descriptions in Weisenberger Investment Companies
    Service (1979).

9. The term "bonds" includes all securities, including cash and
    preferred stocks, that are not equities or common stocks.

10. Fund proportion data for 12/31/72 was obtained from Weisenberger (1979). Data from 3/31/73-3/31/80 from Computer Directions Advisors, Inc.

11. For the market portfolio, we used the return, including dividends, from holding a portfolio replicating all of the stocks on the New York Stock Exchange, weighted by their relative sizes. This data was obtained from the CRSP tapes and from the Wall Street Journal.

12. For Treasury bills, we used one-month holding period returns for T-bills with the shortest time-to-maturity greater than one month. The data was obtained from the CRSP tapes and the Wall Street Journal.

13. For a discussion of this strategy, see Levy (1967) and Jensen (1967). This strategy is inconsistent with any form of the Efficient Markets Hypothesis.

BIBLIOGRAPHY

1. Black, F., 1976. "Why and How Does an Investor Diversify?"
        Markets, Vol. 1, No. 5.

2. Brealey, R. and S. Myers, 1981. Principles of Corporate Finance,
        McGraw-Hill.

3. Fama, E., 1970. "Efficient Capital Markets: A Review of Theory
        and Empirical Work," Journal of Finance 25 (May): pp. 383-
        417.

4. _____, 1972. "Components of Investment Performance," Journal
        of Finance 27 (June): pp. 551-568.

5. Grant, D., 1977. "Portfolio Performance and the 'Cost' of Timing
        Decisions," Journal of Finance 32 (June): pp. 837-846.

6. Jensen, M., 1972a. "Optimal Utilization of Market Forecasts and
        the Evaluation of Investment Performance," in Mathematical
        Models in Investment and Finance, G. Szego and K. Schell
        (eds.), North-Holland: pp. 310-335. .

7. _____, 1972b. "Capital Markets: Theory and Evidence,"
        Bell Journal of Economics and Management Science 3 (Autumn):
        pp. 357-398.

8. Lehmann, E., 1975. Nonparametrics: Statistical Methods Based on
        Ranks, Holden-Day.

9. Merton, R., 1979. "On Market Timing and Investment Performance,
        Part I: An Equilibrium Theory of Value for Market Forecasts,"
        MIT Working Paper #1076-79.

10. Treynor, J. and F. Black, 1973. "How to Use Security Analysis
        to Improve Portfolio Selection," Journal of Business
        (January): pp. 66-86.

11. _____ and K. Mazuy, 1966. "Can Mutual Funds Outguess the
        Market?" Harvard Business Review (July-August): pp. 131-136.

12. Weisenberger Investment Companies Service, 1979. Investment
        Companies, Warren Gorham and Lamont.

(1137,1138 , 131 ? + p__I.)

1145-80