NBER WORKING PAPER SERIES

PERFORMANCE EVALUATION OF MARKET TIMERS

Alex Kane

Stephen Gary Marks

Working Paper No. 2640

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1988

NBER Working Paper #2640
July 1988

PERFORMANCE EVALUATION OF MARKET TIMERS

ABSTRACT

Previous investigators have shown that the Sharpe measure of the performance of a managed portfolio may be flawed when the portfolio manager has market timing ability.

We develop the exact conditions under which the Sharpe measure will completely and correctly order market timers according to ability. The derived conditions are necessary, sufficient, and observable. We compare them to empirical estimates of actual market conditions, and find that the circumstances which can lead to a failure of the Sharpe measure do in fact occur. We show, however, that such failures can be greatly reduced by more frequent sampling.

Alex Kane
University of California,
San Diego
La Jolla, CA 92093-0062

Stephen Gary Marks
School of Management
Boston University
Boston, MA 02215

PERFORMANCE EVALUATION OF MARKET TIMERS*

I. INTRODUCTION

Admati and Ross [1], Dybvig and Ross [3], and Grant [4] demonstrate that the portfolio of a market timer having superior information can appear mean-variance inefficient to an outside observer. In particular, they find that the Sharpe [10] measure of the performance of a portfolio managed by a skilled market timer can indicate performance which is inferior to that of the market. We consider here whether such a measurement error is likely to occur in practice by developing the exact condition under which the Sharpe measure will fail to order timers according to ability. We re-examine the condition specified by Dybvig and Ross, and derive one which is both necessary and sufficient. We then test that requirement against data presented by Merton [8] and show that the Sharpe measure is in fact likely to be deficient under actual market conditions given the current industry practice of using quarterly data.

---

* We would like to thank Alan Marcus for valuable comments.

Note that the source of this measurement error is unobserved shifts in portfolio composition that result in a nonnormal unconditional distribution of returns. See Dybvig-Ross. If information on portfolio shifts is available then it can be used to measure performance. (An interesting technology for using portfolio shifts has been developed and used by Merton [9], Henriksson and Merton [6], Henriksson [5], and Cheng and Lewellen [2] to test market timing ability and by Kane, Lee, and Marks [7] to test yield curve prediction.) Unfortunately, information on portfolio composition is infrequently disclosed. The Sharpe measure remains the industry standard for portfolio performance measurement. Even though our results show that the Sharpe measure calculated on quarterly data is deficient under actual market conditions, we also conclude that more frequent (e.g. monthly or daily) sampling greatly mitigates the problem. That is, with more frequent sampling the Sharpe measure will correctly order timers according to ability, even in the absence of information on portfolio composition.

II. Conditions for a Complete Ordering

Dybvig and Ross [3] posit an environment in which there are two assets. One is riskless with return r. The other is the market portfolio which is risky with return $\tilde{x} = r + \pi + \tilde{s} + \tilde{\varepsilon}$ where π is the market risk premium, $\tilde{s}$ is the signal observed by the portfolio manager but not by an outside observer, and $\tilde{\varepsilon}$ is noise. For convenience, the risk-free rate, is assumed to be zero and $\tilde{s}$ and $\tilde{\varepsilon}$ are assumed to be independent normally distributed with zero means and variances σ_s^2 and σ_ε^2. Uninformed investors estimate the variance of x to be $\sigma^2 = \sigma_s^2 + \sigma_\varepsilon^2$. The manager invests a unit amount under a constant absolute risk aversion utility function $U(\tilde{\omega}) = -\exp[-A\omega]$ with $A > 0$. Dybvig-Ross assume that the manager maximizes client utility, i.e., that there are no principal-agent problems. We make this assumption as well.

Both Grant and Dybvig-Ross compare a market-timer to a nontimer. Our interest is to compare timers to each other in order to rank them according to ability. We assume that manager i does not observe s directly but receives a signal, z_i', that is conditionally normally distributed with mean s and variance σ_i^2. That is, $z_i'|s \sim N(s,\sigma_i^2)$. In Dybvig-Ross, σ_i^2 would be zero. We, however, will allow σ_i^2 to take on all positive values. In this way, we will be able to rank timers; good timers will have lower σ_i^2.

Unconditionally, z_i' will be correlated with s with correlation coefficient

$$\rho_i = \frac{\text{Cov}(z_i', s)}{[\text{Var}(z_i')\sigma_s^2]^{1/2}}$$

Again, in Dybvig-Ross ρ_i would be one. In our model ρ_i can be any number in the unit interval. Better timers will have higher ρ_i.

Since $$\mathrm{Cov}(z_i', s) = E_s E(z_i' s \mid s) = \sigma_s^2$$

and $$\mathrm{Var}(z_i') = \sigma_s^2 + \sigma_i^2$$

we have that

$$\rho_i = \frac{\sigma_s^2}{[(\sigma_s^2 + \sigma_i^2)\sigma_s^2]^{1/2}} = \frac{\sigma_s}{(\sigma_s^2+\sigma_i^2)^{1/2}}$$

or

$$\sigma_i^2 = \sigma_s^2\left(\frac{1-\rho_i^2}{\rho_i^2}\right) .$$

Because the unconditional expectation of z_i' is zero, we can use the above relationships to state the conditional and unconditional distributions of z_i' in terms of s, σ_s^2 and ρ_i as:

$$z_i' \mid s \sim N\left(s,\ \sigma_s^2 \left[\frac{1 - \rho_i^2}{\rho_i^2}\right]\right)$$

$$z_i' \sim N(0,\ \sigma_s^2/\rho_i^2)$$

Given the distributions of z_i', $z_i' \mid s$, and s we can specify the distribution of $s \mid z_i'$. Thus,

$$f(s|z_i') = \frac{f(z_i'|s)f(s)}{f(z_i')}$$

where f refers to the probability density functions. Hence,

$$s|z_i' \sim N(\rho_i^2 z_i',\ (1-\rho_i^2)\sigma_s^2)$$

Although z_i' is not an unbiased estimator of s, we can obtain one by using

$$z_i = \rho_i^2 z_i'$$

instead of z_i', where

$$z_i \sim N(0,\ \rho_i^2 \sigma_s^2)$$

$$z_i|s \sim N(\rho_i^2 s,\ \rho_i^2(1-\rho_i^2)\sigma_s^2)$$

$$s|z_i \sim N(z_i,\ (1-\rho_i^2)\sigma_s^2)$$

Empirically, ρ_i^2 can be estimated from past forecasting errors.

With a zero-risk-free rate, the return on the market can be written as

$$x = \pi + s + \varepsilon$$

and thus the distribution of x conditional on z_i is

$$x \mid z_i \sim N(\pi + z_i,\ (1-\rho_i^2)\sigma_s^2 + \sigma_\varepsilon^2)$$

Portfolio manager i will invest a proportion, γ_i, of fund i in the market and the remaining proportion, $1-\gamma_i$, in the risk-free asset. These proportions will depend on the information, z_i, available to the manager; that is, $\gamma = \gamma_i(z_i)$. The return on the portfolio therefore will be $[\gamma_i(z_i)]x$. Following Dybvig-Ross [3] we assume that γ_i is chosen to maximize the negative exponential utility function

$$U = E[-\exp(-A\gamma x) \mid z_i]$$

where x is normally-distributed conditional on z_i. Using the normal moment-generating function, we may rewrite utility as

$$U = -\exp[-A\gamma E(x \mid s_i) + \tfrac{1}{2} A^2\gamma^2 \mathrm{Var}(x \mid s_i)]$$

Maximizing this is equivalent to maximizing

$$A\gamma(\pi + z_i) - \tfrac{1}{2} A^2\gamma^2[(1-\rho_i^2)\sigma_s^2 + \sigma_\varepsilon^2]$$

where we have substituted for the conditional expectation and variance from the relationships above. At a maximum, then,

$$\gamma_i(z_i) = \frac{\pi + z_i}{A[(1-\rho_i^2)\sigma_s^2 + \sigma_\varepsilon^2]} \tag{1}$$

and we may hereafter drop the subscript i. Note that if $\rho = 1$, we have the case considered by Dybvig-Ross.

Grant and Dybvig-Ross show that, from the viewpoint of an outside observer, a market-timer may appear mean-variance inefficient. That is, a naive investor would expect to obtain the same mean and a lower variance in his or her portfolio by taking a fixed position in the risky and risk-free assets. Equivalently, the unconditional Sharpe [10] measure S for the fund (defined as the risk premium over the standard deviation) may be lower for a timer with $\rho > 0$ than for the market. Consider the square of the Sharpe measure, since it is mathematically more convenient. We can express this as a function of ability, ρ^2, in the form

$$S^2(\rho^2) = \frac{[E(\gamma(\tilde{z})\tilde{x})]^2}{Var(\gamma(\tilde{z})\tilde{x})} = \frac{[\pi^2+\rho^2\sigma_s^2]^2}{\rho^4\sigma_s^4 + \rho^2\sigma_s^2[3\pi^2+\sigma_s^2+\sigma_\varepsilon^2] + \pi^2(\sigma_s^2+\sigma_\varepsilon^2)} . \qquad (2)$$

Comparing the squared Sharpe measure of the best timer ($\rho^2=1$) to a passive strategy, of holding the market portfolio ($\rho^2=0$), we have

$$S^2(1) < S^2(0) \text{ if}$$

$$3\pi^2\sigma_s^2 - \sigma_s^2\sigma_\varepsilon^2 - \sigma_s^4\sigma_\varepsilon^2/\pi^2 - \sigma_s^6/\pi^2 > 0$$

which is the result found by both Grant and Dybvig-Ross. Unfortunately, σ_s^2 is not observable. However, Dybvig-Ross show that this condition will be satisfied if $\pi^2 > \sigma_s^2 + \sigma_\varepsilon^2$ which guarantees that the first term of the inequality dominates the others. The sum $\sigma_s^2 + \sigma_\varepsilon^2$ of course is just the observable market variance, which we denote by σ^2. Thus, we have an observable condition ($\pi^2 > \sigma^2$) which, if satisfied, implies that a timer will be mean-variance inferior to a nontimer (have a lower Sharpe measure). However, this observable condition, while being sufficient, is not necessary. That is, even if $\pi^2 < \sigma^2$ it is possible for a timer to appear mean-variance inferior to a nontimer. In addition, it is possible for superior timers to appear mean-variance inferior to inferior timers.

We may then determine the conditions under which Sharpe measure will provide a complete correct ordering of timers according to ability, and the conditions under which it will fail. That is, we may compare managers of different abilities. The derivative of $S^2(\rho^2)$ with respect to ability, ρ^2, is

$$\frac{dS^2}{d\rho^2} = S^2\left[\frac{2\sigma_s^2}{\pi^2 + \rho^2\sigma_s^2} - \frac{2\rho^2\sigma_s^4 + \sigma_s^2[3\pi^2+\sigma_s^2+\sigma_\varepsilon^2]}{\rho^4\sigma_s^4 + \rho_s^2[3\pi^2 + \sigma_s^2 + \sigma_\varepsilon^2] + \pi^2(\sigma_s^2 + \sigma_\varepsilon^2)}\right] \qquad (3)$$

and from this expression we can obtain several important results.

Proposition 1: Greater ability will yield a higher Sharpe measure than lesser ability for all $\rho^2 \varepsilon [0,1]$ if and only if $\sigma_s^2 + \sigma_\varepsilon^2 > 3\pi^2$. For the proof, see the Appendix.

This guarantees that $S^2(\rho^2)$ will be monotonically increasing in ρ^2 for all $\rho^2 \geq 0$ if and only if $\sigma^2 > 3\pi^2$ and thereby that the Sharpe measure will correctly order ability through the whole range of abilities. The condition for a complete inverse ordering is given by:

Proposition 2. Greater ability will yield a lower Sharpe measure than lesser ability for all $\rho^2 \varepsilon[0,1]$ if $\sigma_s^2 + \sigma_\varepsilon^2 < \pi^2$. The proof is also contained in the Appendix.

It is interesting to note that the condition $\sigma^2 < \pi^2$ not only implies that $S^2(1) < S^2(0)$ as Dybvig-Ross suggested but, more strongly, that $S^2(\rho_1^2) < S^2(\rho_2^2)$ whenever $\rho_1^2 > \rho_2^2$. That is, it implies a complete inverse ordering of performance. Such an ordering is of course just as useful as a complete correct ordering in the absence of principal-agent problems. From these two propositions, we can divide market conditions into three π,σ regions, as depicted in Figure 1.

Figure 1 Goes Here

In region A, utility maximizing managers will be ordered in exactly the wrong way. In region C, ability will be ordered exactly correctly by the Sharpe measure. In region B there will always be misordering over some range of ability. It should be noted that the condition for a complete and correct ordering of ability by the Sharpe measure ($\sigma^2 > 3\pi^2$) is both strong (i.e., necessary and sufficient) and observable. We may also use the estimates supplied by Merton [8] to consider whether this condition has been violated in the past.

III. Empirical Estimates

The problem with using the Sharpe measure to order timers is that although x is normally distributed, the unconditional distribution of the timer's portfolio returns, $[\gamma(z)]x$, is nonnormal (it has a chi-square element [3]). Nevertheless, from Proposition 1 we know that mean and variance will order timers correctly if and only if $\pi/\sigma < .577$ where π and σ are stated in terms of the observation interval.

The usual practice in portfolio performance measurement is to use quarterly data. Therefore we will assess the viability of the Sharpe measure assuming a quarterly measurement interval for measuring fund performance. The most comprehensive estimates of the mean return on a market index are given by Merton [8]. Merton uses a Bayesian prior of a positive risk premium to estimate $\hat{\gamma}_M = \pi/\sigma$ (see his model 2) from monthly data where π is the market risk premium and σ is the market standard deviation. We add the subscript M to Merton's notation to indicate that it is a monthly estimate.

We will apply the condition on the Sharpe measure with quarterly data. Since all rates are continuously compounded, we may convert Merton's monthly estimates of $\hat{\gamma}$ into quarterly estimates (regardless of the level of the risk-free rate) by noting that

$$\pi_Q = t\pi_M = 3\pi_M$$

$$\sigma_Q = \sqrt{t}\sigma_M = \sqrt{3}\sigma_M$$

where t is the number of months in a quarter and thus that

$$\bar{Y}_Q = \sqrt{t}\bar{Y}_M = \sqrt{3}\bar{Y}_M$$

where the subscripts Q and M refer to quarterly and monthly rates respectively.

Table 1 presents Merton's data for thirteen four year sampling periods covering the years 1926-1978. The last column, T, lists the longest corresponding observation interval (in months) for which the Sharpe measure would not fail given Merton's estimates. This "break-even" period is determined by

$$\bar{Y}_T = 1/\sqrt{3}$$

$$\sqrt{T}\ \bar{Y}_M = 1/\sqrt{3}$$

$$T = \frac{1}{3\bar{Y}_M^2}$$

Table 1 Goes Here

As Dybvig-Ross noted, mismeasurement using Sharpe measure has its cause in the nonnormality of the unconditional distribution of returns. The above indicates that for short intervals this nonnormal element is not large enough to affect the validity of the Sharpe measure. As the interval gets longer we reach a point (T) where reversals begin to occur.

It is fascinating that, even though the source of nonnormality is shifts in portfolio composition, the point at which reversals occur (T) does not depend on A in equation (1). (A can be considered to be a shift parameter.

Lower values result in greater shifts.) Rather, T depends only on market conditions as given by $\bar{Y}$, the Sharpe measure of the market. The above result states that the higher the Sharpe measure of the market, the greater the degree of nonnormality and the more difficult it is for timers to distinguish themselves from the market and from each other. That is, reversals occur at smaller intervals.

From Table 1, we see that in six of the 13 subperiods, the Sharpe measure would either fail ($\bar{Y}_Q > 0.577$ or $T < 3$) or come close to failing for a quarterly observation interval (T=3). Merton also estimated $\bar{Y}_M$ over 52 1-year intervals and found the average monthly estimate to be .3719. This results in

$$\bar{Y}_Q = .6441$$

which is again over the threshold of .577. Thus, the conditions for the Sharpe measure to order timers incorrectly according to ability do seem to occur frequently. The table also suggests that monthly observations would greatly mitigate, if not eliminate, the problem.

IV CONCLUSIONS

Grant [4], Dybvig-Ross [3] and Admati-Ross [1] have shown that the Sharpe [10] measure of the performance of the portfolio of a market timer having superior information can be inferior to that of a passive portfolio. Herein we developed the exact conditions under which the Sharpe measure will completely and correctly order market timers according to ability. The condition derived is necessary, sufficient, and observable. We compared this requirement to the empirical estimates of market conditions reported by Merton [8]. We found that the conditions for a failure of the Sharpe measure have in fact occurred. Nevertheless, we show that such failures can be greatly reduced by more frequent sampling.

Appendix: Proofs of Propositions

Proposition 1: Greater ability will yield a higher Sharpe measure than lesser ability for all $\rho^2 \varepsilon[0,1]$ if and only if $\sigma_S^2 + \sigma_\varepsilon^2 > 3\pi^2$.

Proof: From equation (c), we have that

$$\frac{dS^2(0)}{d\rho^2} = S^2(0)\frac{\sigma_S^2}{\pi^2}(\sigma_S^2+\sigma_\varepsilon^2)(\sigma_S^2+\sigma_\varepsilon^2-3\pi^2) \tag{A1}$$

Additionally, from equation (3), $S^2(\rho^2)$ can be written in the form

$$S^2 = \frac{a\rho^4 + b\rho^2 + c}{a\rho^4 + d\rho^2 + e}$$

where a, b, c, d, and e are constants. Thus

$$\frac{dS^2}{d\rho^2} = \frac{a(d-b)\rho^4 + 2a(e-c)\rho^2 + (be-dc)}{(a\rho^4 + d\rho^2 + e^2)}$$

The numerator has, at most, two roots in ρ^2. The denominator is always positive. From equation (3), we see that one root occurs where $S^2 = 0$. From equation (2), we know that $S^2 = 0$ if an only if $\pi^2 + \rho^2\sigma_S^2 = 0$. Thus, one root occurs at $\rho^2 = -\pi^2/\sigma_S^2$. Again from equation (3), we see that the other root occurs at

$$\frac{2\sigma_S^2}{\pi^2 + \rho^2\sigma_S^2} - \frac{2\rho^2\sigma_S^4 + \sigma_S^2[3\pi^2+\sigma_S^2+\sigma_\varepsilon^2]}{\rho^4\sigma_S^4 + \rho_S^2[3\pi^2+\sigma_S^2+\sigma_\varepsilon^2] + \pi^2(\sigma_S^2+\sigma_\varepsilon^2)} = 0$$

Solving this expression for ρ^2 (by finding the common denominator and then setting the numerator equal to zero) results in

$$\rho^2 = (\frac{\pi^2}{\sigma_S^2})(\frac{3\pi^2 - (\sigma_S^2+\sigma_\varepsilon^2)}{\pi^2 + \sigma_S^2 + \sigma_\varepsilon^2}).$$

If $\sigma_S^2 + \sigma_\varepsilon^2 > 3\pi^2$, both roots are negative. If this is the case, then $dS^2/d\rho^2$ is positive at $\rho^2 = 0$ (by A1) and never becomes 0 as ρ^2 increases. Also, $dS^2/d\rho^2$ is continuous. Thus, $dS^2/d\rho^2$ must be positive for $\rho^2 > 0$. This completes the "if" part of the proof. The "only if" follows directly from A1. If $\sigma_S^2 + \sigma_\varepsilon^2 < 3\pi^2$, then $dS^2/d\rho^2$ at $\rho^2 = 0$ is negative and S^2 fails for some values of ρ^2. QED.

Proposition 2. Greater ability will yield a lower Sharpe measure than lesser ability for all $\rho^2 \varepsilon [0,1]$ if $\sigma_S^2 + \sigma_\varepsilon^2 < \pi^2$.

Proof $dS^2/d\rho^2$ is negative at $\rho^2 = 0$ since $\sigma_S^2 + \sigma_\varepsilon^2 < \pi^2$ implies $\sigma_S^2 + \sigma_\varepsilon^2 < 3\pi^2$ and we can apply Proposition 1. Consider then the roots of $dS^2/d\rho^2$. As was stated in the proof of proposition 1, one root in ρ^2 is always negative. The other, equation (A1), can be rewritten as

$$\rho^2 = \frac{\pi^2}{\sigma_S^2}\left[3 - 4\,\frac{\sigma_S^2+\sigma_\varepsilon^2}{\pi^2 + \sigma_S^2+\sigma_\varepsilon^2}\right] .$$

Since $\pi^2 > \sigma_S^2 + \sigma_\varepsilon^2$ it follows that

$$\frac{\sigma_S^2 + \sigma_\varepsilon^2}{\pi^2 + \sigma_S^2 + \sigma_\varepsilon^2} < \frac{1}{2} \quad \text{and} \quad \frac{\pi^2}{\sigma_S^2} > 1$$

which means that this root occurs at $\rho^2 > 1$ and that $S^2(\pi^2)$ never reaches a minimum in [0,1].

Q.E.D.

Figure 1
Market Conditions and Ordering by Sharpe Measure

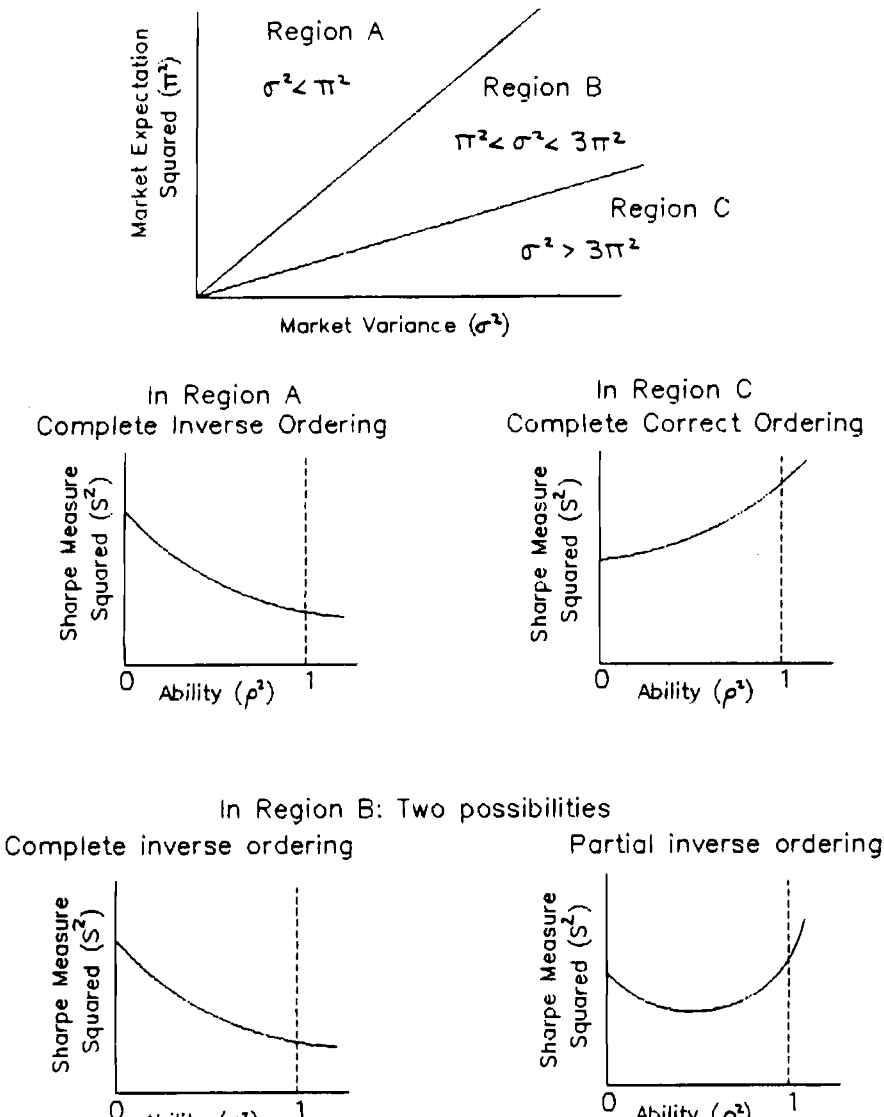

Table 1

Estimates of $\hat{\gamma}_m$, $\hat{\gamma}_Q$ and T for 4-Year Time Intervals

July 1926 - June 1978

	$\hat{\gamma}_M$	$\hat{\gamma}_Q$	$\hat{T}$
7/26 - 6/30	.2768	.4794	4.35
7/30 - 6/34	.1122	.1943	25.48
7/34 - 6/38	.2675	.4633	4.66
7/38 - 6/42	.0790	.1368	53.41
7/42 - 6/46	.5510	.9544	1.10
7/46 - 6/50	.1715	.2970	11.33
7/50 - 6/54	.4119	.7134	1.96
7/54 - 6/58	.3027	.5243	3.64
7/58 - 6/62	.2370	.4103	5.93
7/62 - 6/66	.3336	.5778	3.00
7/66 - 6/70	.1032	.1787	31.30
7/70 - 6/74	.1424	.2466	16.44
7/74 - 6/78	.1547	.2679	13.93

References

1. Anat R. Admati and Stephen A. Ross. "Measuring Investment Performance in a Rational Expectations Equilibrium Model" Journal of Business, 58(1), January, 1985.
2. Eric C. Cheng and Wilbur G. Lewellen. "Market Timing and Mutual Fund Investment Performance" Journal of Business, 57 (January 1984).
3. Philip H. Dybvig and Stephen A. Ross. "Differential Information and Performance Measurement Using a Security Market Line," Journal of Finance, Vol. XL, No. 2, June 1985.
4. Dwight Grant "Market Timing and Portfolio Management," Journal of Finance, Vol. XXXIII, No. 4, September 1978.
5. Roy D. Henriksson. "Market Timing and Mutual Fund Performance" Journal of Business, Vol. 57, 57 (January 1984)
6. Roy D. Henriksson and Robert C. Merton. "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills" Journal of Business, 54 (October 1981)
7. Alex Kane, Young Ki Lee, and Stephen Marks "The Forescasting Ability of Money Market Managers and Its Economic Value" Boston University Working Paper (1987)
8. Robert C. Merton. "On Estimating the Expected Return on the Market," Journal of Financial Economics, Vol. 8, No. 4 (1980).
9. Robert C. Merton, "On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts" Journal of Business, 54 (January 1984).
10. William Sharpe, "Mutual Fund Performance," Journal of Business, January 1966.