Sagi Akron, Doron Nissani, Mahmoud Kaadan

The Evaluation of Mutual Fund Performance on the Tel Aviv Stock Exchange:

The Ranking of Mutual Funds Based Solely on Selection as Opposed to Selection and Timing

Summary

April 2022

1. Introduction

Economic growth and capital market efficiency are intertwined. Economic growth stems from the expansion of the corporate economic activity, which is reflected in an increase in the income held by the general public. The public's savings are a portion of this income, which are then translated into capital market investments, largely through institutional investors. In a perfect market, the investments allocated through institutional investors expand corporate economic activity. In this manner, institutional investors play a crucial role in the allocation of public savings in the capital market.

Following the 2008 global financial crisis, low nominal interest rates have created a reality in capital markets, which has led to the channeling of public savings towards high-yielding, risk-bearing investments. Characterized by professional management and supervised by the Israel Securities Authority (ISA), [Israeli] mutual funds enjoy the trust of institutional investors, and as a result, are an accepted avenue of investment in the stock market.

The Capital Asset Pricing Model (CAPM) describes capital markets as a competitive mechanism possessing complete information attained by quantifying the elements of return and risk for each asset, based on that asset's expected return and standard deviation, respectively (Markowitz, 1952; Sharpe, 1964; Lintner,1965; Mossin, 1966). The definition of an investment portfolio as a characteristic collection of all assets according to their relative weight enables us to outline the efficient frontier – a collection of efficient portfolios composed solely of risky assets. The combination of the efficient frontier and a single risk-free asset allows us to describe the capital market line (CML) as a collection of possible efficient portfolios in the capital market, and according to which, there exists a uniquely-weighted portfolio of risky assets that a risk-averse individual will hold in combination with the risk-free asset (Merton, 1972). Equilibrium dictates that for each asset, there exists a characteristic security market line (SML) which prices the asset in accordance with its particular systematic risk.

The insights of the CAPM radically changed capital markets and elevated the importance of investment portfolios in financial markets. The development of investment portfolios of all kinds accelerated the development of systems to measure fund performance. The major performance measures are derived directly from the CAPM model and include the Sharpe index (Sharpe,1966) – which is defined as the ratio of excess return (risk premium) to the total risk of a portfolio; the Treynor index (Treynor, 1966) – which is defined as the ratio of excess return to the portfolio's systematic risk; and the Modigliani-Modigliani measure (Modigliani & Modigliani, 1997) – which compares the performance of a portfolio relative to a market portfolio bearing the same level of risk. Each one of the measures examines, in relative terms, the performance of the excess returns of an investment portfolio according to the various risk measures. Unlike the above measures, the Jensen index (Jensen, 1968) estimates excess return in absolute terms, that is to say, the difference between the actual return on the asset and its expected return in equilibrium.

The four performance measures presented above are derived from the CAPM, and are valid, therefore, as long as the basic assumption underlying the model, which states that returns on financial assets are derived from a multivariate normal distribution, holds. From a theoretical standpoint, the absence of this assumption means that the efficient frontier is not in fact efficient for risk-averse investments (Rothschild & Stiglitz, 1970). From an operational standpoint, it is easy to test the statistical data through normal distribution tests (Jarque & Bera, 1980). The ability to generalize the CAPM model for complete information, without the constraint of a multivariate normal distribution, remains an open question in the economic literature.

Part of the financial literature suggests that the performance of an investment portfolio or a mutual fund manager can be estimated randomly, through econometric models comprised of time series of a fund's returns relative to those of the market. In practice, the estimation of a fund manager's selection capabilities stems from a change in the portfolio's SML that relates to the estimated risk premium of the fund relative to that of the market, and is described formally as the value at which the regression line is intersected (Jensen, 1968). According to this approach, as the estimate becomes more positive and higher in value, the performance of the fund manager, from the standpoint of asset selection and fund composition, improves, capturing greater excess returns for its holders. Accordingly, a negative estimated value indicates that the fund manager's portfolio selection ability is inferior, and an estimated value of zero means that fund's asset selection does not significantly differ statistically from the market portfolio from the standpoint of performance.

Fama (1972) argued that portfolio selection in and of itself does not explain fund performance and that one must incorporate the fund manager's ability to time investments as well, i.e., whether the fund manager increases the fund's exposure to risky assets (for example, shares) in rising markets and increases its exposure to risk-free assets (riskless bonds) in falling markets. In other words, the timing of trading is no less important than selection when evaluating the performance of fund managers.

Treynor & Mazuy (1966) generalized timing in the least-squares regression model, where selection skill is estimated as the value at the intersection of the regression, and timing skill is estimated through the coefficient of the fund's performance on its curve. As the timing estimate becomes more positive and higher in value, the fund manager's timing is better able to obtain greater returns for its holders. When the value of the timing component equals zero, one receives the linear regression articulated in the Jensen model.

Henriksson & Merton (1981) presented an alternative model for examining the selection and timing skills of mutual fund managers, through a non-continuous linear regression. In this model, selection ability is estimated identically to its estimation under the Jensen model. The estimation of timing skill, however, is divided into two scenarios – a rising market and a falling market. The difference between the timing skill in rising and falling markets expresses the estimated timing score for the fund manager. The greater the difference in the timing ability between the two market states, the greater the fund manager's ability to yield higher returns for its holders.

2. A New Approach to Estimating Mutual Fund Performance – the Weighted Average Between Fund Manager Selection and Timing Skills

The ranking of mutual fund performance sparks great interest among investors, analysts, regulators and all players in the mutual fund industry. Mutual fund management performance is of particular interest when market players seek to understand whether the management fees paid to fund managers are justified and faithfully reflect the excess returns attained by them. At the same time, however, from a practical standpoint, the classic systems used to measure the performance of mutual funds has stagnated. Sometimes, the performance rankings are based on the simplistic ranking of returns for arbitrarily selected periods of time, and at best, classic financial models are employed to calculate excess returns standardized for some sort of risk estimate – most, if not all, classic indices express the fund managers' selection skill. The objective of this research is to cast light on the existing gap in the literature and to propose a new approach to measuring fund performance through the calculation of the weighted average of two estimated management skills – *selection and timing* - that will serve as a vital tool in critical investment decision-making and in the ranking of the managers of interested parties in the capital market, i.e., based on more than a one-dimensional analysis.

For example, classic indices standardized for risk, which measure the *selection* skill of fund managers to attain excess returns, are models that were primarily developed in the second half of the 1960s (Sharpe, 1966; Treynor,1965; Jensen, 1968; Modigliani & Modigliani, 1997). These classic models were more dominant than others because of the belief that they reflect a manager's selection skills, and that it is impossible to estimate timing skills, even though this has an impact, albeit insignificant, on performance measurement. Starting in the 1980s, a new approach began gaining traction among researchers, according to which one could measure a fund manager's *timing* and that this component has a potentially greater impact on fund performance measurement (Brinson, Hood & Beebower, 1986; Bollen & Busse, 2001; Shilling, 1992). Other studies demonstrated that the significance of a fund manager's timing is no less than that of its selection skill (Hensel, Ezra & Ilkiw, 1991; Ibbotson, 2010). In addition, subject to the time horizon of the investment, other researchers claim that timing strategy can be a decisive component in forecasting mutual fund returns – and can, for example, range from 40% to 100% of the explanation of the spread of expected returns (Ibbotson & Kaplan, 2000). To the best of our knowledge, the performance estimation systems commonly used on financial data sites do not include accessible, investor-friendly data, and lack, in particular, the dimension of timing when evaluating a fund manager's investment strategy. Accordingly, the new approach we propose will fill this void, and will provide investors and other parties interested in the mutual fund market with a user-friendly instrument allowing them to breakdown a manager's skills and to rank performance based on a weighted average of these skills' two dimensions – each investor in accordance with the importance he assesses to selection or timing and any given point of time.

We denote the fund manager selection performance indices as follows: SI = Sharpe Index; TI= Treynor index; and M2= Modigliani & Modigliani index. These are calculated respectively as:

$$(1)\ SI_P = \frac{(\mu_P - \mu_F)}{\sigma_P};\quad (2)\ TI_P = \frac{(\mu_P - \mu_F)}{\beta_P};\quad (3)\ M2_P = (\mu_P - \mu_F)\frac{\sigma_M}{\sigma_P} - (\mu_M - \mu_F).$$

Since these indices lack the ability to measure a manager's timing skills, we shall employ an econometric model that was developed in the 1960s (Treynor & Mazuy, 1966) to estimate both fund management skills – selection and timing:

$$(4) \qquad X_{P,t} - X_{F,t} = \zeta_P + \eta_P\left(X_{M,t} - X_{F,t}\right) + \xi_P\left(X_{M,t} - X_{F,t}\right)^2 + \vartheta_t \ .$$

selection　　　　　　　　timing

2.1 Algorithm to Apply a Performance-Ranking System Using a Weight-adjusted Selection-Timing Index

Let us illustrate the proposed innovative ranking system, which takes both fund manager ability measures into account: selection and timing, in the following manner:

1. We first define the reference group for performance ranking. From a classical standpoint, one can select a single group of funds or multiple groups for reference, according to the will of the party interested in ranking fund performance. For example, in a simulation that we will prepare below, we define two groups of flagship funds that have raised significant interest among the members of the ISA's Department of Research, Development and Strategic Economic Consulting with respect to the definition of concurrent relative performance – actively managed "flexible" funds and 90/10 funds, each commanding a 20% market share of the total operating mutual funds.
2. For each fund, we estimate the selection (S) parameter and the timing (T) parameter using the Treynor & Mazuy model (equation 4 above), where higher values of each one of the estimated parameters indicate higher fund management skills. For each fund, we signify cases in which the significance of all the parameters measured in the model (F test), as well as the particular significance for each one of the parameters of fund management skills (selection vs. timing) (t test). In other words, we indicate each fund for which there a statistical significance of (a) 1%, (b) 5% and (c) 10%.
3. We now conduct the "horse race" between all mutual funds and the desired comparison group. For each estimated mutual fund parameter, we calculate a score as follows:

$$Z_{x(S)} = \frac{x-\mu}{\sigma} \ ; \ Z_{x(T)} = \frac{x-\mu}{\sigma} \ .$$

4. Following the scoring, we introduce a probability spread for five states of nature that represent the level of importance that an investor can assess for each pair of estimated skills, selection and timing, normalized for each given fund:
   a. Extreme interest in selection skill: W_S=1.0 W_T=0.0.
   b. Decided interest in selection skill: W_S=0.9 W_T=0.1.
   c. Equal interest in selection and timing skills: W_S=0.5 W_T=0.5.
   d. Decided interest in timing skill: W_S=0.1 W_T=0.9.

e. Extreme interest in timing skill: W_S=0.0 W_T=1.0

Comment: It is understood that one can create a continuum of probabilities, and that the five cases above are meant solely for the sake of clarifying the system.

5. In the next stage, we calculate the weighted average of estimated selection and timing skill scores for each fund. We rank the selection and timing-weighted index performance scores in quintiles, where quintile 1 reflects the highest performance ranking, among the weighted average according to a certain probability series (certain assessed importance between selection and timing at a given time). Similarly, quintile 5 reflects the lowest performance ranking according to the weighted average under a certain probability series.

Ranking of quintiles of performance measurement	Interpretation of the relative ranking of performance parameters – selection and timing according to the three scenarios above
1	normalized parameter ≥ 80 percentile
2	80 percentile > normalized parameter ≥ 60 percentile
3	60 percentile > normalized parameter ≥ 40 percentile
4	40 percentile > normalized parameter ≥20 percentile
5	normalized parameter < 20 percentile

Table 1. Interpretation of the quintile rankings – quintile 1 represents the highest performance, while quintile 5 represents the lowest performance

6. To simplify and to create an investor -friendly ranking device, we differentiate the quintile rankings by color in the ranking table as follows: green symbolizes the highest performance rankings 1 and 2; blue signifies the mid-level quintiles 3 and 4; while light red signifies the lowest performance ranking – 5.
7. We now apply a similar quintile ranking algorithm for the classic performance indices as well - the Sharpe, Treynor & M2 indices – facilitating a comparison between the quintile rankings of the classic indices and the performance ranking index we have created, based on a two-dimensional weighted average of fund management skills – selection and timing – according to the five states of nature spread over the by investor-assessed importance of these two skills.

2.2 Simulation on Mutual Fund Performance in Tel Aviv

We will now demonstrate the ranking performance of the proposed system and the added value it brings to investors, regulators, analysts and all capital market players. First, we create a reference group for the investor, which consists of two types of flagship funds popular among institutional investors, regulators and retail investors: 113 flexible funds (rate of return estimates, January 2008-March 2021) and 141 90/10 funds (rate of return estimates, January 2016 – March 2021). Second, the selection parameter and timing parameter is estimated for each fund (Treynor & Mazuy model, denoted by T-M), which constitute the basis of the proposed weighted average index, under an assumption of assessing extreme importance to selection skill (W_S=1.0, W_T=0.0) and extreme importance to timing skill (W_T=1.0, W_S=0.0), respectively.

Third, we normalize the estimated parameters by calculating a score for each estimated parameter of each fund. In the next stage, we calculate the weighted average of these scores based on the five probabilities of assessed selection and timing importance. We then run the algorithm for ranking quintiles on the weighted average of these scores, based on each one of the five states of nature probabilities of assessed selection-timing importance. In accordance with each selection-timing importance ranking assessed by the investor, there will now be a quintile ranking of mutual fund performance among the reference group. In addition to the proposed new ranking system, we repeat the performance rankings based on the classic performance indices, Sharpe, Treynor and M2, including the systematic risk coefficient, *Beta*. It is now possible to compare the ranking results using the classic performance indicators with the ranking results of the innovative weighted index. Through this comparison, it is possible to primarily examine cases in which the ranking under the classic measures deviates from the proposed index based on the weighted average of selection and timing skills.

The findings of the simulation and quintile rankings are presented in Table 2 below.

Manager Code Num	Weighted Average WS+WT=1 Fund Type	Index (1) T- M Selection 100% (Alpha Z) Quintile rating	Sig. Level	Index (5) T- M Timing 100% (Gamma Z) Quintile rating	Sig. Level	T-M Regression F Sig. Level	T-M Regression F (P-value)	Inv importance Selection 90% Weighted Ave S&T Index (2) T- M	Inv importance Timing 90% Weighted Ave S&T Index (4) T- M	Equal Importance T-M 50% Weighted Ave S&T Index (3) T- M	The classical performance measures: Sharpe ratio Quintile rating	Beta Quintile rating	Treynor ratio Quintile rating	M2 Quintile rating	T-M 50% Vs. Sharpe Quintile rating contrast: Fund Type	Manager Code Num
1	Flexible	2	c	4		a	0	1	3	2	4	1	4	4	Flexible	1
2	Flexible	5		1		a	0	5	1	4	5	2	5	5	Flexible	2
3	Flexible	4		4		a	0	4	4	4	4	2	4	4	Flexible	3
4	Flexible	5		3		a	0.005	5	4	5	5	1	5	4	Flexible	4
5	Flexible	1	a	3		a	0	1	3	2	3	2	2	2	Flexible	5
6	Flexible	2	b	2		a	0	2	2	1	4	1	4	3	Flexible	6
7	Flexible	5		2		a	0	5	2	5	5	2	5	5	Flexible	7
8	Flexible	1	b	5	b	a	0	1	5	4	4	1	4	4	Flexible	8
9	Flexible	5		5		a	0.001	5	5	5	5	1	5	5	Flexible	9
10	Flexible	5	b	1	c	a	0.004	5	1	2	5	1	5	5	Flexible	10
11	Flexible	5		4		a	0.001	5	4	5	5	2	5	5	Flexible	11
12	Flexible	5	b	1		c	0.077	5	2	5	5	5	1	1	Flexible	12
13	Flexible	5	c	5			0.665	5	5	5	5	2	5	5	Flexible	13
14	Flexible	5		2		a	0	5	2	5	5	1	5	5	Flexible	14
15	Flexible	5		5		a	0.001	5	5	5	5	1	5	5	Flexible	15
16	Flexible	5		5			0.388	5	5	5	5	2	5	5	Flexible	16
17	Flexible	5		5			0.163	5	5	5	5	2	5	2	Flexible	17
18	Flexible	5		1		a	0	5	1	3	5	1	5	4	Flexible	18
19	Flexible	5		1			0.124	5	1	5	5	2	5	2	Flexible	19
20	Flexible	5	c	2		a	0	5	2	5	5	1	5	5	Flexible	20
21	Flexible	1	a	5	a	a	0	1	5	2	2	3	2	2	Flexible	21
22	Flexible	5		2		a	0	5	2	5	5	1	5	5	Flexible	22
23	Flexible	4		2		a	0	3	2	2	4	1	5	4	Flexible	23
24	Flexible	5		3		a	0	5	3	5	5	1	5	5	Flexible	24
25	Flexible	5		4		a	0	5	4	5	5	2	5	5	Flexible	25
26	Flexible	5	c	2		a	0	5	2	5	5	1	5	5	Flexible	26
27	Flexible	5		1	c	a	0	5	1	2	5	1	5	4	Flexible	27
28	Flexible	5	b	1	b	a	0	5	1	4	5	2	5	5	Flexible	28
29	Flexible	5		1		a	0	5	1	4	5	1	5	5	Flexible	29

Table 2. Quintile ranking – index based on weighted average between selection-timing vs traditional performance measures.

Comment: "a", "b", and "c" denote the statistical significance level of 1%, 5% and 10%, respectively, for the T-M model.

An examination of Table 2 brings the high added value of the proposed index to light. For example, let us arbitrarily examine a situation in which investors assess equal weight to selection and timing skills. Flexible fund number 28 received the lowest ranking according to the Sharpe index, and is placed accordingly in the lowest quintile – No. 5. An evaluation of management skills under the two-dimensional selection–timing ranking, however, shows that the fund manager has excellent timing skills, which place the fund, based solely on timing, in the highest quintile – No. 1, while, the ranking based solely on selection confirms the fund manager's relatively low ranking, and places the fund once again in the lowest quintile – 5. In other words, for an investor that, at a given time, assesses equal weight to selection and timing skills, the performance of the fund manager is not the worst, but rather average, placed the fund in quintile 4. That is to say, the more an investor gives greater weight to timing, the ranking of the fund manager's performance improves. For example, for an investor that at a given point of time assesses 90% importance to timing skill, the performance of the fund manager's performance will rank in the highest quintile.

Other examples can be found in funds 8 and 10. When the performance of fund 10 is ranked under the selection-biased Sharpe index, they are ranked as the lowest-performing funds, in quintile 5. However, it is saliently clear that, based on the weighted index, for cases in which equal importance is assessed to selection and timing, the fund's performance is ranked among the highest quintiles – quintile 2 for an equal assessment and quintile 1 based on a 90% timing importance assessment. With respect to fund 8, it is also found that when the ranking under the weighted selection-timing index (assuming an equal importance assessment to selection and timing skills) is the same as the ranking according to the Sharpe index, both placing the fund in quintile 4. Investors receive added value through the ability to break down the fund manager's selection and timing skills, as is reflected from the ranking of fund 8 as having very high selection skills (quintile 1), alongside very low timing skills (quintile 5), such that, if one anticipates a sudden shift in market conditions, the investor can adjust its strategy in selecting the fund in accordance with the relevant importance of each estimated skill separately.

3. Empirical Test of Ranking Quintiles According to the Weighted Index and Classic Indices

We now compare the quintile ranking results for the reference group of mutual funds according to the classic indices against the weighted average index, under the five probability scenarios associated with the importance assessed to selection and timing skills respectively. Table 3 displays the correlations of the ranking data according to each performance index. From the table, it is evident that the Sharpe index faithfully represents, at a high level of statistical significance, the Treynor index (0.01>P, 0.92) and the M2 index (P>0.01, 0.81). In addition, one should take note to the interchangeability between the weighted index proposed in this paper, the selection-biased importance ranking and the timing-biased ranking. Finally, one can see that the highest correlation between the weighted selection/timing index proposed in this research and the Sharpe index is attained under the assumption of equal importance assessed by the investor for selection and timing skills (0.01>P, 0.64).

Pearson Correlation Coefficients, N = 254
Prob > |r| under H0: Rho=0

Variable	Selection100%	Selection 90%	Selection 50%	Timing 90%	Timing 100%	Sharpe ratio	Beta	Treynor ratio	M2
Selection100% (T-M)	1	0.95364	0.28198	-0.31527	-0.38383	0.45043	0.11255	0.39946	0.36222
		<.0001	<.0001	<.0001	<.0001	<.0001	0.0733	<.0001	<.0001
Selection 90% (T-M)	0.95364	1	0.36667	-0.23529	-0.3095	0.50834	0.03523	0.4632	0.42395
	<.0001		<.0001	0.0002	<.0001	<.0001	0.5763	<.0001	<.0001
Selection 50% (T-M)	0.28198	0.36667	1	0.63725	0.56416	0.64181	-0.23288	0.66144	0.58096
	<.0001	<.0001		<.0001	<.0001	<.0001	0.0002	<.0001	<.0001
Timing 90% (T-M)	-0.31527	-0.23529	0.63725	1	0.9716	0.29244	-0.23288	0.33759	0.30422
	<.0001	0.0002	<.0001		<.0001	<.0001	0.0002	<.0001	<.0001
Timing 100% (T-M)	-0.38383	-0.3095	0.56416	0.9716	1	0.22745	-0.23853	0.28235	0.26666
	<.0001	<.0001	<.0001	<.0001		0.0003	0.0001	<.0001	<.0001
Sharpe ratio	0.45043	0.50834	0.64181	0.29244	0.22745	1	-0.48778	0.92338	0.81925
	<.0001	<.0001	<.0001	<.0001	0.0003		<.0001	<.0001	<.0001
Beta	0.11255	0.03523	-0.23288	-0.23288	-0.23853	-0.48778	1	-0.56418	-0.55439
	0.0733	0.5763	0.0002	0.0002	0.0001	<.0001		<.0001	<.0001
Treynor ratio	0.39946	0.4632	0.66144	0.33759	0.28235	0.92338	-0.56418	1	0.84872
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001		<.0001
M2	0.36222	0.42395	0.58096	0.30422	0.26666	0.81925	-0.55439	0.84872	1
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	

Table 3. Table of Correlations between the quintile rankings of each system.
Comment: The P-value for each estimated correlation appears underneath the correlation.

Panel A: The full sample descriptive statistics

Variable	N	Mean	Std Dev	Min	5th Pctl	25th Pctl	Median	75th Pctl	95th Pctl	Max	Skewness	Kurtosis
Selection100%	254	3.205	1.422	1	1	2	3	5	5	5	-0.142	-1.363
Selection 90%	254	3.000	1.420	1	1	2	3	4	5	5	0.000	-1.309
Selection 50%	254	3.000	1.420	1	1	2	3	4	5	5	0.000	-1.309
Timing 90%	254	3.000	1.420	1	1	2	3	4	5	5	0.000	-1.309
Timing 100%	254	3.004	1.421	1	1	2	3	4	5	5	-0.007	-1.315
Sharpe ratio	254	3.004	1.418	1	1	2	3	4	5	5	-0.007	-1.303
Beta	254	3.008	1.423	1	1	2	3	4	5	5	-0.006	-1.308
Treynor ratio	254	3.004	1.418	1	1	2	3	4	5	5	-0.007	-1.303
M2	254	3.004	1.418	1	1	2	3	4	5	5	-0.007	-1.303

Panel B: The 90/10 Trust Funds sub-sample descriptive statistics

Variable	N	Mean	Std Dev	Min	5th Pctl	25th Pctl	Median	75th Pctl	95th Pctl	Max	Skewness	Kurtosis
Selection100%	141	3.092	1.055	1	2	2	3	4	5	5	0.110	-1.052
Selection 90%	141	2.787	0.977	1	1	2	3	4	4	5	-0.073	-0.811
Selection 50%	141	2.950	1.161	1	1	2	3	4	5	5	-0.069	-0.835
Timing 90%	141	3.099	1.311	1	1	2	3	4	5	5	-0.225	-1.027
Timing 100%	141	3.092	1.298	1	1	2	3	4	5	5	-0.193	-0.998
Sharpe ratio	141	2.546	1.216	1	1	2	3	3	5	5	0.327	-0.809
Beta	141	3.844	0.936	2	2	3	4	5	5	5	-0.213	-0.994
Treynor ratio	141	2.532	1.216	1	1	1	3	3	5	5	0.287	-0.870
M2	141	2.482	1.251	1	1	1	2	3	5	5	0.418	-0.887

Panel C: The Flexible Trust Funds sub-sample descriptive statistics

Variable	N	Mean	Std Dev	Min	5th Pctl	25th Pctl	Median	75th Pctl	95th Pctl	Max	Skewness	Kurtosis
Selection100%	113	3.345	1.772	1	1	1	4	5	5	5	-0.353	-1.715
Selection 90%	113	3.266	1.798	1	1	1	4	5	5	5	-0.304	-1.760
Selection 50%	113	3.062	1.692	1	1	1	3	5	5	5	-0.031	-1.723
Timing 90%	113	2.876	1.542	1	1	2	2	5	5	5	0.241	-1.476
Timing 100%	113	2.894	1.561	1	1	2	2	5	5	5	0.194	-1.535
Sharpe ratio	113	3.575	1.450	1	1	2	4	5	5	5	-0.635	-1.023
Beta	113	1.965	1.224	1	1	1	2	2	5	5	1.406	1.057
Treynor ratio	113	3.593	1.437	1	1	2	4	5	5	5	-0.612	-1.060
M2	113	3.655	1.348	1	1	3	4	5	5	5	-0.681	-0.731

Table 4. Table of descriptive data of all mutual fund reference groups and subgroups.

The analysis allows us to also scrutinize any group of funds, better or worse, in a statistically significant manner, as presented in Table 4. As the first stage of the "horse race" conducted between the funds, we examine the descriptive fund performance data displayed in Table 4, according to the quintile algorithm for all funds in the sample (Panel A), and according to a cross-section of 90/10 funds (Panel B), as well as the group of flexible funds (Panel C). From Table 4, we learn that in the rankings according to the Sharpe index, the flexible fund group significantly underperforms (Sharpe average 3.57, median 4) the group of 90/10 funds.

In order to draw a statistically acceptable comparison between the performance rankings of flexible funds and those of 90/10 funds, we must conduct a t-test on the difference between the mean quintile rankings of the various performance measures, and particularly to compare the quintile rankings under the Sharpe index as representative for all the classic, selection-biased indices to the quintile rankings of the proposed weighted selection-timing performance index, on the five scenarios of investor assessment of importance. The findings of this analysis are summarized in Table 5, which illustrates that according to all the classic indices, especially the Sharpe index, the quintile performance rankings significantly demonstrate that the quintile rankings of flexible fund performance is higher – i.e., the performance is statistically significantly lower (P>0.01, 1.029). Nonetheless, the picture is not so clearcut, when we examine the weighted selection/timing performance index proposed in this paper. Particularly salient is the fact that in almost all of the five scenarios (with the exception of the scenario that assesses 90% importance to selection), one cannot determine that the performance of flexible funds is inferior to that of 90/10 funds. Accordingly, for example, when investors assess equal importance to selection and timing skills (Selection = 50%), the difference is no longer statistically significant. Practically speaking, as the importance of timing increases, the performance ranking of flexible funds becomes noticeably higher, even though the difference in this case also lacks statistical significance. It appears, therefore, that it is decisively important to test performance rankings on two dimensions – selection and timing, as opposed to settling for a one-dimensional analysis of fund management skill – solely selection.

	(A): 90/10 Trust Funds						(B): Flexible Trust Funds						Pr > |t|		Mean	(Flex) - (90/10)	
Variable	N	Mean	Median	Std Dev	Minimum	Maximum	N	Mean	Median	Std Dev	Minimum	Maximum	Equal (Var)	Unequal (Var)	Difference	Sig.	Diff.
Selection100%	141	3.09	3	1.05	1	5	113	3.35	4	1.77	1	5	0.159	0.182	0.253		+
Selection 90%	141	2.79	3	0.98	1	5	113	3.27	4	1.80	1	5	0.007	0.012	0.478	**	+
Selection 50%	141	2.95	3	1.16	1	5	113	3.06	3	1.69	1	5	0.535	0.551	0.112		+
Timing 90%	141	3.10	3	1.31	1	5	113	2.88	2	1.54	1	5	0.214	0.222	-0.223		-
Timing 100%	141	3.09	3	1.30	1	5	113	2.89	2	1.56	1	5	0.270	0.280	-0.198		-
Sharpe ratio	141	2.55	3	1.22	1	5	113	3.58	4	1.45	1	5	<.0001	<.0001	1.029	***	+
Beta	141	3.84	4	0.94	2	5	113	1.96	2	1.22	1	5	<.0001	<.0001	-1.879	***	-
Treynor ratio	141	2.53	3	1.22	1	5	113	3.59	4	1.44	1	5	<.0001	<.0001	1.061	***	+
M2	141	2.48	2	1.25	1	5	113	3.65	4	1.35	1	5	<.0001	<.0001	1.173	***	+

Table 5. Difference of means test on the quintile rankings of various performance measures – weighted average selection-timing index (with five levels of importance for selection and timing) vs. the quintile rankings of classic performance indices.
Comment: *, ** and *** signify statistical significance levels of 1%, 5% and 10%, respectively.

5. Summary and Conclusions

The systems for ranking the performance of mutual funds have been stagnant for decades. At best, these - including Sharpe (1966); Treynor (1965); Jensen (1968); Modigliani & Modigliani (1997) - are based solely on performance measures biased towards the selection skills exhibited by mutual fund managers. Since these classic measures ignore the consideration of a manager's timing skills, and given the growing recognition of the impact timing has on mutual fund management and of the attainment of excess returns – particularly in the face of anticipated extreme market conditions – there is a need to develop a performance measure that takes two dimensions of mutual fund management skills into consideration.

In this paper, we shed light on this gap between academia and practice, and propose a weighted index for ranking mutual fund management performance, which takes two dimension of fund management skills– selection and timing - , and even allows each investor to incorporate its subjective assessment of the relative importance of each fund management skill , at any given time, necessary for ranking and selecting mutual fund managers, while weighing the commissions they collect for these skills. An innovative weighted index such as this can serve as a key instrument in decision-making by investors, regulators, analysts and all parties interested in the mutual fund industry.

The addition of performance ranking based on quintiles enables the use of this user-friendly and readily available instrument to all players in the mutual fund space, allowing them to adjust their expectations to market conditions through the importance assessment they apply to each type of skill. This analysis demonstrates that the Sharpe index best represents performance ranking in terms of its statistically significant and high correlation with the other selection-biased classic mutual fund management performance measures. The proposed innovative index, which weighs the importance of both selection and timing, exposes the fund managers' average rate of substitution between selection and timing skills. Assessing equal importance (50% to each) generates the highest correlation forecast among the quintile rankings of forecasted performance based on Sharpe.

The proposed selection/timing weighted index clearly reflects the importance of ranking performance on more than the single dimension of selection skill, and it primarily imparts the proper importance for a fund manager's timing skill, subject to a similar investor assessment at a given point of time. This index is an important tool for decision-making based on a two-dimensional analysis, which can be beneficial to all players in the mutual fund industry. Accordingly, it is anticipated that the use of the index will enhance transparency in the mutual fund industry, increase trading volume, as well as the social benefits derived from the capital market overall.

Sources

Bollen, N. P. B., and J. A. Busse. (2001). On the timing ability of mutual fund managers. The Journal of Finance 56 (3): 94–1075.

Brinson, G. P., L. R. Hood, and G. L. Beebower. (1986). Determinants of portfolio performance. Financial Analysts Journal 42 (4): 39–44.

Fama, E.F. (1972). Components of investment performance. Journal of Finance, 27(3): 551–567.

Henriksson, R.D. and R.C. Merton (1981). On market timing and investment performance. II. statistical procedures for evaluating forecasting skills. Journal of Business, 54(4): 513–533.

Hensel, C. R., D. Don Ezra, and J. H. Ilkiw. (1991). The Importance of the asset allocation decision. Financial Analysts Journal 47 (4): 65–72.

Ibbotson, R. G. (2010). The Importance of asset allocation. Financial Analysts Journal 66 (2): 1–3.

Ibbotson, R. G., and P. D. Kaplan (2000). Does asset allocation policy explain 40, 90, or 100 percent of performance?. Financial Analysts Journal, 56(1), 26–33.

Jarque, C.M., and A.K. Bera (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters, 6(3): 255–259.

Jensen, M.C. (1968). The performance of mutual funds in the period 1945-1964. Journal of Finance, 23(2): 389–416.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics, 47(1): 13–37.

Markowitz, H. (1952). Portfolio selection. Journal of Finance, 7(1): 77–91.

Merton, R.C. (1972). An analytic derivation of the efficient portfolio frontier. Journal of Financial and Quantitative Analysis, 7(4): 1851–1872.

Modigliani, F., and L. Modigliani (1997). Risk-adjusted performance. Journal of Portfolio Management, 23(2): 45–54.

Mossin, J. (1966). Equilibrium in a capital asset market. Econometrica, 34(4): 768-783.

Rothschild, M., and J.E. Stiglitz (1970). Increasing risk I: A definition. Journal of Economic Theory, 2(3): 225–243.

Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance, 19(3): 425–442.

Sharpe, W. F. (1966). Mutual Fund Performance. The Journal of Business, 39(1): 119–138.

Shilling, A. Gary. (1992). Market timing: better than a buy-and-hold strategy. Financial Analysts Journal 48 (2): 46–50.

Treynor, J.L. (1965). How to rate management of investment funds. Harvard Business Review, 43(1): 63–75.

Treynor, J.L., and K. Mazuy (1966). Can mutual funds outguess the market? Harvard Business Review, 44(4): 131–136.