Skill, History and Risk-Adjusted Performance

By Arun S. Muralidhar


Arun Muralidhar is a Managing Director at FX Concepts, Inc. – a currency management firm, and the author of “Innovations in Pension Fund Management.” He has developed a risk-adjusted performance measure, published in the Financial Analysts Journal and Journal of Portfolio Management that is ideally suited to institutional investors. He has also published an article in the Journal of Performance Measurement on currency management.

Abstract

There has been a lot of research recently on risk-adjusted performance measures as investors have grown more sophisticated and have peered behind absolute performance measures. An ideal performance measure should adjust for risk, but also provide some indication of the skill of a manager, provide some indication on how portfolios should be constructed and ideally take into account the full performance history of the manager. This note reviews the major risk-adjusted performance measures, highlights deficiencies in the existing range of measures and proposes an alternative that attempts to achieve all three lofty goals in a performance measure.
Skill, History and Risk-Adjusted Performance

**BACKGROUND**

There has been a lot of interest in developing a proper and useful risk-adjusted performance measure. For the longest time, investors satisfied themselves with using the Sharpe ratio (ratio of excess return over the risk-free rate to the standard deviation of that excess) and the Information Ratio (ratio of excess returns vis-a-vis a benchmark to standard deviation of excess returns) to distinguish whether risk-adjusted performance was good or poor. However, a major contribution by Modigliani and Modigliani (1997) highlighted that such measures were inadequate for investors as the ratios said nothing about performance and provided no guidance as to how portfolios should be constructed. By adapting the Sharpe ratio, they were able to truly provide a performance measure (i.e., measured in basis points of performance) that adjusted for differences in volatility between a portfolio and its benchmark. This measure also provided some guidance as to how much should be invested in the risk-free rate and the portfolio being reviewed. This measure came to be called the $M^2$ measure of risk-adjusted performance and the authors were able to demonstrate that rankings of portfolios would be changed when this measured was used as opposed to raw performance. They were able to also show the flaws in using the information ratio – a measure that potentially rejected portfolios with high risk-adjusted performance.

Muralidhar (2000) demonstrated a few flaws in the $M^2$ measure, namely, that while this measure accounted for differences in variance between portfolios and a benchmark, it did not normalize for differences in the correlation of performance between portfolios and a benchmark. As a result, it was not useful for institutional investors who wish to budget tracking error (the standard deviation of excess returns and composed of the variance of the benchmark, portfolio and their
correlation) as it ranked portfolios inefficiently by unduly favoring portfolios with higher tracking error caused by low correlation with the benchmark. It also showed that the M^2 measure could not provide rankings that were consistent with rankings based on skill. This paper provided a new measure called the M^3 measure that (1) controls for differences in standard deviation and correlation; (2) provides guidance on portfolio construction between the risk-free rate, benchmark and active portfolio; and (3) provides rankings consistent with rankings based on skill.

Muralidhar (2001) provides a comparison of the above-mentioned measures and highlights the strengths and weaknesses of each.

However, the one problem with all the measures is that they do not account for differences in data history among portfolios. For example, in order to compare the ratios and or the performance measures, one is required to use the same data period for all portfolios – namely, the lowest common data period. If a manager had 3 years of history and another had only 1 year of performance, an investor is forced to use the one-year history to compare the two managers. This means giving up a lot of data for the manager with the longer history and potentially penalizing managers with the longer data history. One might ask why this is an issue and why not disregard the extra data? The problem with disregarding the additional data is that ignores the fact that the manager with a longer track record has delivered performance through different economic cycles and hence has a more established record. A newer manager may appear to have better performance, but this is over a shorter set of market events, suggesting less experience. However, even a short data history can demonstrate skill. Therefore, the entire performance history of all managers should be considered and not abbreviated data histories.

In the next section we provide a more technical discussion on the creation of a measure that achieves all the positive attributes of the M^3 measure, but which also goes the additional step of providing guidance to investors that accounts for differences in data history. For simplicity we
will refer to this measure as the Skill, History and Risk-Adjusted (SHARAD) measure. However, first we must review other measures and establish notation.

**LUCK VERSUS SKILL**

Outperformance over a benchmark unfortunately does not tell the investor whether the external manager or mutual fund manager is skillful. It does not provide the investor with a measure of confidence that they can have that the alpha is generated by skill-based processes. Critical factors involved in answering the luck versus skill question, include time, desired degree of confidence, investment returns of the portfolio and benchmark, standard deviation of the portfolio and the benchmark, and the degree of correlation between the two. The problem is there can be a lot of noise in performance data, and the more volatile the portfolio and the excess return series of a manager, the greater the noise and, hence, greater the time needed to resolve this question. Ambarish and Seigel (1996) demonstrate that the minimum number of data points or time History $H$, should be large enough for skill to emerge from the noise or, equivalently,

$$H > \frac{S^2(\sigma_I^2 - 2\rho\sigma_I\sigma_B + \sigma_B^2)}{\left[\left(\frac{\sigma_I^2}{2}\right)^2\left(\frac{\sigma_B^2}{2}\right)^2\right]^2} \quad (1)$$

where $I$ is the manager, $B$ is the benchmark, $r$ denotes the return, $\sigma$ denotes the standard deviation, $\rho$ is the correlation of returns between the manager and the benchmark, and $S$ is the number of standard deviations for a given confidence level. One can see from their paper that it is theoretically more appropriate and simpler than statistical process control alternatives such as Philips and Yashchin (1999). The details for the derivation of equation (1) are provided in Appendix 1 for the interested reader.
If the tracking error of portfolio 1 versus the benchmark is defined as the standard deviation of excess returns, it is trivial to define $\text{TE}(1)$ as follows:

$$
\text{TE}(1) = \sqrt{(\sigma_1^2 - 2\rho\sigma_1\sigma_B + \sigma_B^2)}
$$

(2)

Note: the second term in the numerator of the confidence in skill calculation or equation (1) is the same as the square of the tracking error of portfolio 1 versus the benchmark.

Equation (1) suggests that even a 300 basis point outperformance may require 175 years of data to claim with 84% confidence that the manager is skillful.\(^v\) Muralidhar and U (1997) and Muralidhar (1999) recognize that $H$ is often given by performance history and solve for the degree of confidence, $S$ instead.

Equation (1) makes it clear that the confidence in skill is intricately linked to the information ratio (IR). The annualized information ratio is equal to the annualized excess return divided by the annualized tracking error or, alternatively,

$$
\text{IR}(1) = \frac{\{r(1)-r(B)\}}{\text{TE}(1)}
$$

(3)

As a result and using (2) and (3), equation (1) can be rewritten in terms of $S$, where $S$ is a function of IR.

$$
S < \sqrt{H} \left[ \text{IR}(1) - \left( \frac{\sigma_i^2 - \sigma_B^2}{2\text{TE}(1)} \right) \right]
$$

(4)

The confidence in skill is derived from converting $S$ to percentage terms for a normal distribution or alternative the cumulative probability of a unit normal with standard deviation of $S$. Define
$C(S_t)$ as the cumulative probability of a unit normal with standard deviation of $S$ for fund $1$, and this will be the measure of confidence in skill. This measure will lie between 0% and 100% and hence acts as a probability measure. For example, when $S$ is equal to 1, then $C(S) = 84\%$. Also, when the second term in equation (4), or $\left( \frac{\sigma_t - \sigma_B}{2TE(1)} \right)$ is generally small or insignificant, the IR and length of data history will largely determine the confidence in skill. This is the case when tracking error is substantial and driven largely by a low correlation between the portfolio and the benchmark (i.e., $\sigma_t \cong \sigma_B$). As a result, two portfolios with identical variances, information ratios and tracking errors but differing only in length of history will have different confidence in skill – longer the history, greater the confidence.

**RISK-ADJUSTED PERFORMANCE**

Practitioners and academics recognize that performance unadjusted for risk is not very meaningful. Various risk measures are used to adjust performance. The two most commonly used measures of risk adjustment are the Sharpe ratio and the information ratio (also known as the differential Sharpe ratio).

Newer measures have been proposed that are variations on these measures. This section develops and evaluates a series of measures that are inter-related.

**The Sharpe Ratio and the Information Ratio**

The Sharpe ratio effectively adjusts performance above the risk-free rate by the volatility of the excess returns (where excess return is the portfolio return minus the risk-free rate), and the information ratio, demonstrated above, adjusts excess of benchmark performance by the volatility of the excess return series. Logue and Rader (1997) suggest that the Sharpe ratio is the best way
to adjust for risk. The information ratio (described above) is a variation of the Sharpe and is based on excess returns and the volatility of the excess returns.

**The M² Measure**

Modigliani and Modigliani (1997) make an important contribution by showing that the portfolio and the benchmark must have the same risk to compare them in terms of basis points of risk-adjusted performance. They propose that the portfolio be leveraged or deleveraged using the risk-free asset. If \( B \) is the benchmark being compared to portfolio \( I \), the leverage factor, \( d \), is defined as follows:

\[
d = \frac{\sigma_B}{\sigma_I}
\]  

Figure 1 demonstrates this transformation. It creates a new portfolio, called the risk-adjusted portfolio (RAP), whose return \( r(RAP) \) is equal to the leverage factor multiplied by the original return plus one minus the leverage factor multiplied by the risk-free rate. Thus, if portfolio \( F \) is the risk-less asset with zero standard deviation and is uncorrelated with other portfolios, the risk-adjusted return

\[
r(RAP) = d \times r(\text{actual portfolio}) + (1 - d) r(F),
\]  

where \( \sigma_{RAP} = \sigma_B \)

The correlation of the original portfolio to the benchmark is identical to the correlation of the RAP to the benchmark, as “leverage or deleverage” using the risk-free rate does not change the
correlation characteristics. The correlation is normally less than unity. If the correlation = 1, it could lead to a riskless arbitrage.

Figure 1: An evaluation of the M² measure

Figure 1 demonstrates four regions using the M² measure:

I. Portfolio outperformance on an absolute and risk-adjusted basis.

II. Portfolio outperformance on an absolute basis and underperformance on a risk-adjusted basis.

III. Portfolio underperformance on an absolute and risk-adjusted basis.

IV. Portfolio underperformance on an absolute basis and outperformance on a risk-adjusted basis.
The paper suggests that this $M^2$ adjustment allows for a comparison of “apples to apples”, namely, returns from the benchmark and the RAP have the same volatility. It shows that making this adjustment can reverse peer rankings of mutual funds or managers. The rankings are shown to be identical using the Sharpe ratio measure as the principle is similar. The $M^2$ measure, however, was preferred as it expressed risk-adjusted performance in terms of basis points of outperformance and provides guidance on assets allocated to the external manager (allocation of $d$) and the risk free asset (allocation of $1-d$). The paper also discards the use of the information ratio as it could lead to incorrect decisions. For example, portfolios in region IV would have a negative information ratio, but would have a risk-adjusted performance greater than the benchmark. Graham and Harvey (1997) propose a variation of this method, assuming the riskless asset need not be an asset uncorrelated with other assets. This only leads to different allocations across funds rather than suggesting a new approach.

The $M^2$ adjustment made the comparison in terms of basis points of outperformance by ensuring all portfolios had the same variance as the benchmark. The one major shortcoming was that two funds, normalized for the benchmark volatility, could have different correlations with the benchmarks and hence different tracking errors (see equation (2)). Tracking error is important to investors, especially institutional investors, because it provides a measure of the variability of a manager’s returns around the benchmark. Investors would prefer, all else being equal, funds with lower tracking error (and hence greater predictability in returns). Hence these rankings could provide investors with incorrect information about the relative risk-adjusted performance of funds.
The $M^3$ methodology – Adjusting for Differences in Correlations

An investor has to rely on available data to make projections for the future. Assuming historical distributions are preserved in the future, the three-dimensional problem of a comparisons of return, standard deviations, and correlations has to be synthesized into a simple two-dimensional space of return and risk. In mean-variance space, the risk-less asset is portfolio $F$ (with returns $r(F)$) and it can be used to leverage or deleverage the desired mutual fund/manager. In tracking error space, the only portfolio with zero tracking error is the benchmark portfolio as it is perfectly correlated with itself (where $\rho = 1$, $TE = 0$, as $\sigma_B = \sigma_i$). Therefore, combining active mutual funds/managers with passive benchmarks and the risk-less asset can be used to alter the overall portfolio’s standard deviation and its correlation with the benchmark.

To create measures of correlation-adjusted performance, the investor needs to invest in the mutual fund, the risk-less asset and benchmark to ensure: (a) the volatility of this composite is equal to that of the benchmark (Modigliani and Modigliani 1997); and (b) the tracking error of this composite is equal to the target tracking error (Muralidhar 2000). The $M^3$ measure recognizes that the investor has to consider basis points of risk-adjusted performance after ensuring that correlations of various funds versus the benchmark are also equal.

The $M^3$ model as Applied to a Retail Investor

While the principle is identical for any investor hiring an external manager, this section assumes that an investor is invested in a Defined Contribution pension plan and must evaluate several mutual funds. Hammond (1997) states that to establish performance-related thresholds for managers, the investor must set a target tracking error and compare funds to the target. A similar approach is proposed by Litterman, Longerstaey, Rosengartern and Winkelman (2001).
Assume that the investor is willing to tolerate a certain target annualized tracking error around the benchmark, say 300 bps \( (TE(target)) \). The investor essentially wants to earn the highest risk-adjusted alpha for given tracking error and variance of the portfolio. Now define \( a, b, \) and \( (1-a-b) \) as the proportions invested in the mutual fund, the benchmark, and the risk-less asset. Let \( \text{CAP} \) be the correlation adjusted portfolio and therefore the returns of a CAP,

\[
r(CAP) = a*r(\text{mutual fund}) + b*r(B) + (1-a-b)*r(F) \quad (8)
\]

As one can see, this is an extension of the \( M^2 \) measure. Further, the investor must hold appropriate proportions of each to ensure the final portfolio has the target tracking error and the standard deviation of the benchmark. For a specific mutual fund, say mutual fund \( 1 \), with a risk adjusted return \( r(CAP-1) \), equation (8) can be re-written as:

\[
r(CAP-1) = a*r(1) + (1-a-b)*r(F) + b*r(B) \quad (8')
\]

where the coefficients of each portfolio represent the optimal weight of that specific portfolio to ensure complete risk adjustment. In addition, from the constraint on tracking error, there is a unique target correlation between the \( \text{CAP} \) and benchmark \( B \). As demonstrated in Muralidhar (2000), this target correlation of the portfolio with that of the benchmark \( (\rho_{T,B}) \) is given by the equation for tracking error (equation (2)) when \( \sigma_B = \sigma_1 \); namely,

\[
\rho_{T,B} = 1 - \frac{TE(\text{target})^2}{2\sigma_B^2} \quad (9)
\]
By maximizing the $r(CAP)$ subject to the condition that the variance be identical to the benchmark, and its correlation to the benchmark equal to the target correlation, we find that for mutual fund $1$,

$$a = \frac{\sigma_B^2 (1 - \rho_{T,B}^2)}{\sigma_1^2 (1 - \rho_{1,B}^2)} = \frac{\sigma_B}{\sigma_1} \sqrt{\frac{(1 - \rho_{T,B}^2)}{(1 - \rho_{1,B}^2)}} \quad (10)$$

$$b = \rho_{T,B} - a \frac{\sigma_1}{\sigma_B} \rho_{1,B} \quad (11)$$

The details of these calculations are provided in the Appendix 2. If you substitute for “$a$” in equation (11), the allocation to the benchmark is independent of variances and is only a function of the correlation terms. While $b$ and $(1-a-b)$ may be greater than or less than zero (negative coefficients being equivalent to shorting the futures contract relating to the benchmark and borrowing at the risk-free rate), $a$ is constrained to being positive as it is not currently possible to short mutual funds.

This method is preferred to the $M^2$ as it: (a) expressed risk-adjusted performance in basis points; (b) gives advice on portfolio construction – specifically between the risk-free asset, the benchmark (passive investing) and the active portfolio (active management); and (c) provides rankings that are identical with rankings based on skill for equal histories. This measure also has the attractive property of keeping a constant annualized tracking error target over all time horizons. However, the $M^3$ falls short when two funds have different time periods of data.
In this section, we exploit the fact that the \( C(S) \) measure is a probability measure that has dependence on the history of the data series and that the \( M^3 \) measure provides appropriate risk adjustment and guidance on portfolio construction and expressed risk-adjusted performance in basis points. As a result, we can look at the SHARAD measure as a probability-adjusted measure (or as an expected risk-adjusted measure) by defining it for a fund 1 very simply as

\[
SHARAD (I) = C(S_I) * r(CAP-1)
\]

Where the “\( S \)” measure relates to the confidence that the \( CAP \) returns are skill-based (as opposed to using raw returns), and \( C(S_I) \) is the cumulative probability of a unit normal with standard deviation of \( S \). This revised measure now will have all the attractive properties of the \( M^3 \), but through the \( C(S) \) term accounts for time in a manner that is consistent with the skill evaluation. For managers with identical data histories, this adjustment will have no impact in changing orderings, but for different data histories, one should see interesting results. In the next section, we examine some data from mutual funds to see what the impact is of such adjustments.
Testing the Measure with Data

We examine data for 10 US equity mutual funds (all ranked 5-star by Morningstar) with different inception dates through August 1999. These funds are ordered based on their annualized return since inception and shown in Table 1. These funds are compared to the S&P500 index (benchmark) for the respective start date of the fund. This data set was chosen to broadly represent funds that have a high rating by the traditional rating agencies and to demonstrate some of the deficiencies of each of the measures.

The data ranges from as much as 20 years (Fund 4 and 10) to as little as 11.5 years (Fund 6). Normally, all additional data beyond 11.5 years would need to be excluded to compare these funds. We have assumed a fixed risk free rate of 5% for all funds to estimate the $M^2$, $M^3$ and the Sharpe ratios. In addition, we assume a $M^3$ target tracking error of an annualized 10% for all funds regardless of data history. We also provide details on the allocation to various funds using the $M^3$ methodology. The lower part of Table 1 provides the various returns, ratios and statistics using the measures described above. Table 2 provides the ranks of each fund for the various parameters by which a fund may be evaluated and Table 3 provides the correlation of ranks.

To help the reader who is unfamiliar with these techniques understand the impact of different histories, we provide the same three tables in the Appendix using 10 years of data through August 1999 for all funds (i.e., identical data histories). The key results about the superiority of the $M^3$ methodology are demonstrated in these tables as the risk-adjusted performance is expressed in performance units, rankings are identical to those on skill, and information is provided on portfolio construction. In this setup, the SHARAD rankings are identical to the $M^3$ rankings.
The following critical observations can be made from Tables 1-3, where the inception dates for the different funds are as per their respective data histories:

- The rankings differ between the Sharpe ($M^2$) and information ratio, and further between the Sharpe and the $M^3$. All three measures show a healthy disregard for absolute performance or excess performance over the benchmark as evidenced by the low correlation of ranks.

- Even though these funds are rated 5 star by Morningstar, the $M^3$ excess is negative for a number of funds (Funds 1, 4, 7, 9 and 10). Therefore, on a risk-adjusted basis these are funds that underperform.

- A poor fund cannot be saved by a long data history alone as demonstrated by Funds 4 and 10. The SHARAD of these funds rank poorly in this subset, even though they have the longest data history of the funds. The converse is true for Fund 6, where the short history is not a constraint on the confidence in skill.

- The $M^3$ no longer provide ranks that are identical to those based on skill (Fund 4).

- All previous measures, with the exception of confidence in risk-adjusted performance, unfortunately do not provide ranks that are identical with the SHARAD. Even the confidence in risk-adjusted excess is coincidently identical to the SHARAD. For example, if the data history for Fund 7 is changed to 20 years, then this is no longer the case (even though the two measures are highly correlated). Hence, in situations where there is a difference in performance history, other measures will need to be replaced by the SHARAD.

- Inspite of the relative superiority of the $M^3$ over the $M^2$/Sharpe or Information Ratio, rankings can be reversed when adjustment is made for time. Fund 4 has a lower rankings with SHARAD than with $M^3$ excess, whereas Fund 7 improves its rank.
• The fund with the best SHARAD, Fund 3, does not have the highest raw excess, or Sharpe, or even the longest data history. This reflects the complex inter-relationships between return, volatility, correlation and time.

• In this setup, the information ratio seems to have the greatest correlation of ranks with other risk adjusted performance measures. One can see why the IR correlates highly with the confidence in skill and as a result with $M^3$. This may explain its reasonably strong correlation with the SHARAD measure. While this was not the case in the tables in the Appendix where the data histories are identical, the relatively high correlation of ranks may also explain some of the attractive feature of the information ratio as highlighted in Gupta, Prajogi and Stubbs (1999).

Table 4 provides a comparison of the various methods in achieving the preferred objectives of the ideal measure. One can see that the SHARAD measure meets most of the criteria one would use to identify a useful risk-adjusted performance measure.

**Caveats and Extensions**

This is by no means the final word on performance measures and one issue that worries clients is whether increasing the size of assets under management corrupts the investment process beyond the impact of transactions costs. The SHARAD measure ignores the impact of changes in size and relegates any such impact to the confidence in skill measure. However, this measure can also be used for multiple fund portfolios and hence lends itself to use by institutional investors who normally hire more than one manager. The key issue is whether the history is an accurate predictor of the future and this is an open question as changes to the investment process or personnel can affect future performance. One issue that the user will need to deal with is whether firms/funds have skill inherent in their processes or whether the skill lies with the individuals
managing the particular portfolio. This information will be revealed by examining the consistency in skill through portfolio manager changes. We anticipate that if these measures are run periodically that any deterioration in either skill of the manager or inherent in the investment process will be captured by the way in which returns are generated.

One interesting issue that is raised is whether this measure is restricted to comparisons within an asset class or can be used across a portfolio of asset classes (and by construct managers). Dowd (2000) examines this in more detail and provides an enhanced Sharpe ratio to capture covariance of managers across asset classes. However, one possible use of the SHARAD measure in a portfolio context would be to consider a mix of managers across asset classes versus a benchmark that is a complex combination of multiple asset classes. Muralidhar (2001) demonstrates how optimal multiple manager portfolios can be selected using the $M^3$ measure in a single asset class context. The extension to multiple asset classes for the SHARAD is feasible, but complex for investors to analyze as results will reflect many new parameters (i.e., correlations across asset classes and managers across asset classes). This will be considered in future research.

**CONCLUSIONS**

This paper sought to address a shortcoming of the existing range of risk-adjusted performance measures, namely their inability to rank funds on an equivalent basis if the funds had different lengths of performance histories. In addition, it has been argued that a good risk-adjusted performance measure should (a) be expressed in basis points so that clients can relate to the measure; (b) provide some instruction on portfolio construction; and (c) adjust for the major market risk (i.e., tracking error through its components) an investor is exposed to; and (d) provide rankings that are consistent with skill. We introduce a new measure called the SHARAD measure, which is in effect, a probability adjusted risk-adjusted performance measure that
achieves the first three and directly incorporates the skill evaluation as part of the measure. There
will always be difficulties in distinguishing actual skill from presumed skill, and further ensuring
consistency in skill, but the periodic application of this measure will demonstrate whether ex-ante
measures are good predictors of true skill. The SHARAD measure is an attempt to achieve these
lofty goals and provides rankings that are different from the others. This paper demonstrates that
the final ranking is a complex evaluation of return, variance, correlation with the benchmark and
time and hence simpler measures, while easy to calculate, may provide investors with incorrect
recommendations.
## Tables

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Fund 1</th>
<th>Fund 2</th>
<th>Fund 3</th>
<th>Fund 4</th>
<th>Fund 5</th>
<th>Fund 6</th>
<th>Fund 7</th>
<th>Fund 8</th>
<th>Fund 9</th>
<th>Fund 10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fund Return</strong></td>
<td>25.16%</td>
<td>23.43%</td>
<td>22.46%</td>
<td>21.93%</td>
<td>21.39%</td>
<td>21.34%</td>
<td>20.97%</td>
<td>20.76%</td>
<td>19.22%</td>
<td>18.21%</td>
</tr>
<tr>
<td><strong>Fund Std. Dev.</strong></td>
<td>25.55%</td>
<td>25.36%</td>
<td>15.61%</td>
<td>22.09%</td>
<td>15.98%</td>
<td>13.71%</td>
<td>18.98%</td>
<td>14.69%</td>
<td>19.90%</td>
<td>16.52%</td>
</tr>
<tr>
<td><strong>BM Return</strong></td>
<td>19.18%</td>
<td>15.66%</td>
<td>18.76%</td>
<td>17.24%</td>
<td>19.91%</td>
<td>17.99%</td>
<td>18.26%</td>
<td>18.60%</td>
<td>17.53%</td>
<td>17.24%</td>
</tr>
<tr>
<td><strong>BM Std. Dev.</strong></td>
<td>13.36%</td>
<td>14.87%</td>
<td>15.00%</td>
<td>14.99%</td>
<td>14.90%</td>
<td>13.06%</td>
<td>15.01%</td>
<td>13.11%</td>
<td>14.83%</td>
<td>14.99%</td>
</tr>
<tr>
<td><strong>Risk Free Rate</strong></td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
<td>0.70</td>
<td>0.79</td>
<td>0.86</td>
<td>0.83</td>
<td>0.74</td>
<td>0.83</td>
<td>0.86</td>
<td>0.92</td>
<td>0.89</td>
<td>0.91</td>
</tr>
<tr>
<td><strong>Target Correlation</strong></td>
<td>0.72</td>
<td>0.77</td>
<td>0.78</td>
<td>0.78</td>
<td>0.77</td>
<td>0.71</td>
<td>0.78</td>
<td>0.71</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Tracking Error</strong></td>
<td>12.19%</td>
<td>10.48%</td>
<td>0.61%</td>
<td>7.10%</td>
<td>1.08%</td>
<td>0.65%</td>
<td>3.98%</td>
<td>1.58%</td>
<td>5.07%</td>
<td>1.53%</td>
</tr>
<tr>
<td><strong>a-Allocation to Fund</strong></td>
<td>49%</td>
<td>63%</td>
<td>147%</td>
<td>88%</td>
<td>84%</td>
<td>151%</td>
<td>117%</td>
<td>294%</td>
<td>140%</td>
<td>201%</td>
</tr>
<tr>
<td><strong>b-Allocation to Benchmark</strong></td>
<td>7%</td>
<td>-8%</td>
<td>-53%</td>
<td>-30%</td>
<td>11%</td>
<td>-50%</td>
<td>-49%</td>
<td>-23%</td>
<td>-89%</td>
<td>-123%</td>
</tr>
<tr>
<td><strong>c-Allocation to Risk-Free</strong></td>
<td>44%</td>
<td>45%</td>
<td>7%</td>
<td>42%</td>
<td>6%</td>
<td>9%</td>
<td>32%</td>
<td>39%</td>
<td>49%</td>
<td>22%</td>
</tr>
<tr>
<td><strong>Time Horizon (Yrs)</strong></td>
<td>11.75</td>
<td>11.92</td>
<td>15.25</td>
<td>20.00</td>
<td>15.58</td>
<td>11.50</td>
<td>14.33</td>
<td>11.67</td>
<td>15.75</td>
<td>20.00</td>
</tr>
</tbody>
</table>

### Returns/Ratios

<table>
<thead>
<tr>
<th></th>
<th>Raw excess</th>
<th>Sharpe</th>
<th>M²</th>
<th>Information Ratio</th>
<th>r(CAP) - 10% Tracking Error</th>
<th>M² excess</th>
<th>Confidence in Skill-Raw</th>
<th>Confidence in Skill-M²</th>
<th>SHARAD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fund Return</strong></td>
<td>5.98%</td>
<td>7.77%</td>
<td>3.70%</td>
<td>4.69%</td>
<td>1.48%</td>
<td>3.34%</td>
<td>2.72%</td>
<td>2.16%</td>
<td>1.69%</td>
</tr>
<tr>
<td><strong>Sharpe</strong></td>
<td>0.79</td>
<td>0.73</td>
<td>1.12</td>
<td>0.77</td>
<td>1.03</td>
<td>1.19</td>
<td>0.84</td>
<td>1.07</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>M²</strong></td>
<td>15.54%</td>
<td>15.81%</td>
<td>21.78%</td>
<td>16.49%</td>
<td>20.28%</td>
<td>20.56%</td>
<td>17.63%</td>
<td>19.06%</td>
<td>15.60%</td>
</tr>
<tr>
<td><strong>Information Ratio</strong></td>
<td>0.49</td>
<td>0.74</td>
<td>6.08</td>
<td>0.66</td>
<td>1.36</td>
<td>5.13</td>
<td>0.63</td>
<td>1.36</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>r(CAP) - 10% Tracking Error</strong></td>
<td>15.84%</td>
<td>15.80%</td>
<td>23.23%</td>
<td>16.19%</td>
<td>20.30%</td>
<td>21.80%</td>
<td>17.20%</td>
<td>19.69%</td>
<td>13.72%</td>
</tr>
<tr>
<td><strong>M² excess</strong></td>
<td>-3.34%</td>
<td>0.14%</td>
<td>4.47%</td>
<td>-1.06%</td>
<td>0.39%</td>
<td>3.81%</td>
<td>-1.06%</td>
<td>1.08%</td>
<td>-3.81%</td>
</tr>
<tr>
<td><strong>Confidence in Skill-Raw</strong></td>
<td>48.68%</td>
<td>79.74%</td>
<td>100.00%</td>
<td>76.33%</td>
<td>99.97%</td>
<td>100.00%</td>
<td>75.58%</td>
<td>99.84%</td>
<td>25.44%</td>
</tr>
<tr>
<td><strong>Confidence in Skill-M²</strong></td>
<td>12.63%</td>
<td>51.88%</td>
<td>95.97%</td>
<td>31.84%</td>
<td>56.06%</td>
<td>90.18%</td>
<td>34.42%</td>
<td>64.43%</td>
<td>63.53%</td>
</tr>
<tr>
<td><strong>SHARAD</strong></td>
<td>2.00%</td>
<td>8.20%</td>
<td>22.29%</td>
<td>5.16%</td>
<td>11.38%</td>
<td>19.66%</td>
<td>5.92%</td>
<td>12.68%</td>
<td>0.90%</td>
</tr>
</tbody>
</table>

### Rank

<table>
<thead>
<tr>
<th></th>
<th>Raw Return</th>
<th>Sharpe</th>
<th>Information Ratio</th>
<th>r(CAP) - 10% Tracking Error</th>
<th>M² excess</th>
<th>Confidence-M²</th>
<th>SHARAD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rank</strong></td>
<td>Raw Return</td>
<td>Sharpe</td>
<td>Information Ratio</td>
<td>r(CAP) - 10% Tracking Error</td>
<td>M² excess</td>
<td>Confidence-M²</td>
<td>SHARAD</td>
</tr>
<tr>
<td><strong>Raw Return</strong></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sharpe</strong></td>
<td>0.85</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information Ratio</strong></td>
<td>0.19</td>
<td>0.10</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>r(CAP) - 10% Tracking Error</strong></td>
<td>-0.02</td>
<td>-0.21</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M² excess</strong></td>
<td>0.12</td>
<td>0.01</td>
<td>0.81</td>
<td></td>
<td>0.95</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td><strong>Confidence in Skill-M²</strong></td>
<td>0.08</td>
<td>0.03</td>
<td>0.84</td>
<td></td>
<td>0.96</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td><strong>SHARAD</strong></td>
<td>0.08</td>
<td>0.03</td>
<td>0.84</td>
<td></td>
<td>0.96</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td><strong>Time Horizon (Yrs)</strong></td>
<td>8.00</td>
<td>5.00</td>
<td>1.00</td>
<td></td>
<td>0.97</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Raw Return</th>
<th>Raw excess</th>
<th>Sharpe</th>
<th>Information Ratio</th>
<th>r(CAP) - 10% Tracking Error</th>
<th>M² excess</th>
<th>Confidence-M²</th>
<th>SHARAD</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correlation of ranks</strong></td>
<td>Raw Return</td>
<td>Raw excess</td>
<td>Sharpe</td>
<td>Information Ratio</td>
<td>r(CAP) - 10% Tracking Error</td>
<td>M² excess</td>
<td>Confidence-M²</td>
<td>SHARAD</td>
<td>Time</td>
</tr>
<tr>
<td><strong>Raw Return</strong></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Raw excess</strong></td>
<td>0.85</td>
<td>-0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sharpe</strong></td>
<td>-0.05</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information Ratio</strong></td>
<td>0.19</td>
<td>0.10</td>
<td>0.83</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>r(CAP) - 10% Tracking Error</strong></td>
<td>-0.02</td>
<td>-0.21</td>
<td>0.96</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M² excess</strong></td>
<td>0.12</td>
<td>0.01</td>
<td>0.81</td>
<td>0.95</td>
<td>0.83</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Confidence in Skill-M²</strong></td>
<td>0.08</td>
<td>0.03</td>
<td>0.84</td>
<td>0.96</td>
<td>0.85</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SHARAD</strong></td>
<td>0.08</td>
<td>0.03</td>
<td>0.84</td>
<td>0.96</td>
<td>0.85</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time Horizon (Yrs)</strong></td>
<td>-0.29</td>
<td>-0.37</td>
<td>-0.47</td>
<td>-0.47</td>
<td>-0.29</td>
<td>-0.36</td>
<td>-0.42</td>
<td>-0.42</td>
<td>1.00</td>
</tr>
<tr>
<td>Measure</td>
<td>Information Ratio</td>
<td>Sharpe Ratio</td>
<td>M²</td>
<td>M³</td>
<td>Confidence in Skill</td>
<td>SHARAD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>----</td>
<td>----</td>
<td>---------------------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance in %</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank based on Skill</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes (for same History)</td>
<td>Yes (for same History)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjust for Data History</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidance on Portfolio Selection</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control for Varying Risk of Manager</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 1

THE LUCK OR SKILL TECHNIQUE.

Assume that one wants to determine how a fund has performed relative to a benchmark using a data series that extends over many periods of time. In order to establish luck from skill, Ambarish and Seigel (1996) assume that one is evaluating a portfolio (P) versus a benchmark (B). One would expect that a measure of relative risk (per unit of time) would be the variance per unit of time of \( \frac{dP}{P} - \frac{dB}{B} \) or variance per unit of time of \([r(P) - r(B)]\). Assume that P and B follow a generalized Weiner process, so that:

\[
\begin{align*}
\frac{dP}{P} &= \mu_p \, dh + \sigma_p \, dz_p \\
\frac{dB}{B} &= \mu_B \, dh + \sigma_B \, dz_B
\end{align*}
\]  

(A.1.1)

where (\( \mu_p, \sigma_p \)) and (\( \mu_B, \sigma_B \)) are instantaneous mean and volatility parameters of the portfolio and benchmark, respectively. The parameter \( dh \) is the change with respect to time and \( dz_i = \varepsilon(s)(dh)^{1/2} \) is the increment on standard Brownian Motion for \( i.e. (h) \) has zero mean and unit standard deviation. \( E[dz_i] = 0 \) and \( E[(dz_i)^2] = dh \).

If \( \rho_{P,B} \) is the coefficient of correlation between \( P \) and \( B \), then the dynamics of \( R(h) = \frac{P(h)}{B(h)} \) can be discovered by applying Itô's Lemma such that

\[
\frac{dR}{R} = (\mu_p - \mu_B + \sigma_B^2 - \sigma_p \sigma_B \rho_{P,B})dh + \sigma_p dz_p - \sigma_B dz_B
\]  

(A.1.2)
Now define the stochastic variable $dw$ such that
\[
\sigma_R dw = \sigma_p dz_p - \sigma_B dz_B
\]  
(A.1.3)

where
\[
\sigma_R^2 = \sigma_B^2 + \sigma_p^2 - 2\sigma_P \sigma_B \rho_{P,B}
\]  
(A.1.4)

which is the square of the tracking error ($TE(P)$) of portfolio $P$ versus the benchmark, $B$.

Then, as in any simple Brownian motion, $R(s)$ can be defined in the following way:
\[
R(h) = R(0) \exp[\sigma_R \sqrt{h} \cdot \exp\left(\left(\frac{\mu_p - \sigma_p^2}{2}\right) - \left(\frac{\mu_B - \sigma_B^2}{2}\right)\right) \cdot h]
\]  
(A.1.5)

where $\epsilon$ is the standard normal variable, and $\exp[]$ stands for exponential. The first exponential term in (A.1.5) is the noise (or luck) and the second exponential term is the signal (or skill). Therefore, for skill embedded in managing portfolio $P$ to dominate noise associated with returns from luck, it must be the case that the history of the fund (or number of data observations), $S$, satisfies the following equation; namely that

\[
H > \frac{S^2 \left(\sigma_p^2 - 2\rho_{P,B}\sigma_p\sigma_B + \sigma_B^2\right)}{\left(\frac{\mu_p}{\sigma_p^2} \cdot \frac{\sigma_p}{2}\right) \left(\frac{\mu_B}{\sigma_B^2} \cdot \frac{\sigma_B}{2}\right)^2}
\]  
(A.1.6)

Where $S$ is the number of standard deviations for a given confidence level.
APPENDIX 2

DETERMINING a AND b IN THE CAPEQUATION.

Consider \( r(CAP) \) for mutual fund 1,

\[
r(CAP-1) = ar(1) + (1-a-b)r(F) + br(B)
\]  
(A.2.1)

The investors wants to select a, and b, such that

\[
\sigma_{CAP-1}^2 = \sigma_B^2
\]  
(A.2.2)

and

\[
TE(CAP) = TE(target)
\]  
(A.2.3)

which can be re-written as

\[
\rho_{CAP-1,B} = \rho_{T,B}
\]  
(A.2.4)

Expanding on (A.2.2),

\[
\sigma_{CAP-1}^2 = \sigma_B^2 = a^2\sigma_1^2 + b^2\sigma_B^2 + 2ab\sigma_1\sigma_B\rho_{1,B}
\]  
(A.2.5)

Also, the covariance of \( r(CAP-1) \) and the benchmark \( B \) is
\[ \rho_{CAP-1,B} \sigma_{CAP-1} \sigma_B = a \sigma_I \sigma_B \rho_{1,B} + b \sigma_B^2 \]  
(A.2.6)

Using (A.2.2) and (A.2.4), equation (A.2.6) can be re-written as

\[ \rho_{T,B} \sigma_B^2 = a \sigma_I \sigma_B \rho_{1,B} + b \sigma_B^2 \]  
(A.2.7)

\[ b = \rho_{T,B} - a^* \frac{\sigma_I}{\sigma_B} \rho_{1,B} \]  
(A.2.8)

Substituting (A.2.8) in (A.2.5)

\[ a^2 \sigma_I^2 + (\rho_{T,B} - a^* \frac{\sigma_I}{\sigma_B} \rho_{1,B})^2 \sigma_B^2 + 2a(\rho_{T,B} - a^* \frac{\sigma_I}{\sigma_B} \rho_{1,B}) \sigma_I \sigma_B \rho_{1,B} = \sigma_B^2 \]  
(A.2.9)

Solving for a,

\[ a = \pm \sqrt{\frac{\sigma_B^2 (1 - \rho_{T,B}^2)}{\sigma_I^2 (1 - \rho_{1,B}^2)}} \]  
(A.2.10)

Substituting for a in A.2.8

\[ b = \rho_{T,B} - \rho_{1,B} \sqrt{\frac{(1 - \rho_{T,B}^2)}{(1 - \rho_{1,B}^2)}} \]  
(A.2.11)
## Table A.1

<table>
<thead>
<tr>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
<th>Fund</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Fund Return</td>
<td>22.54%</td>
<td>25.63%</td>
<td>21.90%</td>
<td>24.08%</td>
<td>21.00%</td>
<td>21.61%</td>
<td>22.40%</td>
<td>20.56%</td>
<td>22.47%</td>
</tr>
<tr>
<td>Fund Std. Dev.</td>
<td>25.50%</td>
<td>24.93%</td>
<td>13.84%</td>
<td>21.33%</td>
<td>17.28%</td>
<td>14.37%</td>
<td>18.73%</td>
<td>14.79%</td>
<td>17.41%</td>
</tr>
<tr>
<td>BM Return</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
<td>17.09%</td>
</tr>
<tr>
<td>BM Std. Dev.</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
<td>13.27%</td>
</tr>
<tr>
<td>Risk Free Rate</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.00%</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.69</td>
<td>0.75</td>
<td>0.83</td>
<td>0.80</td>
<td>0.79</td>
<td>0.97</td>
<td>0.83</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>Target Correlation</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>Tracking Error</td>
<td>18.95%</td>
<td>17.33%</td>
<td>7.84%</td>
<td>13.33%</td>
<td>10.59%</td>
<td>3.79%</td>
<td>10.65%</td>
<td>5.32%</td>
<td>7.48%</td>
</tr>
<tr>
<td>a-Allocation to Fund</td>
<td>48%</td>
<td>60%</td>
<td>153%</td>
<td>85%</td>
<td>100%</td>
<td>663%</td>
<td>112%</td>
<td>342%</td>
<td>231%</td>
</tr>
<tr>
<td>b-Allocation to Benchmark</td>
<td>7%</td>
<td>-13%</td>
<td>-61%</td>
<td>-37%</td>
<td>-31%</td>
<td>-622%</td>
<td>-60%</td>
<td>-285%</td>
<td>-206%</td>
</tr>
<tr>
<td>c-Allocation to Risk-Free</td>
<td>44%</td>
<td>53%</td>
<td>8%</td>
<td>53%</td>
<td>31%</td>
<td>59%</td>
<td>48%</td>
<td>43%</td>
<td>75%</td>
</tr>
<tr>
<td>Time Horizon (Yrs)</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
</tr>
</tbody>
</table>

### Returns/Ratios

- **Raw excess**: 5.45% 8.54% 4.81% 6.99% 3.91% 4.52% 5.31% 3.47% 5.38% 3.66%
- **Sharpe**: 0.69 0.83 1.22 0.89 0.93 1.16 0.93 1.05 1.00 1.13
- **M³**: 14.13% 15.98% 21.19% 16.87% 17.28% 20.34% 17.32% 18.95% 18.31% 19.94%
- **Information Ratio**: 0.29 0.49 0.61 0.52 0.37 1.19 0.50 0.65 0.72 0.95
- **M³ excess**: -2.71% -1.32% 6.33% -0.47% 0.11% 22.86% 0.15% 6.71% 3.35% 5.82%
- **Confidence in Skill**: 45.58% 68.92% 96.57% 75.86% 71.47% 99.95% 77.78% 94.61% 91.91% 93.98%
- **Confidence in Skill-M³**: 10.13% 19.59% 33.82% 97.73% 44.15% 44.15% 51.33% 51.33% 51.33% 51.33%
- **SHARAD**: 2.82% 5.33% 22.88% 7.34% 8.83% 39.95% 8.95% 23.40% 17.48% 22.16%

### Table A.2

<table>
<thead>
<tr>
<th>Rank</th>
<th>Raw Return</th>
<th>Raw excess</th>
<th>Sharpe</th>
<th>Information Ratio</th>
<th>r(CAP) - 10% Tracking Error</th>
<th>M³ excess</th>
<th>Confidence in Skill-M³</th>
<th>SHARAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>8</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>9</td>
<td>3</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>9</td>
<td>3</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table A.3

<table>
<thead>
<tr>
<th>Correlation of ranks</th>
<th>Raw Return</th>
<th>Raw excess</th>
<th>Sharpe</th>
<th>Information Ratio</th>
<th>r(CAP) - 10% Tracking Error</th>
<th>M³ excess</th>
<th>Confidence in Skill-M³</th>
<th>SHARAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Return</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Raw excess</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Sharpe</td>
<td>-0.66</td>
<td>-0.66</td>
<td>1.00</td>
<td>0.94</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Information Ratio</td>
<td>-0.38</td>
<td>-0.38</td>
<td>0.79</td>
<td>1.00</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>r(CAP) - 10% Tracking Error</td>
<td>-0.73</td>
<td>-0.73</td>
<td>0.94</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>M³ excess</td>
<td>-0.73</td>
<td>-0.73</td>
<td>0.94</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Confidence in Skill-M³</td>
<td>-0.73</td>
<td>-0.73</td>
<td>0.94</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>SHARAD</td>
<td>-0.73</td>
<td>-0.73</td>
<td>0.94</td>
<td>0.87</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Bibliography


i Thanks to Kurtay Ogunc for planting the seeds of this research with his probing questions and to Franco Modigliani, Srinivas Pullavarti, Sudhir Krishnamurthi and Lester Seigel for their invaluable assistance. Also, thanks to Joseph Silvera for the many discussions on these topics. Finally, thanks to two anonymous referees for comments.


iii Where the risk-free rate has no zero variance, the denominator is that standard deviation of the portfolio.

iv One of the problems of Philips and Yashchin (1999) is that the user is required to specify an information ratio above which funds would be rated good. In the technique employed here, no such classification is required.

v This result is from outperformance engendered through 13.2% basis points of tracking error, where the benchmark standard deviation = 15%, the actual standard deviation = 25% and the correlation between the two was 0.9.

vi Sharpe (1994).

vii In addition, when benchmark returns are negative (e.g., in currency mandates), the M² measure will incorrectly rank underperforming funds with high volatility are preferred to funds with low volatility. This is a quirk of the method as it is generally assumed that benchmarks have positive returns.

viii These are heroic assumptions to say the least. Some forecast needs to be made on expected outperformance, variability of performance to achieve this outperformance and correlations between portfolio and benchmark returns. Historical performance is one way of making forecasts, but the M-3 measure is independent of the forecasting technique. In addition, one must believe that markets are inefficient to conduct such analyses.

ix This measure is independent of the level of tracking error and hence is applicable across all tracking error targets.

x This may change with the development of exchange traded funds (ETFs) on active portfolios. In some cases it may be difficult to short the benchmark as well and then b will need to constrained to being greater than or equal to zero. This would not change the analysis of the measure. In general though, most benchmarks can be shorted either through their futures contract or through a swap.

xi We obviously take a few liberties here with the use of S by assuming that S is an equality rather than an inequality. In addition, the rankings need not be exactly identical with the skill rankings as demonstrated later.

xii This translates into different target correlations for each fund because of the differences in benchmark standard deviation over their respective histories. If two funds had an identical time History, the target correlation would be the same for both.

xiii Thanks to Srinivas Pullavarti for this observation.