Financial Analysts Journal
Volume 64 • Number 3
©2008, CFA Institute

Does the Measure Matter in the Mutual Fund Industry?

Martin Eling

A frequent comment is that investment funds with a nonnormal return distribution cannot be adequately evaluated by using the classic Sharpe ratio. Research on hedge fund data that compared the Sharpe ratio with other performance measures, however, found virtually identical rank ordering by the various measures. The study reported here analyzed a dataset of 38,954 funds investing in seven asset classes over 1996–2005 and found that the previous result is true not only for hedge funds but also for mutual funds investing in stocks, bonds, real estate, funds of hedge funds, commodity trading advisers, and commodity pool operators. In short, choosing a performance measure is not critical to fund evaluation and the Sharpe ratio is generally adequate.

The most widely known risk-adjusted performance measure is the Sharpe ratio. It measures the relationship between the risk premium (mean excess returns) and the standard deviation of the returns generated by the fund, portfolio, or asset being measured (Sharpe 1966). Hedge funds and other alternative investments, however, are prone to generating returns that have a nonnormal distribution. For this reason, Brooks and Kat (2002), Mahdavi (2004), Sharma (2004), and Sharpe (2007), among others, have claimed that these funds cannot be adequately evaluated by using the Sharpe ratio. This problem motivated the development of numerous new performance measures, including Omega, the Sortino ratio, the Calmar ratio, and the modified Sharpe ratio, all of which are currently being debated as measures of performance in the hedge fund literature (for an overview, see Lhabitant 2004).

In a recent study, Eling and Schuhmacher (2007) compared these new performance measures with the Sharpe ratio by using the data of 2,763 hedge funds. Despite hedge fund returns' significant deviation from a normal distribution, the Sharpe ratio and the other measures in their study resulted in virtually identical rank ordering for the hedge funds. Eling and Schuhmacher analyzed only hedge funds, however, and thus did not consider whether this result is also true for funds investing in other asset classes.

Martin Eling is at the University of St. Gallen, Switzerland, and currently visiting professor at the University of Wisconsin–Madison.

The aim of the study reported here was to address this issue. Combining two large datasets, I analyzed 38,954 investment funds concentrated in a large number of asset classes, including stocks, bonds, real estate, hedge funds, funds of hedge funds, commodity trading advisers (CTAs), and commodity pool operators (CPOs).

Performance Measures

In risk-adjusted performance measurement, the fund return is adjusted in relation to a suitable risk measure. In investment fund analysis, the Sharpe ratio is often chosen to be the performance measure and the analyst compares the Sharpe ratio of the fund of interest with the Sharpe ratios of other funds or market indices (see, for example, Ackermann, McEnally, and Ravenscraft 1999; Schneeweis, Kazemi, and Martin 2002).

In the context of hedge funds, use of the Sharpe ratio has been strongly criticized because hedge fund returns do not exhibit a normal distribution.[1] For example, use of derivative instruments results in an asymmetrical return distribution, and fat tails, which leads to the danger that use of standard risk and performance measures will underestimate risk and overestimate risk-adjusted performance.[2] To avoid this problem, some researchers recommend the use of newer performance measures that illustrate the risk of loss (Pedersen and Rudholm-Alfvin 2003; Lhabitant 2004).

The newer performance measures differ from the Sharpe ratio in that standard deviation is replaced by an alternative risk measure. The alternative risk measures considered in this study are the lower partial moments (LPMs) of orders 1, 2, and 3; three variants based on the drawdown; and three value-at-risk (VaR) approaches. The risk measures, the performance measures, and references that contain more information about all the measures are in **Exhibit 1** (for a brief overview of all measures, see Eling and Schuhmacher 2007).

The standard deviation involves both positive and negative deviations of return from its expected value, which is not the general understanding of risk. In contrast, LPMs consider only negative deviations of returns from a minimal acceptable return, the situation that most investors would like to avoid. Thus, LPMs might seem to be the more appropriate measure of risk. Using the lower partial moments of orders 1, 2, or 3, one can define the Omega, Sortino ratio, and Kappa 3 performance measures. Excess return is used as a return measure for these three measures, not in relation to the risk-free interest rate but, rather, in relation to the given minimal acceptable return, τ. Of course, other measures could be used for return; an example is the higher partial moments (HPMs) of order 1, as is the case with the upside potential ratio. With this measure, the first-order HPM is combined with the second-order LPM.

The drawdown of a fund measures the loss incurred over a given investment period. As Exhibit 1 shows, the Calmar, Sterling, and Burke ratios use as risk measures, respectively, the maximum drawdown, an average of a certain number of drawdowns, and a type of standard deviation of a number of the largest drawdowns.

Exhibit 1. Performance Measures

Risk Measure	Performance Measure	Reference
Standard deviation	$Sharpe\ ratio_i = \left(r_i^a - r_f\right)/\sigma_i$	Sharpe (1966)
Lower partial moment of order		
1	$Omega_i = \left(r_i^a - \tau\right)/LPM_{i1}\left(\tau\right) + 1$	Shadwick and Keating (2002)
2	$Sortino\ ratio_i = \left(r_i^a - \tau\right)/\sqrt[2]{LPM_{i2}\left(\tau\right)}$	Sortino and Van der Meer (1991)
2	$Upside\ potential\ ratio_i = HPM_{i1}\left(\tau\right)/\sqrt[2]{LPM_{i2}\left(\tau\right)}$	Sortino, Van der Meer, and Plantinga (1999)
3	$Kappa\ 3_i = \left(r_i^a - \tau\right)/\sqrt[3]{LPM_{i3}\left(\tau\right)}$	Kaplan and Knowles (2004)
Drawdown		
Maximum	$Calmar\ ratio_i = \left(r_i^a - r_f\right)/-D_{i1}$	Young (1991)
Average	$Sterling\ ratio_i = \left(r_i^a - r_f\right)/\left[(1/K)\sum_{k=1}^{K} - D_{ik}\right]$	Kestner (1996)
Standard deviation	$Burke\ ratio_i = \left(r_i^a - r_f\right)/\sqrt[2]{\sum_{k=1}^{K} D_{ik}^2}$	Burke (1994)
Value at risk		
Standard	$Excess\ return\ on\ value\ at\ risk_i = \left(r_i^a - r_f\right)/VaR_i$	Dowd (2000)
Conditional	$Conditional\ Sharpe\ ratio_i = \left(r_i^a - r_f\right)/CVaR_i$	Agarwal and Naik (2004)
Modified	$Modified\ Sharpe\ ratio_i = \left(r_i^a - r_f\right)/MVaR_i$	Gregoriou and Gueyie (2003)

Notes:

r_i^a = mean return, equal to $(1/T)\sum_{t=1}^{T} r_{it}$, with r_{it} as discrete return of fund i in month t ($t = 1, \ldots, T$) and T as number of months;

r_f = (constant) risk-free interest rate;

σ_i = standard deviation, equal to $\sqrt{(1/T-1)\sum_{t=1}^{T}\left(r_{it} - r_i^a\right)^2}$;

LPM_{in} = lower partial moment of order n, equal to $(1/T)\sum_{t=1}^{T}\max\left(\tau - r_{it}, 0\right)^n$, with τ as the minimum acceptable return;

HPM_{in} = higher partial moment of order n, equal to $(1/T)\sum_{t=1}^{T}\max\left(r_{it} - \tau, 0\right)^n$;

D_{ik} = drawdown of fund i;

K = number of drawdowns ($k = 1$: maximum drawdown; $k = 2$: second-largest drawdown; $k = 3$: third-largest drawdown; . . .);

VaR_i = value at risk, equal to $-(r_i^a + z_\alpha \cdot \sigma_i)$, with z_α the α-quantile of the standard normal distribution;

$CVaR_i$ = conditional value at risk, equal to $E(-r_{it} \mid r_{it} \leq -VaR_i)$;

$MVaR_i$ = modified value at risk, equal to $-\{r_i^a + \sigma_i\,[z_\alpha + (z_\alpha^2 - 1) \times S_i/6 + (z_\alpha^3 - 3z_\alpha) \times E_i/24 - (2z_\alpha^3 - 5z_\alpha) \times S_i^2/36]\}$, with S_i as skewness $\left[\text{equal to} \left(1/T\sum_{t=1}^{T}(r_{it} - r_i^a)^3\right)/\sigma_i^3\right]$ and E_i as excess kurtosis $\left\{\text{equal to} \left[1/T\sum_{t=1}^{T}\left(r_{it} - r_i^a\right)^4\right]/\sigma_i^4 - 3\right\}$.

VaR is the possible loss of an investment that is not exceeded with a given probability of 1 – α in a certain period. To take into account the distribution of returns below the VaR, the literature frequently considers expected loss if the VaR is exceeded. This consideration leads to the conditional VaR. To include skewness and kurtosis in computing VaR, the Cornish–Fisher expansion can be used, which leads to the modified VaR.

Data and Descriptive Statistics

In the empirical investigation, I used two large datasets. I obtained return data for 17,817 stock funds, 12,279 bond funds, and 751 real estate funds from the Thomson Datastream database.[3] Return data for 4,048 hedge funds, 1,949 funds of hedge funds, 1,076 CTAs, and 1,034 CPOs were taken from the Center for International Securities and Derivatives Markets (CISDM) database.[4] For all funds, I obtained monthly net-of-fee returns for the period from January 1996 to December 2005.

The return distributions of all funds are set out in **Table 1**, which provides the mean, the median, the standard deviation, the minimum, and the maximum of the first four moments of the return distribution (mean value, standard deviation, skewness, and excess kurtosis). For example, for the sample of the 17,817 stock funds, the standard deviation row in Panel A means that, for this period, across the 17,817 funds, the standard deviation had a mean of 4.70 percent (second column in that row) with a standard deviation of 2.43 percent (fourth column in that row).[5] Table 1 also provides the results of the Jacque–Bera test, which gives the portion of funds for which the assumption of normally distributed returns must be rejected at the 1 percent (5 percent) significance level, and the table shows the average correlation among the funds in each sample.

According to capital market theory, a functional relationship exists between the risk and the return of an investment: Taking higher risk is rewarded with a higher return. Using the mean value as a measure of return and the standard deviation as a measure of risk, I found that this relationship is generally true for the analyzed funds. For example, the asset class with the lowest risk (bonds) also provides the lowest return. When risk and return for the various asset classes are compared, hedge funds and funds of hedge funds appear to be the most attractive. Hedge funds provide the highest return but do not have the highest risk, and funds of hedge funds have a noticeably low standard deviation for the level of return generated, which might be a result of their higher degree of diversification in comparison with single funds.

Although some investors are primarily concerned with the central tendencies of the return distribution (mean value, standard deviation), others may care more about the extreme values. For these investors, skewness, kurtosis, and the results of the Jacque–Bera test are interesting to consider. Of particular note is the fact that the rejection rate for the Jacque–Bera test is high for hedge funds and for other asset classes. At a 1 percent significance level, the rejection rate varies from 19.84 percent for stock funds to 45.54 percent for real estate funds.

These strong deviations from normally distributed returns appear to imply that use of the Sharpe ratio is inappropriate not only for measuring hedge fund performance but also for measuring the performance of other asset classes.

Note also the high average correlations between the stock funds, 0.57, and the funds of funds, 0.55. In contrast, the sample of hedge funds is diverse, with an average correlation of 0.16.

Like other databases, the Datastream and CISDM databases suffer from survivorship bias. **Table 2** shows the attrition rate and survivorship bias for the analyzed funds. Survivorship bias was calculated as the difference in fund returns between all funds and the surviving funds.

The survivorship bias is 0.08 percent per month for hedge funds, which is comparable to other values found in the literature (e.g., Ackermann, McEnally, and Ravenscraft 1999; Liang 2000). The fact that the attrition rate and the survivorship bias are lower for traditional investments, such as stocks and bonds, than for commodity funds is well documented in the literature (see Liang 2000). In this sample, survivorship bias amounted to only 0.01 percent for stock funds and 0.0034 percent for bond funds; for CTAs and CPOs, the bias is, respectively, 0.10 percent and 0.09 percent.[6]

Performance Measurement

The findings reported in this section were generated by first using the performance measures discussed to determine the fund performance in each asset class. Next, for each performance measure, the funds were ranked on the basis of the measured values. Finally, the rank correlations between the performance measures were calculated.[7] For the LPM-based performance measures, I assumed that the minimal acceptable return was equal to the risk-free monthly interest rate [$\tau = r_f = 0.35$ percent, which is the interest rate on 10-year U.S. T-bonds as of 30 December 2005 (4.28 percent per year)]. For the Sterling and Burke ratios, the five largest drawdowns were considered ($N = 5$). The VaR-based performance measures were calculated by using a significance level of $\alpha = 0.05$.

Table 1. Return Distribution, 1996–2005

Time-Series Analysis by Fund Type	Cross-Sectional Analysis (across funds) Mean	Median	Standard Deviation	Minimum	Maximum
A. Stocks (17,817 funds)					
JB rejection: 19.84% (26.73%) at 1% (5%) level; average correlation between funds: 0.57					
Mean value (%)	0.53	0.49	1.19	–9.52	9.79
Standard deviation (%)	4.70	4.50	2.43	0.06	29.31
Skewness	–0.29	–0.32	0.76	–9.50	9.38
Excess kurtosis	0.76	0.11	4.35	–7.19	100.83
B. Bonds (12,279 funds)					
JB rejection: 25.60% (31.89%) at 1% (5%) level; average correlation between funds: 0.28					
Mean value (%)	0.37	0.34	0.58	–3.94	6.23
Standard deviation (%)	1.91	1.36	1.69	0.01	17.17
Skewness	–0.38	–0.32	1.04	–10.67	10.00
Excess kurtosis	1.53	0.20	7.11	–7.99	119.65
C. Real estate (751 funds)					
JB rejection: 45.54% (53.66%) at 1% (5%) level; average correlation between funds: 0.30					
Mean value (%)	0.90	0.86	0.84	–3.54	4.60
Standard deviation (%)	3.49	3.65	2.44	0.01	22.77
Skewness	–0.45	–0.53	1.20	–6.77	6.80
Excess kurtosis	2.44	1.06	6.45	–5.99	61.93
D. Hedge funds (4,048 funds)					
JB rejection: 37.67% (43.60%) at 1% (5%) level; average correlation between funds: 0.16					
Mean value (%)	0.97	0.86	1.48	–18.96	19.58
Standard deviation (%)	4.37	3.01	4.32	0.03	49.50
Skewness	0.01	0.00	1.15	–9.21	6.23
Excess kurtosis	2.45	0.91	6.13	–4.71	95.00
E. Funds of hedge funds (1,949 funds)					
JB rejection: 29.66% (34.89%) at 1% (5%) significance level; average correlation between funds: 0.55					
Mean value (%)	0.67	0.64	0.59	–7.95	11.89
Standard deviation (%)	1.94	1.43	1.71	0.06	21.75
Skewness	–0.26	–0.27	0.96	–8.00	6.60
Excess kurtosis	1.81	0.39	5.23	–3.99	79.08
F. CTAs (1,076 funds)					
JB rejection: 31.42% (37.95%) at 1% (5%) level; average correlation between funds: 0.13					
Mean value (%)	0.80	0.70	1.40	–7.96	11.16
Standard deviation (%)	5.89	4.78	4.46	0.01	35.16
Skewness	0.28	0.26	0.87	–3.96	5.87
Excess kurtosis	1.49	0.59	3.65	–7.14	40.75
G. CPOs (1,034 funds)					
JB rejection: 26.86% (32.45%) at 1% (5%) level; average correlation between funds: 0.23					
Mean value (%)	0.48	0.52	1.40	–13.87	14.68
Standard deviation (%)	5.16	4.48	3.72	0.07	35.45
Skewness	0.16	0.19	0.87	–4.92	4.61
Excess kurtosis	1.40	0.45	4.06	–6.90	33.59

Note: JB = Jacque–Bera.

Sources: Stock, bond, and real estate data are from Datastream; hedge fund, fund-of-hedge-fund, CTA, and CPO data are from the CISDM.

Table 2. Attrition Rate and Survivorship Bias by Fund Type, 1996–2005

	Attrition Rate (%)							Survivorship Bias (%)						
Year	Stocks	Bonds	Real Estate	Hedge Funds	Funds of Hedge Funds	CTAs	CPOs	Stocks	Bonds	Real Estate	Hedge Funds	Funds of Hedge Funds	CTAs	CPOs
1996	2.14	2.10	6.33	5.06	2.92	2.64	14.10	0.00	0.00	0.02	0.07	0.02	0.02	0.10
1997	2.72	3.97	1.05	10.79	5.01	15.62	20.22	0.00	–0.01	0.00	0.05	0.02	0.10	0.07
1998	4.17	5.43	0.00	13.81	8.01	22.26	18.82	0.00	0.00	0.00	0.17	0.03	0.25	0.16
1999	5.34	7.46	6.11	14.42	6.75	17.94	22.47	0.02	0.00	0.01	0.13	0.02	0.07	0.07
2000	3.52	5.22	6.03	11.00	7.49	19.42	15.58	0.00	0.01	–0.01	0.06	0.02	0.23	0.18
2001	6.17	9.09	5.60	12.54	8.46	13.82	14.62	0.02	0.01	0.00	0.10	0.03	0.07	0.07
2002	9.15	9.25	5.17	12.61	4.24	10.48	17.01	0.01	0.00	0.00	0.09	0.01	0.08	0.13
2003	9.17	8.29	6.78	12.21	4.85	11.81	19.63	0.03	0.00	0.02	0.05	0.01	0.08	0.02
2004	7.35	7.45	3.48	13.23	7.90	13.43	12.54	0.02	0.01	0.01	0.05	0.01	0.07	0.02
2005	7.90	8.95	6.30	14.80	7.88	18.74	18.96	0.03	0.01	0.01	0.05	0.01	0.06	0.07
Average	5.76	6.72	4.68	12.05	6.35	14.62	17.40	0.01	0.00	0.00	0.08	0.02	0.10	0.09

Table 3 shows the rank correlations of the Sharpe ratio in relation to the other performance measures. All the performance measures display a high rank correlation with respect to the Sharpe ratio. For hedge funds, the rank correlation coefficient with the Sharpe ratio varies between 0.94 (Sterling ratio) and 1.00 (excess return on VaR). On average, the rank correlation of the Sharpe ratio in relation to the other performance measures for hedge funds is 0.97. The correlation is also high between the Sharpe ratio, Omega, the Sortino ratio, Kappa 3, and the conditional Sharpe ratio (rank correlation greater than 0.98 in each case). These findings regarding hedge funds clearly confirm the results of Eling and Schuhmacher (2007).

I also found high rank correlations for all the other asset classes. The highest rank correlations are for the stock funds. On average for stock mutual funds, the rank correlation of the Sharpe ratio with the other performance measures is 0.99. Real estate has the lowest. Moreover, there appears to be a negative relationship between the rejection rate for the Jacque–Bera test and rank correlation: The asset class with the highest rejection rate in Table 1 (real estate) has the lowest rank correlation, and the asset class with the lowest rejection rate (stocks) has the highest rank correlation. Even when the returns of more than half of all funds deviate significantly from normally distributed returns (which is the case with real estate), however, only slight changes in rankings and rank correlation occur. I also found high rank correlations when comparing the new performance measures with each other; this comparison is available as supplemental material in the *FAJ* area of www.cfapubs.org.

Two test statistics can be used to check the significance of the rank correlations (see Eling and Schuhmacher 2007). The first is a standardized version of the Hotelling–Pabst statistic. In this test, the hypothesis of independence of the two related

Table 3. Rank Correlation Based on Various Performance Measures by Fund Type, 1996–2005

Performance Measure	Stocks	Bonds	Real Estate	Hedge Funds	Funds of Hedge Funds	CTAs	CPOs
Omega	1.00	0.99	0.98	0.99	0.99	1.00	1.00
Sortino ratio	1.00	1.00	0.98	0.99	0.99	1.00	1.00
Kappa 3	1.00	1.00	0.98	0.98	0.98	0.99	1.00
Upside potential ratio	0.98	0.97	0.95	0.96	0.95	0.95	0.96
Calmar ratio	0.99	0.95	0.96	0.95	0.93	0.98	0.98
Sterling ratio	0.98	0.95	0.94	0.94	0.91	0.96	0.97
Burke ratio	0.99	0.95	0.95	0.95	0.93	0.98	0.98
Excess return on VaR	0.97	0.95	0.96	1.00	0.99	0.97	0.99
Conditional Sharpe ratio	0.98	0.97	0.96	0.98	0.97	0.98	0.99
Modified Sharpe ratio	1.00	0.99	0.97	0.97	0.97	0.99	0.99
Average	0.99	0.97	0.96	0.97	0.96	0.98	0.99

rankings is checked for all correlation coefficients. At the significance level of $\alpha = 0.01$, however, in no case could I confirm the hypothesis of independence. Therefore, the hypothesis of independence of the measurement series must be rejected for all correlation coefficients presented in Table 3. In addition to testing whether the rankings are independent (in other words, the rank correlation is zero), I also checked the hypothesis that the rank correlation is smaller than a certain given rank correlation x. For this second test, I used the Fisher transformation and found that for a significance level of $\alpha = 0.01$, the hypothesis that the rank correlation is smaller than x is rejected for all x smaller than 0.896 (see Rees 1987, p. 383, for the t-statistic).

In conclusion, on the basis of the data studied, none of the new performance measures results in significant changes in the ranking of investment funds from rankings found by using the Sharpe ratio. Thus, which of the numerous measures is used to assess the performance of the various funds does not matter. Because the newer performance measures result in rankings that are practically the same, and thus give a similar assessment of the funds, use of the Sharpe ratio is justified, at least from a practical perspective.

Robustness

The various robustness tests I carried out are important because the findings presented in the previous section are valid only for the subject being examined, the time period considered, and several other given parameters (e.g., the minimal acceptable return).

From the robustness tests, I found that the main result to be robust with respect to

- variations in the investigation period (I broke down the full 1996–2005 period into five periods of two years each),
- variations of the exogenously fixed parameters (for the LPM-based measures, the minimal acceptable return was varied between 0 percent and 1 percent; for the drawdown-based measures, the number of drawdowns was varied between 1 and 10; for the VaR-based measures, the significance level was varied between 1 percent and 20 percent),
- an elimination of outliers (I eliminated between 1 and 10 of the highest and lowest returns from the time series), and
- a separate consideration of surviving funds and dissolved funds (to account for a potential survivorship bias in the results).

For all these tests, high rank correlations comparable to those presented in the previous section were found. **Table 4** shows the robustness results for stocks. Results for the other asset classes are available as supplemental material in the *FAJ* area of www.cfapubs.org; see Table S3.

As an additional robustness test, I split the samples of stock funds, bond funds, and hedge funds into various strategy groups (e.g., the total sample of 4,048 hedge funds contains diverse funds, so this sample was split into groups by hedge fund strategy, such as convertible arbitrage or distressed securities). Results are presented in **Table 5**. Again, I found high rank correlations among the performance measures.

Table 4. Results of Robustness Tests: Stock Mutual Funds

	Basic	Investigative Period					Parameters			Outliers	Bias	
Performance Measure	Correlation	1996–97	1998–99	2000–01	2002–03	2004–05	A	B	C	D	E	F
Omega	1.00	1.00	0.99	0.99	1.00	1.00	1.00	—	—	1.00	1.00	1.00
Sortino ratio	1.00	0.99	0.99	0.99	1.00	0.99	1.00	—	—	1.00	1.00	1.00
Kappa 3	1.00	0.99	0.98	0.99	1.00	0.99	1.00	—	—	1.00	1.00	1.00
Upside potential ratio	0.98	0.97	0.97	0.94	0.97	0.97	0.95	—	—	0.99	0.97	0.96
Calmar ratio	0.99	0.97	0.97	0.85	0.98	0.95	—	0.99	—	0.99	0.98	0.97
Sterling ratio	0.98	0.95	0.96	0.85	0.99	0.95	—	0.98	—	0.99	0.98	0.97
Burke ratio	0.99	0.96	0.97	0.88	0.99	0.97	—	0.99	—	0.99	0.98	0.97
Excess return on VaR	0.97	0.99	1.00	1.00	0.99	0.99	—	—	1.00	0.98	0.97	1.00
Conditional Sharpe ratio	0.98	0.99	0.97	0.97	0.99	0.96	—	—	0.98	0.97	0.98	0.99
Modified Sharpe ratio	1.00	0.99	0.99	0.99	1.00	0.99	—	—	0.99	0.99	0.99	0.99
Average	0.99	0.98	0.98	0.95	0.99	0.98	—	—	—	0.99	0.99	0.98

Notes: Column A represents variation of minimum acceptable return between 0 percent and 1 percent. Column B represents variation of number of drawdowns between 1 and 10. Column C represents variation of significance level between 1 percent and 20 percent. Column D represents elimination of 1 up to 10 of the highest and lowest returns. In these four tests, the rank correlations presented here are average values above the various robustness tests. Column E reports results for separate consideration of surviving funds (N = 13,039). Column F reports results for separate consideration of dissolved funds (N = 4,778).

Table 5. Rank Correlation: Other Measures with Sharpe Ratio for Different Strategy Groups, 1996–2005

Fund Strategy	No. of Funds	Omega	Sortino Ratio	Kappa 3	Upside Potential Ratio	Calmar Ratio	Sterling Ratio	Burke Ratio	Excess Return on VaR	Conditional Sharpe Ratio	Modified Sharpe Ratio	Average
A. Hedge funds (4,048)												
Convertible arbitrage	195	1.00	0.99	0.99	0.97	0.93	0.92	0.93	0.99	0.99	0.98	0.97
Distressed securities	141	0.99	0.98	0.97	0.95	0.91	0.91	0.92	1.00	0.98	0.96	0.96
Emerging markets	378	0.99	0.99	0.99	0.95	0.97	0.96	0.97	1.00	0.98	0.98	0.98
Equity long only	84	0.99	0.99	0.98	0.96	0.97	0.95	0.96	0.99	0.98	0.98	0.98
Equity long–short	1,553	0.99	0.99	0.98	0.96	0.95	0.93	0.95	1.00	0.98	0.97	0.97
Equity Market Neutral	231	1.00	1.00	0.99	0.97	0.97	0.97	0.97	1.00	0.99	0.99	0.98
Event-driven multistrategy	155	0.99	0.98	0.97	0.94	0.88	0.88	0.89	1.00	0.95	0.95	0.94
Fixed income	107	1.00	0.99	0.99	0.98	0.96	0.96	0.96	0.99	0.99	0.98	0.98
Fixed income—MBS	72	0.97	0.98	0.97	0.93	0.93	0.95	0.94	0.98	0.97	0.96	0.96
Fixed income arbitrage	158	0.99	0.99	0.99	0.95	0.96	0.96	0.96	0.99	0.98	0.98	0.97
Global macro	207	0.99	0.99	0.99	0.96	0.97	0.96	0.97	1.00	0.99	0.98	0.98
Merger arbitrage	120	0.99	0.99	0.98	0.96	0.94	0.93	0.94	0.99	0.94	0.97	0.96
Multistrategy	72	0.97	0.95	0.94	0.92	0.84	0.84	0.84	1.00	0.96	0.89	0.91
No strategy	59	1.00	0.99	0.99	0.97	0.97	0.96	0.97	0.99	0.97	0.99	0.98
Other[a]	60	0.99	0.98	0.98	0.98	0.92	0.92	0.92	0.93	0.94	0.93	0.95
Relative value multistrategy	72	0.98	0.95	0.93	0.92	0.90	0.92	0.91	0.97	0.95	0.93	0.94
Sector	342	0.99	0.98	0.98	0.94	0.95	0.92	0.95	1.00	0.98	0.95	0.96
Short bias	42	1.00	1.00	1.00	0.96	0.99	0.98	0.99	1.00	0.99	1.00	0.99
All	4,048	0.99	0.99	0.98	0.96	0.95	0.94	0.95	1.00	0.98	0.97	0.97
B. Stock funds (17,817 funds)												
Fund investment style												
Value	1,944	1.00	1.00	1.00	0.97	0.98	0.98	0.99	0.95	0.98	1.00	0.98
Growth	3,310	1.00	1.00	1.00	0.97	0.99	0.99	0.99	0.98	0.98	1.00	0.99
Other	12,563	1.00	1.00	1.00	0.97	0.98	0.98	0.98	0.97	0.98	1.00	0.98
Market cap												
Large	1,414	1.00	1.00	1.00	0.97	0.98	0.99	0.99	0.90	0.90	1.00	0.97
Small	2,026	1.00	1.00	1.00	0.97	0.98	0.98	0.98	0.96	0.99	1.00	0.99
Other	14,377	1.00	1.00	1.00	0.97	0.98	0.98	0.98	0.98	0.99	1.00	0.99
Currency												
US$	7,242	1.00	1.00	1.00	0.97	0.99	0.98	0.99	1.00	1.00	1.00	0.99
Euro	5,156	1.00	1.00	1.00	0.98	0.99	0.99	0.99	0.92	0.97	1.00	0.98
Other	5,419	1.00	1.00	1.00	0.97	0.99	0.98	0.99	0.98	0.97	0.99	0.99
All	17,817	1.00	1.00	1.00	0.98	0.99	0.98	0.99	0.97	0.98	1.00	0.99

(continued)

Table 5. Rank Correlation: Other Measures with Sharpe Ratio for Different Strategy Groups, 1996–2005 (continued)

Fund Strategy	No. of Funds	Omega	Sortino Ratio	Kappa 3	Upside Potential Ratio	Calmar Ratio	Sterling Ratio	Burke Ratio	Excess Return on VaR	Conditional Sharpe Ratio	Modified Sharpe Ratio	Average
C. Bond funds (12,279 funds)												
Fund investment style												
Government	534	0.98	1.00	1.00	0.95	0.86	0.87	0.87	0.99	0.98	0.98	0.95
Corporate	459	1.00	1.00	1.00	0.98	0.96	0.96	0.97	0.99	0.99	0.99	0.98
Mixed	11,286	0.99	1.00	1.00	0.97	0.95	0.96	0.96	0.96	0.97	0.99	0.97
Currency												
US$	4,079	0.99	1.00	0.99	0.95	0.89	0.90	0.90	0.99	0.98	0.99	0.96
Euro	4,656	0.99	1.00	1.00	0.96	0.96	0.97	0.97	0.90	0.96	0.99	0.97
Other	3,544	1.00	1.00	1.00	0.97	0.97	0.97	0.98	0.96	0.96	0.99	0.98
All	12,279	0.99	1.00	1.00	0.97	0.95	0.95	0.95	0.95	0.97	0.99	0.97

[a]The CISDM database contains information on 22 strategy groups, but because of the small size of the following groups, we combined them into the one group called "Other": capital structure arbitrage, market timing, option arbitrage, other relative value, and Regulation D.

Explanation for the High Rank Correlations

From a practical point of view, one could argue that the high rank correlations are simply a result of using similar performance measures; that is, the numerator is excess return for 10 of the 11 measures and the denominator contains a more or less comparable risk measure. I also found high rank correlations when comparing the risk measures and the return measures, which resulted in high rank correlations when I compared the performance measures.[8]

Eling and Schuhmacher (2007) suggested that one possible explanation for the high rank correlations is that fund returns are elliptically distributed. The distributions that permit mean–variance analysis can be elliptical rather than the multivariate normal distributions (see Ingersoll 1987). Lhabitant (2004), as well as Eling and Schuhmacher, found evidence for elliptically distributed hedge funds returns. Both studies found a good statistical fit by using the lognormal, the logistic, the Weibull, or the generalized beta distribution. I determined the underlying distribution for each fund on the basis of historical returns by using the distribution-fitting software BestFit. The results are presented in **Table 6**. The parametric distribution that best fits the empirical distribution is in most cases a logistic, a Weibull, or a normal distribution. I thus confirmed preceding findings that fund returns are often elliptically distributed.

To further explore the link between the fund's return distribution and rank correlation, I analyzed a series of synthetic returns produced by a Monte Carlo simulation. **Table 7** presents the rank correlations for 1,000 simulated funds with 120 monthly returns under five distributional assumptions (normal, lognormal, logistic, Weibull, and generalized beta distribution; I used the @RISK simulation software). All the simulated funds were calibrated to produce equal means and standard deviations, but under the various distributional assumptions, they have different values for skewness and kurtosis. Nevertheless, the simulated time series exhibit high rank correlations. The only exception is the drawdown-based performance measures—perhaps because I did not correlate the simulated returns, so these time series contain no strong common factor. Thus, apparently, the higher the correlation between the funds, the higher the rank correlations for the drawdown-based measures. Reconsidering the results in Table 3 confirms this finding: Stock funds have the highest rank correlations when the drawdown-based measures were used, whereas hedge funds exhibit relatively low rank correlations. The results for funds of hedge funds, however, do not confirm this evidence.[9]

I conclude that the reasons for the high rank correlations are that the performance measures are relatively similar (i.e., the risk and return measures are comparable) and that the fund returns are relatively similar (i.e., the returns are elliptically distributed and correlated).[10, 11]

Why the Sharpe Ratio Is Right

When analyzing either hedge funds and mutual funds, why is the Sharpe ratio the right measure for investors? From a practitioner's point of view, the Sharpe ratio might be considered superior to other performance measures for the following reasons:

- It is widely used in the investment industry and is the best known performance measure (Modigliani and Modigliani 1997). Most asset allocation analyses use the mean–variance approach in analyzing the trade-off between risk and return (Leland 1999; Sharpe 2007). The Sharpe ratio is also reported by most providers of financial information, such as Morningstar and Yahoo! Finance.
- It provides a convenient summary of two important aspects (risk and return) of any investment strategy (Sharpe 1994) and is probably the best understood performance measure (Lo 2002). Compared with other, more complex

Table 6. Best Fit Analysis

Fund Type	Logistic	Weibull	Normal	Generalized Beta	Loglogistic	Lognormal	Other
Stock funds	30.18%	26.96%	12.27%	8.05%	3.22%	0.80%	18.51%
Bond funds	37.12	18.06	13.38	13.71	3.01	1.34	13.38
Real estate funds	40.40	14.52	11.85	14.36	9.68	0.83	8.35
Hedge funds	36.75	8.43	11.45	7.63	15.66	1.20	18.88
Funds of hedge funds	37.40	11.19	10.35	7.68	12.52	1.34	19.53
CTAs	30.87	5.70	10.40	4.70	23.32	1.51	23.49
CPOs	30.25	7.56	11.76	5.38	20.17	2.18	22.69

Table 7. Simulation Analysis

Measure	Normal—@RISK Function: RiskNormal (0.0053;0.047)	Lognormal[a]—@RISK Function: RiskLognorm [0.1;0.047; Risk Shift(–0.0947)]	Logistic—@RISK Function: RiskLogistic (0.0053;0.02559)	Weibull[a]—@RISK Function: RiskWeibull 8.1842;0.34282; RiskShift(–0.31787)	Generalized Beta—@RISK Function: RiskBetaGeneral (9.2401;3.4794; –0.2782;0.1121)
Mean value (%)	0.53%	0.53%	0.53%	0.53%	0.53%
Standard deviation (%)	4.70%	4.70%	4.70%	4.70%	4.70%
Skewness	0.00	1.43	0.02	–0.54	–0.51
Excess kurtosis	–0.04	3.69	1.12	0.34	–0.05
JB rejection at 1% (5%) level	0.80% (1.50%)	100% (100%)	25.50% (34.80%)	17.40% (40.10%)	6.30% (21.20%)
	Rank Correlation Compared with the Sharpe Ratio				
Omega	1.00	1.00	1.00	1.00	1.00
Sortino ratio	1.00	1.00	1.00	1.00	1.00
Kappa 3	1.00	1.00	1.00	1.00	1.00
Upside potential ratio	0.96	1.00	0.95	0.97	0.98
Calmar ratio	0.39	0.59	0.59	0.42	0.40
Sterling ratio	0.42	0.66	0.63	0.46	0.38
Burke ratio	0.53	0.75	0.73	0.56	0.54
Excess return on VaR	1.00	1.00	1.00	1.00	1.00
Conditional Sharpe ratio	0.98	0.93	0.99	0.99	0.99
Modified Sharpe ratio	1.00	0.99	1.00	1.00	1.00

[a]To generate negative returns, I needed to shift the lognormal and Weibull distributions.

performance metrics, such as the drawdown-based measures, the Sharpe ratio is simple to calculate. And it is easily communicated to other professionals and even nonprofessionals. Furthermore, the data requirements are fewer than for measures that require the calculation of higher moments.

- A wide range of statistical tests are available for the Sharpe ratio (see, for example, Jobson and Korkie 1981; Memmel 2003), which is not the case for the other performance measures. Additionally, the Sharpe ratio has been the subject of much research; thus, its strengths and weaknesses are well known to researchers and practitioners, which also is not the case for the other performance measures.
- As I have shown, when analyzing either hedge funds or mutual funds, the choice of performance measure does not critical influence the relative evaluation of funds.

From a theoretical point of view, the Sharpe ratio is consistent with expected utility maximization under the assumption of elliptically distributed returns (Ingersoll 1987). Even without the assumption of elliptically distributed returns, mean–variance analysis of mutual funds and hedge funds approximately preserves the ranking of preferences in standard utility functions.[12] Furthermore, if an investor maximizes the expected utility of portfolio return and considers utility a quadratic function of portfolio return, only mean–variance-efficient portfolios need to be considered (Sharpe 2007). The Sharpe ratio thus builds on a sound theoretical framework, which cannot be said of many of the other performance measures examined in this article (e.g., the drawdown-based performance measures). The Sharpe ratio is closely connected to the separation theorem derived by Tobin (1958) and the efficient frontier derived by Markowitz (1952), which are the theoretical foundations of many other important applications in financial theory and practice, such as the capital asset pricing model or the Fama and French (1993) three-factor model. Finally, as Dowd (1999, 2000) showed, the Sharpe ratio can be the right measure when a fund represents the entire risky investment or when it represents only a portion of the investor's risky investment (thus requiring that correlations be taken into account).

In conclusion, from a practical and a theoretical point of view, the Sharpe ratio is adequate for analyzing both hedge funds and mutual funds. This statement does not mean that the Sharpe ratio is the *only* right measure and that all the other measures are useless. Of course, I am aware of the important differences between the Sharpe ratio and the other measures. What I showed, however, is that in almost all practical decision-making problems, the results of using the Sharpe ratio and of using the other measures are so close that which of the different measures is used makes almost no difference. Why not, then, use the simplest measure with the best theoretical foundation, namely, the Sharpe ratio?

Conclusion

The main result from the empirical investigation is that the choice of performance measure does not affect the ranking of hedge funds and mutual funds. I found a slight negative relationship between the rejection rate for the Jacque–Bera test and the rank correlation: The asset class with the highest rejection rate (real estate) has the lowest rank correlation and the asset class with the lowest rejection rate (stocks) has the highest rank correlation. Even for fund returns that usually display a strong deviation from a normal distribution, however, I found only small changes in rankings and rank correlation.

Drawing precise statistical inferences about fund performance is generally difficult because of the low signal-to-noise ratio (Kritzman 1986)— that is, the small value of return relative to the level of risk found for many funds. Therefore, a sample that is both large and covers an extensive period of time is needed to verify statistically whether the results are genuine or spurious (see Blake and Timmermann 2003 for a related discussion). The results that I have presented are based on a large data sample (38,954 funds investing in seven asset classes) and cover a lengthy period (1996–2005, which is as long as possible, especially for the sample of hedge funds). And the results are confirmed by numerous robustness tests, which should allow sound conclusions to be drawn.

From a practical point of view, the choice of performance measure is not critical to the relative evaluation of hedge funds and mutual funds. The Sharpe ratio is the best known and best understood performance measure and might thus be considered superior to other performance measures from a practitioner's point of view. From a theoretical point of view, the Sharpe ratio could also be considered superior to the other performance measures because it is consistent with expected utility maximization. I conclude that the Sharpe ratio is adequate for analyzing the returns of hedge funds and the returns of mutual funds.

I am grateful to Thomas Parnitzke, Hato Schmeiser, and Denis Toplek for valuable comments and suggestions.

This article qualifies for 1 CE credit.

Notes

1. See Kao (2002); Amin and Kat (2003); Gregoriou and Gueyie (2003).
2. See Geman and Kharoubi (2003); Kat (2003); Lamm (2003).
3. The Datastream investment funds database contains no strategy descriptions and does not categorize funds into various strategy groups. The only information available is the fund name (with the International Securities Identification Number and the Stock Exchange Daily Official List number), the country of issue, and the underlying currency. Therefore, I classified funds according to their names; that is, all the selected funds have the words "stock," "bond," "real estate," or a similar expression in their names. The underlying assumption was that a fund having such words in its title invests in the particular securities named. To reduce misclassification, I cross-checked all funds by examining their return distributions. For example, for all bond funds with a standard deviation of monthly returns three times higher than the average (i.e., three times 1.91 percent), I checked on the internet whether this fund really did have a focus on bonds or was misclassified. Altogether, I did 538 of these tests, corrected misclassified funds, and eliminated all ambiguous cases. This plausibility check should have reduced the danger of misclassification to a minimum. As shown in later robustness tests, the selected funds included a wide variety of countries and investment styles (e.g., value and growth, small capitalization and large capitalization). Other academic studies that used the Datastream data to study the performance measurement of mutual funds include Gemmill and Thomas (2002) and Otten and Bams (2002).
4. The CISDM database has been the subject of many academic studies (e.g., Capocci and Hübner 2004; Ding and Shawky 2007). The full database contains information on 8,542 funds. I eliminated 435 funds, however, because they appeared twice in the database, had less than four monthly returns, or reported returns only on a quarterly basis.
5. Note that, in light of the minimums and maximums (the fifth and sixth columns in the standard deviation row of Panel A in Table 1), the standard deviation of 2.43 percent is relatively small because outliers in the data resulted in a highly skewed distribution of the standard deviation across all funds.
6. I also calculated estimators for the backfilling bias by stepwise deleting the first 12, 24, 36, 48, and 60 months of returns (see Brown, Goetzmann, and Ibbotson 1999; Fung and Hsieh 2000; Capocci and Hübner 2004). For example, the monthly return of a portfolio invested in all the hedge funds would be 0.97 percent and a portfolio invested in all the funds of hedge funds, 0.67 percent. Eliminating the first 12 (24, 36, 48, 60) months of returns for each fund reduced these returns about 0.25 (0.41, 0.30, 0.43, 0.34) percentage points for hedge funds and by 0.02 (0.05, 0.02, 0.11, 0.10) percentage points for funds of hedge funds. More complete results are available as supplemental material in the *FAJ* area of www.cfapubs.org. These values are comparable to other values in the literature. For example, Fung and Hsieh found that the backfilling bias is noticeably lower for funds of funds than for hedge funds. I found comparable results for CTAs and CPOs, but for stock, bond, and real estate funds, the extent of the backfilling bias is low and its direction unclear. Other types of bias, such as the self-selection bias, should be negligible for mutual funds. For example, mutual funds entail no selection bias because they must publicly disclose their performance. For alternative

investment vehicles (hedge funds, funds of hedge funds, CTAs, CPOs) that do not make such a disclosure, the magnitude of the self-selection bias is limited and its direction unclear (Fung and Hsieh 2000, p. 299).

7. I calculated Spearman's (1904) rank correlation coefficient, r_s, which is a nonparametric measure of correlation. Unlike the Pearson product-moment correlation coefficient, Spearman's rank correlation coefficient requires neither that the relationship between the variables be linear nor that the variables be measured on interval scales; it can be used for variables measured at ordinal level. I converted the performance measurement results to ranks and calculated the differences, d_i, between the ranks of each fund i on two measures, as $r_s = 1 - \left(6\sum_{i=1}^{N} d_i^2\right) / \left(N^3 - N\right)$, where N denotes the total number of funds considered. Rank correlation matters in the context of this study because the performance of funds is regularly ranked in order to benchmark the success of the fund compared with that of other funds and to serve as the basis for investment decisions.
8. High rank correlations when comparing risk measures were also reported in a different context by Pfingsten, Wagner, and Wolferink (2004).
9. I also took the approach of removing the common factor and then testing whether performance as determined by various measures can or cannot be differentiated. I calculated fund returns in excess of the beta-adjusted mean return of all funds, r_{mt}, for each asset class and month (the excess return of fund i in time period t was calculated as $er_{it} = r_{it} - \beta_i \cdot r_{mt}$). I used various definitions of beta (beta = 1, constant beta, rolling 24-month beta). Again, I found high rank correlations between the performance measures. The results of these tests are available upon request.
10. I also carried out some other numerical tests, which showed that the result of high rank correlation between different measures is robust even for diverse funds. These tests are available upon request.
11. Another supposition might be that the high rank correlations are a result of the monthly measurement interval; low-frequency data usually show relatively little skewness and excess kurtosis (see Bollen and Busse 2001; Malkiel and Saha 2005; Kosowski, Timmermann, Wermers, and White 2006). I used weekly and daily data, however, to calculate the performance measures for a randomly selected sample of 1,000 stock funds and again found high rank correlations. These tests are reported in the supplemental material in the *FAJ* area of www.cfapubs.org; see Table S4. When weekly (daily) data were used, the average rank correlation was 0.98 (0.93).
12. See Levy and Markowitz (1979) and Hlawitschka (1994) for mutual funds, and for hedge funds, see Fung and Hsieh (1999).

References

Ackermann, C., R. McEnally, and D. Ravenscraft. 1999. "The Performance of Hedge Funds: Risk, Return, and Incentives." *Journal of Finance*, vol. 54, no. 3 (June):833–874.

Agarwal, V., and N.Y. Naik. 2004. "Risk and Portfolio Decisions Involving Hedge Funds." *Review of Financial Studies*, vol. 17, no. 1 (Spring):63–98.

Amin, G.S., and H.M. Kat. 2003. "Hedge Fund Performance, 1990–2000: Do the 'Money Machines' Really Add Value?" *Journal of Financial and Quantitative Analysis*, vol. 38, no. 2 (June):251–274.

Blake, D., and A. Timmermann. 2003. *Performance Persistence in Mutual Funds*. Study prepared for the Financial Services Authority (April 8).

Bollen, N., and J. Busse. 2001. "On the Timing Ability of Mutual Fund Managers." *Journal of Finance*, vol. 56, no. 3 (June):1075–1094.

Brooks, C., and H.M. Kat. 2002. "The Statistical Properties of Hedge Fund Index Returns and Their Implications for Investors." *Journal of Alternative Investments*, vol. 5, no. 2 (Fall):26–44.

Brown, S.J., W.N. Goetzmann, and R.G. Ibbotson. 1999. "Offshore Hedge Funds: Survival and Performance, 1989–1995." *Journal of Business*, vol. 72, no. 1 (January):91–117.

Burke, G. 1994. "A Sharper Sharpe Ratio." *Futures*, vol. 23, no. 3 (March):56.

Capocci, D., and G. Hübner. 2004. "Analysis of Hedge Fund Performance." *Journal of Empirical Finance*, vol. 11, no. 1 (January): 55–89.

Ding, B., and H.A. Shawky. 2007. "The Performance of Hedge Fund Strategies and the Asymmetry of Return Distributions." *European Financial Management*, vol. 13, no. 2 (March):309–331.

Dowd, K. 1999. "Financial Risk Management." *Financial Analysts Journal*, vol. 55, no. 4 (July/August):65–71.

———. 2000. "Adjusting for Risk: An Improved Sharpe Ratio." *International Review of Economics & Finance*, vol. 9, no. 3 (July): 209–222.

Eling, M., and F. Schuhmacher. 2007. "Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds?" *Journal of Banking & Finance*, vol. 31, no. 9 (September):2632–2647.

Fama, E.F., and K.R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, vol. 33, no. 1 (February):3–56.

Fung, W., and D.A. Hsieh. 1999. "Is Mean–Variance Analysis Applicable to Hedge Funds?" *Economics Letters*, vol. 62, no. 1 (January):53–58.

———. 2000. "Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases." *Journal of Financial and Quantitative Analysis*, vol. 35, no. 3 (September):291–307.

Geman, H., and C. Kharoubi. 2003. "Hedge Funds Revisited: Distributional Characteristics, Dependence Structure and Diversification." *Journal of Risk*, vol. 5, no. 4 (Summer):55–73.

Gemmill, G., and D.C. Thomas. 2002. "Noise Trading, Costly Arbitrage, and Asset Prices: Evidence from Closed-end Funds." *Journal of Finance*, vol. 57, no. 6 (December):2571–2594.

Gregoriou, G.N., and J.P. Gueyie. 2003. "Risk-Adjusted Performance of Funds of Hedge Funds Using a Modified Sharpe Ratio." *Journal of Wealth Management*, vol. 6, no. 3 (Winter):77–83.

Hlawitschka, W. 1994. "The Empirical Nature of Taylor-Series Approximations to Expected Utility." *American Economic Review*, vol. 84, no. 3 (June):713–719.

Ingersoll, J.E. 1987. *Theory of Financial Decision Making*. Totowa, NJ: Rowman & Littlefield.

Jobson, D., and B. Korkie. 1981. "Performance Hypothesis Testing with the Sharpe and Treynor Measures." *Journal of Finance*, vol. 36, no. 4 (September):889–908.

Kao, D.L. 2002. "Battle for Alphas: Hedge Funds versus Long-Only Portfolios." *Financial Analysts Journal*, vol. 58, no. 2 (March/April):16–36.

Kaplan, P.D., and J.A. Knowles. 2004. *Kappa: A Generalized Downside Risk-Adjusted Performance Measure*. Morningstar Associates and York Hedge Fund Strategies (January).

Kat, H.M. 2003. "10 Things That Investors Should Know about Hedge Funds." *Journal of Wealth Management*, vol. 5, no. 4 (Spring): 72–81.

Kestner, L.N. 1996. "Getting a Handle on True Performance." *Futures*, vol. 25, no. 1 (January):44–46.

Kosowski, R., A. Timmermann, R. Wermers, and H. White. 2006. "Can Mutual Fund 'Stars' Really Pick Stocks? New Evidence from a Bootstrap Analysis." *Journal of Finance*, vol. 61, no. 6 (December):2551–2595.

Kritzman, M. 1986. "How to Detect Skill in Management Performance." *Journal of Portfolio Management*, vol. 12, no. 2 (Winter):16–20.

Lamm, R. McFall. 2003. "Asymmetric Returns and Optimal Hedge Fund Portfolios." *Journal of Alternative Investments*, vol. 6, no. 2 (Fall):9–21.

Leland, H.E. 1999. "Beyond Mean–Variance: Performance Measurement in a Nonsymmetrical World." *Financial Analysts Journal*, vol. 55, no. 1 (January/February):27–36.

Levy, H., and H.M. Markowitz. 1979. "Approximating Expected Utility by a Function of Mean and Variance." *American Economic Review*, vol. 69, no. 3 (June):308–317.

Lhabitant, F.S. 2004. *Hedge Funds: Quantitative Insights*. Chichester, U.K.: Wiley.

Liang, B. 2000. "Hedge Funds: The Living and the Dead." *Journal of Financial and Quantitative Analysis*, vol. 35, no. 3 (September):309–326.

Lo, A.W. 2002. "The Statistics of Sharpe Ratios." *Financial Analysts Journal*, vol. 58, no. 4 (July/August):36–52.

Mahdavi, M. 2004. "Risk-Adjusted Return When Returns Are Not Normally Distributed: Adjusted Sharpe Ratio." *Journal of Alternative Investments*, vol. 6, no. 4 (Spring):47–57.

Malkiel, B.G., and A. Saha. 2005. "Hedge Funds: Risk and Return." *Financial Analysts Journal*, vol. 61, no. 6 (November/December):80–88.

Markowitz, H.M. 1952. "Portfolio Selection." *Journal of Finance*, vol. 7, no. 1 (March):77–91.

Memmel, C. 2003. "Performance Hypothesis Testing with the Sharpe Ratio." *Finance Letters*, vol. 1, no. 1 (February):21–23.

Modigliani, F., and L. Modigliani. 1997. "Risk-Adjusted Performance: How to Measure It and Why." *Journal of Portfolio Management*, vol. 23, no. 2 (Winter):45–54.

Otten, R., and D. Bams. 2002. "European Mutual Fund Performance." *European Financial Management*, vol. 8, no. 1 (March):75–101.

Pedersen, C.S., and T. Rudholm-Alfvin. 2003. "Selecting a Risk-Adjusted Shareholder Performance Measure." *Journal of Asset Management*, vol. 4, no. 3 (September):152–172.

Pfingsten, A., P. Wagner, and C. Wolferink. 2004. "An Empirical Investigation of the Rank Correlation between Different Risk Measures." *Journal of Risk*, vol. 6, no. 4 (Summer):55–74.

Rees, D.G. 1987. *Foundations of Statistics*. London: Chapman and Hall.

Schneeweis, T., H. Kazemi, and G. Martin. 2002. "Understanding Hedge Fund Performance: Research Issues Revisited—Part 1." *Journal of Alternative Investments*, vol. 5, no. 3 (Winter):6–22.

Shadwick, W.F., and C. Keating. 2002. "A Universal Performance Measure." *Journal of Performance Measurement*, vol. 6, no. 3 (Spring):59–84.

Sharma, M. 2004. "AIRAP—Alternative RAPMs for Alternative Investments." *Journal of Investment Management*, vol. 2, no. 4 (First Quarter):106–129.

Sharpe, W.F. 1966. "Mutual Fund Performance, Part 2: Supplement on Security Prices." *Journal of Business*, vol. 39, no. 1 (January):119–138.

———. 1994. "The Sharpe Ratio." *Journal of Portfolio Management*, vol. 21, no. 1 (Fall):49–59.

———. 2007. "Expected Utility Asset Allocation." *Financial Analysts Journal*, vol. 63, no. 5 (September/October):18–30.

Sortino, F.A., and R. van der Meer. 1991. "Downside Risk." *Journal of Portfolio Management*, vol. 17, no. 4 (Summer):27–31.

Sortino, F.A., R. van der Meer, and A. Plantinga. 1999. "The Dutch Triangle." *Journal of Portfolio Management*, vol. 26, no. 1 (Fall):50–58.

Spearman, C. 1904. "The Proof and Measurement of Association between Two Things." *The American Journal of Psychology*, vol. 15, no. 1 (January):72–101.

Tobin, J. 1958. "Liquidity Preference as Behavior toward Risk." *Review of Economic Studies*, vol. 25, no. 2 (February):65–86.

Young, T.W. 1991. "Calmar Ratio: A Smoother Tool." *Futures*, vol. 20, no. 1 (October):40.