Performance Measurement Manipulation

Gaming the current Morningstar star system.

Kenneth Winston

It has been known since Markowitz that efficient portfolio construction involves the simultaneous optimization of expected risk and expected return. In practice, though, investors often follow more simplified processes. There is widespread use of a two-step process: choosing a policy portfolio to embody a desired risk structure, and then selecting securities or funds within the policy portfolio using a quality measure assumed to be independent of the risk structure. This practice has been the subject of debate (see Bernstein [2003]).

A more extreme simplification involves the use of a single scalar measure that incorporates both risk and return. The literature provides numerous examples of scalar measures such as Sharpe ratios, Jensen alphas, information (or "selection Sharpe") ratios, and Treynor ratios for portfolio assessment. Unfortunately, no one number suffices to describe portfolio behavior over the horizons most investors use to evaluate performance. There are too many significant systematic factors.

Single scalar measures are moreover subject to manipulation. A number of authors have investigated the ability to manipulate various measures by informationless techniques, where the manipulator has no special insight about the future but takes advantage of peculiarities of the measure to make it appear that he or she does. Many but not all of these investigations concentrate on hedge fund techniques like Weisman [2002].

Spurgin [2001] reviews some techniques for gaming Sharpe ratios and introduces a new Sharpe ratio

Kenneth Winston is global chief risk officer at Morgan Stanley Investment Management in New York City.

manipulation method. Goetzmann et al. [2002] construct a Sharpe ratio maximizing distribution, assuming complete markets in contingent claims. They show further that manipulation is possible without complete markets, given only one call and one put.

In many cases, single scalar measures that are designed to provide simplified ex post performance analysis are used for ex ante portfolio selection, despite the absence of evidence that they have any persistence. These measures may be used in a two-step process as a prediction of quality, or they may simply be used alone with little thought given to portfolio structure.

In 1985, Morningstar, Inc., introduced a single scalar method for rating mutual funds that has been credited with influencing up to trillions of dollars of flows. It awarded from one (worst) to five (best) stars to a mutual fund according to a score derived from the fund's past performance and risk. The scoring was changed in July 2002 to the current method (see "The Morningstar Rating Methodology" [2002] and "The New Morningstar Rating Methodology" [2003]).

It appears that Morningstar has improved its measure, and its attempt to grade funds using risk adjustment is laudable. As many people use Morningstar's star system as a forward-looking selection device, it is useful to know the inevitable weaknesses of this popular scalar measure. I will show some informationless techniques that can be used to manipulate star ratings, so a buyer can attempt to distinguish star ratings attained through skill from star ratings attained through manipulation. Finally, I return to basic investment fundamentals to suggest how to choose funds.

MORNINGSTAR SCORE

Sharpe's [1998] analysis of the old Morningstar star system concludes that "Morningstar's... measure has a number of drawbacks." Among those noted were:

- The stand-alone nature of the measure. Since many or most investors do not buy mutual funds in isolation, but rather as part of a portfolio of investments, a measure that does not take into account interactions with the rest of the investor's portfolio is inherently flawed.
- The heroic assumption that statistics from the past are good estimators of future behavior. "Heroic" is too mild a word—"wrong" is more accurate. See Morey [2002] or a study by Morningstar itself (Ptak [2003]).[1]
- The tendency of the measure to reward extreme leverage (borrowing or lending at the risk-free rate).

The current method places funds in one of about 62 categories and ranks them according to a score against the rest of the category. There are many adjustments to the score, but the basic form is derived from the standard constant relative risk aversion utility function. It is an estimate of

$$MRAR(\gamma) = E[(1 + r_G)^{-\gamma}]^{-\frac{1}{\gamma}} - 1$$

where γ is later set to 2, so that

$$MRAR(2) = \frac{1}{\sqrt{E[\frac{1}{(1 + r_G)^2}]}} - 1$$

where r_G is the geometric excess (over the risk-free rate) return of the fund. The highest-ranking 10% of funds within a peer group get five stars, the next highest 22.5% four stars, 35% three stars, 22.5% two stars, and 10% one star.

Discussing the desirable aspects of a score, Morningstar indicates that the new methodology was among other things aimed at risk-adjusting fund ratings. One desirable property of risk adjustment—which the Sharpe ratio has—is that "a fund's score is not sensitive to its proportion of risk-free assets or its amount of leverage" ("The New Morningstar Rating Methodology" [2003, p. 3]). The Morningstar score, while adjusting for risk, does not have this property, leaving open the possibility it can be manipulated by using leverage.

PEER GROUPS

While the old Morningstar method had an auxiliary category ranking, the basic method grouped together very broad asset classes such as domestic equity. This led to complaints that the old method tended to award stars to funds whose styles were in favor, obscuring individual fund performance. The current method, by incorporating more finely drawn categories into the basic system, partially addresses this complaint.

To address the portfolio construction issue, Morningstar assumes that funds within categories are substitutes for each other except for manager skill, saying "the rela-

EXHIBIT 1

Characteristics of Selected Morningstar Categories

Category	Mean Return	Mean Std Dev	Median Std Dev	Min Std Dev	Max Std Dev	Count	Lvg-Neutral Return
Precious Metals	35.98	36.72	36.82	31.62	50.65	34	44.10
Specialty-Technology	-12.87	35.10	35.13	16.12	62.77	264	39.85
Latin America Stock	5.51	27.69	28.48	25.42	28.81	18	25.56
Communications	-12.48	29.41	27.09	15.62	51.41	34	23.03
Small Growth	0.14	23.98	23.47	13.89	49.26	516	17.10
Divrsfied Emging Mkts	11.53	22.61	22.80	17.66	29.33	138	16.11
Specialty-Natural Res	5.02	22.54	21.46	16.73	32.43	65	14.21
Mid-Cap Growth	-5.17	21.81	21.28	13.42	38.77	626	13.97
Small Blend	8.57	20.67	20.27	11.45	43.07	314	12.65
Small Value	12.06	19.57	19.01	10.54	34.71	191	11.09
Mid-Cap Blend	4.77	18.84	18.97	7.81	37.51	245	11.04
Large Growth	-8.12	19.05	18.11	10.54	51.90	974	10.04
Foreign Large Growth	-5.04	18.04	17.92	13.67	25.83	168	9.82
Foreign Large Value	1.59	18.20	17.86	14.53	22.76	99	9.76
Foreign Large Blend	-3.32	17.25	17.38	13.34	21.68	366	9.23
World Stock	-2.84	17.84	17.36	4.90	57	273	9.22
Large Blend	-4.13	16.69	16.85	2.83	34.41	1127	8.67
Large Value	-0.94	16.58	16.49	9.85	26.17	675	8.30
Mid-Cap Value	5.83	17.21	16.29	7.68	27.33	185	8.10
Emg Markets Bond	13.47	12.05	11.75	7.75	15.75	43	4.18
Specialty-Real Estate	16.22	11.84	11.36	3.74	25.36	136	3.90
High Yield Bond	3.37	9.04	9.11	2.45	15.52	342	2.50
World Bond	6.77	6.41	6.48	1.41	10.30	116	1.26
Long-Term Bond	5.42	6.41	5.92	4.24	10.91	67	1.05
Multisector Bond	5.28	5.50	5.39	0	12.00	136	0.87
Muni National Long	5.95	5.22	5.29	3.87	7.28	244	0.84
Muni National Interm	5.52	4.72	4.58	2.45	13.60	175	0.63
Intermd-Term Bond	4.01	4.32	4.24	2.45	7.42	675	0.54
Intermd Government	3.38	4.08	4.12	2.24	7.28	298	0.51
High Yield Muni	6.78	3.81	3.61	1.41	9.27	71	0.39
Short Government	2.49	2.48	2.55	1	4.80	139	0.19
Bank Loan	2.10	3.01	7.72	1.41	5.74	36	0.15
Muni National Short	3.69	2.02	2	0	4.12	75	0.12
Short-Term Bond	2.42	2.13	2	1	4	202	0.12
Ultrashort Bond	1.04	0.62	1	0	1.41	68	0.03

tive star ratings of two funds should be affected more by manager skill than by market circumstances or events that lie beyond the fund managers' control" [2003, p.16].

The peer group model that Morningstar and others use can be made more precise as follows. There are n securities (or funds; we use the word *securities* inclusively) and k peer groups, where $k << n$. There is a $k \times n$ incidence matrix M, where the (i, j) entry of M is 0 if security j is not in peer group i, and 1 if it is. Since peer groups are disjoint, we have $MM' = D$, where $'$ denotes transpose, and D is a $k \times k$ diagonal matrix. The (i, i) element of D is the number of securities in the i-th peer group.

The peer group model assumes there are benchmarks that capture the systematic behavior of each peer group. If R_t is the n-vector of security returns in period t, and b_t is the k-vector of benchmark returns, then the peer group model assumes $R_t = M'b_t + \varepsilon_t$. Here ε_t is an n-vector of idiosyncratic returns uncorrelated with each other and with the peer group benchmark returns in b_t.

For mutual funds, this implies that the only relevant information about interactions between funds is embodied in the generic description of the category. Manager skill incorporated in ε_t is assumed to be orthogonal to portfolio construction criteria.

There are many different mandates among the thousands of available mutual funds. Some funds in the world bond category have a mandate to produce currency-hedged returns; others tactically hedge currencies; and others do not hedge. Exhibit 1 shows that the highest standard deviation of a fund in the Morningstar world bond category was seven times the size of the lowest.

But even assuming that peer groups were truly groups of peers, one would want to use an estimate of the expected value of ε_t. Morningstar's measure for awarding stars makes no reference to a peer group benchmark, other than grouping funds in peer groups in order to award stars. If portfolios are constructed by assuming that all funds in a peer group have a beta of one to the peer group benchmark, the aggregate risk will not be what is desired. The MRAR(2) quality measure interacts with the risk criteria in a way that is not consistent with the two-step portfolio construction process.

EXPECTED VALUE ASSUMING LOGNORMALITY

Assume portfolio excess return relatives X are generated by the Itô process:[2]

$$dX = \alpha X dt + \sigma X dB(t) \tag{1}$$

where α is a drift parameter and σ is the standard deviation of the diffusion. $B(t)$ is a standard Brownian motion with mean zero and standard deviation $\sqrt{t}$ at time t. The familiar (Black-Scholes) application of Itô's lemma tells us that:

$$d\ln(X) = (\alpha - \sigma^2/2)dt + \sigma dB(t) \tag{2}$$

that is, this lognormal process has a growth rate $g = \alpha - \sigma^2/2$.

If $x = \ln(X)$ is the normal variate, then to evaluate MRAR(2) we want first to find the expected value of $(1 + r_G)^{-2} = \exp(-2x)$, which is the integral:

$$(MRAR(2)+1)^{-2} = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-2x)\exp\left(\frac{-(x-g)^2}{2\sigma^2}\right)dx \tag{3}$$

We obtain

$$MRAR(2) = \exp\left(g - \sigma^2\right) - 1 = \exp\left(\alpha - 3\sigma^2/2\right) - 1 \tag{4}$$

Under the lognormal assumption, we see that MRAR(2) is simply a function of compound return (g) minus a variance penalty.

ESTIMATING THE CURRENT SCORE FROM A TAYLOR SERIES

Without the assumption of (log)normality, we can look directly at properties of the time series average of

$$\frac{1}{(1+r_G)^2}$$

As Young and Trent [1969] use a Taylor series to estimate geometric mean as a function of the first four moments of the distribution, we can similarly write:

$$\frac{1}{(1+r_G)^2} \approx 1 - 2r_G + 3r_G^2 - 4r_G^3 + 5r_G^4 - \ldots \tag{5}$$

If μ is the arithmetic mean of r_G, we can rewrite this as:

$$\frac{1}{(1+\mu)^2(1+z_G)^2} \approx \frac{1}{(1+\mu)^2}\left(1 - 2z_G + 3z_G^2 - 4z_G^3 + 5z_G^4 - \ldots\right) \tag{6}$$

where:

$$z_G = \frac{r_G - \mu}{1+\mu}$$

Periodic (monthly) excess returns over the mean are generally low, so this series converges. Define the higher normalized central moments

$$\mu_i = \frac{E[(r_G - \mu)^i]}{\sigma^i}$$

where $\sigma^2 = E(r_G{}^2) - \mu^2$. We can then write:

$$E\left(\frac{1}{(1+r_G)^2}\right) \approx \frac{1}{(1+\mu)^2}\left(1 + 3\frac{\sigma^2}{(1+\mu)^2} - 4\frac{\sigma^3\mu_3}{(1+\mu)^3} + 5\frac{\sigma^4\mu_4}{(1+\mu)^4} - \ldots\right) \tag{7}$$

For most distributions of mutual fund returns, the quantity $\sigma/(1+\mu)$—monthly sample standard deviation divided by one plus monthly arithmetic mean return—will be significantly lower than one, and the effects of higher moments (multiplied by third, fourth, and higher powers of $\sigma/[1+\mu]$) will be slight. My focus is the effect of leverage on MRAR(2), but options techniques similar to those used by Goetzmann et al. [2002] to improve the Sharpe ratio also can be used to improve MRAR(2). Options change the exp($-2x$) term in different regions of Equation (3). While closed-form solutions are difficult, it appears that a simple strategy of selling calls improves MRAR(2) in most cases.

I hereafter drop the higher moments from (7) and concentrate on the mean and standard deviation. This gives a good approximation of Morningstar's MRAR(2):

$$MRAR(2) \approx \frac{(1+\mu)^2}{\sqrt{\left((1+\mu)^2 + 3\sigma^2\right)}} - 1 \tag{8}$$

LEVERAGING AND DELEVERAGING

Consider the effect on this score of leveraging or deleveraging a portfolio. If we put a fraction $(1-\beta)$ in the risk-free asset (with return r_f), and β in the portfolio, we get:

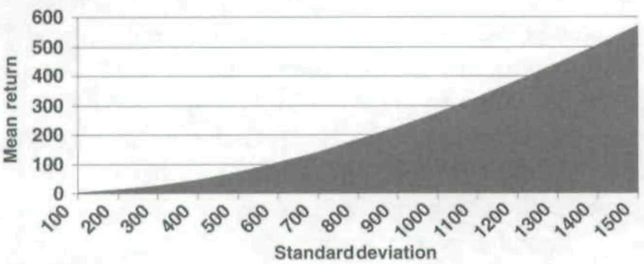
$$(1-\beta)r_f - \beta\left((1+r_f)(1+r_G) - 1\right)$$

Adding one to this, dividing by $1 + r_f$, and subtracting one gives $r_{G,leverage} = \beta r_G$. Thus leverage changes the mean μ and standard deviation σ linearly by the leverage factor β. Adding leverage to (8), we find that:

$$(MRAR(2)_\beta + 1)^2 \approx \frac{(1+\beta\mu)^4}{(1+\beta\mu)^2 + 3\beta^2\sigma^2} \tag{9}$$

EXHIBIT 2

Leverage-Neutral Monthly Arithmetic Mean Return (bp)



The numerator of the derivative of Equation (9) is

$$\beta^2[\mu(\mu^2 + 3\sigma^2)] + \beta(2\mu^2 - 3\sigma^2) + \mu \tag{10}$$

while the denominator is a positive term.

Setting (10) to zero and solving for β, we find

$$\beta_{\max} = \frac{3\sigma^2 - 2\mu^2 - \sigma\sqrt{3(3\sigma^2 - 8\mu^2)}}{2\mu(\mu^2 + 3\sigma^2)} \tag{11}$$

Equation (11) gives the leverage that maximizes MRAR(2), given a known μ and a known σ. Ex ante, we do not know what these moments will be and thus cannot leverage optimally to maximize Morningstar stars. The general range of σ is usually known in advance, though, so the direction in which leverage or deleverage is desirable is generally discernible in advance. For example, for a limited-term government bond fund, we know that σ will be quite low, while for an emerging technology fund it will be comparatively high.

The direction of leverage can be determined by setting $\beta = 1$ in (10) and solving for the μ that makes the expression equal to zero for a given σ. This gives us a cubic equation with a solution:

$$\mu = -\frac{2}{3} + (R + \sqrt{D})^{1/3} + (R - \sqrt{D})^{1/3} \tag{12}$$

where

$$R = \frac{5}{2}\sigma^2 + \frac{1}{27}$$

and $D = \sigma^2\left(\sigma^4 + \frac{71}{12}\sigma^2 + \frac{6}{27}\right)$

(see Weisstein [2004]).

Equation (12) is the mean return that is leverage-neutral for a given standard deviation. If the mean return is below that value, leveraging down improves Morningstar stars. If the mean return is above that value, leveraging up improves Morningstar stars. In Exhibit 2, the shaded area is where leveraging down improves Morningstar stars; the unshaded area is where leveraging up improves Morningstar stars. Recall that mean return here is monthly arithmetic mean return.

This tells us that for low-risk funds we should leverage up, and for higher-risk funds we should leverage down.[3] Even if we do not know the optimal leverage, a step in the right direction can be beneficial.[4]

For example, Oppenheimer Emerging Technologies fund, which by mandate and design is a high-volatility fund, would have to produce monthly excess returns of 368 basis points or more (annual excess returns of 54%) to avoid benefiting from leveraging down, according to Morningstar stars. The fund actually returned 64% in 2003, but it is unlikely to maintain that pace over multi-year periods.

Leveraging down can be accomplished in an obvious way by holding cash, or more subtly by futures and options techniques. Similarly, leveraging up can be accomplished by borrowing or by such methods as buying Treasury futures instead of Treasuries, or mortgage TBAs instead of mortgages.

The only case in the samples I checked in which my calculations would have led us astray over the last three years was the Oppenheimer Gold & Special Minerals Fund. With high risk, we would have assumed that leveraging down was appropriate—the leverage-neutral return was an annual rate of 29.48%. In fact, the fund returned a monthly average that annualized to 42.66% over the three years. This is highly unusual. Over the five years ending December 31, 2003 (which includes the spectacular last three years), the annualized monthly average was 22.25%.

EMPIRICAL RESULTS OF LEVERAGE

I looked at funds with widely varying characteristics to see the effects of leverage on the Morningstar score. Exhibit 3 shows that approximation (8) agrees closely with computed MRAR(2). From the leverage-neutral returns in Exhibit 3, we can see that funds that have a mandate to take high risk are encouraged to leverage down, possibly leading to inappropriately low levels of risk, while funds with a mandate to take low risk are encouraged to leverage up, leading to inappropriately high levels of risk.[5]

Exhibit 3 is based on return histories of funds and benchmarks, using the formulas above, without the many adjustments that Morningstar makes. I also obtain from Morningstar's Datalab product three-year returns and risks, and apply them to expression (8) for some 16,280 funds (many with multiple share classes), although not all had data.

Looking at the correspondence between expression (8) and the three-year stars that Morningstar awarded for the period ending February 29, 2004, I found excellent alignment. The correlation between the figure Morningstar gives in Datalab for three-year MRAR(2) and our expression (8) is 99% over the 12,048 funds with data (including multiple share classes) in Datalab.

Thus I created Exhibit 4, indicating what leverage would have improved expression (8) for the funds in Exhibit 3 to the level where the highest possible number of stars would be awarded. The max-star β is the fund's optimal leverage computed as in Exhibit 3—except for the slightly different time period of Exhibit 4—when five stars are not achievable by leverage, and the least extreme leverage to achieve five stars when five stars are achievable.

The leverage-neutral returns for selected Morningstar categories are in the last column of Exhibit 1.[6] I would have liked that exhibit to indicate how likely it is that leveraging up or down will raise Morningstar stars, but to generate a probability, I would have to have a prior expectation of return in a given category. Without such priors, I present the leverage-neutral returns in Exhibit 1 and leave it to the reader to decide whether it is likely that such a return will be attained.

EXHIBIT 3

Leverage-Neutral Returns

Fund/Portfolio	Mean Return μ	Volatility σ	MRAR (2)	Exprsn (8)	Optimal β	Leveraged MRAR (2)	Leverage -Neutral Return	Distance to Lvg-Neutral
Emerging Tech	-161.67	1169.48	-364.26	-363.78	0.00	0.00	54.23%	72.00%
Nasdaq 100	-85.41	1118.10	-279.43	-269.30	0.00	0.00	49.18%	58.97%
Emrgng Growth	90.65	1008.01	-57.34	-57.09	0.30	13.90	39.34%	27.91%
Gold & Spcl Mn	300.51	879.86	186.17	189.59	1.48	189.54	29.48%	-13.18%
Global Opptnty	5.29	842.88	-100.20	-99.55	0.02	0.07	26.93%	26.30%
Intl Growth	-31.19	760.35	-120.98	-117.06	0.00	0.00	21.71%	25.39%
S&P 500	-40.25	530.21	-82.06	-82.32	0.00	0.00	10.33%	15.06%
Main Street	-32.83	449.86	-63.09	-63.15	0.00	0.00	7.40%	11.26%
Intl Bond	106.33	222.93	99.22	98.96	9.67	421.27	1.80%	-11.74%
Strategic Inc	60.80	170.49	56.54	56.47	8.07	189.11	1.05%	-5.20%
Roch Municipal	32.94	145.17	29.83	29.79	5.50	80.12	0.76%	-3.27%
US Govt	34.09	122.90	31.87	31.83	8.18	119.08	0.54%	-3.62%
Citigroup BIG	41.89	121.63	39.71	39.68	10.80	143.75	0.53%	-4.61%
Senior Flt Rate	17.20	83.60	16.18	16.16	8.58	68.46	0.25%	-1.83%
Ltd Govt	20.09	47.61	19.76	19.75	36.79	334.75	0.08%	-2.36%

AD ASTRA PER ASPERA

Morningstar's star system has become popular because of the natural human desire to simplify the decision-making process. Morningstar fund analysts provide in-depth information about funds, and a web site with many sophisticated analytical and portfolio construction tools—but fund flows are still overwhelmingly determined simply by stars.

Simplification can be overdone. Miller says:

> The span of absolute judgment and the span of immediate memory impose severe limitations on the amount of information that we are able to receive, process, and remember. By organizing the stimulus input simultaneously into several dimensions and successively into a sequence or chunks, we manage to break (or at least stretch) this informational bottleneck [1956, p. 96].

Several dimensions are not one dimension, leaving even limited human processing ability more latitude in assessing mutual funds.[7]

Predicting which funds will perform well in the future is beyond my objective, but a careful evaluation of past behavior is certainly possible by going back to investment basics. The relevant risk in a prospective fund is its marginal effect on an investor's entire portfolio. Careful assessment of covariance characteristics—beyond category membership—can be made by financial analysts who are able to look at an investor's entire portfolio, including hard-to-quantify factors such as life circumstances and risk tolerance. Morningstar and others provide quantitative portfolio construction tools for this purpose as well.

Institutional investors can evaluate portfolios by constructing a benchmark that represents the manager's style. Such a benchmark is an informationless portfolio—in some cases an index or combination of indexes, in other cases a more customized asset list—that is an investable alternative to the portfolio it benchmarks. As Bailey, Richards, and Tierney [1989] explain, a properly chosen benchmark captures a fund's general characteristics but is orthogonal to the process that the manager uses to add value.

One important characteristic of a

EXHIBIT 4

Leverage Needed to Maximize Stars

Fund	Category	Mstar 3yr Stars	Max Stars w/Leverage	Max-Star β
Emerging Tech	Specialty Tech	2	5	0.45
Emerging Growth	Small Growth	3	4	0.30
Gold & Spcl Mn	Precious Metals	2	5	2.04
Global Opptnty	World Stock	2	4	0.29
Intl Growth	Foreign Lg Growth	2	5	0.39
Main Street	Large Blend	4	5	0.81
Intl Bond	World Bond	5	5	1.00
Strategic Inc	Multisector Bond	4	5	1.22
Roch Municipal	Muni NY Long	4	5	1.12
US Govt	Interm Bond	3	5	1.22
Senior Flt Rate	Bank Loan	3	5	2.30
Ltd Govt	Short Govt	3	5	1.35

properly chosen benchmark is that the beta of the fund to its benchmark is one, so that there is no systematic up- or down-market bias to fund/benchmark performance. This characteristic is clearly rarely attained by using the risk-free rate as a benchmark, and in fact is generally not attained by using category averages or category benchmarks. Institutional-quality benchmarks would avoid the oversimplification of categories, while still providing investors with information about fund style.

If such benchmarks were available for mutual funds—produced, perhaps, by an independent organization like Morningstar—then portfolio construction and performance evaluation would be far more accurate.

CONCLUSIONS

The current Morningstar star system reduces multidimensional fund performance to a single number, and then awards stars on the basis of this single number. The goal of measuring past performance adjusted for risk is a good one, but all one-dimensional scoring systems are inherently flawed. Users of Morningstar stars should be aware of some of the informationless strategies that can be used to increase stars without any management skill. In addition, funds whose managers do have skill but that are grouped with other funds that have varying risk mandates may have fewer stars than is warranted by their value.

Any single measure is subject to error. Investors should not use Morningstar stars, information ratios, or any other scalar performance measure in isolation. In addition to the characteristics necessary to determine a fund's likely interaction with the rest of an investor's portfolio, users of risk-adjusted performance measures should be aware of:

- Leverage, including both direct actions such as borrowing, lending, and holding cash, and indirect leverage and deleverage through derivatives.
- Optionality, through the explicit use of either options or dynamic option replication that will tend to throw off other measures. Portfolios that use optionality may need to have a full set of Greeks to describe their behavior.

There is no substitute for the hard work of determining multiple aspects of portfolio performance. Morningstar stars—or any other scalar measure of fund performance—should be used only as part of a complete balanced diet of information.

ENDNOTES

The author thanks for helpful comments Paul Kaplan, Tom Philips, Rudi Schadt, William Sharpe, Charles Trzcinka, and Andrew Weisman. This article was written and submitted when the author was a senior investment officer and director of risk management at OppenheimerFunds, Inc.

[1]An October 6, 2003, *Wall Street Journal* article quotes Morningstar analyst Ptak as saying: "There's almost a blithe acceptance of the predictive power of past performance out there. Whether through fund ads, the way managers are touted in the media, or the way investors research funds, past performance is put on a pedestal, and that assertion just [doesn't] hold up well under scrutiny."

[2]Return relatives are one plus returns. We assume throughout that one month is a single period, so unless otherwise noted all quantities are expressed in monthly terms.

[3]By saying "we should" leverage up or down, I do not mean to advocate taking any of the gaming actions described in this article. I am describing what a Morningstar star maximizer with no information "should" do. I advocate the use of superior information and portfolio construction to manage portfolios, not manipulation.

[4]Depending on how much data are available, Morningstar takes combinations of three-year, five-year, and ten-year star ratings to assign an overall rating. There are also adjustments if funds have changed categories. But since the MRAR(2) formula is at the heart of every score, and since funds, especially funds in very high or very low risk categories, know the general level of their volatility in advance, the direction a fund should leverage (up or down) in order to get more Morningstar stars is usually quite clear in advance.

[5]For Exhibit 3 I use the 36 months ending December 31, 2003. Geometric returns over the one month T-bill rate are used for monthly returns r_G. MRAR(2) is computed by taking the sum of the 36 inverse squares of $(1 + r_G)$, and then taking the inverse square root of that sum and subtracting one. Since an

all-cash portfolio has an MRAR(2) of zero, any portfolio with a negative MRAR(2) could have been improved in retrospect by setting β to zero—I show the optimal β from Equation (11) or zero, whichever is higher. Leveraged MRAR(2) shows the improved score using the indicated β. All periodic figures are given in basis points per month, except for the leverage-neutral return and distance to leverage-neutral. To put these figures on a more familiar annual scale, I take the monthly averages, add one, and raise them to the 12th power. These are not actual compound returns. Except for indexes, funds shown in Exhibits 3 and 4 are OppenheimerFunds.

There is also a kind of moral hazard that can attach at the end of a measurement period, where most of the data to compute μ and σ are known. A manager on the borderline might be encouraged to take extreme leverage measures to move into another star category.

[6]Exhibit 1 uses data from Morningstar's Datalab for the three years ending February 29, 2004. The mean return, and mean, median, minimum, and maximum standard deviation figures are annualized percentages. The count includes multiple share classes. Some minimum standard deviations are zero due to loss of precision at low values. The leverage-neutral return is computed from the median standard deviation using Equation (12); as in Exhibit 3, I put this on a more familiar annualized basis by taking one plus the monthly figure to the 12th power and subtracting one.

[7]Charles Trzcinka has pointed out that it is curious that consumers of food seem to have no trouble reading multi-measure nutrition tables, but that consumers of mutual funds seem to have the urge to use single measures.

REFERENCES

Bailey, J.V., T.M. Richards, and D.E. Tierney. "Benchmark Portfolio: Concept and Design." In Frank Fabozzi, ed., *Managing Institutional Assets*. Chicago: Probus, 1989, pp. 245-275.

Bernstein, Peter. "Economics and Portfolio Strategy." Peter L. Bernstein, Inc., March 2003.

Goetzmann, William, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch. "Sharpening Sharpe Ratios." Yale ICF Working Paper No. 02-08, February 2002. Available at http://www.stern.nyu.edu/fin/pdfs/seminars/033f-goetzmann.pdf.

Miller, George A. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *The Psychological Review,* 63 (1956), pp. 81-97.

Morey, Matthew. "Rating the Raters: An Investigation of Mutual Fund Rating Services." *Journal of Investment Consulting,* November/December 2002.

"The Morningstar Rating™ Methodology." Morningstar, Inc., October 1, 2003.

"The New Morningstar Rating™ Methodology." Morningstar, Inc., April 22, 2002.

Ptak, Jeffrey. "Back to the Future?" *Morningstar Mutual Funds,* October 21, 2003.

Sharpe, William. "Morningstar's Risk-Adjusted Ratings." *Financial Analysts Journal*, July/August 1998, pp. 21-33.

Spurgin, R.B. "How to Game Your Sharpe Ratio." *The Journal of Alternative Investments,* 4(3) (2001), pp. 38-46.

Weisman, A. "Informationless Investing and Hedge Fund Performance Measurement Bias." *The Journal of Portfolio Management,* 28(4) (2002), pp. 80-92.

Weisstein, Eric. "Cubic Equation." *MathWorld*—A Wolfram Web Resource, 2004. Available at http://mathworld.wolfram.com/CubicEquation.html

Young, William, and Robert Trent. "Geometric Mean Approximations of Individual Security and Portfolio Performance." *Journal of Financial and Quantitative Analysis*, 4 (1969), pp. 179-199.