Portfolio Performance Manipulation and Manipulation-proof Performance Measures

Jonathan Ingersoll, Matthew Spiegel, William Goetzmann
Yale School of Management, PO Box 208200, New Haven, CT 06520-8200

Ivo Welch
Brown University, Department of Economics Box B, Providence, RI 02912

Numerous measures have been proposed to gauge the performance of active management. Unfortunately, these measures can be gamed. Our article shows that gaming can have a substantial impact on popular measures even in the presence of high transactions costs. Our article shows there are conditions under which a manipulation-proof measure exists and fully characterizes it. This measure looks like the average of a power utility function, calculated over the return history. The case for using our alternative ranking metric is particularly compelling for hedge funds whose use of derivatives is unconstrained and whose managers' compensation itself induces a nonlinear payoff. (*JEL* G11, G23, G24)

Money managers often claim that they can provide superior performance in their funds. Investors must rank and then select managers based upon these claims, and any other information at their disposal. Ideally, manager evaluation requires a consideration of the inputs to the investment process as well as the resulting outcomes. Inputs include answers to questions such as: What is a fund's style and investment philosophy? Does the fund practice closet indexing, does it have high expenses or high turnover, does it employ window dressing, or engage in behavior which might result in portfolios with aberrant performance?[1] An assessment of the inputs allows the returns that are earned to be evaluated in light of the risks undertaken and the expenses incurred.

Along with this input information, output information measured by periodic rates of return is readily available for most—but not all—sectors

This article has benefited from the comments of an anonymous referee and Maureen O'Hara at the *The Review of Financial Studies* and numerous colleagues including the participants at the Five Star Conference hosted by New York University, the Hedge Fund Conference hosted by Borsa Italiana, and the Berkeley Program in Finance. Address correspondence to Jonathan Ingersoll, Yale School of Management, PO Box 208200, New Haven, CT 06520–8200, or e-mail: jonathan.ingersoll@yale.edu.

[1] See, for example, Brown and Goetzmann (1997) on style investing, Cremers and Petajisto (2006) on closet indexing, and Lakonishok et al. (1991) on window dressing. "Aberrant" is used here only in the statistical sense of a sample not representing the population. An example of a practice that can result in aberrant performance is writing covered calls. This was a common portfolio technique in the 1980s and resulted in superior performance provided the underlying portfolio or stocks did not have very large gains. Of course, the realized performance is aberrant only if the actions and its consequences were unanticipated or unusual.

doi:10.1093/rfs/hhm025 Advance Access publication 17 May 2007

of the investment industry. Funds investing in nonmarketable assets typically do not measure returns periodically. Conversely funds that hold marketable securities generally report returns based on observed security prices and dividend payments. Reporting requirements for open-end mutual funds in the United States, for example, are quite explicit. Every mutual fund must report its own performance by pricing individual securities at end of day prices and use this to calculate daily net asset values per share. Rating services such as Morningstar, Lipper, Reuters, and Business Week are leading vendors of detailed information about past mutual fund returns, expenses, and turnover. In some cases, particularly in the hedge fund industry, returns may calculated over different time intervals, or subject to chosen methods of valuation, but these are typically available on a monthly basis to potential investors, and are often used for the quantitative aspects of performance measurement and manager evaluation.

Modern performance measurement dates to 1966 when William Sharpe, using mean-variance theory, introduced the Sharpe ratio to provide a quantifiable, one-dimensional measure for performance. It is still probably the best-known and most widely used of the numerical ratings.[2] Subsequently, Jensen (1969) introduced alpha, the first benchmark-based measure. These, and their close analogues the information ratio, M-squared, and others are now widely used to rank investment managers and to evaluate the attractiveness of investment strategies in general. Other measures, such as the Henriksson and Merton (1981) market timing measure evaluate other aspects of performance but are based on the same or similar theories.

Inherent in this quantitative evaluation process is the potential for moral hazard. If investors use such scalar performance measures to rank or select money managers, either directly or indirectly through reputation, then money managers have an obvious incentive to take actions that enhance these measures—either through effort and application of skill, or through "information-free" activities that amount to manipulation. More formally, manipulation is action taken to increase a fund's performance measure that does not actually add value for the fund's investor. This has been referred to as "information-free" investing as a way to emphasize that it is not based upon the production and deployment of value-relevant information about the underlying assets in the portfolio.[3]

Our article contributes in two ways to the literature on performance evaluation. First, it points out the vulnerability of traditional measures to a number of simple dynamic manipulation strategies. Second, it offers

[2] The Sharpe ratio is used, for example, in the Schwab Select List and the Standard and Poor's Select Funds mutual fund ratings and in the *Hulbert Financial Digest* newsletter ratings.

[3] See, for example, Brown et al. (2004) and Weisman (2002).

a formal definition of the properties that a manipulation-proof measure should have and derives such a measure.

For a measure to be manipulation-proof it must not reward information-free trading. In this regard the existing set of performance measures suffers from two weaknesses. First, most were designed to be used in a world where asset and hence portfolio returns have "nice" distributions like the normal or the lognormal. However, this ignores the fact that managers can potentially use derivatives (or dynamic trading strategies) to radically alter return distributions. Hedge fund returns, in particular, have distributions that can deviate substantially from normality, and the hedge fund industry is one in which performance measures like the Sharpe ratio or information ratio are most commonly employed. Second, "exact" performance measures can be calculated only in theoretical studies (e.g., Leland (1999) and Ferson and Siegel (2001)); in general, they must be estimated. Since standard statistical techniques are designed for independent and identically distributed variables, it is possible to manipulate the existing measures by using dynamic strategies resulting in portfolios whose returns are not so distributed even though they may come from standard "nice" distributions.

The article describes three general strategies for manipulating a performance measure. The first is the manipulation of the underlying distribution to influence the measure. This can enhance measured performance even if the evaluator calculates the measure without any estimation error whatsoever. The second is the dynamic manipulation that induces time variation into the return distribution in order to influence measures that assume stationarity. This can enhance a portfolio's score even if the measure is calculated without any "estimation error" but without regard to the (unobservable) time variation in the portfolio's return distribution. The third type encompasses dynamic manipulation strategies that focus on inducing estimation error. Since measures have to be estimated with real world data, there are strategies that can induce positive biases in the resulting values. As a simple example, imagine an evaluator who estimates a fund's Sharpe ratio by calculating its excess return's standard deviation and average excess return over a 36-month period using monthly data. Further assume the fund manager simply wishes to maximize the expected value of the calculated Sharpe ratio. A simple strategy that accomplishes this has the fund sell an out of the money option in the first month, while investing the remaining funds in the risk free asset. If the option expires worthless the portfolio is then invested wholly in the risk free asset for the remaining 35 months. Whenever this happens the portfolio has a zero standard deviation, a positive excess return, and thus an infinite Sharpe ratio. Since there is a strictly positive probability that the option will expire worthless the expected value of the calculated Sharpe ratio must be infinity. Worse yet, as the article shows, even with goals other than maximizing the expected calculated performance value and concern for other moments of the test statistic, a manager can deliberately induce large economically significant measurement errors to his advantage. Also notice that increasing the frequency with which returns are observed does not shield the set of existing measures from this type of dynamic manipulation.

If the set of current measures is vulnerable to manipulation the question naturally arises as to whether a manipulation-proof measure can be constructed. To answer this question, one first needs to determine what "manipulation-proof" means. This article defines a manipulation-proof performance measure (MPPM) as one that has four properties:

1. The measure should produce a single valued score with which to rank each subject.
2. The score's value should not depend upon the portfolio's dollar value.
3. An uninformed investor cannot expect to enhance his *estimated* score by deviating from the benchmark portfolio. At the same time informed investors should be able to produce higher scoring portfolios and can always do so by taking advantage of arbitrage opportunities.
4. The measure should be consistent with standard financial market equilibrium conditions.

It turns out that these four requirements are enough to uniquely identify a manipulation-proof measure. The first condition eliminates measures that produce only partial rankings. It also eliminates useless measures like one that simply lists the set of available returns. The second condition simply implies that returns are sufficient statistics rather than dollar gains or losses. The third and fourth provide the actual structure. As the earlier discussion indicates there are several ways to manipulate a performance measure and the second condition implies that a MPPM must be immune to all of them. In particular that means the manager cannot expect to benefit by trying to alter the score's estimation based upon observable data. To accomplish this goal, the score must be (1) increasing in returns (to recognize arbitrage opportunities), (2) concave (to avoid increasing the score via simple leverage or adding unpriced risk), (3) time separable to prevent dynamic manipulation of the estimated statistic, and (4) have a power form to be consistent with an economic equilibrium.

If these four conditions are not met, then we show that there does not exist any MPPM. This negative result is potentially as important as the positive result, in that it implies no relative superiority among existing measures. Given any numeric statistic evaluating performance, unless the four above conditions are satisfied, there is always a way to enhance it without superior information. Hence there is no first-best solution to the problem of the manager's moral hazard—information-free manipulation is always possible.

The MPPM derived in this article is

$$\hat{\Theta} \equiv \frac{1}{(1-\rho)\Delta t} \ell\text{n} \left(\frac{1}{T} \sum_{t=1}^{T} [(1+r_t)/(1+r_{ft})]^{1-\rho} \right). \tag{1}$$

The $\hat{\Theta}$ statistic is an estimate of the portfolio's premium return after adjusting for risk. That is, the portfolio has the same score as does a risk-free asset whose continuously-compounded return exceeds the interest rate by $\hat{\Theta}$. Here T is the total number of observations, and Δt is the length of time between observations. These two variables serve to annualize the measure. The portfolio's (un-annualized) rate of return at time t is r_t, and the risk-free rate is r_{ft}. The coefficient ρ should be selected to make holding the benchmark optimal for an uninformed manager.

The measure is easy to calculate. As an example, consider a fund with monthly (*i.e.*, $\Delta^+ = 1/12$) returns of -10, 5, 17, and -2% when the monthly risk-free rate is 1%. If ρ is 2, then $\hat{\Theta} = 6.6\%$, and the fund has the same score as a risk-free asset with an annual rate of return of 20.4%. For ρ equal to 3, $\hat{\Theta} = 1.2\%$, and the fund is equivalent to a risk-free asset returning 14.0%. The score is higher when $\rho = 2$ because its risk is not so heavily penalized.

Our article was originally motivated by the question of whether the existing set of performance measures is sufficiently manipulation-proof for practical use. This is particularly relevant in the presence of transactions costs, which may offset whatever "performance" gains a manager might hope to generate from trading for the purpose of manipulating the measure. Therefore, our article explores how difficult it is to game the existing measures meaningfully. For the seven measures examined here—four ratios (Sharpe (1966), Sortino and van der Meer (1991), Leland (1999), and Sortino et al. (1999)), and three regression intercepts (the CAPM alpha, Treynor and Mazuy (1966), and Henriksson and Merton (1981))—the answer is not encouraging. Simple dynamic strategies that only relever the portfolio each measurement period or buy (very liquid) at the money options can produce seemingly spectacular results, even in the presence of very high transactions costs. For example, consider the Henriksson and Merton (1981) measure. A simple options trading scheme in the presence of transactions costs equal to 20% produces "very good" results. The final regression statistics report that the portfolio has returns that are superior to the market's nearly 65% of the time, and (using a 5% critical value) statistically significantly better 9% of the time. Obviously, lower and more realistic transactions costs only make matters worse. The other measures analyzed here are similarly susceptible to having their estimated values gamed.

While many of the measures currently in common use can be manipulated, there is one measure that is manipulation-proof under the necessary conditions we specify: the Morningstar Risk Adjusted Rating (Morningstar, 2006), introduced in July 2002. Morningstar evidently developed it to provide a measure that was more generally applicable and robust than its previous rating tools, and hence chose a measure resembling a representative utility function. While they were not seeking a manipulation-proof measure they found one.[4]

Our results have a number of implications for investment management. Hedge funds and other alternative investment vehicles have broad latitude to invest in a range of instruments, including derivatives. Mitchell and Pulvino (2001) document that merger arbitrage, a common hedge fund strategy, generates returns that resemble a short put-short call payoff. Recent research by Agarwal and Naik (2000) shows that hedge fund managers in general follow a number of different styles that are nonlinear in the returns to relevant indices. In a manner similar to Henriksson and Merton, Agarwal and Naik use option-like payoffs as regressors to capture these nonlinearities. In fact, option-like payoffs are inherent in the compensation-structure of the typical hedge fund contract. Goetzmann et al. (2003) show that the high water mark contract, the most common type in the hedge fund industry, effectively leaves the investor short 20% of a call option. The call is at-the-money each time it is "reset" by a payment and out-of-the money otherwise. Given the nonlinearity resulting from these kinds of strategies it is particularly useful to have a performance measurement tool that is not subject to informationless manipulation through the use of derivatives.

The article is structured as follows. Section 1 discusses various ways in which the Sharpe ratio and alpha can be manipulated. Section 2 discusses the manipulation of a number of other measures including those using reward to variability (Section 2.1) and the Henriksson-Merton and Treynor-Mazuy market timing measures (Section 2.2). Section 3 derives the MPPM, and Section 4 concludes.

1. Manipulation of the Sharpe Ratio and Alpha

The Sharpe ratio and Jensen's alpha were the first and probably are still the most widely used performance measures of the risk-reward and benchmark types. Although both are known to be subject to manipulation, they remain widely used and often cited performance measures. They, therefore, make

[4] Paul Kaplan, personal communication.

good examples to introduce performance manipulation in more detail. Other measures are examined in the next section.

1.1 Static analysis of the Sharpe ratio

A number of articles have shown that by altering the distribution function governing returns the statistical mean and variance can be manipulated to increase the Sharpe ratio to some degree. Ferson and Siegel (2001) look at a fairly general case (that includes potentially private information) while Lhabitant (2000) shows how using just a couple of options can generate seemingly impressive values. Since the focus of this article is on dynamic strategies and the development of a MPPM to prevent all sources of manipulation the only static result presented here is the characterization of the return distribution that maximizes the Sharpe ratio. Readers interested in additional details should consult one of the above cited articles.

The Appendix shows that a Sharpe ratio maximizing portfolio in a complete market is characterized by a state i return in excess of the interest rate (x_i^{MSR}) of

$$x_i^{\mathrm{MSR}} = \overline{x}_{\mathrm{MSR}} \left[1 + \frac{1 - \hat{p}_i / p_i}{S_{\mathrm{MSR}}^2} \right]. \tag{2}$$

Here $S_{\mathrm{MSR}} \equiv \left[\sum \hat{p}_i^2 / p_i - 1 \right]^{1/2}$ is the maximum possible Sharpe ratio, and p_i and $\hat{p}_i$ are the true and risk-neutral probabilities of state i. Since the Sharpe ratio is invariant to leverage, we are free to specify the portfolio's mean excess return,$\overline{x}_{\mathrm{MSR}}$, at any level desired.

While the maximal-Sharpe-ratio (MSR) has a number of interesting properties it is also true that one way to mitigate its impact on a fund's apparent performance is to sample returns more frequently. As one does so the percentage difference between the market portfolio's Sharpe ratio and that of the MSR goes to zero. However, as will be seen below, frequent sampling does not help when managers use dynamic manipulation strategies.

1.2 Dynamic analysis of the Sharpe ratio

Calculated values of the Sharpe ratio and virtually all other performance measures use statistics based on the assumption that the reported returns are independent and identically distributed. While this may be a good description of typical portfolio returns in an efficient market, it clearly can be violated if a portfolio's holdings are varied dynamically depending on its performance.

Consider a money manager who has been lucky or unlucky with an average realized return high or low relative to the portfolio's realized variance. To maximize the Sharpe ratio the portfolio can now be modified to take into account the difference in the distributions between the realized and future expected returns, just as if the returns distribution was not *ex ante* identically distributed each period. This dynamic manipulation makes the past and future returns dependent when computed by an unconditional measure.

To illustrate, suppose the manager has thus far achieved an historical average excess return of $\overline{x}_h$ with a standard deviation of σ_h. The portfolio's average excess return and standard deviation in the future are denoted by $\overline{x}_f$ and σ_f. Then its measured Sharpe ratio over the entire period will be will be

$$S = \frac{\gamma\overline{x}_h + (1-\gamma)\overline{x}_f}{\sqrt{\gamma(\overline{x}_h^2 + \sigma_h^2) + (1-\gamma)(\overline{x}_f^2 + \sigma_f^2) - [\gamma\overline{x}_h + (1-\gamma)\overline{x}_f]^2}}$$
$$= \frac{\gamma\overline{x}_h + (1-\gamma)\overline{x}_f}{\sqrt{\gamma\overline{x}_h^2(1 + 1/S_h^2) + (1-\gamma)\overline{x}_f^2(1 + 1/S_f^2) - [\gamma\overline{x}_h + (1-\gamma)\overline{x}_f]^2}} \quad (3)$$

where γ is the fraction of the total time period that has already passed, and S_h and S_f are the past and future Sharpe ratios.

By inspection, the overall Sharpe ratio is maximized by holding in the future a portfolio that maximizes the future Sharpe ratio, $S_f = S_{\mathrm{MSR}}$. However, regardless of what Sharpe ratio can be achieved in the future, leverage is no longer irrelevant as it was in the static case. Maximizing the Sharpe ratio in Equation (3), we see that the optimal leverage gives a target mean excess return of

$$\overline{x}_f^* = \begin{cases} \overline{x}_h(1 + S_h^{-2})/(1 + S_f^{-2}) & \text{for} \quad \overline{x}_h > 0 \\ \infty & \text{for} \quad \overline{x}_h \leqslant 0. \end{cases} \quad (4)$$

If the manager has been lucky in the past and achieved a higher than anticipated Sharpe ratio, $S_h > S_f$, then the portfolio should be targeted in the future at a lower mean excess return (and lower variance) than it has realized. This allows the past good fortune to weigh more heavily in the overall measure.[5] Conversely, if the manager has been unlucky and $S_h < S_f$, then in the future, the portfolio should be targeted at a higher mean excess return than that so far realized. In the extreme, if the average excess return realized has been negative, then the manager should use as much leverage as possible to minimize the impact of the poor history.

[5] For a fixed Sharpe ratio, the overall mean is linear in the future leverage while the overall standard deviation is convex. The proportional changes in the mean and standard deviation with respect to future leverage are equal when the future and historical Sharpe ratios are equal. Therefore, when the historical Sharpe ratio is less than the future Sharpe ratio, increasing leverage increases the overall mean at a faster rate than the standard deviation and vice versa.

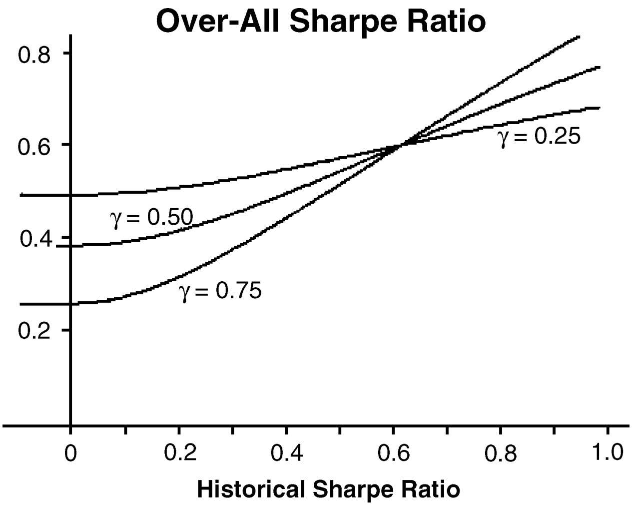

Figure 1
Two-period Sharpe ratio optimization when the evaluator knows the portfolio's overall mean and variance
The horizontal axis represents the portfolio's historical Sharpe ratio during the time interval 0 to γ. After seeing these results, the manager can then rebalance the portfolio to optimize the Sharpe ratio calculated over the entire interval 0 to 1. The vertical axis displays the resulting expected Sharpe ratio from time 0 to time 1.

The Sharpe ratio that can be achieved over the entire period is[6]

$$S = \begin{cases} \sqrt{\dfrac{S_{\mathrm{MSR}}^2 S_h^2 + \gamma S_h^2 + (1-\gamma) S_{\mathrm{MSR}}^2}{1 + (1-\gamma) S_h^2 + \gamma S_{\mathrm{MSR}}^2}} & \text{for} \quad S_h > 0 \\ S_{\mathrm{MSR}} \sqrt{\dfrac{1-\gamma}{1+\gamma S_{\mathrm{MSR}}^2}} & \text{for} \quad S_h \leqslant 0. \end{cases} \tag{5}$$

Figure 1 plots the overall Sharpe ratio as a function of the realized historical Sharpe ratio for histories of different durations. The overall Sharpe ratio can, on average, be maintained above (is forced below) the theoretical maximum, S_{MSR}, whenever the past performance has been good (bad). The realized Sharpe ratio, of course, has the most impact when the past history is long ($\gamma \approx 1$).

As can be seen in the graph or from Equation (5), the over-all Sharpe ratio is increasing and convex in the historical Sharpe ratio. Therefore, dynamic-Sharpe-ratio-maximizing strategies should, on average, be able to produce a Sharpe ratio higher than Σ_{MSR}.[7]

---

[6] Because the optimal future leverage is infinite when $\overline{x}_h \leqslant 0$, the past Sharpe ratio does not affect the overall Sharpe ratio, and the value of the overall Sharpe ratio for $S_h < 0$ is the same as for$S_h = 0$.

[7] Brown et al. (2004) provide evidence that some Australian money managers engage in a pattern of trading consistent with the behavior described here.

Simulations show that dynamic manipulation of the Sharpe ratio can have a substantial effect. In a lognormal market with a risk premium of 12%, a manager re-levering his portfolio after 30 months of a 5-year evaluation period can achieve a Sharpe ratio of 0.714. This is 18% higher than the static MSR (and 13% higher than the standard biased estimate of the ratio). In addition, dynamic manipulation works in exactly the same way for portfolios that have not been Sharpe optimized each period. Any portfolio with a Sharpe ratio of Σ_f can be levered as shown in Equation (4) to achieve a higher average Sharpe ratio than this on average. For example, in the same lognormal market, if the market is re-levered after 30 months, its Sharpe ratio improves on average from 0.597 to 0.672, an increase of 13%.

In addition, we have assumed here that a single measurement period is shorter than the time between rebalancings. If the portfolio manager can alter his portfolio within a single measurement period, then the distribution of a single return can be affected. For example, if a portfolio increases sharply in value during the first part of a measurement period, switching to a less aggressive position may make the ultimate measured return smaller and keep the Sharpe ratio high. Goetzmann et al. (2000) analyze the problem of performance measurement when the investment rebalancing period is shorter than the measurement period.

Particular circumstances may permit additional opportunities to manipulate the Sharpe ratio of a portfolio dynamically. Smoothing returns over time will leave the portfolio's mean return unchanged but decrease its variance, so smoothing returns will increase the Sharpe ratio. Funds with illiquid assets whose prices are only reported occasionally may benefit here. Hedge funds or other portfolios with high-water mark performance fees will also benefit. A standard performance contract calls for expensing the performance fee monthly but paying it only annually, but unpaid fees can be "lost" based on poor later fund performance. This contractual provision automatically moves recorded returns from periods of good performance to periods of poor performance thereby smoothing the reported returns and increasing the Sharpe ratio.

1.3 Jensen's alpha and related measures

Several benchmark performance measures are related to the Sharpe ratio. The most familiar of these is Jensen's alpha, which is the intercept of the market model regression, $x_t = \alpha + \beta x_{mt} + \varepsilon_t$. If the CAPM holds, this intercept term should be zero when the regression is expressed in excess return form.

Jensen's alpha measures the marginal impact to the Sharpe ratio of combining a small amount of an asset with the market portfolio.[8]

[8] If asset i is combined with the market into a portfolio $wx_i + (1 - w)x_{\mathrm{mkt}}$, then $(\partial S/\partial w)|_{w=0} = \alpha_i/\sigma_{\mathrm{mkt}}$.

Consequently, any portfolio with a Sharpe ratio in excess of the market's must have a positive alpha. In particular, the alphas of the maximal Sharpe ratio portfolios described above are[9]

$$\alpha_{\mathrm{MSR}} = \overline{x}_{\mathrm{MSR}}(1 - S^2_{\mathrm{mkt}}/S^2_{\mathrm{MSR}}) > 0\,. \tag{6}$$

Clearly, alpha is subject to severe manipulation. A MSRP can be created with any desired leverage, so its alpha can in theory be made as large as desired by levering.

The Treynor (1965) ratio and the Treynor appraisal ratio were introduced in part to negate this leverage effect on alpha. The Treynor measure is the ratio of alpha to beta[10] while the appraisal measure is the ratio of alpha to residual standard deviation. Like the Sharpe measure, and unlike alpha, the two Treynor measures are unaffected by leverage.

Both Treynor measures indicate superior performance for the MSRP. The MSRP's two Treynor ratios are[11]

$$\begin{aligned} T_{\mathrm{MSR}} &\equiv \tfrac{\alpha_{\mathrm{MSR}}}{\beta_{\mathrm{MSR}}} = (S^2_{\mathrm{MSR}}/S^2_{\mathrm{mkt}} - 1)\overline{x}_{\mathrm{mkt}} > 0 \\ A_{\mathrm{MSR}} &\equiv \frac{\alpha_{\mathrm{MSR}}}{\sqrt{\mathrm{Var}[x_{\mathrm{MSR}} - \beta_{\mathrm{MSR}} \cdot x_{\mathrm{mkt}}]}} = \sqrt{S^2_{\mathrm{MSR}} - S^2_{\mathrm{mkt}}} > 0\,. \end{aligned} \tag{7}$$

Any MSRP has the largest possible Treynor appraisal ratio, but Treynor ratios in excess of the MSRP's can be achieved by forming portfolios with positive alphas and betas close to zero.[12]

In addition, since the Sharpe ratio is subject to dynamic manipulation, alpha and both Treynor measures are as well and can be increased above these statically achieved values. The manipulation is illustrated in Figure 2.

---

9 The beta of the maximal-Sharpe-ratio portfolio, x_{MSR}, is

$$\beta_{\mathrm{MSR}} \equiv \frac{\mathrm{Cov}[x_{\mathrm{MSR}}, x_{\mathrm{mkt}}]}{\mathrm{Var}[x_{\mathrm{mkt}}]} = \frac{\mathrm{Cov}[x_{\mathrm{MSR}}, x_{\mathrm{mkt}}]}{\mathrm{Var}[x_{\mathrm{MSR}}]}\frac{\mathrm{Var}[x_{\mathrm{MSR}}]}{\mathrm{Var}[x_{\mathrm{mkt}}]} = \frac{\overline{x}_{\mathrm{mkt}}}{\overline{x}_{\mathrm{MSR}}}\frac{\mathrm{Var}[x_{\mathrm{MSR}}]}{\mathrm{Var}[x_{\mathrm{mkt}}]} = \frac{\overline{x}_{\mathrm{MSR}}}{\overline{x}_{\mathrm{mkt}}}\frac{S^2_{\mathrm{mkt}}}{S^2_{\mathrm{MSR}}}$$

where S_{mkt} is the Sharpe ratio of the market. The third equality follows since the MSRP is mean-variance efficient by definition so $\overline{x}_{\mathrm{mkt}} = \overline{x}_{\mathrm{MSR}} \cdot \mathrm{Cov}[x_{\mathrm{MSR}}, x_{\mathrm{mkt}}]/\mathrm{Var}[x_{\mathrm{MSR}}]$.

10 Treynor (1965) original definition of his measure was $r_f - \alpha/\beta$; however, α/β is now the commonly accepted definition.

11 The Treynor ratio can be computed from the MSRP's α in Equation (6) and its β in footnote 9. To compute the Treynor appraisal ratio, we use α and

$$\begin{aligned} &\mathrm{Var}[x_{\mathrm{MSR}} - \beta_{\mathrm{MSR}} x_{\mathrm{mkt}}] = \mathrm{Var}[x_{\mathrm{MSR}}] - 2\beta_{\mathrm{MSR}}\mathrm{Cov}[x_{\mathrm{MSR}}, x_{\mathrm{mkt}}] + \beta^2_{\mathrm{MSR}}\mathrm{Var}[x_{\mathrm{mkt}}] \\ &\quad = \mathrm{Var}[x_{\mathrm{MSR}}] - \beta^2_{\mathrm{MSR}}\mathrm{Var}[x_{\mathrm{mkt}}] = \frac{\overline{x}^2_{\mathrm{MSR}}}{S^2_{\mathrm{MSR}}} - \left(\frac{\overline{x}_{\mathrm{MSR}}}{\overline{x}_{\mathrm{mkt}}}\frac{S^2_{\mathrm{mkt}}}{S^2_{\mathrm{MSR}}}\right)^2 \frac{\overline{x}^2_{\mathrm{mkt}}}{S^2_{\mathrm{mkt}}} = \frac{\overline{x}^2_{\mathrm{MSR}}}{S^2_{\mathrm{MSR}}}\left(1 - \frac{S^2_{\mathrm{mkt}}}{S^2_{\mathrm{MSR}}}\right). \end{aligned}$$

12 A portfolio that shorts β_{MSR} dollars in the market for every dollar in the MSRP will have a beta of zero and a positive alpha. Therefore, its Treynor ratio will be $+\infty$.

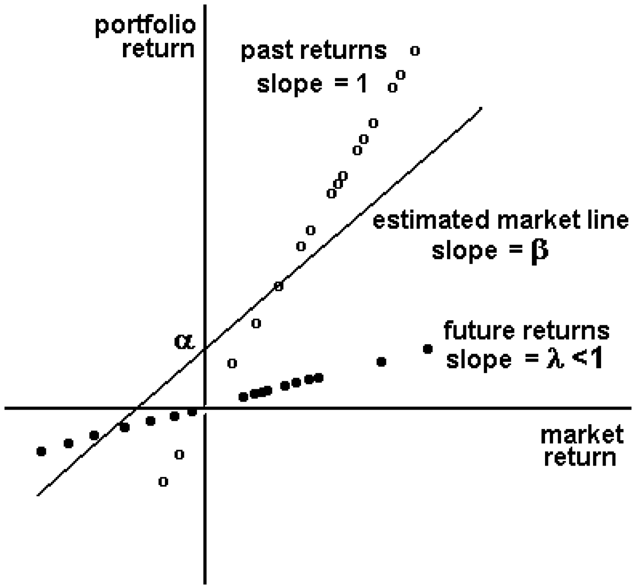

Figure 2
Illustration of two-period alpha manipulation
The horizontal and vertical axes are the market's and the portfolio's excess returns. The portfolio is originally fully invested in the market (i.e., with a market exposure of $\lambda = 1$). In the past, the market return (∘) has been higher than expected. To manipulate alpha, the market exposure in the future should be decreased to $\lambda > 1$ so future returns (•) will lie on a shallower line. The estimated market line using all the data will have a slope between 1 and λ and a positive α. Conversely, if history had produced below-average returns, the market exposure should be increased.

If returns have been above average, then decreasing the leverage in the future will generate a market line with a positive alpha from all the data for the portfolio. Not surprisingly, this is very similar to the dynamic manipulation that produces a superior Sharpe ratio—leverage is decreased (increased) after good (bad) returns.

Alpha-like measures can also be computed from models other than the CAPM. Under quite general conditions, the generalized alpha of an asset or portfolio in a single-period model is[13]

$$\alpha_p^{\text{gen}} = \overline{x}_p - B_p\overline{x}_m \qquad \text{where} \quad B_p = \frac{\text{Cov}[u'(1 + r_f + \tilde{x}_m), \tilde{x}]}{\text{Cov}[u'(1 + r_f + \tilde{x}_m), \tilde{x}_m]}. \quad (8)$$

B_p is a generalized measure of systematic risk, r_f is the per-period (not-annualized) interest rate, and $u(\cdot)$ is the utility function of the representative investor holding the market. The systematic risk coefficient, B_p, can be estimated by regressing x on x_m using u' $(\cdot)$ as an instrumental variable. Ingersoll (1987) and Leland (1999) suggest using a power utility function

[13] See Ingersoll (1987) for the derivation of this general measure of systematic risk.

where

$$u'(1+r_f+x_m) = (1+r_f+x_m)^{-\rho} \quad \text{with}$$
$$\rho = \frac{\ell\mathrm{n}[\mathbb{E}(1+r_f+\tilde{x}_m)] - \ell\mathrm{n}(1+r_f)}{\mathrm{Var}[\ell\mathrm{n}(1+r_f+\tilde{x}_m)]}. \qquad (9)$$

This generalized alpha gives a correct measure of mispricing assuming that the representative utility function is correctly matched to the market portfolio; that is, if the market portfolio does maximize the utility function employed. But this statement, tautological as it is, only applies to single-period or static manipulation. Even if the utility is correctly matched to the market, strategically rebalancing the portfolio over time can give an apparent positive alpha due to the deviation between the average and a properly conditioned expectation. The basic technique is the same—decrease (increase) leverage after good (bad) returns.

Beyond the single-factor CAPM, there are several alpha-like performance measures in use. These models augment the single factor CAPM model with additional risk factors such as the Fama-French factors. However, within the simulations discussed in the next sub-section, no such factors exist. The market returns are generated in an environment where Equation (8) with $\alpha^{\text{gen}} = 0$ is the correct way to price. This single factor model should do at least as well in the simulated environment as any model with additional, but within the simulation, unpriced, factors. Similarly, if a multifactor model were simulated, a multidimensional rebalancing of the portfolio should produce positive alphas. For this reason, the Chen and Knez (1996) measure and similar measures have not been examined.

1.4 Manipulating the Sharpe ratio and alpha with transactions costs

In practice, a manager can change a portfolio's characteristics far more frequently than once during the typical measurement period. On the other hand, the costs of transacting may eliminate the apparent advantage of many manipulation strategies. As we show below even very high transactions costs cannot prevent managers from manipulating the Sharpe ratio or other performance measures.

Determining the optimal dynamic manipulation strategies for all of the popular performance measures is beyond the scope of our article. The optimal manipulation strategy depends on the size of the transactions costs, the complete set of returns to date within the evaluation period, the distribution of future returns, and the number of periods remaining in the evaluation period. However, for our purposes, an optimal strategy is not necessary. We only want to establish whether reasonably simple trading strategies can distort the existing measures even in the presence of transactions costs, and if so, by how much.

Table 1 shows the Sharpe ratio performance of a dynamically rebalanced portfolio. The portfolio is invested only in the market; no derivatives are used to alter the distribution. Only the leverage is changed over time. Initially the portfolio is invested fully in the market. After the first year, the portfolio is levered at the beginning of each month so that the target mean is given by Equation (4); however, leverage is restricted so that the market exposure is constrained between 50 and 150% of the portfolio's value. Levering is achieved by buying or selling synthetic forward contracts consisting of a long position in calls and a short position in puts that are at-the-money in present value, that is, the strike price per dollar invested in the market is $e^{r/12}$. Trades in these two options are assessed a round-trip transactions cost of 0, 10, or 20%. The simulation consisted of 10,000 repetitions.

On average the dynamically-manipulated portfolio's Sharpe ratio is 13% higher than the market's and 7% higher than the static MSRP's. In the 10,000 trials, the manipulated portfolio had a Sharpe ratio higher than the market's 82.6% of the time in the absence of transactions costs.[14] Even with a 20% round-trip transactions cost, the dynamically-manipulated portfolio still beat the market almost three-quarters of the time.

The Sharpe-manipulated portfolio beats the market frequently, but is it substantially better statistically; that is, is the difference in the Sharpe ratios significant? In practice this question is seldom asked because Sharpe ratios are at least as difficult to estimate precisely as are mean returns.[15]

However, differences between two portfolios' Sharpe ratios can be more precisely estimated than either of the individual ratios if the underlying returns are correlated, as would generally be true and is certainly true in our example. Statistical tests of portfolio returns generally assume that the returns are independent and identically distributed over time. Such is not the case here for the manipulated portfolio and this makes deriving the distribution of the difference in the Sharpe ratios a difficult task.

[14] The frequency with which one portfolio beats another must be interpreted with caution when the winning portfolios hold derivatives. To illustrate, consider two portfolios that are almost identical. The only difference is that the second portfolio sells deep out-of-the-money options (or similar rare payoff derivatives). The proceeds are invested in bonds. Whenever the option expires out-of-the-money, the second portfolio's returns will be greater than the first portfolio's returns. This will lead to a "better" outcome for the second portfolio by almost any performance measure. Since the payoff event can be made as rare as desired, any portfolio, even an *ex ante* optimal one, can be beaten often. In our simulations here only the leverage is changed so they are not subject to this problem. In addition, the appropriate null hypothesis is not that the benchmark's Sharpe ratio exceeds that of another portfolio half the time. If the benchmark is more mean-variance efficient than the manipulated portfolio, its true Sharpe ratio will be larger and, with symmetric measurement errors, its sample Sharpe should be greater more than 50% of the time. This is in contrast to alpha, for which $\text{Prob}\{\alpha > 0\} = 50\%$ is an appropriate null hypothesis.

[15] Lo (2002) has shown that the asymptotic standard error of the Sharpe ratio of a portfolio with iid returns over T periods is $[(1 + S^2)/T]^{1/2}$ so the coefficient of determination is $S^{-1}[(1 + S^2)/T]^{1/2}$. The coefficient of determination for the mean return is $\sigma T^{-1/2}\overline{x}^{-1} = S^{-1}T^{-1/2}$. The former is larger by a factor of $(1 + S^2)^{1/2}$so the same difference in percentage terms is less significant for a Sharpe ratio than it is for an average return.

Table 1
Dynamic manipulation of the Sharpe ratio and related performance measures

	Sharpe ratio Σ		Jensen α		Treynor ratio T		Appraisal ratio A		General alpha α^{gen}		Portfolio characteristics	
Trans. cost (%)	Mean (std dev)	Freq $S > S_{\text{mkt}}$ signif (%)	Mean (std dev) (%)	Freq. $\alpha > 0$ signif. (%)	Mean (std dev) (%)	Freq. T > 0 signif. (%)	Mean (std dev)	Freq. A > 0 signif. (%)	Mean (std dev) (%)	Freq. $\alpha^{\text{gen}} > 0$ signif. (%)	Market exposure	Beta
0	0.673	82.6	2.05	92.4	2.39	92.4	0.365	92.4	1.94	90.6	0.955	0.939
	(0.439)	**20.4**	(1.51)	**37.6**	(1.77)	**37.6**	(0.259)	**40.7**	(1.54)	**34.0**	(0.323)	(0.314)
10	0.649	74.6	1.66	86.4	1.91	86.4	0.294	86.4	1.50	82.7	0.960	0.943
	(0.425)	**13.3**	(1.56)	**27.9**	(1.77)	**28.5**	(0.266)	**30.3**	(1.60)	**25.3**	(0.321)	(0.312)
20	0.647	74.0	1.61	85.8	1.86	85.8	0.288	85.8	1.53	83.1	0.965	0.947
	(0.427)	**13.1**	(1.55)	**27.2**	(1.75)	**27.6**	(0.263)	**29.0**	(1.59)	**24.3**	(0.321)	(0.312)
Mkt	0.597	—				True value of market Sharpe ratio = 0.597						
	(0.454)											

This table presents the annualized Sharpe ratio, Jensen alpha, Treynor ratio, Treynor appraisal ratio, and generalized alpha for dynamically manipulated portfolios over a 60-month period. The Sharpe ratio and Treynor appraisal ratio are annualized by multiplying the monthly measure by the square root of 12. The alpha, generalized alpha, and Treynor ratio are annualized by multiplying by 12. The measures are computed from 10,000 simulations. The average and the standard deviation of each measure are reported. The "freq." columns have two numbers showing the fraction of times that the portfolio showed superior performance ($S > S_{\text{mkt}}$, $\alpha > 0$, $T > 0$, $A > 0$, or $\alpha^{\text{gen}} > 0$) and the approximate fraction of times the portfolio's performance was significantly superior at the 5% level. The latter number was estimated as the frequency with which the number was more than 1.65 standard deviations above the null hypothesis value.

The portfolio is always invested in the market. Derivatives are used only to change the market exposure. Initially, the portfolio is fully invested in the market. After 12 months, the leverage is adjusted each month based on the performance history as described in Equation (4) to set the target mean to $\overline{x} = \overline{x}_h(1 + S_h^{-2})(1 + S_{\text{mkt}}^{-2})$; adjustments are restricted to keep the market exposure between 50% and 150% of the portfolio's value. The average and standard deviation of the market exposure and resulting beta are also reported. Each portfolio adjustment requires paying a transaction cost computed as 0%, 10%, or 20% (round trip) of the value of the synthetic forward contracts (a call plus a put struck at-the-money in present value terms) needed to change the market exposure.

The market parameters are: risk free rate 5% per annum, market premium 12%, market standard deviation 20%.

Furthermore, there is little point in such a derivation as our simulation is a constructed example, and the statistic we derive would be applicable only to that case.

Fortunately we do not need to know the distribution. Our simulations give us a sample distribution of the differences, so we can easily determine how often the difference is k standard deviations above or below zero. For example, with no transactions costs, the difference between the manipulated Sharpe ratio and the market Sharpe ratio was more than 1.65 standard deviations above zero 20.4% of the time. It was never more than 1.65 standard deviations below zero. Were the difference normally distributed, each of these should have occurred 5% of the time. Of course, the differences between the Sharpe ratios is not normal, but this is still evidence that a properly constructed test would conclude that the difference was significant more often than chance would prescribe if the true expected difference were zero. Table 1 gives the percentages of times that the Sharpe ratio difference exceeds 1.65 standard deviations.

Using Jensen's alpha or one of the Treynor measures to evaluate performance gives even "better" results. The dynamic portfolio has an average annualized alpha over 2% and the alpha is positive more than 92% of the time in the absence of transactions costs. Even with 20% round-trip costs, the average alpha is still 1.6% and is positive over 85% of the time. The generalized alpha shows only slightly weaker performance for the manipulated portfolio. In the simulations we had the benefit of knowing the correct utility function, but superior performance as measured by the generalized alpha is insensitive to the value assumed for the risk aversion parameter. Similar superior performance was found for risk aversion parameters throughout the range 2 to 4.

It is not surprising that the alphas can be manipulated more easily than can the Sharpe ratios. The CAPM null hypothesis is that $\alpha = 0$. With measurement error, positive and negative deviations are approximately equally likely. However, the CAPM null hypothesis on the Sharpe ratio is that it is less than the market's. The manipulation on the Sharpe ratio has first to make up this difference before beating it.

2. Manipulating Other Measures

2.1 Reward-to-variability measures

Many other performance measures have been proposed over the years to correct perceived flaws in the Sharpe ratio or to extend or modify its measurement. Some of these measures use a benchmark portfolio—usually some market index. All of them are subject to the same type of manipulation that can be used on the Sharpe ratio. In this and the next sections we examine many of the more popular alternatives.

Modigliani and Modigliani (1997) M-squared score is simply a restatement of the Sharpe ratio. The M-squared measure is the expected excess return that would be earned on a portfolio if it were levered so that its standard deviation was equal to that on the benchmark. Clearly maximizing the Sharpe ratio also maximizes the M-squared measure relative to any benchmark. Therefore, it is also subject to the same manipulation.

Sharpe's information ratio is another simple variation on the original Sharpe ratio. The difference is that the excess returns are calculated relative to a risky benchmark portfolio rather than the risk-free rate. If x and x_b are the excess returns on the portfolio and the benchmark, the information ratio is

$$S_{\text{infomation}} = \frac{\overline{x} - \overline{x}_b}{\sqrt{\mathrm{Var}(\tilde{x} - \tilde{x}_b)}} \,. \tag{10}$$

The information ratio is the primary measure in Roll (1992) tracking-error (TEV) rating. Since excess returns are the returns on zero net cost (or arbitrage) portfolios, they can be combined simply by adding them together without any weighting. Clearly the arbitrage portfolio that is a combination of the excess returns $\tilde{x}$ and $\tilde{x}_b$ has an information ratio with respect to the benchmark, $\tilde{x}_b$, numerically equal to the Sharpe ratio of $\tilde{x}$ alone; therefore, the information ratio is subject to the same manipulation as the Sharpe ratio. In particular, adding the MSRP excess returns to the market portfolio will achieve the highest possible information ratio relative to the market.

One criticism commonly leveled against the Sharpe ratio is that very high returns are penalized because they increase the standard deviation more than the average. This is the reason the MSRP has bounded returns. To overcome this problem, it has been suggested to measure risk using only "bad" returns. In particular, Sortino and van der Meer (1991) and others have measured risk as the root-mean-square deviation below some minimum acceptable return. Sortino et al. (1999) have further suggested that the "reward" in the numerator should only count "good" returns. Sortino's downside-risk and Sortino, van der Meer, and Plantinga's (SVP) upside-potential Sharpe-like measures are

$$D = \frac{\mathbb{E}[\tilde{x}]}{(\mathbb{E}[\mathrm{Min}^2(\tilde{x} - \underline{x}, 0)])^{1/2}} \qquad U = \frac{\mathbb{E}[\mathrm{Max}(\tilde{x} - \underline{x}, 0)]}{(\mathbb{E}[\mathrm{Min}^2(\tilde{x} - \underline{x}, 0)])^{1/2}} \,. \tag{11}$$

The minimal acceptable excess return, $\underline{x}$, is commonly chosen to be zero.[16]

While these two measures do avoid the problem inherent in the Sharpe ratio of penalizing very good outcomes, they do this *too well* so that

[16] The minimum acceptable return is sometimes set to the portfolio's average return for the Sortino measure. This makes the denominator the semi-standard deviation. The average return should not be used like this for the SVP measure as the ratio is then invariant to a uniform upward shift in the entire probability distribution.

the highest possible returns are sought in lieu of all others. As shown in the Appendix, the Sortino downside-risk and SVP upside-potential maximizing portfolios are very similar. They both hold the MSRP, an extra very large investment in the state security for the state with the highest market return (the lowest likelihood ratio $\hat{p}_I/p_I \leqslant \hat{p}_i/p_i$), and bonds.

$$x_i^{\text{MDR}} = \frac{\overline{x}}{D_{\text{MDR}}^2}\left[1-(1+S_{\text{MSR}}^2)\frac{p_I}{\hat{p}_I}\right] + \frac{S_{\text{MSR}}^2}{D_{\text{MDR}}^2}\frac{p_I}{\hat{p}_I}x_i^{\text{MSR}}\; i < I$$

$$x_I^{\text{MDR}} = \frac{\overline{x}}{D_{\text{MDR}}^2}\left[1-\frac{\hat{p}_I}{p_I}(1-D^2)\right]$$

$$x_i^{\text{MUP}} = \left[\frac{\hat{p}_I}{p_I}U^2-(1-\hat{p}_I)\right]^{-1} S_{\text{MSR}}^2[x_i^{\text{MSR}}-\overline{x}(1+S_{\text{MSR}}^2)]\; i < I$$

$$x_I^{\text{MUP}} = \frac{\overline{x}}{p_I}\left[1-\frac{p_I(1-\hat{p}_I)}{\hat{p}_I U_{\text{MUP}}^2}\right] \tag{12}$$

where $D_{\text{MDR}}^2 = 1+\frac{p_I^2}{\hat{p}_I^2}\sum_{i=1}^{I}\hat{p}_i^2/p_i - 2\frac{p_I}{\hat{p}_I}$ and $U_{\text{MUP}}^2 = \frac{p_I^2}{\hat{p}_I^2}\sum_{i=1}^{I-1}\hat{p}_i^2/p_i$. The D_{MDR} and U_{MUP} are the largest possible Sortino downside-risk and SVP upside-potential ratios.

Figure 3 illustrates the maximal Sortino and SVP ratio portfolios. The 1-month return on the market has a twenty-period binomial distribution that approximates a lognormal market with a continuously-compounded risk premium of 12% and a logarithmic volatility of $\sigma = 20\%$.[17] There are twenty periods over the 1 month and twenty-one final states. Both portfolios achieve a very high return in the best market outcome state and have a negative excess return in all but this best state.[18] The annualized Sortino and SVP ratios of the market are 1.00 and 2.86. The scores for the optimized Sortino ratio portfolio are over twice as high, at 2.83 and 6.04. The same is true for the optimized SVP portfolio; it scores 2.76 and 6.17.

Clearly the maximal Sortino and SVP ratio portfolios are far from optimal, and even a cursory examination of the period-by-period returns would likely attract very negative attention. For example, both of these portfolios have Sharpe ratios that are less than a tenth that of the market. Conversely the MSRP has a Sortino and SVP ratio just a bit below that of the market.

[17] A binomial market environment is used as an illustration here because in a lognormal market or any market with a continuous distribution, the maximal Sortino and SVP portfolios have an infinitesimal negative excess return except in the highest return "state." In this state, which occurs with zero probability in a market with a continuum of states, the return is infinite.

[18] For any choice of $\underline{x}$, both portfolios have returns less than $\underline{x}$ in all but the best state.

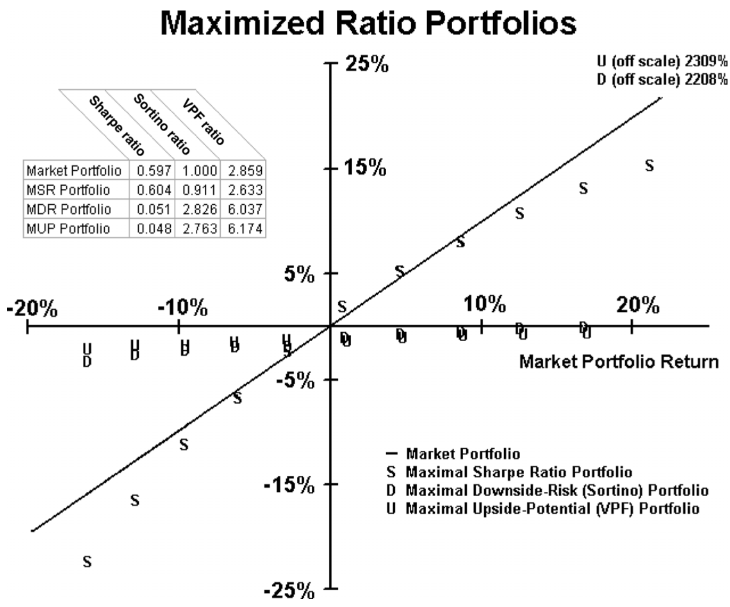

Figure 3
Comparison of Maximal Sortino and SVP ratio portfolios to the market portfolio
The x-axis is the market return. S, D, and U are the returns on maximal-Sharpe-ratio, maximal-downside-risk-ratio (Sortino), and maximal-upside-potential-ratio (SVP) portfolios. The market environment is a binomial model approximation to a lognormal market with a continuously compounded risk premium of 12% and a logarithmic volatility of $\sigma = 20\%$. There are twenty periods over the 1 month.

The maximal-ratio strategies are also risky in a sampling sense. The best market state is likely to have only a small probability of occurring, and should it fail to occur during the evaluation period, the sample averages and therefore the ratios would be negative even in a bull market. On the other hand, should no below-$\underline{x}$ outcomes occur, both ratios would be infinite. Of course, the maximal-ratio portfolios need not be employed to manipulate the measures. The payoffs in Figure 2 indicate that a high ratio can probably be achieved merely by purchasing calls—the further out-of-the-money the better. For example, in our simulation environment, simply investing the entire portfolio in at-the-money market index calls each month has a Sortino ratio of 1.20 and a SVP ratio of 3.68—substantially higher than the market's. Using 10% out-of-the-money calls increases these ratios to 1.79 and 5.10.

As with the Sharpe ratio, dynamic manipulation can also increase the measure or increase the measure further when used along with static manipulations. The following simple scheme can often achieve an infinite ratio when the minimal acceptable excess return is zero (or negative). If the first excess return is positive (which should be true more than one-half of the time), then holding the risk-free asset for every other period will give a positive numerator and a denominator of zero resulting in an infinite measure.

To illustrate the dynamic manipulation possibilities of these two measures even with very restricted portfolios, we employ simulations similar to those used for the Sharpe ratio. If the historical returns up through time t_0 have an average excess return of $\overline{x}_h$ and a Sortino measure of D_h, then the portfolio's measured Sortino ratio over the entire period will be

$$\begin{aligned} D &= \frac{\gamma \overline{x}_h + (1-\gamma)\overline{x}_f}{\sqrt{\gamma \sum_{t<t_0} \mathrm{Max}^2(-x_t, 0) + (1-\gamma)\sum_{t>t_0} \mathrm{Max}^2(-x_t, 0)}} \\ &= \frac{\gamma \overline{x}_h + (1-\gamma)\overline{x}_f}{\sqrt{\gamma \overline{x}_h^2 D_h^{-2} + (1-\gamma)\overline{x}_f^2 D_f^{-2}}} \end{aligned} \tag{13}$$

where the subscript f denotes future performance. As with the Sharpe ratio, the overall Sortino ratio is maximized by maximizing the Sortino ratio in the future and selecting the leverage so that

$$\overline{x}_f^* = \begin{cases} \overline{x}_h D_f^2 / D_h^2 & for \quad \overline{x}_h > 0 \\ \infty & for \quad \overline{x}_h \leqslant 0. \end{cases} \tag{14}$$

Table 2 shows the Sortino and SVP ratio performance of a dynamically rebalanced portfolio under the same conditions used for the Sharpe ratio. Again no derivative assets, which could create static manipulation of the ratios, were used to change the return distribution. The portfolio was fully invested in the market, and each month after the 12th month, leverage was changed using synthetic forwards to target the expected return as described in Equation (14). Again the market exposure was constrained to the range of 50 to 150% of the portfolio's value. The manipulated portfolios produce statistics that are superior to simply holding the market portfolio for both modified measures as well as the Sharpe ratio.[19]

The Sortino ratio is higher than the market's more than 82% of the time even with 20% round-trip transactions costs. It is significantly higher at the 5% confidence level almost 14% of the time.

The Sharpe and SVP ratios are also higher nearly three-quarters of the time for this manipulation. The same manipulation produces good results for all three ratios because each of the ratios has the same general dynamic manipulation rule—decrease leverage after good luck (after the historical score has been high).

[19] The simulations also show that for a base lognormal distribution, the sample Sortino and SVP ratios are almost certainly biased (like the Sharpe ratio for the MSRP and unlike (approximately) the Sharpe ratio for the market). The true Sortino and SVP ratios for the market are 1.002 and 2.844. The sample averages were 1.117 and 2.972 with standard errors of 0.00935 and 0.00831. So the sample averages were higher than the true values by more than 12 and 15 standard errors, respectively.

Table 2
Dynamic manipulation of the Sortino and van der Meer (1991) and Sortino et al. (1999) ratios

	Trans. cost (%)	Sortino ratio: Mean (std dev)	Sortino ratio: Frequency $D > D_{mkt}$ signif. (%)	SVP ratio: Mean (std dev)	SVP ratio: Frequency: $U > U_{mkt}$ signif. (%)	Sharpe ratio: Mean (std dev)	Sharpe ratio: Frequency: $S > S_{mkt}$ signif. (%)	Portfolio characteristics: Mkt exposure	Portfolio characteristics: Beta
Sortino portfolios	0	1.338	83.6	3.155	76.2	0.667	77.4	0.885	0.876
		(0.932)	**16.6**	(0.831)	**12.1**	(0.419)	**15.3**	(0.318)	(0.308)
	10	1.314	83.1	3.137	75.4	0.659	76.4	0.890	0.879
		(0.914)	**15.4**	(0.813)	**11.8**	(0.415)	**14.7**	(0.320)	(0.309)
	20	1.312	82.4	3.134	74.8	0.656	75.7	0.891	0.881
		(0.921)	**13.8**	(0.821)	**10.7**	(0.419)	**13.2**	(0.320)	(0.309)
SVP portfolios	0	1.337	86.7	3.178	83.3	0.672	81.8	0.896	0.881
		(0.994)	**13.5**	(0.832)	**20.0**	(0.422)	**16.4**	(0.304)	(0.293)
	10	1.304	87.1	3.152	83.6	0.685	81.9	0.909	0.894
		(0.930)	**13.2**	(0.817)	**21.7**	(0.421)	**15.4**	(0.306)	(0.295)
	20	1.307	86.4	3.154	82.5	0.659	81.2	0.903	0.887
		(0.936)	**12.3**	(0.824)	**21.1**	(0.426)	**14.5**	(0.304)	(0.294)
Market		1.117	—	2.972	—	0.598	—		
		(0.935)		(0.831)		(0.454)			
True mkt		1.002	—	2.844	—	0.597	—		

The table presents the annualized Sortino and van der Meer (1991), Sortino et al. (1999) and Sharpe ratios for dynamically manipulated portfolios over a 60-month (5-year) period. The ratios are annualized by multiplying the monthly ratios by the square root of 12. The measures are computed in 10,000 simulations. The average and the standard deviation of each measure are reported. The "freq." columns have two numbers showing the fraction of times that the portfolio showed superior performance with the ratio larger than the market's ratio and the approximate fraction of times the portfolio showed superior performance that was significant at the 5% level. The latter number was estimated as the frequency with which the difference between the portfolio's ratio and the market's ratio was more than 1.65 standard deviations above zero.

The portfolio is always invested in the market. Derivatives are used only to change the market exposure. Initially the portfolio is fully invested in the market. After 12 months, the leverage is adjusted each month based on the performance history as described in Equation (14) to set the target mean to $\overline{x} = \overline{x}_h D^2_{mkt}/D^2_h$ for the Sortino portfolios and the target upside potential to $\mathbb{E}[\mathrm{Max}(\tilde{x}_f, 0)] = \mathrm{Avg}[\mathrm{Max}(\tilde{x}_h, 0)]U^2_{mkt}/U^2_h$ for the SVP portfolios. Adjustments are restricted to keep the market exposure between 50 and 150% of the portfolio's value. The average and standard deviation of the market exposure and resulting beta are also reported. Each portfolio adjustment requires paying a transaction cost computed as 0%, 10%, or 20% (round trip) of the value of the synthetic forward contracts (a call plus a put struck at-the-money in present value terms) needed to change the market exposure.

The market parameters are: risk free rate 5%, market premium 12%, market standard deviation 20%.

The best dynamic manipulation SVP ratio is almost the same. The target in the future is proportional to the upside realization in the past and the ratio of the squares of the expected and realized SVP measures, that is, $\mathbb{E}[\mathrm{Max}(\tilde{x}_f, 0)] = \mathrm{Avg}[\mathrm{Max}(x_h, 0)] \times U_f^2/U_h^2$. The simulation results for the SVP measure manipulation are not presented here as they were very similar. Understandably, this manipulation produces somewhat higher average SVP ratios and beats the market's SVP ratio more often while producing somewhat lower Sortino ratios.

2.2 Henriksson-Merton and Treynor-Mazuy timing measures

The timing measures of Treynor and Mazuy (1966) and Henriksson and Merton (1981) are bivariate regressions that use an extra market factor to capture managers' timing abilities (rather than to capture an additional source of risk). The regressions used are

$$\tilde{x}_t = \gamma_0 + \gamma_1 \tilde{m}_t + \gamma_2 \tilde{w}_t + \tilde{\varepsilon}_t \tag{15}$$

where $\tilde{x}_t$ and $\tilde{m}_t$ are the excess returns on the portfolio and the market and $\tilde{w}_t$ is the variable used to capture timing ability. The timing variables are $\tilde{w}_t \equiv \mathrm{Max}(-\tilde{m}_t, 0)$ and $\tilde{w}_t \equiv \tilde{m}_t^2$ for HM and TM, respectively. The HM model is consistent with a manager who chooses one of two amounts of leverage for his portfolio depending on whether he forecasts that the market return will exceed the risk-free rate or not. As shown by Admati et al. (1986), the TM model is consistent with a manager whose target beta varies linearly with his forecast for the expected market rate of return.

Unlike the measures we have already considered, performance here is characterized by two numbers. In each case a positive γ_2 is an indication of market timing ability and a positive γ_0 is considered a sign of superior stock selection. However, as shown by Jagannathan and Korajczyk (1986), simple trading strategies including or mimicking options can produce positive γ_2s and negative γ_0s or vice versa.[20] An unequivocal demonstration of timing ability must, therefore, satisfy a joint test on the two coefficients.

Merton (1981) has shown that the total contribution of the manager's timing and selectivity in the HM model is $\gamma_0 \mathrm{e}^{-r\Delta t} + \gamma_2 P(1, \Delta t, \mathrm{e}^{r\Delta t})$ per dollar invested where $P(S, \tau, K)$ is the value of a τ-period put option on the market with a strike price of K. The intuition for this conclusion is that a market timer who could forecast perfectly whether or not the market return would exceed the interest rate and who adjusted his portfolio to be fully invested in the market or bonds would essentially provide such a put for free. In addition his selectivity ability would provide the present value

[20] For example, buying (writing) calls produces a positive (negative) γ_2 and a negative (positive) γ_0. The MSRP also has a negative γ_2 and a positive γ_0. It should be noted that due to the γ_2 term in these regressions, γ_0 and γ_1 are not the standard Jensen's alpha and CAPM beta.

of γ_0 per dollar invested. The total contribution is the amount by which the value of the protective put exceeds its average "cost" measured by the lowered present value of the extra average return.

Similarly, under the TM model, the total contribution of the manager's investment ability per dollar invested is given by the same formula in which the value of a derivative contract that pays the square of the market's excess return replaces the put's value. This value is derived in Appendix D. The total contributed value, V, of the money manager's contribution to timing and selectivity is

$$\begin{aligned} V_{\mathrm{HM}} &= \gamma_0 e^{-r\Delta t} + \gamma_2 P(1, \Delta t, e^{r\Delta t}) \\ V_{\mathrm{TM}} &= \gamma_0 e^{-r\Delta t} + \gamma_2 e^{r\Delta t}(e^{\sigma^2 \Delta t} - 1) \,. \end{aligned} \tag{16}$$

Unfortunately with a complete market of derivative assets, the TM and HM measures can be manipulated to any degree desired; that is, any values for the γs can be achieved with many different portfolios. The minimum-variance portfolio that achieves any particular (positive) target values for γ_0 and γ_2 will be long in the MSRP and the "timing" put struck at-the-money in present value terms.[21] It might be short or long in the market and bonds depending on the targeted mean excess return and market beta.

Figure 4 illustrates the payoffs on the minimum-variance zero-beta portfolios with a timing target of $\gamma_2 = 0.1$ and a selectivity target of $\gamma_0 = 10$ basis points per month. The environment is the same as before—a lognormal market with a risk premium of 12% per year and a volatility of 20%. The graph displays approximately the middle 95% of the return distribution. The timing returns fluctuate from local extremes near plus and minus 20% over the market range from −7 to +7%. In practice, of course, only scattered points on the curves would be seen, and the portfolio would probably be described as somewhat volatile but otherwise "normal."

Reexamining Figure 2 we see that the alpha-manipulating strategy that decreases market exposure after good returns and increases market exposure after poor returns has a general convex shape. This strategy will likely produce a false positive timing performance. Simulations verify that this is true; however, while γ_2 and the total contribution as measured in Equation (16) are positive, simply changing the market exposure also yields a (false) negative selection ability. Both γ_0 and γ_2 can be made positive if a call position is written against the portfolio since this produces an extra return when the market has a small return.

Table 3 shows the performance of a manipulated portfolio that holds the market and writes a 10% out-of-the-money call on 2% of its holdings;

[21] See section D of the Appendix for a proof. For the Treynor-Mazuy measure the portfolio needs to be long in an exotic derivative contract paying the square of the market price rather than put options.

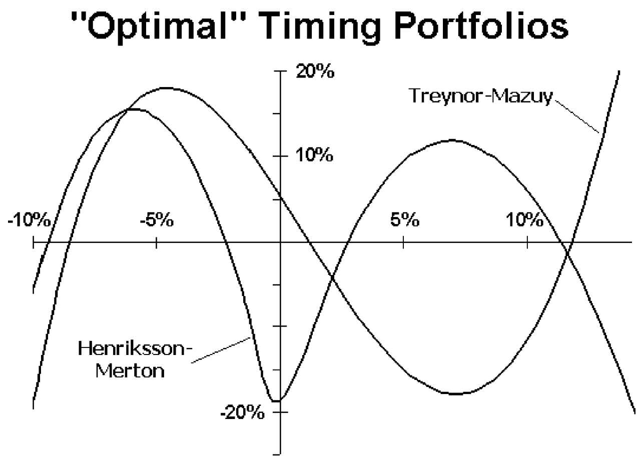

Figure 4
Payoffs on "optimal" timing portfolios
The horizontal axis is the monthly excess market return. The two curves show the monthly excess returns on the Treynor-Mazuy and Henriksson-Merton optimized portfolios with minimum-variance, zero-beta, a selectivity coefficient of γ_0= 0.1%, and a timing coefficient of γ_2= 0.1 from the regression: $x_t = \gamma_0 + \gamma_1 m_t + \gamma_2 w_t + \varepsilon_t$. The market environment is a lognormal market with a continuously-compounded risk premium of 12% and a logarithmic volatility of $\sigma = 20\%$.

that is, the portfolio shares only 98% of the market returns above 10% in any month. After 1 year, the leverage is changed every month so that the market exposure is equal to 50% or 150% of the portfolio's value if the fraction of months in which the market's return was more than average was more or less than half, respectively. The resulting portfolio produces a total contributed value of around 98 to 156 basis points per year depending on transactions costs as measured by either HM or TM. Our simulations show that the total contributed value was positive about 65 to 70% of the time and significantly so (at the 5% level) 9 to 14% of the time. The results are not as strong as for the other measures, but this is not surprising. The HM and TM measures are timing measures specifically designed for timing ability and should be able to better reject our dynamic manipulations than the other measures discussed so far.

The components of performance, selectivity and timing, do not have as strong an individual showing because the simulation was designed to produce over all good performance and the two measures are negatively correlated. Increasing the out-of-the-money call sales, for example, would increase γ_0 and decrease γ_2. However, each component is positive more than half of the time and significantly so more often than chance would allow.

Table 3
Dynamic manipulation of the Henriksson and Merton (1981) and Treynor and Mazuy (1966) timing coefficients

		Total value (ς)		Selectivity (γ_0)		Timing (γ_2)		Portfolio characteristics	
	Trans. cost (%)	Mean (std dev) (%)	Frequency $V > 0$ signif (%)	Mean (std dev) (%)	Freq: $\gamma_0 > 0$ signif. (%)	Mean (std dev)	Freq: $\gamma_2 > 0$ signif. (%)	Mkt exposure	Beta
Henriksson-Merton	0	1.56	71.0	0.63	56.1	0.033	55.7	1.07	1.05
		(3.03)	**14.0**	(5.01)	**5.9**	(0.245)	**6.7**	(0.71)	(0.28)
	10	1.30	69.0	0.32	53.2	0.036	56.0	1.09	1.06
		(2.92)	**11.4**	(5.07)	**5.7**	(0.248)	**7.0**	(0.71)	(0.27)
	20	0.98	64.9	0.00	50.6	0.035	56.5	1.09	1.07
		(2.84)	**9.0**	(5.10)	**5.0**	(0.248)	**6.9**	(0.71)	(0.27)
Treynor-Mazuy	0	1.56	70.6	1.24	65.2	0.007	52.6	1.07	1.05
		(3.12)	**13.67**	(3.50)	**10.2**	(0.096)	**6.3**	(0.71)	(0.28)
	10	1.30	68.3	0.95	61.8	0.007	53.6	1.09	1.06
		(3.01)	**11.1**	(3.47)	**8.6**	(0.097)	**6.3**	(0.71)	(0.27)
	20	0.98	64.5	0.63	57.6	0.007	53.5	1.09	1.07
		(2.93)	**9.0**	(3.42)	**7.3**	(0.097)	**5.9**	(0.71)	(0.27)

The table presents the Henriksson-Merton and Treynor-Mazuy coefficients and total contributed value for dynamically manipulated portfolios over a 60-month period. The coefficients are computed from the regression: $\tilde{x}_t = \gamma_0 + \gamma_1 \tilde{m}_t + \gamma_2 \tilde{w}_t + \tilde{\varepsilon}_t$ where $\tilde{x}_t$ and $\tilde{m}_t$ are the excess returns on the portfolio and the market, respectively. The timing variable is $\tilde{w}_t = \text{Max}(-\tilde{m}_t, 0)$ for HM and $\tilde{w}_t = \tilde{m}_t^2$ for TM. The total contributed value for the HM model is $\gamma_0 \mathrm{e}^{-r\Delta t} + \gamma_2 P$ where P is the value of a put on \$1 in the market at-the-money in present value terms. The total contributed value for the TM model is $\gamma_0 e^{-r\Delta t} + \gamma_2 e^{r\Delta t}(e^{\sigma^2 \Delta t} - 1)$.

The measures are computed in 10,000 simulations, and the average and the standard deviation are reported. (The standard errors of the averages are one hundredth of the standard deviation reported.) Also given is the fraction of times that the portfolio showed superior performance for each measure.

Each month, the portfolio is invested in the market with a 10% out-of-the-money call sold on 2% of the position. (That is, the portfolio will achieve only 98% of any market appreciation above 10% each month.) The remaining money is fully invested in the market initially. After 12 months, the market exposure is adjusted each month to equal 50% or 150% of the portfolio's value depending on whether the past market returns were above or below average. The average and standard deviation of the market exposure (ignoring the call) and resulting beta (including the call) are also reported. Each adjustment in exposure requires paying a transaction cost which is computed as 0% 10% or 20% (round trip) of the value of the synthetic forward contract (a call plus a put struck at-the-money in present value terms) needed to change the leverage. Each call sold every month also requires paying a transaction cost.

The market parameters are: risk free rate 5% per annum, market premium 12%, market standard deviation 20%.

3. Manipulation-proof Performance Measures

In the preceding sections of this article, we have established that popular measures of performance are susceptible to manipulation even when transactions costs are high. We have shown that manipulation requires no superior information and little technical sophistication. Simple schemes like trading puts and calls or simply altering a portfolio's leverage can often create performance that looks superior. Furthermore, even if money managers are not seeking to manipulate their performance numbers, competitive evolution may favor those managers whose strategies happen to look like our manipulation strategies if these measures are widely used. We have shown that these problems can be of concern even when evaluating well-diversified portfolios of equities, but increasingly money is being invested in hedge funds and other venues whose return distributions differ substantially from those historically found in all-equity funds. For such funds even more problems may arise.

All of the above issues can be avoided with a manipulation-free measure. But what does it mean for a measure to be manipulation-free? What exactly should it encourage a manager to do or not do? Intuitively, if a manager has no private information and markets are efficient, then holding some benchmark portfolio, possibly levered, should maximize the measure's expected value. The benchmark might be the market, but in some contexts other benchmarks could be appropriate. Static manipulation is the tilting of the portfolio away from the (levered) benchmark even when there is no informational reason to do so. Dynamic manipulation is altering the portfolio over time based on past performance rather than on new information. Our goal is to characterize a measure that punishes uninformed manipulation of both types.

Formally a performance measure is a function of the portfolio's probability distribution across the outcome states. That is, if $\mathbf{r}$ is the vector of returns across all the possible outcomes, then the performance measure is a real-valued function, $\Theta(\mathbf{r})$, of those returns.[22] In practice, of course, we do not know the true distribution, and the estimated performance measure must be a function of the returns realized over time. We will denote the estimated performance measure as $\hat{\Theta}(\langle r_t, s_t \rangle_{t=1}^{T})$.[23] Each time

[22] In principle, a performance measure could be a pair or more of numbers (like γ_0 and γ_2 in the TM and HM models or mean and variance). In general, of course, such performance measures would not provide a complete ranking of portfolios. In one sense multivalued performance measures would be harder to manipulate since they would provide more information. In another more practical sense, multivalued performance measures would be easier to manipulate since without a complete ranking there would be more portfolios that the market benchmark did not beat. We confine our attention to single-valued performance measures like the Sharpe ratio.

[23] Using the time series of returns in the estimated performance measure $\hat{\Theta}$ to represent the true performance measure assumes that the realizations of portfolio returns (and states) are sufficiently independent and identically distributed over time so that the frequencies will be equal to the *ex ante* probabilities in a sufficiently large sample.

period and its return (r_t) can, but need not, be identified with the state (s_t) that occurred. For example, estimating the Sharpe ratio requires knowing only the returns while to estimate Jensen's alpha we need to know the contemporaneous market return for each return.

One property that any MPPM must have is that it recognizes arbitrage opportunities as good. If the portfolio-return distribution $\mathbf{r}_1$ dominates the portfolio return distribution $\mathbf{r}_2$, then the performance measure must rank it higher, $\Theta(\mathbf{r}_1) > \Theta(\mathbf{r}_2)$. This same property should hold for the estimated performance measure, though the dominance in this case may only be apparent as not all states may have been realized. We already know that the Sharpe ratio fails this simple test—if the returns in one or more states are increased sufficiently, the ratio will decrease even though the return distribution unequivocally improves based on the observed evidence.

However, this property, while necessary, is not sufficient to make a measure manipulation-proof. To have a performance measure that is proof against dynamic manipulation the function must have a strong independence property. This requires that altering some of the components does not affect the relative ranking based on other components. A performance measure cannot be dynamically manipulated only if the relative rankings of different futures are unaffected by the particular history that has occurred.[24] Note that this does *not* say that the history does not affect rankings; it does say that different potential futures can be compared independently of the past.

This independence property is known from utility theory where it is called (strong) utility independence or strong utility separation. Furthermore, a ranking (i.e., ordinal) function has this independence property if and only if it has an additive representation.[25]

That is, the only possible MPPMs can be expressed as

$$\hat{\Theta}(\langle r_t, s_t \rangle_{t=1}^{T}) = \Upsilon\left(\frac{1}{T}\sum_{t=1}^{T} \theta_t(r_t, s_t)\right). \tag{17}$$

The function $\Upsilon(\cdot)$ is any increasing function and is irrelevant for ranking; therefore, it can be taken to be the identity function except when otherwise more convenient. We choose to write the argument as an average rather than a simple sum because most performance measures are based on some kind of average. Each function θ_t must be increasing if $\hat{\Theta}$ is to satisfy the "arbitrage is good" criterion.

[24] Precisely, the measure is proof against dynamic manipulation only if $\hat{\Theta}(\langle r_t^0 \rangle_1^\tau, \langle r_t' \rangle_{\tau+1}^T) > \hat{\Theta}(\langle r_t^0 \rangle_1^\tau, \langle r_t'' \rangle_{\tau+1}^T)$ implies $\hat{\Theta}(\langle r_t \rangle_1^\tau, \langle r_t' \rangle_{\tau+1}^T) > \hat{\Theta}(\langle r_t \rangle_1^\tau, \langle r_t'' \rangle_{\tau+1}^T)$ for any history $\langle r_t \rangle_1^\tau$. This independence property must also hold for all conditioning subsets.

[25] See, for example, Debreu (1960).

Most performance measures treat each time period identically, apart from any possible state dependence. In this case, all θ_t's are the same and any MPPM must be equivalent to a simple average of some, possibly state-dependent, function of returns.[26] Some simple performance measures that cannot be dynamically manipulated are the average return, $\theta(r_t) = r_t$; the average market-adjusted return, $\theta(r_t, s_t) = r_t - m_t(s_t)$;[27] and the geometric average return, $\theta(r_t) = \ell\text{n}\,(1 + r_t)$.

Any function θ will create a performance measure, Θ, that is proof against dynamic manipulation. To prevent static manipulation, θ must be strictly increasing, as previously noted, and also concave; otherwise simply adding unpriced risk to a portfolio could increase its measured performance. Concavity is a natural property to demand since it is also necessary (and sufficient) if we want to rate as better the less risky (in a Rothschild–Stiglitz sense) of two portfolios with the same average return.

Since any measure with θ_t an increasing, concave function of the return is a MPPM, it might seem that we have simply rediscovered the von Neumann Morgenstern utility functions. But this is not the case—the set of MPPMs overlaps the set of von Neumann Morgenstern utility functions only for a power form. The necessity of the power form comes from the use of the returns at different times to proxy for the returns in different outcome states. Even if we consider θ to be a utility function, the formula in Equation (17) is not an average (proxying for an expectation) achieved level of utility. It is instead a time-series average of assessment of the increase in utility. This means that while the measure cannot be manipulated, some choices of $\hat{\Theta}$ will rank higher portfolios that are stochastically dominated.

Consider the simplest case of selecting a portfolio over a T-period horizon when returns are independent and identically distributed. Any portfolio that is not first-order stochastically dominated must maximize the expectation of $u(\prod(1 + r_t)) = u\left[\exp\left(\sum \ell\text{n}\,(1 + r_t)\right)\right]$ for some strictly increasing function u. If we want the MPPM to select a stochastically nondominated portfolio, we must then have

$$u\left[\exp\left(\sum \ell\text{n}(1 + r_t)\right)\right] = \Upsilon\left(\frac{1}{T}\sum \theta_t(r_t)\right)$$
$$\Rightarrow \sum \ell\text{n}(1 + r_t) = \Upsilon'\left(\frac{1}{T}\sum \theta_t'[\ell\text{n}(1 + r_t)]\right) \tag{18}$$

[26] An identical treatment of each period's return is not a requirement for the measure to be manipulation-proof. For example, the time-dependent average $\theta_t(r_t) = \delta^{T-t} r_t$, which puts extra emphasis on recent returns, could be used.

[27] Note that, as already shown, Jensen's alpha, $\theta(r_t, s_t) = r_t - \beta m_t(s_t)$, is not manipulation-proof. The problem with alpha is that beta must be estimated and the estimated value of beta can be easily manipulated. Any fixed (data-independent) value of β could be used in an MPPM, but these would not, of course, be Jensen's alpha.

where $\theta_t'(x) \equiv \theta_t(e^x - 1)$ and $\Upsilon'(x) = \ell\text{n}(u^{-1}[\Upsilon(x)])$. But this second equality can be true only if θ_t' and Υ' are linear functions or, equivalently, only if θ is of the power form, $\theta(r) = (1+r)^\delta$.

The specific MPPM we propose here is

$$\hat{\Theta} = \frac{1}{(1-\rho)\Delta t}\ell\text{n}\left(\frac{1}{T}\sum_{t=1}^{T}[(1+r_t)/(1+r_{ft})]^{1-\rho}\right) \tag{19}$$

where r_{ft} and r_t are the per-period (not annualized) interest rate and the rate of return on the portfolio over period t. We have transformed the measure so that $\hat{\Theta}$ can be interpreted as the annualized continuously compounded excess return certainty equivalent of the portfolio. That is, a risk-free portfolio earning $\exp[\ell\text{n}(1+r_{ft}) + \hat{\Theta}\Delta t]$ each period would have a measured performance of $\hat{\Theta}$.

Finally, we want to associate the MPPM with some benchmark portfolio. This would typically be some market index. In the absence of any private information, we want the MPPM to score the chosen benchmark highly. If the benchmark portfolio has a lognormal return, $1 + r_b$, then the parameter ρ should be selected so that

$$\rho = \frac{\ell\text{n}[\mathbb{E}(1+\tilde{r}_b)] - \ell\text{n}(1+r_f)}{\text{Var}[\ell\text{n}(1+\tilde{r}_b)]}. \tag{20}$$

Historically this number is around 2 to 4 for the CRSP value-weighted market portfolio depending on the time period and frequency of data used. We have used a relative risk aversion of three in our simulations.

It is interesting to note that the measure proposed here is identical in substance and nearly in form to the Morningstar Risk Adjusted Rating (Morningstar, 2006) which was introduced in July, 2002.[28] The Morningstar ranking is motivated directly as a representative utility function rather than from the manipulation-proof properties we addressed. The manipulation-proof feature of the measure is an apparently unintended and evidently unrecognized consequence of the functional form Morningstar adopted to evaluate funds. Its performance relative to other ranking measures was analyzed in Stutzer (2005).

Table 4 shows the performance for our manipulated portfolios in our previous simulations. The MPPMs are determined for three different risk

[28] Additional information in this paragraph is based on personal communication with Paul Kaplan. In the notation of Equation (19), the Morningstar rating is

$$\text{MRAR}(\gamma) = \left[(1/T)\sum[(1+r_t)/(1+r_{ft})]^{-\gamma}\right]^{-1/\gamma\cdot\Delta t} - 1$$

so $\hat{\Theta} \equiv \ell\text{n}[1 + \text{MRAR}(\rho - 1)]$. Morningstar uses $\gamma = 2$ ($\rho = 3$) and monthly returns ($\Delta t = 1/12$) in their ratings.

Table 4
The manipulation-proof performance measure

Measure the portfolio was designed to manipulate	Rating by own measure			Rating by manipulation-proof performance measure Θ: MPPM Θ ($\rho = 2$)			MPPM Θ ($\rho = 3$)			MPPM Θ ($\rho = 4$)		
	Freq won (%)	Freq signif + (%)	Freq signif − (%)	Avg $\Theta_{portf} - \Theta_{mkt}$ (%)	Freq $\Theta_{portf} >$ Θ_{mkt} (%)	Freq signif +/− (%)	Avg $\Theta_{portf} - \Theta_{mkt}$ (%)	Freq $\Theta_{portf} >$ Θ_{mkt} (%)	Freq signif +/− (%)	Avg $\Theta_{portf} - \Theta_{mkt}$ (%)	Freq $\Theta_{portf} >$ Θ_{mkt} (%)	Freq signif +/− (%)
Sharpe S	82.6	20.4	1.4									
Alpha α	92.4	37.6	0.2	−0.84	46.3	0.2	−0.96	46.3	0.4	−1.08	46.0	0.5
Gen α^{gen}	90.6	34.0	0.3			9.1			9.6			9.9
Sortino D	83.6	16.6	1.6	−1.14	42.0	0.6	−1.04	42.9	1.0	−0.92	44.9	1.3
						9.9			9.7			8.9
SVP U	83.3	20.0	2.9	−1.00	46.5	0.5	−0.87	48.1	0.9	−0.74	49.7	1.4
						10.3			9.9			9.2
HM value V_{HM}	71.0	14.0	2.2	−0.62	45.4	3.1	−1.26	38.5	2.4	−1.91	33.6	1.7
TM value V_{TM}	70.6	13.7	2.2			6.9			9.2			11.6

The table compares the manipulation-proof performance measure to the Sharpe, Sortino, and SVP ratios, the alphas, and the Henriksson-Merton and Treynor-Mazuy timing measures for the various portfolios designed to dynamically manipulate those measures as presented in Tables 1–3. The difference between the market's and the portfolio's MPPM are given for three representative risk aversions. The frequencies with which the portfolio beat the market according to the various measures including the MPPM are given along with the approximate frequencies with which the portfolio significantly (5%) outperformed or underperformed the market. The latter numbers are estimated as the frequency with which the difference in measures was more than 1.65 standard deviations above the null value.

The MPPM is defined in Equation (19) as $\hat{\Theta} \equiv \ell n \left(\frac{1}{T} \sum [(1+r_{ft})^{-1}(1+r_{ft}+x_t)]^{1-\rho} \right)(1-\rho)\Delta t$.

The market parameters are: risk free rate 5% per annum, market premium 12%, market standard deviation 20%. This is consistent with a risk aversion of $\rho = 3$.

aversions, 2, 3, and 4.[29] The simulation was constructed with a market portfolio in which the risk aversion was 3, so the numbers in that column are correct. In practice we would not know the correct risk aversion to use so the other MPPMs are given to check the sensitivity.

The computed MPPMs show correctly that the manipulated portfolios are not as good as the market.[30] For example, the Sharpe-manipulated portfolio had a Sharpe ratio exceeding the market's 82.6% of the time and was significantly larger 20.4% of the time at the 5% confidence level. On the other hand, the MPPM measure shows that this portfolio beat the market only 46.3% of the time and was significantly (5% confidence) better only 0.4% of the time. Conversely it significantly underperformed the market 9.6% of the time as measured by the MPPM but only 1.4% of the time as measured by the Sharpe ratio. The average performance was 96 basis points below the market on a certainty equivalent basis.

The other portfolios show similar results. Their underperformances range from 87 to 126 basis points and they never beat the market more than 50% of the time. They significantly underperform the market more than 9% of the time and outperform it from 0.4 to 2.4% of the time. The results are similar for MPPMs computed assuming risk aversions of 2 or 4.

Of course, as well as penalizing poor performance, we would like an MPPM to recognize good performance when it occurs. Tables 5 and 6 show the performance of money managers who try to provide stock selection and market timing, respectively.

Table 5 compares the annualized Sharpe ratio, alpha, and MPPM for informed and uninformed money managers. The informed money manager can provide an annual alpha of 1% at the cost of holding a portfolio with undiversified risk. The three panels show portfolios with annual residual risks of 20, 2, and 0.2%.[31] These are portfolios of approximately one, one hundred, and ten thousand component stocks. The uniformed manager holds similar under-diversified portfolios but has an alpha of zero and does not engage in any manipulation. The table shows the average, standard deviation, and frequencies for the difference between the portfolio's and market's Sharpe ratios, alphas, and MPPMs.

---

[29] The range of relative risk aversions in the market might be much wider than 2 to 4; however, this range is meant to capture the range of risk aversions that would be relevant for the construction of mutual funds or other well diversified portfolios. Relative risk aversions of 2 to 4 correspond to portfolios that would range from leveraged positions of 1.5 down to 0.75. This should encompass most portfolios that would be ranked.

[30] Only the zero-transactions-cost cases are reported. With transactions costs, the MPPM scores would be even lower.

[31] The logarithmic market return over a period of length Δt is $[\mu + \alpha - \frac{1}{2}(\sigma^2 + \upsilon^2)]\Delta t + (\sigma\tilde{\varepsilon} + \upsilon\tilde{\eta})\sqrt{\Delta t}$ where $\tilde{\varepsilon}$ and $\tilde{\eta}$ are standard normal variables. The market's expected rate of return and log variance per unit time are μ and σ^2, respectively. The informed trader provides an α of 1% at a cost of a residual variation of υ of 20, 2, or 0.2%. The uninformed trader has the same residual variation with an α of zero.

Table 5
The manipulation-proof performance measure

Residual risk	Informed trader ($\alpha = 1\%$) Avg excess	Stand dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)	Uninformed trader ($\alpha = 0\%$) Avg excess	Stand dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)
Annual logarithmic residual standard deviation = 20%										
Sharpe	−0.142	0.350	34.0	2.0	10.5	−0.183	0.349	30.5	1.6	12.6
Alpha	1.01%	9.15%	56.0	6.9	3.6	−0.01%	9.15%	48.9	4.9	4.6
MPPM	−4.82%	9.02%	29.2	1.4	13.4	−5.90%	9.12%	25.8	1.1	15.4
Annual logarithmic residual standard deviation = 2%										
Sharpe	0.048	0.045	85.0	25.6	0.4	−0.004	0.046	47.1	4.5	5.4
Alpha	1.00%	0.93%	90.5	36.5	0.2	−0.00%	0.93%	49.0	4.8	4.8
MPPM	0.95%	0.90%	85.3	27.1	0.4	−0.06%	0.91%	47.7	4.4	5.6
Annual logarithmic residual standard deviation = 0.2%										
Sharpe	0.048	0.006	100	100	0.0	−0.000	0.005	49.6	5.1	4.8
Alpha	1.00%	0.18%	100	100	0.0	−0.00%	0.18%	49.0	5.1	4.7
MPPM	1.00%	0.09%	100	100	0.0	−0.00%	0.09%	49.8	5.1	5.0

The table compares the manipulation-proof performance measure to the Sharpe ratio and alpha for an informed trader who can create a portfolio with a positive alpha by taking on various levels of increased unsystematic risk.
The difference between the portfolio's and market's Sharpe ratios and the portfolio's alpha and MPPM, Θ, are reported. The frequencies with which the portfolio beat the market according to each are given along with the approximate frequencies with which the portfolio significantly (5%) outperformed or underperformed the market. The latter numbers are estimated as the frequency with which the measure was more than 1.65 standard deviations positive or negative.
The MPPM is defined in Equation (19) as $\hat{\Theta} \equiv \ell n\left(\frac{1}{T}\sum[(1+r_{ft})^{-1}(1+r_{ft}+x_t)]^{1-\rho}\right)(1-\rho)\Delta t$.
The market parameters are: risk free rate 5% per annum, market premium 12%, market standard deviation 20%. This is consistent with a risk aversion of $\rho = 3$.

The portfolio alphas are just as expected. They average 1% and 0% for the informed and uninformed managers. For the uninformed manager, they are significantly positive or negative just about the predicted 5% of the time. For the informed manager they are significantly positive (negative) more (less) than 5% of the time regardless of the diversification—again as expected since alpha does not penalize for under-diversification.

For 10,000 stock portfolios all three of the measures show that the informed manager's portfolio is better than the market and the uninformed manager's portfolio is essentially identical to the market. For one hundred stock portfolios, the MPPM and Sharpe have almost the same results though the MPPM does marginally better at recognizing both informed and uninformed managers. For single-stock portfolios, the MPPM does substantially better at illustrating that the 1% alpha does not properly compensate for the lack of diversification.

Table 6 compares the MPPM and Henriksson-Merton and Treynor-Mazuy timing measures for an informed and uninformed market timer. The informed market timer knows that a time varying mean explains a portion of the observed market variance. In the two panels, the informed trader's information explains 0.1% and 1% of the market's variation,

Table 6
The manipulation-proof performance measure

	Informed timer (0.1%)					Random timer				
	Avg	St dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)	Avg	St dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)
HM γ_2	0.052	0.185	61.5	8.6	2.7	−0.002	0.187	49.3	4.8	4.8
HM V	0.07%	0.20%	63.2	8.9	2.4	0.001%	0.203%	50.0	5.0	5.1
TM γ_2	0.208	0.865	60.2	7.8	2.9	−0.013	0.870	49.2	4.9	5.0
TM V	0.07%	0.20%	62.9	9.0	2.5	0.001%	0.207%	50.1	5.0	5.0
MPPM Θ	0.40%	2.35%	57.2	6.8	3.4	−6.42%	9.23%	24.2	0.9	16.9
	Informed timer (1%)					Random timer				
	Avg	St dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)	Avg	St dev	Freq won (%)	Freq signif + (%)	Freq signif − (%)
HM γ_2	0.512	0.591	81.3	21.4	0.7	−0.003	0.598	49.7	5.0	5.1
HM V	0.68%	0.64%	85.8	27.0	0.4	−0.01%	0.65%	49.1	4.8	4.9
TM γ_2	2.037	2.747	77.8	17.9	1.0	−0.004	2.78	49.9	4.9	5.0
TM V	0.69%	0.66%	85.8	26.8	0.4	−0.01%	0.67%	49.3	4.7	4.9
MPPM Θ	4.20%	7.53%	71.8	13.2	1.7	−10.5%	12.2%	18.9	0.5	20.8

The table compares the manipulation-proof performance measure (MPPM) to the Henriksson-Merton and Treynor-Mazuy timing measures for an informed market timer, whose information about a changing mean explains 0.1% or 1% of the market's variance, and an uninformed market timer who varies leverage randomly to the same degree. The average timing coefficient, γ_2 as defined in Equation (15), and the average contributed value, V as defined in Equation (16), are reported along with the MPPM, Θ. The frequencies with which the portfolio beat the market according to the three are given along with the approximate frequencies with which the portfolio significantly (5%) outperformed or underperformed the market. The latter numbers are estimated as the frequency with which the measure was more than 1.65 standard deviations positive or negative.

The MPPM is defined in Equation (19) as $\hat{\Theta} \equiv \ell n\left(\frac{1}{T}\sum[(1+r_{ft})^{-1}(1+r_{ft}+x_t)]^{1-\rho}\right)(1-\rho)\Delta t$.

The market parameters are: risk free rate 5% per annum, market premium 12%, market standard deviation 20%. This is consistent with a risk aversion of $\rho = 3$.

and he adjusts his leverage optimally.[32] The uninformed trader believes incorrectly that he has similar quality information, so he adjusts his leverage to the same degree, but does so randomly relative to actual outcomes.

The table shows the average, standard deviation, and frequencies for the timing coefficients, γ_2, and the total contributed value (as given in Equation (16)) and the difference between the portfolio's and market's MPPMs. The HM and TM models are designed to recognize informed

[32] The logarithmic market return over a period of length Δt is $(\mu + \tilde{s} - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{(1-\delta^2)\Delta t}\,\tilde{\varepsilon}$,where $\tilde{\varepsilon}$ is a standard normal variable. The market's unconditional expected rate of return and logarithmic variance per unit time are μ and σ^2, respectively. The information about the changing mean is in the signal, $\tilde{s}$, which is normally distributed with mean zero and variance $\delta^2\sigma^2\Delta t$ where δ is the fraction of variation known to the informed trader. In the simulations, δ is set to 0.1% and 1%. The optimal market holding conditional on a signal, s, is equal to the conditional risk premium divided by the relative risk aversion times the conditional variance, $(\mu + s - r)\rho(1-\delta^2)\sigma^2$. Since the unconditional risk premium is equal to the relative risk aversion times the unconditional variance, the optimal leverage conditional on a signal s is $(1 + s/\rho\sigma^2)(1-\delta^2)$.

timing, but not to penalize uninformed timing. Thus, they show results consistent with the null hypothesis for the uninformed timer just as they would for a manager not trying to time at all. The MPPM, on the other hand, shows that the uninformed timer is definitely hurting the portfolio's performance, and that the uninformed timer who thinks he has better information hurts the portfolio more. The HM and TM measures recognize the informed trader more frequently because, unlike the MPPM, they do not penalize the portfolio's performance for the induced lack of complete intertemporal diversification.

4. Conclusion

Portfolio performance measurement has been a topic of interest within both the practitioner and academic communities for a long time. However, a reasonable concern among those that use a particular measure is whether or not the manager being evaluated might react by attempting to manipulate it. Several articles have noted that even when the evaluator knows the moments of the return distribution it is still possible to use informationless trades to boost the expected Sharpe ratio.

What has not been recognized is that because any performance measure must be estimated a fund manager can potentially game the estimation procedure as well. As this article shows, a manager that seeks to manipulate many of the more popular measures can indeed produce very impressive performance statistics. For example, a simple rebalancing strategy can yield a market beating Jensen's alpha 86% of the time even when transactions costs are set at an unrealistically high 20% of each dollar traded. These statistics, however, are fraudulent in that they are produced entirely without private information. Indeed, most investors would probably prefer to hold the market index rather than the portfolios that manipulate the various measures tested in this article.

If the current set of measures can be manipulated, can one then find a measure that cannot be and is simultaneously of some value to investors? This article has shown that indeed such a measure exists. MPPMs can be completely characterized (up to a monotonic transformation) as the weighted average of a utility-like function. Not only is it manipulation-proof but it is no more difficult to calculate than the Sharpe ratio, and considerably easier to calculate than many of the other performance measures that have been proposed in the past to remedy the Sharpe ratio's shortcomings.

Appendices

A: The Maximal-Sharpe-Ratio Portfolio in a Complete Market

Consider a portfolio with excess return x_i in state i. The probability of state i is p_i. The Sharpe ratio of this portfolio is

$$S = \frac{\Sigma\, p_i x_i}{\sqrt{\Sigma\, p_i x_i^2 - (\Sigma\, p_i x_i)^2}} \tag{A1}$$

This ratio is invariant to scaling so when maximizing it, we can fix the expected excess return at any positive value, $\Sigma\, p_i x_i = \overline{x} > 0$, with no loss of generality. Then maximizing the Sharpe ratio is equivalent to minimizing the mean squared return subject to an expected return of $\overline{x}$ with a cost of zero.

Form the Lagrangian[33]

$$L = \tfrac{1}{2} \sum p_i x_i^2 + \lambda(\overline{x} - \sum p_i x_i) + \gamma(0 - \sum \hat{p}_i x_i) \tag{A2}$$

The first-order conditions for a minimum are

$$0 = \frac{\partial L}{\partial x_i} = p_i x_i^* - \lambda p_i + \gamma \hat{p}_i \qquad 0 = \frac{\partial L}{\partial \lambda} = \overline{x} - \sum p_i x_i \qquad 0 = \frac{\partial L}{\partial \gamma} = -\sum \hat{p}_i x_i \tag{A3}$$

The second-order condition for an interior minimum is also met.

Solving the first equation in (A3) gives the MSRP excess return in state i as

$$x_i^* = \lambda + \gamma \hat{p}_i / p_i\ . \tag{A4}$$

To determine the Lagrange multipliers, multiply Equation (A4) by p_i and $\hat{p}_i$ and sum over states. Recognizing that $\Sigma\, p_i = \Sigma\, \hat{p}_i = 1$, gives

$$\begin{aligned} \overline{x} &= \sum p_i x_i^* = \lambda \sum p_i + \gamma \sum \hat{p}_i = \lambda + \gamma \\ 0 &= \sum \hat{p}_i x_i^* = \lambda \sum \hat{p}_i + \gamma \sum \hat{p}_i^2 / p_i = \lambda + \gamma \sum \hat{p}_i^2 / p_i \end{aligned} \tag{A5}$$

These equations can be solved to determine the multipliers' values

$$\gamma = \frac{-\overline{x}}{\Sigma\, \hat{p}_i^2 / p_i - 1} \qquad \lambda = \overline{x} - \gamma \tag{A6}$$

So the MSR portfolio is

$$x_i^* = \overline{x} + \gamma(\hat{p}_i / p_i - 1)\ . \tag{A7}$$

[33] The zero-net-wealth budget constraint is expressed here using the risk-neutral probabilities in place of the state prices. Since the state price is $e^{-rT}\, \hat{p}_i$, a portfolio with a risk-neutral expected excess return of zero has a zero cost.

The variance of the maximal-Sharpe ratio portfolio and the square of its Sharpe ratio are

$$\sigma_x^2 = \sum p(x_i^* - \overline{x})^2 = \sum p_i \gamma^2 (\hat{p}_i/p_i - 1)^2 = \gamma^2 \left[\sum \hat{p}_i^2/p_i - 2\sum \hat{p}_i + \sum p_i\right]$$
$$= \gamma^2 \left[\sum \hat{p}_i^2/p_i - 1\right] = \overline{x}^2 \left[\sum \hat{p}_i^2/p_i - 1\right]^{-1} \tag{A8}$$

$$S_{\text{MSR}}^2 = \frac{\overline{x}^2}{\sigma_x^2} = \sum \hat{p}_i^2/p_i - 1\ . \tag{A9}$$

The maximal Sharp ratio portfolio is therefore

$$x_i^* = \overline{x} + \overline{x} S_{\text{MSR}}^{-2}(1 - \hat{p}_i/p_i) \qquad with \quad S_{\text{MSR}} = \left[\sum \hat{p}_i^2/p_i - 1\right]^{1/2} \tag{A10}$$

The sum in Equation (A9) can be expressed as $\Sigma p_i(\hat{p}_i/p_i)^2 = \mathbb{E}[\hat{p}^2/p^2]$, so the square of the maximal Sharpe ratio is one less than the expectation of the square of the realized probability likelihood ratio.

B: Sortino (Downside Risk) and Sortino, Van der Meer, and Plantinga (Upside Potential) Measures

The Sortino and van der Meer (1991) downside risk measure and Sortino et al. (1999) upside potential measure both compute the risk term in the denominator by considering only excess returns below some minimum acceptable return, $\underline{x}$. The upside potential measure also only counts returns above this in the numerator. The two measures are

$$D = \frac{\mathbb{E}[x]}{(\mathbb{E}[\text{Min}^2(x_i - \underline{x}, 0)])^{1/2}} \qquad U = \frac{\mathbb{E}[\text{Max}(x_i - \underline{x}, 0)]}{(\mathbb{E}[\text{Min}^2(x_i - \underline{x}, 0)])^{1/2}}\ . \tag{B1}$$

The usual choice for$\underline{x}$ is zero. In this case both ratios, like the Sharpe ratio, are invariant to leverage.[34] This is the choice we use in this article.

The Sortino downside-risk measure

For $\underline{x} = 0$, the Sortino measure can be maximized exactly like the Sharpe ratio by fixing the mean excess return and minimizing the risk measure in the denominator. Forming the Lagrangian

$$L = \frac{1}{2}\sum p_i \text{Min}^2(x_i, 0) + \lambda\left(\overline{x} - \sum p_i x_i\right) + \gamma\left(0 - \sum \hat{p}_i x_i\right) \tag{B2}$$

and differentiating gives the first-order conditions as

$$0 = \frac{\partial L}{\partial x_i} = p_i \text{Min}(x_i, 0) - \lambda p_i + \gamma \hat{p}_i \qquad 0 = \frac{\partial L}{\partial \lambda} = \overline{x} - \sum p_i x_i \qquad 0 = \frac{\partial L}{\partial \gamma} = -\sum \hat{p}_i x_i \tag{B3}$$

The second-order condition for an interior minimum is also met.

Note that the first condition in Equation (B3) is independent of x_i when $x_i > 0$; therefore, it can be satisfied only if just a single state has a positive excess return (or two or more

[34] The measures are also invariant to leverage if $\underline{x}$ is set proportional to the average return. If it is equal to the mean return, the denominator of each measure is the semi-standard deviation, and this is sometimes done for the Sortino ratio. This is a poor choice for the SVP upside measure, though, as U is then invariant to a uniform upward shift in the entire portfolio distribution.

states with the same likelihood ratio have the same positive excess return). The state with a positive excess return must be that state with the smallest likelihood ratio, $\hat{p}_i/p_i$. Since the likelihood ratio is inversely proportional to the marginal utility in a state, the single state with the positive excess return is the state with the highest market return—denoted as state I. Using the first order conditions to solve for the Lagrange multipliers and the return in the highest return state, x_I, gives

$$\begin{aligned} \overline{x} &= \sum_{i=1}^{I} p_i x_i = p_I x_i + \lambda \sum_{i=1}^{I-1} p_i + \gamma \sum_{i=1}^{I-1} \hat{p}_i = p_I x_I + \lambda(1-p_I) + \gamma(1-\hat{p}_I) \\ 0 &= \sum_{i=1}^{I} \hat{p}_i x_i = \hat{p}_I x_I + \lambda \sum_{i=1}^{I-1} \hat{p}_i + \gamma \sum_{i=1}^{I-1} \hat{p}_i^2/p_i = p_I x_I + \lambda(1-\hat{p}_I) + \gamma \sum_{i=1}^{I-1} \hat{p}_i^2/p_i \\ 0 &= -\lambda p_I - \gamma \hat{p}_I. \end{aligned} \tag{B4}$$

Solving these three equations for λ, γ, and x_I^{MDR} , and substituting back into the first equation in (B3) gives the down-side risk maximizing portfolio

$$x_i^{\mathrm{MDR}} = \gamma(\hat{p}_I/p_I - \hat{p}_i/p_i) \qquad i < I \qquad\qquad x_I^{\mathrm{MDR}} = \gamma\Big[\hat{p}_I/p_I - \sum \hat{p}_i^2/p_i\Big]/\hat{p}_I \tag{B5}$$

where $\gamma \equiv \overline{x}\left[2 - \hat{p}_I/p_I - (p_I/\hat{p}_I)\sum \hat{p}_i^2/p_i\right]^{-1}$. The square of the minimized downside risk in the denominator is

$$\begin{aligned} \mathbb{E}[\mathrm{Min}^2(\tilde{x},0)] &= \sum_{i=1}^{I-1} p_i x_i^2 = \sum_{i=1}^{I-1} p_i(\lambda + \gamma\hat{p}_i/p_i)^2 = \lambda^2 \sum_{i=1}^{I-1} p_i + 2\lambda\gamma \sum_{i=1}^{I-1} \hat{p}_i + \gamma^2 \sum_{i=1}^{I-1} \hat{p}_i^2/p_i \\ &= \gamma^2 \frac{\hat{p}_I^2}{p_I^2}(1-p_I) - 2\gamma^2 \frac{\hat{p}_I}{p_I}(1-\hat{p}_I) + \gamma^2 \sum_{i=1}^{I-1} \hat{p}_i^2/p_i = \gamma^2 \left[\frac{\hat{p}_I^2}{p_I^2} - 2\frac{\hat{p}_I}{p_I} + \sum_{i=1}^{I} \hat{p}_i^2/p_i\right] \\ &= \overline{x}^2 \frac{\hat{p}_I}{p_I}\left[\frac{\hat{p}_I}{p_I} - 2 + \frac{p_I}{\hat{p}_I}\sum_{i=1}^{I} \hat{p}_i^2/p_i\right]^{-1}. \end{aligned} \tag{B6}$$

So the maximized squared Sortino downside risk measure is

$$D_{\mathrm{MDR}}^2 = \frac{\overline{x}^2}{\mathbb{E}[\mathrm{Min}^2(\tilde{x},0)]} = 1 + \frac{p_I^2}{\hat{p}_I^2}\sum_{i=1}^{I} \hat{p}_i^2/p_i - 2\frac{p_I}{\hat{p}_I}\,. \tag{B7}$$

The Sortino, van der Meer, and Plantinga (upside potential) measures

Maximizing the upside potential measure is very similar. Again if the minimal acceptable return, $\underline{x}$,is zero the SVP upside-potential measure is invariant to leverage so to maximize it, we can fix the numerator, $\mathbb{E}[\mathrm{Max}(x_i,0)]$, and minimize the square of the denominator, $\mathbb{E}[\mathrm{Min}^2(\tilde{x},0)]$. We form the Lagrangian and optimize

$$\begin{aligned} L &\equiv \frac{1}{2}\sum p_i \mathrm{Min}^2(x_i,0) + \lambda\Big[N - \sum p_i \mathrm{Max}(x_i,0)\Big] + \gamma\Big(0 - \sum \hat{p}_i x_i\Big) \\ 0 &= \frac{\partial L}{\partial x_i} = \begin{cases} p_i x_i - \gamma\hat{p}_i & for \quad x_i < 0 \\ -\lambda p_i - \gamma \hat{p}_i & for \quad x_i > 0\,. \end{cases} \end{aligned} \tag{B8}$$

As with the maximized Sortino measure only one state can have a positive excess return, and it must be the state with the lowest likelihood ratio and the highest market return.

Using the first order conditions to solve for the Lagrange multiplier and the return in the highest return state, x_I^{MUP}, gives

$$0 = \sum_{i=1}^{I} \hat{p}_i x_i^{\text{MUP}} = \hat{p}_I x_I^{\text{MUP}} + \gamma \sum_{i=1}^{I-1} \hat{p}_i^2 / p_i$$

$$\Rightarrow \quad \gamma = -\hat{p}_I x_I \left[\sum_{i=1}^{I-1} \hat{p}_i^2 / p_i \right]^{-1} = -N \frac{\hat{p}_I}{p_I} \left[\sum_{i=1}^{I-1} \hat{p}_i^2 / p_i \right]^{-1} . \tag{B9}$$

The denominator of the squared SVP ratio is

$$\mathbb{E}[\text{Min}^2(\tilde{x}, 0)] = \sum_{i=1}^{I-1} p_i x_i^2 = \gamma^2 \sum_{i=1}^{I-1} p_i (\hat{p}_i / p_i)^2 = N^2 \frac{\hat{p}_I^2}{p_I^2} \left[\sum_{i=1}^{I-1} \hat{p}_i^2 / p_i \right]^{-1} , \tag{B10}$$

and the squared ratio itself is

$$U^2 = \frac{N^2}{\mathbb{E}[\text{Min}^2(\tilde{x}, 0)]} = \frac{p_I^2}{\hat{p}_I^2} \sum_{i=1}^{I-1} \hat{p}_i^2 / p_i \ . \tag{B11}$$

For comparison purposes, it is better to express results in terms of the mean excess return rather than the numerator using

$$\overline{x} = \sum_{i=1}^{I} p_i x_i = p_I x_I + \sum_{i=1}^{I-1} p_i x_i = N + \gamma \sum_{i=1}^{I-1} p_i (\hat{p}_i / p_i)$$

$$= N + \gamma (1 - \hat{p}_I) = N - \frac{N}{U^2} \frac{p_I}{\hat{p}_I} (1 - \hat{p}_I) \ . \tag{B12}$$

So the maximal upside potential portfolio is

$$x_i^{\text{MUP}} = -\overline{x} [U^2 (\hat{p}_I / p_I) - (1 - \hat{p}_I)]^{-1} \hat{p}_i / p_i \ \ i < I$$

$$x_I^{\text{MUP}} = \frac{\overline{x}}{p_I} \left[1 - \frac{p_I (1 - \hat{p}_I)}{\hat{p}_I U^2} \right]^{-1} . \tag{B13}$$

Comparing the maximal Sortino and SVP portfolios to the maximal Sharpe portfolio, we see that for all but for the highest return state the excess returns are linearly related:

$$\begin{aligned} x_i^{\text{MDR}} &= \frac{\overline{x}}{D_{\text{MDR}}^2} \left[1 - (1 + S_{\text{MSR}}^2) \frac{p_I}{\hat{p}_I} \right] + \frac{S_{\text{MSR}}^2}{D_{\text{MDR}}^2} \frac{p_I}{\hat{p}_I} x_i^{\text{MSR}} \qquad &\text{for } i < I \\ x_i^{\text{MUP}} &= \left[\frac{\hat{p}_I}{p_I} U^2 - (1 - \hat{p}_I) \right]^{-1} S_{\text{MSR}}^2 [x_i^{\text{MSR}} - \overline{x}(1 + S_{\text{MSR}}^2)] \qquad &\text{for } i < I \ . \end{aligned} \tag{B14}$$

Therefore the maximal-Sortino and maximal-SVP ratio portfolios are long positions in the MSRP with an extra long position in the state security paying off in the highest market return state combined with borrowing or lending.

C: Sharpe, Sortino, and SVP Ratios in a Lognormal Market

In this section of the appendix, we derive the Sharpe, Sortino, and SVP ratios in a lognormal market. We denote one plus the realized rate of return on the market benchmark portfolio by b and use this to index the continuous state space.

If the market return has a continuously compounded expected rate of return μ and logarithmic variance of σ^2, the probability density for the return over a period of length Δt is

$$p(b) = \frac{1}{b\sigma\sqrt{\Delta t}}\phi\left(\frac{\ell\text{n}\, b - (\mu - \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right) \tag{C1}$$

where $\phi(\cdot)$ is the standard normal density function. The risk-neutral probability density, $\hat{p}(b)$, is the same with μ replaced by r, the continuously-compounded interest rate.

The Sharpe ratio of the market is

$$S_{\text{mkt}} = \frac{\mathbb{E}[b] - e^{r\Delta t}}{(\text{Var}[b])^{1/2}} = \frac{e^{\mu\Delta t} - e^{r\Delta t}}{[e^{2\mu\Delta t}(e^{\sigma^2\Delta t} - 1)]^{1/2}} = \frac{1 - e^{-(\mu - r)\Delta t}}{(e^{\sigma^2\Delta t} - 1)^{1/2}}\,. \tag{C2}$$

The square of the downside risk in the denominator of the Sortino and SVP ratios is[35]

$$\begin{aligned}
\mathbb{E}[\text{Min}^2(b - e^{r\Delta t}, 0)] &= \int_0^{e^{r\Delta t}} (e^{\ell\text{n}b} - e^{r\Delta t})^2 p(b)\,\text{d}b \\
&= \int_0^{e^{r\Delta t}} (e^{2\ell\text{n}b} - 2e^{r\Delta t}e^{\ell\text{n}b} + e^{2r\Delta t})p(b)\,\text{d}b \\
&= e^{(2\mu+\sigma^2)\Delta t}\Phi\left(\frac{(r - \mu - \frac{3}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right) - 2e^{(\mu+r)\Delta t}\Phi\left(\frac{(r - \mu - \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right) \\
&\quad + e^{2r\Delta t}\Phi\left(\frac{(r - \mu + \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right)\,.
\end{aligned} \tag{C3}$$

The upside potential in the numerator of the SVP measures is

$$\begin{aligned}
\mathbb{E}[\text{Max}(b - e^{r\Delta t}, 0)] &= \int_{e^{r\Delta t}}^{\infty} (e^{\ell\text{n}b} - e^{r\Delta t})\,\text{d}F(b) \\
&= e^{\mu\Delta t}\Phi\left(\frac{(\mu - r + \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right) - e^{r\Delta t}\Phi\left(\frac{(\mu - r - \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}}\right).
\end{aligned} \tag{C4}$$

For the maximal-ratio portfolios, analysis similar to that in Appendix A yields a maximal squared Sharpe ratio of

$$S^2_{\text{MSR}} = \int \hat{p}^2(b)/p(b)\,\text{d}b - 1 = \mathbb{E}[\hat{p}^2(b)/p^2(b)] - 1 \tag{C5}$$

[35] For a normal random variable, z, with mean $\bar{z}$ and variance v^2, the upper and lower partial exponential moments are

$$\int_K^{\infty} e^{\xi z}\,\text{d}F(z) = \exp(\xi\bar{z} + \tfrac{1}{2}\xi^2 v^2)\Phi\left(\frac{K + \bar{z} + \xi v^2 - K}{v}\right)$$

$$\int_0^{K} e^{\xi z}\,\text{d}F(z) = \exp(\xi\bar{z} + \tfrac{1}{2}\xi^2 v^2)\Phi\left(\frac{K - \bar{z} - \xi v^2}{v}\right)\,.$$

for a portfolio with an excess return of $x^*(b) = \overline{x} + \gamma[\hat{p}(b)/p(b) - 1]$ where $\gamma = -\overline{x}/S^2_{\mathrm{MSR}}$ as before. To determine the maximal Sharpe ratio portfolio, we need the likelihood ratio

$$\frac{\hat{p}(b)}{p(b)} = \exp\left\{-\frac{[\ell\mathrm{n}b - (r - \frac{1}{2}\sigma^2)\Delta t]^2}{2\sigma^2\Delta t} + \frac{[\ell\mathrm{n}b - (\mu - \frac{1}{2}\sigma^2)\Delta t]^2}{2\sigma^2\Delta t}\right\}$$
$$= b^{-(\mu-r)/\sigma^2}\exp\left[\left(\frac{\mu + r}{\sigma^2} - 1\right)\frac{\mu - r}{2}\Delta t\right]. \tag{C6}$$

The square of the maximal Sharpe ratio is

$$S^2_{\mathrm{MSR}} = \mathbb{E}[\hat{p}^2(b)/p^2(b)] - 1 = \exp\left[\left(\frac{\mu + r}{\sigma^2} - 1\right)(\mu - r)\Delta t\right]\mathbb{E}[b^{-2(\mu-r)/\sigma^2}] - 1. \tag{C7}$$

Because $\ell\mathrm{n}\, b$ is normally distributed with mean $(\mu - \frac{1}{2}\sigma^2)\Delta t$ and variance $\sigma^2\Delta t$, $\ell\mathrm{n}b^{\xi}$ is also normally distributed with $\mathbb{E}[\ell\mathrm{n}b^{\xi}] = \xi(\mu - \frac{1}{2}\sigma^2)\Delta t$ and variance $\mathrm{Var}[\ell\mathrm{n}b^{\xi}] = \xi^2\sigma^2\Delta t$. Using the moment generating function for a normal random variable,[36] the maximal Sharpe ratio is

$$S^2_{\mathrm{MSR}} = \exp\left[\left(\frac{\mu + r}{\sigma^2} - 1\right)(\mu - r)\Delta t\right] \times \exp\left[-2\frac{\mu - r}{\sigma^2}(\mu - \frac{1}{2}\sigma^2)\Delta t + \frac{2(\mu - r)^2}{\sigma^2}\Delta t\right] - 1$$
$$= \exp\left[\frac{(\mu - r)}{\sigma^2}\Delta t(\mu + r - \sigma^2 - 2r + \sigma^2)\right] - 1 = \exp\left[\frac{(\mu - r)^2}{\sigma^2}\Delta t\right] - 1\,. \tag{C8}$$

The maximal-Sortino ratio and maximal-SVP ratio portfolios will be degenerate with an infinite holding of the security for the highest return "state" and an infinitesimal holding of all other "state" securities. Both maximized ratios will be infinite as well. However, if we restrict attention to "well-formed" portfolios then the ratios can be computed.

For example, consider a call option with a strike price of H written on one dollar invested in the market. The payoff on this portfolio is $C(b) = \mathrm{Max}(b - H, 0)$

The Sharpe, Sortino, and SVP ratios of the portfolio are

$$S = \frac{\mathbb{E}[C(b)] - C_0 e^{r\Delta t}}{\sqrt{\mathbb{E}[C^2(b)] - \mathbb{E}^2[C(b)]}}$$
$$D = \frac{\mathbb{E}[C(b)] - C_0 e^{r\Delta t}}{\sqrt{\mathbb{E}[\mathrm{Max}^2(C_0 e^{r\Delta t} - C(b), 0)]}} \qquad U = \frac{\mathbb{E}[\mathrm{Max}(C(b) - C_0 e^{r\Delta t}, 0)]}{\sqrt{\mathbb{E}[\mathrm{Max}^2(C_0 e^{r\Delta t} - P(b), 0)]}} \tag{C9}$$

where C_0 is the cost of the option. The expectations in the Sharpe ratio are

$$\mathbb{E}[C(b)] = \int_H^{\infty}(b - H)p(b)\,\mathrm{d}b = e^{\mu\Delta t}\Phi(h_H) - H\Phi(h_H^-)$$
$$\mathbb{E}[C^2(b)] = \int_H^{\infty}(b - H)^2 p(b)\,\mathrm{d}b = e^{(2\mu+\sigma^2)\Delta t}\Phi(h_H^+) - 2He^{\mu\Delta t}\Phi(h_H) + H^2(h_H^-)$$
$$\text{where} \quad h_K \equiv \frac{-\ell\mathrm{n}K + (\mu + \frac{1}{2}\sigma^2)\Delta t}{\sigma\sqrt{\Delta t}} \qquad h_K^{\pm} \equiv h_K \pm \sigma\sqrt{\Delta t}\,. \tag{C10}$$

[36] The moment generating function for a normal random variable, z, with mean $\overline{z}$ and variance σ_z^2 is $\psi(s) \equiv \mathbb{E}[\exp(sz)] = \exp(s\overline{z} + \frac{1}{2}s^2\sigma_z^2)$.

The square of the risk measure in the denominators of the Sortino and SVP ratios is

$$
\begin{aligned}
&\int_0^{C_0 e^{r\Delta t}+H} [C_0 e^{r\Delta t} - \mathrm{Max}(b-H,0)]^2 p(b)\,\mathrm{d}b = C_0^2 e^{2r\Delta t}\int_0^H p(b)\,\mathrm{d}b \\
&\quad + \int_H^{C_0 e^{r\Delta t}+H} (C_0 e^{r\Delta t} - b + H)^2 p(b)\,\mathrm{d}b \\
&\quad = C_0^2 e^{2r\Delta t}\Phi(-h_H^-) + (C_0 e^{r\Delta t}+H)^2[\Phi(h_H^-) - \Phi(h^-_{C_0 e^{r\Delta t}+H})] \qquad \text{(C11)} \\
&\qquad -2(C_0 e^{r\Delta t}+H)e^{\mu\Delta t}[\Phi(h_H) - \Phi(h_{C_0 e^{r\Delta t}+H})] + e^{(2\mu+\sigma^2)\Delta t}[\Phi(h_H^+) - \Phi(h^+_{C_0 e^{r\Delta t}+H})]\,.
\end{aligned}
$$

The numerator for the Sortino ratio is the same as for the Sharpe ratio. For the SVP ratio, it is

$$
\int_{H+C_0 e^{r\Delta t}}^{\infty} (b - H - C_0 e^{r\Delta t})p(b)\,\mathrm{d}b = e^{\mu\Delta t}\Phi(h_{H+C_0 e^{r\Delta t}}) - (H + C_0 e^{r\Delta t})\Phi(h^-_{H+C_0 e^{r\Delta t}})\,. \tag{C12}
$$

D: Timing Measures: Henriksson-Merton and Treynor-Mazuy

The Henriksson-Merton (HM) and Treynor-Mazuy (TM) measures are based on the regression

$$
\tilde{x}_t = \gamma_0 + \gamma_1\tilde{m}_t + \gamma_2\tilde{w}_t + \tilde{\varepsilon}_t \tag{D1}
$$

where $\tilde{x}_t$ and $\tilde{m}_t$ are the excess returns on the portfolio and the market and $\tilde{w}_t$ is a variable used to capture timing ability. The variables used by HM and TM are $\tilde{w}_t \equiv \mathrm{Max}(-\tilde{m}_t, 0)$ and $\tilde{w}_t \equiv \tilde{m}_t^2$, respectively.

To achieve a good score on these measures we want both γ_0 and γ_2 as large as possible. For an OLS regression, the coefficients will be

$$
\gamma_1 = \frac{\sigma_{xm}\sigma_w^2 - \sigma_{xw}\sigma_{mw}}{\sigma_m^2\sigma_w^2 - \sigma_{mw}^2} \qquad \gamma_2 = \frac{\sigma_{xw}\sigma_m^2 - \sigma_{xm}\sigma_{mw}}{\sigma_m^2\sigma_w^2 - \sigma_{mw}^2} \qquad \gamma_0 = \overline{x} - \gamma_1\overline{m} - \gamma_2\overline{w}\,. \tag{D2}
$$

The parameters, $\overline{m}, \overline{w}, \sigma_m^2, \sigma_w^2$, and σ_{mw}, are fixed by the environment so the regression coefficients are determined by the portfolio's expected return and its covariances with the market and the timing variable, w. Since the expected return and the covariances are linear in the individual state returns, any values for $\overline{x}, \sigma_{xm}$, and σ_{xw} and hence the regression coefficients can be achieved with four or more assets.[37]

With more than four assets, we have extra degrees of freedom. It makes sense to find the minimum-variance portfolio that achieves a given HM or TM score. Therefore, we fix σ_{xm}, σ_{xw}, and $\overline{x}$ (which fixes γ_0, γ_1, and γ_2) and minimize the portfolio's variance. The optimization problem is

$$
\mathrm{Min} \quad \frac{1}{2}\sum p_i x_i^2
$$

[37] If the CAPM holds, then the mean excess return for each asset would be proportional to its covariance with the market, and we would not be able to set them independently for a portfolio. As we have seen, the CAPM cannot hold in a market with sufficient number of derivative assets.

$$\text{subject to} \quad \sigma_{xm} = \sum p_i x_i (m_i - \overline{m}) \quad \sigma_{xw} = \sum p_i x_i (w_i - \overline{w}) \quad \overline{x} = \sum p_i x_i \quad 0 = \sum \hat{p}_i x_i \tag{D3}$$

where the last constraint fixes the cost of the portfolio's excess returns to zero. Forming the Lagrangian and minimizing gives

$$\begin{aligned} L \equiv \tfrac{1}{2} \textstyle\sum p_i x_i^2 + \lambda_1 \left[\sigma_{xm} - \sum p_i x_i (m_i - \overline{m})\right] + \lambda_2 \left[\sigma_{xw} - \sum p_i x_i (w_i - \overline{w})\right] \\ + \lambda_3 \left[\overline{x} - \textstyle\sum p_i x_i\right] + \lambda_4 \left[0 - \textstyle\sum \hat{p}_i x_i\right] \\ 0 = \frac{\partial L}{\partial x_i} = p_i x_i - \lambda_1 p_i (m_i - \overline{m}) - \lambda_2 p_i (w_i - \overline{w}) - \lambda_3 p_i - \lambda_4 \hat{p}_i \\ \Rightarrow \quad x_i = \lambda_1 (m_i - \overline{m}) + \lambda_2 (w_i - \overline{w}) + \lambda_3 + \lambda_4 \hat{p}_i / p_i \ . \end{aligned} \tag{D4}$$

The optimizing portfolio in a complete market can be split into four parts. It will hold bonds, λ_1 units in the market, $-\lambda_4$ in the MSRP, and λ_2 units in a contract that pays w. For HM, the last will be a put option that is at-the-money in present value terms. For TM, it will be a derivative contract with paying off the squared excess market return.

The Lagrange multipliers that determine these positions are the solutions to the constraints[38]

$$\begin{aligned} \overline{x} &= \sum p_i x_i = \sum p_i [\lambda_1 (m_i - \overline{m}) + \lambda_2 (w_i - \overline{w}) + \lambda_3 + \lambda_4 \hat{p}_i / p_i] = \lambda_3 + \lambda_4 \\ \sigma_{xm} &= \sum p_i (m_i - \overline{m}) x_i = \lambda_1 \sum p_i (m_i - \overline{m})^2 + \lambda_2 \sum p_i (m_i - \overline{m})(w_i - \overline{w}) \\ &\quad + \lambda_3 \sum p_i (m_i - \overline{m}) + \lambda_4 \sum \hat{p}_i (m_i - \overline{m}) \\ &= \lambda_1 \sigma_m^2 + \lambda_2 \sigma_{mw} - \lambda_4 \overline{m} \\ \sigma_{xw} &= \sum p_i (w_i - \overline{w}) x_i = \lambda_1 \sum p_i (m_i - \overline{m})(w_i - \overline{w}) + \lambda_2 \sum p_i (w_i - \overline{w})^2 \\ &\quad + \lambda_3 \sum p_i (w_i - \overline{w}) + \lambda_4 \sum \hat{p}_i (w_i - \overline{w}) \\ &= \lambda_1 \sigma_{mw} + \lambda_2 \sigma_w^2 + \lambda_4 \sum \hat{p}_i (w_i - \overline{w}) \\ 0 &= \sum \hat{p}_i x_s = \sum \hat{p}_i [\lambda_1 (m_i - \overline{m}) + \lambda_2 (w_i - \overline{w}) + \lambda_3 + \lambda_4 \hat{p}_i / p_i] \\ &= -\lambda_1 \overline{m} + \lambda_2 \sum \hat{p}_i (w_i - \overline{w}) + \lambda_3 + \lambda_4 \sum \hat{p}_i^2 / p_i \end{aligned} \tag{D5}$$

Solving for the Lagrange multipliers we have

$$\begin{aligned} x_i &= \overline{x} + \lambda_4 (\hat{p}_i / p_i - 1) + \lambda_1 (m_i - \overline{m}) + \lambda_2 w_i \\ \lambda_1 &= \frac{\sigma_{xm} (S_{\text{MSR}}^2 \sigma_w^2 - \Pi_w^2) - \sigma_{xw} (S_{\text{MSR}}^2 \sigma_{mw} - \overline{m} \Pi_w) + \overline{x} (\sigma_{mw} \Pi_w - \overline{m} \sigma_w^2)}{\sigma_m^2 (S_{\text{MSR}}^2 \sigma_w^2 - \Pi_w^2) - \sigma_{mw} (S_{\text{MSR}}^2 \sigma_{mw} - \overline{m} \Pi_w) + \overline{m} (\sigma_{mw} \Pi_w - \overline{m} \sigma_w^2)} \\ \lambda_2 &= \frac{-\sigma_{xm} (S_{\text{MSR}}^2 \sigma_{mw} - \overline{m} \Pi_w) + \sigma_{xw} (S_{\text{MSR}}^2 \sigma_m^2 - \overline{m}^2) - \overline{x} (\sigma_m^2 \Pi_w - \overline{m} \sigma_{mw})}{\sigma_m^2 (S_{\text{MSR}}^2 \sigma_w^2 - \Pi_w^2) - \sigma_{mw} (S_{\text{MSR}}^2 \sigma_{mw} - \overline{m} \Pi_w) + \overline{m} (\sigma_{mw} \Pi_w - \overline{m} \sigma_w^2)} \end{aligned} \tag{D6}$$

[38] Recall that the risk-neutral expectation of the excess market return m, $\sum \hat{p}_i m_i$ is zero. This is not true of w as it is not the value of a marketed asset. The sum $\sum \hat{p}_i (w_i - \overline{w})$ is the difference between the risk-neutral and true expectations of w, $\hat{\mathbb{E}}[\tilde{w}] - \mathbb{E}[\tilde{w}]$.This is the negative of the forward risk-premium.

where $\lambda_3 = \overline{x} - \lambda_4$

$$\lambda_4 = \frac{\sigma_{xm}(-\Pi_w \sigma_{mw} + \overline{m}\sigma_w^2) + \sigma_{xw}(\sigma_m^2 \Pi_w - \overline{m}\sigma_{mw}) - \overline{x}(\sigma_m^2 \sigma_w^2 - \sigma_{mw}^2)}{\sigma_m^2(S_{\mathrm{MSR}}^2 \sigma_w^2 - \Pi_w^2) - \sigma_{mw}(S_{\mathrm{MSR}}^2 \sigma_{mw} - \overline{m}\Pi_w) + \overline{m}(\sigma_{mw}\Pi_w - \overline{m}\sigma_w^2)}$$

$$\Pi_w \equiv \hat{\mathbb{E}}[\tilde{w}] - \mathbb{E}[\tilde{w}] .$$

The coefficient γ_2 is a measure of the manager's timing ability while γ_0 measures stock selection ability.

Merton (1981) has shown that in the context of the HM model, the total contribution of the manager's timing and selectivity ability per dollar invested over a time interval Δt is $\gamma_0 e^{-r\Delta t} + \gamma_2 P(1, \Delta t; e^{r\Delta t})$ where $P(S, \tau; K)$ is the value of a put option with strike price K and time to maturity τ written on the market. The manager provides an extra average return of γ_0. The first term is the present value of this. (It is discounted at the risk-free rate since the extra return is not correlated with the market.) In addition, through timing the manager provides the fraction γ_2 of a protective put. The total contribution is the amount by which the value of the protective put exceeds its average "cost" measured by the lowered present value of the extra average return.

For the TM model the total contribution of timing and selectivity is

$$\text{total contribution}_{\mathrm{TM}} = \gamma_0 e^{-r\Delta t} + \gamma_2 e^{r\Delta t}(e^{\sigma^2 \Delta t} - 1) . \tag{D7}$$

The first term is the same as for the HM model. The second term is γ_2 multiplied by the present value of the squared excess return on the market. This is not an asset price, but its present value can be determined as follows. Let M_t denote the value of the market portfolio with $M_0 = 1$. We are interested in $w_t \equiv (M_t - e^{rt})^2 = M_t^2 - 2M_t e^{rt} + e^{2rt}$. For a lognormal market, the present value of w_t is

$$\begin{aligned} &e^{-r\Delta t}(\hat{\mathbb{E}}[\tilde{M}_{\Delta t}^2] - 2e^{r\Delta t}\hat{\mathbb{E}}[\tilde{M}_{\Delta t}] + e^{2r\Delta t}) \\ &\quad = e^{-r\Delta t}(e^{(2r+\sigma^2)\Delta t} - 2e^{r\Delta t}e^{r\Delta t} + e^{2r\Delta t}) = e^{r\Delta t}(e^{\sigma^2 \Delta t} - 1) \end{aligned} \tag{D8}$$

where $\hat{\mathbb{E}}$ denotes the risk-neutral expectation. As in the HM model, the manager's timing provides the fraction γ_2 of a derivative security; this time one that pays the square of the excess market return.

References

Admati, A., S. Bhattacharya, P. Pfleiderer, and S. Ross. 1986. On Timing and Selectivity. *Journal of Finance* 41:715–32.

Agarwal, V., and N. Naik. 2000. Multi-Period Performance Persistence Analysis of Hedge Funds. *Journal of Financial and Quantitative Analysis* 35:327–42.

Brown, S., D. R. Gallagher, O. W. Steenbeek, and P. L. Swan. 2004. Double or Nothing: Patterns of Equity Fund Holdings and Transactions. Working Paper, Stern School of Management, NYU.

Brown, S., and W. Goetzmann. 1997. Mutual Fund Styles. *The Journal of Financial Economics* 43:373–99.

Chen, Z. W., and P. J. Knez. 1996. Portfolio Performance Measurement: Theory and Evidence. *Review of Financial Studies* 9:551–6.

Cremers, M., and A. Petajisto. 2006. How Active Is Your Fund Manager? A New Measure That Predicts Performance. Working Paper, Yale School of Management.

Debreu, G. 1960. Topological Methods in Cardinal Utility Theory, in K. J. Arrow, S. Karlin, and P. Suppes (eds), *Mathematical Methods in the Social Sciences, 1959, Proceedings of the First Stanford Symposium*. Stanford CA: Stanford University Press.

Ferson, W., and A. Siegel. 2001. The Efficient Use of Conditioning Information in Portfolios. *Journal of Finance* 3:967–82.

Goetzmann, W., J. Ingersoll, and Z. Ivkovic. 2000. Monthly Measurement of Daily Timers. *Journal of Financial and Quantitative Analysis* 35:257–90.

Goetzmann, W. N., J. E. Ingersoll Jr, and S. A. Ross. 2003. High Water Marks. *Journal of Finance* 58:1685–717.

Henriksson, R. D., and R. C. Merton. 1981. On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *Journal of Business* 54:513–33.

Ingersoll, J. 1987. *Theory of Financial Decision Making, Rowman and Littlefield*, Totowa, NJ.

Jagannathan, R., and R. Korajczyk. 1986. Assessing the Market Timing Performance of Managed Portfolios. *Journal of Business* 59:217–35.

Jensen, M. 1969. Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios. *Journal of Business* 42:167–247.

Lakonishok, J., A. Shleifer, R. Thaler, and R. Vishny. 1991. Window Dressing by Pension Fund Managers. *American Economic Review* 81:227–31.

Leland, H. 1999. Beyond Mean-Variance: Risk and Performance Measurement in a Nonsymmetrical World. *Financial Analysts Journal* 55:27–36.

Lhabitant, Francois-Serge. 2000. Derivatives in Portfolio Management: Why Beating the Market is Easy. *Derivatives Quarterly* 6:39–45.

Lo, A. 2002. The Statistics of Sharpe Ratios. *Financial Analysts Journal* 58(July/August):36–52.

Merton, R. C. 1981. On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts. *Journal of Business* 54:363–406.

Mitchell, M., and T. Pulvino. 2001. Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance* 56:2135–75.

Modigliani, F., and L. Modigliani. 1997. Risk-Adjusted Performance. *Journal of Portfolio Management* 23(Winter):45–54.

Morningstar. 2006. The Morningstar Rating™ Methodology Morningstar Methodology Paper available online at http://corporate.morningstar.com/US/documents/Methodology Documents/MethodologyPapers/MorningstarFundRating_Methodology.pdf

Roll, R. 1992. A Mean/Variance Analysis of Tracking Error. *Journal of Portfolio Management* 18(summer):13–22.

Sharpe, W. F. 1966. Mutual Fund Performance. *Journal of Business* 39:119–38.

Sortino, F., and R. van der Meer. 1991. Downside Risk. *Journal of Portfolio Management* 17(summer):27–31.

Sortino, F., R. van der Meer, and A. Plantinga. 1999. The Dutch Triangle. *Journal of Portfolio Management* 26(fall):50–8.

Stutzer, M. 2005. Mutual Fund Ratings: What is the Risk in Risk-Adjusted Fund Returns? Working Paper, University of Colorado, Boulder.

Treynor, J. 1965. How to Rate Management of Investment Funds. *Harvard Business Review* 43(January–February):63–75.

Treynor, J., and K. Mazuy. 1966. Can Mutual Funds Outguess the Market? *Harvard Business Review* 44(July–August):131–6.

Weisman, A. 2002. Informationless Investing and Hedge Fund Performance Measurement Bias. *Journal of Portfolio Management* 26(summer):81–91.