Rubicon Global Advisors White Paper

Evaluating Investment Performance of Portfolio Managers

Dr. Shafiqur Rahman, Ph.D., MBA
Professor of Banking and Finance
Portland State University
Portland, OR 97207-0751
rahmans@pdx.edu

May 7th, 2015

Rubicon Global Advisors
5200 Meadows Road • Suite 150
Lake Oswego, OR 97035
503.997.2007 (Direct)
503.726.5210 (Office)
503.726.5211 (Fax)
www.rubiconglobalgroup.com

Evaluating Investment Performance of Portfolio Managers

© Dr. Shafiqur Rahman, Ph.D., MBA
Professor of Banking and Finance
Portland State University
Portland, OR 97207-0751
rahmans@pdx.edu

The investment performance of mutual fund managers has been extensively examined in the finance literature.[1] Beginning with the work of Sharpe (1966), Treynor (1965), and Jensen (1968, 1969), all these studies have been concerned with measuring performance in two dimensions: risk and return. These methods compare a particular manager's performance with that of a benchmark portfolio (such as S&P 500 Index, Russell 3000 Index) after adjusting for difference in risk across portfolios and the benchmark. This white paper discusses these various risk-adjusted measures that can be employed to evaluate investment performance of portfolio managers in general.

Alternative Measures of Risk

Before discussing the alternative measures, we first discuss the various statistical measures used in computing performance measures. **Standard deviation**, a statistical measure of dispersion of periodic returns around the average return, depicts how widely a portfolio's returns may vary over time. Standard deviation is a measure of total risk an investor faces from an investment portfolio. Investors use the standard deviation of historical returns to predict the range of returns that are most likely for a given portfolio. A high standard deviation indicates that the future returns are likely to vary widely implying greater volatility

[1] See Lee and Rahman (1994) for a survey of these studies.

or risk. The standard deviation is computed using the trailing monthly total returns for a large number of months.

Beta is a measure of how sensitive is a portfolio's return to market movements. Beta is the slope of regression of returns on a security or portfolio on the market returns:

$$r_{it} = \alpha_i + \beta_i r_{mt} + e_{it}$$

where r_{it} is the return on ith security or portfolio, r_{mt} is the return on the market index, α_i is the intercept term, β_i is the slope term (beta), and e_{it} is the random error term. It is important to note that a low beta for a portfolio does not necessarily imply that the portfolio has a low level of volatility. A low beta signifies only that the portfolio's market-related risk is low. While a portfolio's total risk or absolute volatility is measured by its standard deviation, beta measures its systematic risk, i.e., a part of the total risk of a given portfolio that is due to its relationship or covariability with the market. For example, a gold fund usually has a low beta, as its performance is tied more closely to the price of gold and gold-mining stocks than to the overall stock market. Thus, a gold fund might fluctuate wildly because of rapid changes in gold prices, but its beta will remain low.

Semivariance is a measure of the dispersion of all observations that fall below the mean or target value of a data set. It is an average of the squared deviations (from the mean or target value) of observations that are less than the mean or target value. Semivariance is similar to variance; however, it only considers observations below the mean. A useful tool in portfolio analysis, semivariance provides a measure of downside risk. While standard deviation and variance provide measures of volatility, semivariance only looks at the negative fluctuations of an asset or portfolio. This measure allows potential investors to understand how much downside volatility the portfolio may experience, versus standard deviation, which examines all volatility - upside and downside. It's widely accepted that investors tend to view downside risk (losses) differently from upside potential (gains) in analyzing risk.

Semivariance avoids many of the limitations that plague other measures of risk such as variance and standard deviation. It is an asymmetric measure that focuses on the downside of the probability distribution and avoids penalizing superior performance. It is a relatively complete measure in that it uses all values of the shortfall with their associated probabilities. It is also nonlinear in that it penalizes larger values more than smaller values because of the squaring of the deviations in the calculation. This is more consistent with observed investor behavior because most investors perceive infrequent but large losses as more likely than more frequent but small losses. However, there is a general lack of understanding of semivariance as a measure of risk as opposed to variance and standard deviation that are better known and more integrated into the formal structure of investment decision making. Furthermore, mathematical optimization algorithm used by most practitioners to achieve trade-offs between risk and return is generally not set up to form portfolios taking into consideration semivariance as a measure of risk. Finally, the analytics for bundling securities into portfolios are more difficult using semivariance than they are using variance as a measure of risk. That is, the pairwise relationship between securities (in the form of covariance) cannot be aggregated to calculate portfolio semivariance as in case of portfolio variance. As a result, the portfolio has to be treated as a whole rather than building up the risk measures from component securities (Dobbs, 2015).

Maximum drawdown is one of the key indicators to assess the risk of a given portfolio. It is the worst loss that a portfolio suffers during a given period. It measures the largest peak-to-trough decline in the value of a portfolio (before a new peak appears). It is calculated as a percentage as being: (peak value before largest drop minus lowest value before new high established) divided by the peak value before the largest drop. For example, if a portfolio starts being worth $100,000, increases in value to $150,000, decreases to $90,000, increases to $125,000, then decreases to $80,000, then increases to $225,000, the max drawdown is ($150,000- $80,000) / $150,000 = 46.67%. The highest peak of $225,000 is not considered in the calculation because the drawdown began at a peak of $150,000. Likewise the increase to $125,000 before the drop to $80,000 is ignored, because $125,000 was not a new peak.

R-squared (**R^2**) is a statistical measure that represents the percentage of movements of portfolio's return that can be explained by movements in the return of a benchmark. It is simply a measure of the relationship of the portfolio's returns with the benchmark returns. R-squared values range from 0.0 to 1.0. An R-squared of 1.0 means that all movements of a portfolio return are completely explained by movements in the index. A high R-squared (between .85 and 1.0) indicates the portfolio's performance patterns have been in line with the index. A portfolio with a low R-squared (.70 or less) doesn't follow the index much.

Alternative Performance Measures

Now we turn to various risk-adjusted performance measures. **Sharpe ratio** is a risk-adjusted performance measure developed by Nobel Laureate William Sharpe. It is computed by dividing the average excess return (average return minus risk-free rate) for the portfolio by its standard deviation of returns to determine reward per unit of risk. This is called the reward-to-variability ratio. It is simply the ratio of reward (which is good) to variability (which is bad). The reward-to-variability ratio provides an absolute measure of performance of a portfolio. The higher the Sharpe ratio, the better the portfolio's historical risk-adjusted performance. To determine the quality of performance, a portfolio's Sharpe ratio is compared with that of the market portfolio. A portfolio with a Sharpe ratio higher than that of the market would indicate that the portfolio has outperformed the market, while a lower Sharpe ratio would indicate underperformance. (Morningstar calculates the Sharpe ratio for the past 36-month period by dividing a portfolio's annualized excess returns over the risk-free rate by its annualized standard deviation.) The Sharpe ratio is an industry standard measure for evaluating investment performance of managed portfolios because of its simplicity and ease of calculation.

The Sharpe ratio is based on variance presumed to be the only relevant summary statistic for capturing total risk of the portfolio. This measure makes an important implicit assumption. Since the measure is based on total risk, and thus aggregates two components of total risk – nondiversifiable systematic or market risk and diversifiable nonsystematic or idiosyncratic risk, this measure of performance is most appropriate when the investor has all or most of his or her wealth in the portfolio under review. In other words, the Sharpe ratio does not take into account the fact that the investor may hold other portfolios to diversify nonsystematic risk across portfolios (Dobbs, 2015).

Tracking error is defined as the annualized standard deviation of excess returns (that is, the portfolio return less than the returns on the appropriate benchmark or index portfolio). To calculate an annual tracking error, multiply the observed tracking error times the square root of the number of periods in one year. If monthly observations are used to calculate monthly tracking error, the annualized tracking error is calculated by multiplying the observed monthly tracking error times the square root of 12. The resulting statistic reveals the total risk of the portfolio under consideration, controlling for common factors influencing both the actual portfolio and the benchmark. In other words, the tracking error of a portfolio reveals the total risk in excess of the risk of the benchmark (Dobbs, 2015). R^2 and tracking error appear to be negatively related. A portfolio with high R^2 will show lower tracking error and vice versa. However, whether combination of high R^2 and low tracking error is preferable to low R^2 and high tracking error is debatable. As discussed later, $1 - R^2$ is a proxy for the Jensen's alpha. Low R^2 (and high tracking error) may imply that the manager moves away from the index by under- and/or over-weighting sectors/securities in the index in search of mispriced assets to raise alpha.

Sortino ratio, a variation of the Sharpe ratio, differentiates harmful volatility from volatility in general by using a value for downside deviation. The ratio is named after Dr. Frank A. Sortino, who popularized downside risk optimization. The Sortino ratio is the portfolio's excess return over the risk-free rate divided by the semistandard deviation (that is, square root of semivariance), and so it measures the return to "bad" volatility. Volatility caused by negative returns is considered bad or undesirable by an investor, while volatility caused by positive returns is good or acceptable. The Sortino ratio removes the effects of upward price movements on standard deviation to measure only return against downward price volatility by using the semistandard deviation in the denominator. In this way, the Sortino ratio can help an investor assess risk in a better manner than simply looking at excess returns to total volatility, as such a measure does not consider how often returns are positive as opposed to how often they're negative.

As with the Sharpe ratio, higher Sortino ratios indicate more favorable risk/return relations, and the Sortino ratio of a particular fund is most useful when compared to Sortino ratios of comparable portfolios or the benchmark. The Sortino ratio is essentially the same as the Sharpe ratio with one important difference – total risk is defined as downside risk rather than portfolio variance. For this reason, the Sortino ratio is more attractive than the Sharpe ratio when measuring the performance of portfolios whose returns are asymmetric. Like the Sharpe ratio, the Sortino ratio is inappropriate for investors whose total holdings are diversified and reflect no real nonsystematic risk. Choosing between the Sharpe ratio and the Sortino ratio depends on whether the investor wants to focus on upside as well as downside variability or just downside variability. Sharpe ratio is better at analyzing portfolios with low volatility because the Sortino ratio won't have enough observations to derive reliable semivariance. This makes Sortino ratio appropriate for analyzing highly volatile portfolios (Dobbs, 2015).

Alpha is a measure of the difference between a portfolio's realized return and its return required to compensate for the systematic risk (measured by beta) undertaken by the portfolio. Alpha was developed by Jensen (1968, 1969) and is popularly known as Jensen's alpha. Alpha is based on a least-squares regression of the portfolio's return over Treasury bills return (called excess return) on the excess return of the benchmark or index:

$$R_{pt} = \alpha_p + \beta_p R_{mt} + \varepsilon_{pt}$$

where R_{pt} is the excess return (actual return minus risk-free return) on the pth portfolio, α_p is the Jensen's alpha, β_p is the beta of the portfolio, R_{mt} is the excess return on the benchmark or index, and ε_{pt} is a random error term that has an expected value of zero. A portfolio is expected to earn $\beta_p R_{mt}$ based on its systematic risk. If a portfolio actually earns more than $\beta_p R_{mt}$, α_p will be positive. A portfolio following a passive strategy (random buy-and-hold policy) can be expected to yield a zero alpha. If the portfolio is not doing well, α_p will be negative. A positive alpha indicates that the portfolio has performed better than its beta would predict. In contrast, a negative alpha indicates the portfolio's underperformance, given the expectations established by the portfolio's beta. Such results might be attributable to incurring excessive expenses in unsuccessful attempt to detect undervalued securities. Thus α_p is a measure of superior or inferior performance of a portfolio.

Jensen's original alpha was based on standard capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) which dictates how security or portfolio returns are generated based on beta and market return:

$$R_{it} = a_i + b_i R_{mt} + e_{it}$$

where R_{it} is the excess return on security or portfolio i for period t, R_{mt} is the excess return on the value-weighted market portfolio, and e_{it} is a zero-mean disturbance term. Alpha has evolved since its original development by Jensen (1968, 1969). Standard CAPM is an *ex ante* model to explain the cross-sectional variation in asset returns. Having recognized the apparent inability of this standard model to explain cross-sectional variation in asset returns, Fama and French (1993) took a different approach. Using the relationship between asset prices and observed variables representing firm characteristics and associated risks, they developed a model to explain cross-sectional variation in asset returns. Their model is an *ex post* or empirical model that looks at the behavior of stock prices and reverse-engineers them to identify which factors explain and predict variation in stock returns.[2] In their model, the excess return on a portfolio or security is explained by sensitivity to three factors: broad market excess return, a factor based on market capitalization, and a factor based on book-to-market equity ratio (B/M):

$$R_{it} = a_i + b_i R_{mt} + c_i SMB_t + d_i HML_t + e_{it}$$

[2] DeMuth (2014) compared this approach to building a transcontinental railroad starting from the final destination, San Francisco, and working backwards.

where SMB_t is the difference between the returns on diversified portfolios of small and large cap stocks, and HML_t is the difference between the returns on diversified portfolios of high B/M stocks (value stocks) and low B/M stocks (growth stocks).

Jegadeesh and Titman (1993) showed that intermediate-term returns tend to persist; stocks with higher returns in the prior year tend to have higher future returns. Carhart (1997) augmented Fama and French's (1993) three-factor model by adding a factor to incorporate Jegadeesh and Titman's (1993) one-year momentum anomaly into Fama and French's (1993) three-factor model:

$$R_{itt} = a_i + b_i R_{mt} + c_i SMB_t + d_i HML_t + s_t MOM_t + e_{it}$$

where MOM_t is the difference between the returns on diversified portfolios of past 12-month winners (i.e., stocks with high returns) and losers (i.e., stocks with low returns). This model is a response to the inability of Fama and French's (1993) three-factor model to capture the part of the variation in asset returns resulting from a momentum investment strategy, as Fama and French (1996) acknowledged. In Carhart's (1997) four-factor model, the excess portfolio returns are regressed on the excess market returns, and portfolios representing SMB, HML, and momentum and the estimated intercept of this regression is the newer version of Jensen's alpha:

$$R_{pt} = \alpha_p + \beta_p R_{mt} + \lambda_{\mathbf{p}} SMB_t + \gamma_{\mathbf{p}} HML_t + \eta_{\mathbf{p}} MOM_t + \varepsilon_{pt}$$

Treynor ratio: Treynor (1965) developed the Treynor ratio that is based on systematic risk as measured by portfolio beta coefficient. The Treynor ratio is computed by dividing the average excess return (net of risk-free return) for the portfolio by its beta. It is also known as the "reward-to-volatility ratio." It measures the risk premium earned per unit of systematic risk. Similar to the Sharpe ratio, Treynor ratio is a measure of reward per unit of risk. Unlike Sharpe ratio, Treynor ratio utilizes "market" risk (beta) instead of total risk (standard deviation). The measure distinguishes between total risk and systematic risk, implicitly assuming that portfolio under review is well diversified, that is, diversifiable nonsystematic has been eliminated and can be ignored. A high Treynor ratio indicates superior performance. Performance relative to the market is measured with reference to the market's Treynor ratio. A Treynor ratio higher (lower) than that of the market portfolio would be indicative of superior (inferior) performance.

Given their similarity, when should the Sharpe or the Treynor ratio be used and why? Actually, given the assumptions underlying each measure, both can be said to be correct. Therefore, it is desirable to calculate both measures for the set of portfolios being evaluated. However, investors who have all or substantially all of their assets in a portfolio should rely more on the Sharpe ratio because it assesses the portfolio's total return in relation to total risk, which includes any nonsystematic risk assumed by the investor. On the other hand, for those investors whose portfolio constitutes only one relatively small part of their total assets – that is, they have numerous other assets – systematic risk may well be the relevant risk. In these circumstances, the Treynor ratio is appropriate because it considers only systematic or nondiversifiable risk (Jones, 2002). Sharpe ratio has an apparent advantage over the Treynor ratio though. While the Treynor ratio depends on the validity of the CAPM and the reliability of beta estimate, the

Sharpe ratio is independent of the CAPM and less prone to associated measurement error.

Appraisal or information ratio is another measure based on the CAPM. It is defined as the ratio of Jensen's alpha and the standard deviation of the residuals from the CAPM regression. The former is an estimate of the average excess return on a portfolio over and above the excess return that exactly compensates investors for the systematic risk of the portfolio and the latter is a proxy for the nonsystematic or diversifiable risk of the portfolio. The appraisal ratio thus reveals the average value added by managers (above the systematic risk-based excess return) per unit of nonsystematic risk. The information ratio is often used to gauge the skill of managers of mutual funds, hedge funds, etc. In this case, it measures the active return of the manager's portfolio divided by the amount of risk that the manager undertakes relative to the benchmark. The denominator is often referred to as the portfolio's active risk. The higher the information ratio, the higher the active return of the portfolio, given the amount of risk taken, and the better the manager (Dobbs, 2015).

$1 - R^2$: Amihud and Goyenko (2013) show that R^2 from regression of a fund's returns on factor returns in a multifactor model can be used as a proxy for a mutual fund's performance usually measured by regression intercept or alpha. R^2 is the proportion of the fund return variance that is explained by the variation in these factors; thus, lower R^2 means that the fund tracks them less closely. If lower R^2 helps improve the mutual fund managers' performance, R^2 should be negatively related to alpha, and Amihud and Goyenko (2013) find statistically significant negative correlation between R^2 and alpha. Performance is thus measured by $1 - R^2$, the proportion of the fund's variance that is not captured by the relationship between the return on the portfolio and the benchmark return.

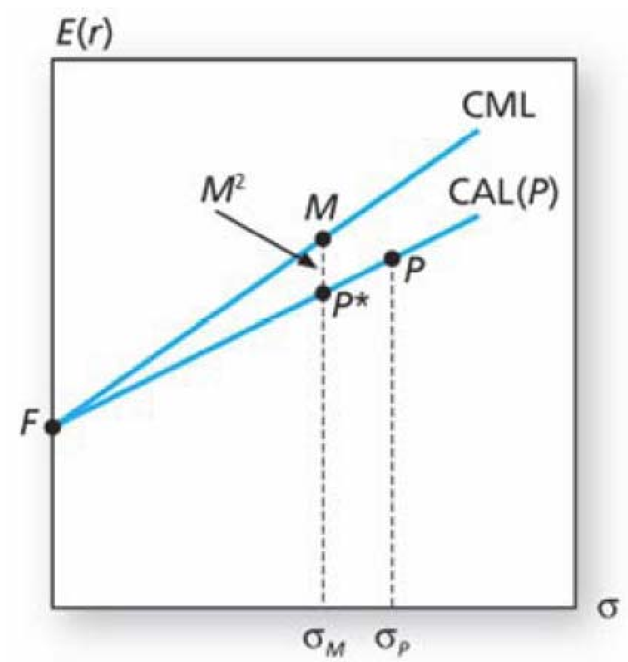
M-squared (**M^2**) is a performance measure that adjusts for the total risk of a managed portfolio. M^2 is named after Nobel Laureate Franco Modigliani and his grand daughter Leah Modigliani who jointly developed the measure (Modigliani and Modigliani, 1997). This measure simply takes a portfolio's average return and determines what it would have been if the portfolio had had the same degree of total risk as the market portfolio usually represented by the S&P 500 Index. A hypothetical portfolio (P*) made up of the T-bills and the managed portfolio (P) is created with the same standard deviation as the market or benchmark portfolio (M). The return on the hypothetical portfolio is

$$R_{P*} = (1 - \sigma_M/\sigma_P) \times R_F + (\sigma_M/\sigma_P) \times R_P$$

where R_{P*} is return on the hypothetical portfolio, R_P is return on the managed portfolio, R_F is return on the risk-free asset or T-bills, σ_M is standard deviation of market return, and σ_P is standard deviation of managed portfolio return. R_{P*} measures the return an investor would have earned if the portfolio had been altered by use of the risk-free rate to match the market portfolio's risk level as measured by standard deviation. Because the market index and *P** have the same standard deviation, their returns are comparable. M^2 is the difference between the return on P* and the return on the market portfolio:

$$M^2 = R_{P*} - R_M$$

where R_M is the return on the market portfolio. The managed portfolio is a superior (inferior) performer if M^2 is positive (negative). M^2 is directly related to the Sharpe ratio. Therefore, a higher Sharpe ratio implies a higher M^2, and a lower Sharpe ratio implies a lower M^2 (Dobbs, 2015). M^2 is graphically presented in the table that follows.



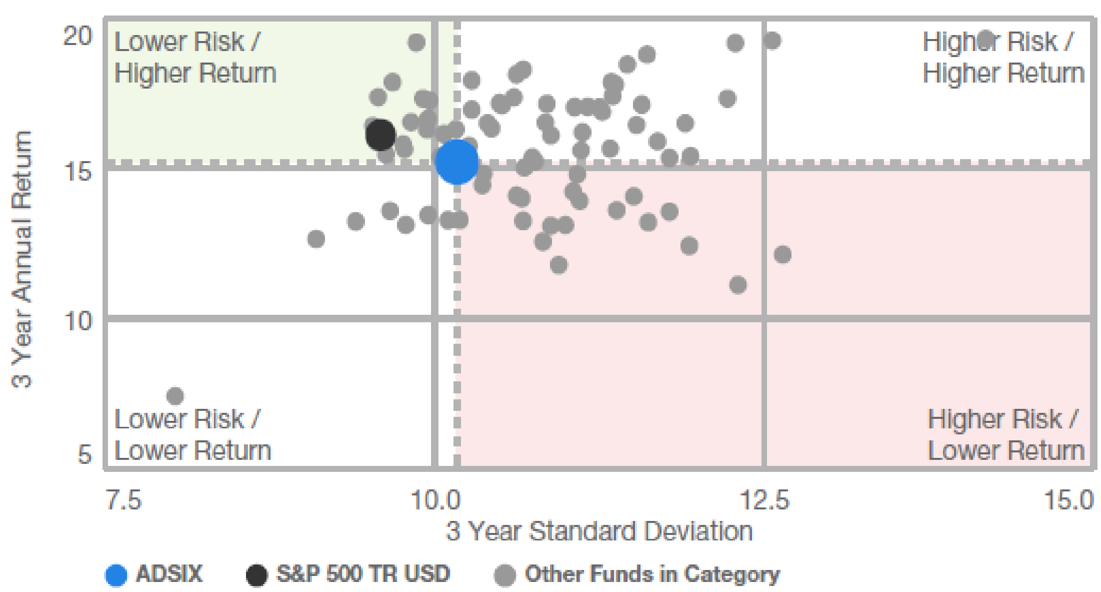
M^2 will evaluate the skill of a manager exactly does the Sharpe ratio. The Jensen's alpha and the Treynor ratio will produce the same conclusions regarding the existence of manager skill. However, it is possible for the Sharpe ratio and M^2 to identify a manager as not skillful, although the alpha and the Treynor ratio may come to the opposite conclusion. This outcome is likely to occur in instances where the manager takes on a large nonsystematic risk in the portfolio relative to the portfolio's systematic risk. In that case, while the numerator remains the same in the Sharpe as well as Treynor ratio, increased nonsystematic risk will lower the

Sharpe ratio but leave the Treynor ratio unaffected. As the market index, by definition, has no nonsystematic risk, the portfolio's performance will look weaker relative to the market index under the Sharpe ratio than under the Treynor ratio and Jensen's alpha (Bailey, Richards, and Tierney, 2009).

Position in risk-return space: Portfolio performance can be measured by examining the position of a given portfolio in risk-return space. A portfolio can be placed in one of the four quadrants in risk-return space. When portfolio returns are plotted against their standard deviation, portfolios may have low return and low risk, low return and high risk, high return and low risk, and high return and high risk. Portfolios lying on the northwest quadrant of the risk-return space are the ones generating highest return while undertaking lowest level of risk and are most desirable. The table that follows shows an example of how portfolios are scattered in risk-return space and the location of S&P 500 Index and the fund ADSIX on the northwest quadrant. The northwest quadrant portfolios are considered superior performer compared to those lying in other three quadrants. These portfolios are the ones with the highest Sharpe ratio.

RoMaD (return over maximum drawdown) is a risk-adjusted performance measure used as an alternative to the Sharpe ratio or the Sortino ratio, applied primarily when evaluating performance of hedge funds (Hayes, 2015). It is expressed as portfolio return divided by maximum drawdown. Drawdown is the difference between a portfolio's point of maximum return (the "high-water" mark), and any subsequent low point of performance. Maximum drawdown is the largest difference between a high-water mark and a subsequent low level.

Maximum drawdown is becoming the preferred way of expressing the risk of a hedge fund portfolio for investors who believe that observed loss patterns over longer periods of time are the best available proxy for actual exposure. If the maximum achieved value for a portfolio to-date was $1,000 and the subsequent minimum level was $900, the maximum drawdown is 10% = [($1000 − $900) / $1000]. Return over maximum drawdown is the average return in a given period for a portfolio, expressed as a proportion of the maximum drawdown level. It enables investors to ask the question: are they willing to accept an occasional drawdown of x percent in order to generate an average return of y percent? An investment with a maximum drawdown of 20 percent and an average return of 10 percent (RoMaD = 0.50) would be considered the more attractive investment than one with a maximum drawdown of 40 percent and an average return of 10 percent (RoMaD = 0.25).



References

Amihud, Y., and Goyenko, R. (2013). Mutual Fund's R^2 as Predictor of Performance. *Review of Financial Studies*, 26, 667-694.

Bailey, J., Richards, T., and Tierney, D. (2009). Evaluating Portfolio Performance, in *Investment Performance Measurement*, Philip Lawton and Todd Jankowski (Editors). New Jersey, John Wiley & Sons, Inc.

Carhart, M. (1997). On Persistence in Mutual Fund Performance, *Journal of Finance*, 52, 57-82.

DeMuth, P. (2014). What's Up with Fama and French's New 5-Factor Model? The Mysterious New Factor V. *Forbes Online*, January 20. Retrieved from http://www.forbes.com/sites/phildemuth/2014/01/20/whats-up-with-fama-frenchs-new-5-factor-model-the-mysterious-new-factor-v/

Dobbs, J. (2015). The Investment Advisor Body of Knowledge: Readings for the CIMA Certification. New Jersey, John Wiley & Sons, Inc.

Fama, E. F., and French, K. R. (1993). Common Risk Factors in the Returns on Bonds and Stocks, *Journal of Financial Economics*, 33, 3-53.

Fama, E. F., and French, K. R. (1996). Multifactor Explanation of Asset Pricing Anomalies. *Journal of Finance*, 51, 55-84.

Hayes, Adam (2015). Return over Maximum Drawdown (RoMaD). Retrieved from http://www.investopedia.com/terms/r/return-over-maximum-drawdown-romad.asp?view=print

Jegadeesh, N., and Titman, S. (1993). Returns to Buying winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48, 93-130.

Jensen, M. (1968). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, 23, 389-416.

Jensen, M. (1969). Risk, the Pricing of Capital Assets, and the Evaluations of Investment Portfolios, *Journal of Business*,42, 167-247.

Jones, C. (2002). Investments Analysis and Management, 8th Ed. New Jersey, John Wiley & Sons, Inc.

Lee, C. F., and Rahman, S. (1990). Market Timing, Selectivity, and Mutual Fund Performance: An Empirical Investigation. *Journal of Business*, 63, 261-278.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics*, 47, 13-37.

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

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34, 768-783.

Sharpe, W. (1964). Capital Asset Prices, A Theory of Market Equilibrium Under Conditions of Risk, *Journal of Finance*, 19, 425-42.

Sharpe, W. (1966). Mutual Fund Performance. *Journal of Business*, 39, 119-138.

Treynor, J. L. (1965). How to Rate Management of Investment Fund. *Harvard Business Review*, 13, 63-75.