CHAPTER TWENTY-FOUR

Portfolio Performance Evaluation

MOST FINANCIAL ASSETS are managed by professional investors, who thus at least indirectly allocate the lion's share of capital across firms. Efficient allocation therefore depends on the quality of these professionals and the propensity of financial markets to direct capital to the best stewards. Therefore, if capital markets are to be reasonably efficient, investors must be able to measure the performance of their asset managers. And such measurement must be sufficiently accurate to allow a proper rank ordering of their ability. At a minimum, efficiency demands that performance evaluation be accurate enough to distinguish managers who (net of fees) are inferior to a randomly chosen, diversified portfolio. Thus, the social benefit of a reasonably reliable performance evaluation method is as large as that of market efficiency.

How can we evaluate the performance of a portfolio manager? It turns out that even an average portfolio return is not as straightforward to measure as it might seem. In addition, adjusting average returns for risk presents a host of other problems. In the end, performance evaluation is far from trivial.

We begin with the measurement of portfolio returns. From there we move on to conventional approaches to risk adjustment. We identify the problems with these approaches when applied in various real-life situations. We then turn to some practical procedures for performance evaluation in the field such as style analysis, the Morningstar Star Ratings, and in-house performance attribution.

24.1 The Conventional Theory of Performance Evaluation

Average Rates of Return

We defined the holding-period return (HPR) in Section 5.1 of Chapter 5 and explained the differences between arithmetic and geometric averages. Suppose we evaluate the performance of a portfolio over a period of 5 years from 20 quarterly rates of return. The arithmetic average of this sample of returns would be the best estimate of the expected rate of return of the portfolio for the next quarter. In contrast, the geometric average is the constant quarterly return over the 20 quarters that would yield the same total cumulative return. Therefore, the geometric average, r_G, is defined by

$$(1 + r_G)^{20} = (1 + r_1)(1 + r_2)\cdots(1 + r_{20})$$

The right-hand side of this equation is the compounded final value of a \$1 investment earning the 20 quarterly rates of return over the 5-year observation period. The left-hand side is the compounded value of a \$1 investment earning r_G *each* quarter. We solve for $1 + r_G$ as[1]

$$1 + r_G = [(1 + r_1)(1 + r_2)\cdots(1 + r_{20})]^{1/20}$$

Each return has an equal weight in the geometric average. For this reason, the geometric average is referred to as a **time-weighted average.**

To set the stage for discussing the more subtle issues that follow, let us start with a trivial example. Consider a stock paying a dividend of \$2 annually that currently sells for \$50. You purchase the stock today, collect the \$2 dividend, and then sell the stock for \$53 at year-end. Your rate of return is

$$\frac{\text{Total proceeds}}{\text{Initial investment}} = \frac{\text{Income} + \text{Capital gain}}{50} = \frac{2 + 3}{50} = .10, \text{ or } 10\%$$

Another way to derive the rate of return that is useful in the more difficult multiperiod case is to set up the investment as a discounted cash flow problem. Call r the rate of return that equates the present value of all cash flows from the investment with the initial outlay. In our example the stock is purchased for \$50 and generates cash flows at year-end of \$2 (dividend) plus \$53 (sale of stock). Therefore, we solve $50 = (2 + 53)/(1 + r)$ to find again that $r = 10\%$.

Time-Weighted Returns versus Dollar-Weighted Returns

When we consider investments over a period during which cash was added to or withdrawn from the portfolio, measuring the rate of return becomes more difficult. To continue our example, suppose that you were to purchase a second share of the same stock at the end of the first year, and hold both shares until the end of year 2, at which point you sell each share for \$54.

Total cash outlays are

Time	Outlay
0	\$50 to purchase first share
1	\$53 to purchase second share a year later
	Proceeds
1	\$2 dividend from initially purchased share
2	\$4 dividend from the 2 shares held in the second year, plus \$108 received from selling both shares at \$54 each

[1]This formula gives the geometric average as a quarterly rate of return, consistent with the quarterly rates used to compute it. When the observation period is of length h years (1/4 in this example), the *annualized* compounded rate is defined by $1 + r_{GA} = (1 + r_{Gh})^{1/h}$. In general, the annualized geometric average of T observations, each of length h, is $1 + r_{GA} = \left(\prod_{t=1}^{T}(1 + r_t)\right)^{1/hT}$ where $\prod$ is the product operator. In our example with $T = 20$ quarterly observations, each of length $h = ¼$ year, $1/hT = 1/5$, so to find the annualized geometric average, we would take the fifth root of the cumulative return over the 5-year investment period.

Using the discounted cash flow (DCF) approach, we can solve for the average return over the 2 years by equating the present values of the cash inflows and outflows:

$$50 + \frac{53}{1 + r} = \frac{2}{1 + r} + \frac{112}{(1 + r)^2}$$

resulting in $r = 7.117\%$.

This value is called the internal rate of return, or the **dollar-weighted rate of return** on the investment. It is "dollar weighted" because the stock's performance in the second year, when two shares of stock are held, has a greater influence on the average overall return than the first-year return, when only one share is held.

The time-weighted (geometric average) return is 7.81%:

$$r_1 = \frac{53 + 2 - 50}{50} = .10 = 10\% \qquad r_2 = \frac{54 + 2 - 53}{53} = .0566 = 5.66\%$$

$$r_G = (1.10 \times 1.0566)^{1/2} - 1 = .0781 = 7.81\%$$

The dollar-weighted average is less than the time-weighted average in this example because the return in the second year, when more money was invested, is lower.

CONCEPT CHECK 24.1

Shares of XYZ Corp. pay a $2 dividend at the end of every year on December 31. An investor buys two shares of the stock on January 1 at a price of $20 each, sells one of those shares for $22 a year later on the next January 1, and sells the second share an additional year later for $19. Find the dollar- and time-weighted rates of return on the 2-year investment.

Dollar-Weighted Return and Investment Performance

Every household faces several daunting saving goals, for example, the education of children and retirement. Many of these goals allow for tax-sheltered savings, for example, IRAs or 401(k) plans for retirement and 529 plans for college expenses. These accounts are, by their nature, separated from other household assets.

Households have considerable latitude in choices of investment venues and will want to check results from time to time. How should they do this? The answer here is quite simple. First, the household must maintain a spreadsheet of time-dated cash inflows and outflows. It is a simple task to record the current value of the investment account. In this setting, the dollar-weighted average over any investment period will yield the effective rate of return earned for the period.[2]

To guarantee that you can accomplish this important task, be sure to solve Problem 1 at the end of the chapter. See also the nearby Excel Application box.

Adjusting Returns for Risk

Evaluating performance based on average return alone is not very useful. Returns must be adjusted for risk before they can be compared meaningfully. The simplest and most popular way to adjust returns for portfolio risk is to compare rates of return with those of other investment funds with similar risk characteristics. For example, high-yield bond portfolios

[2]Excel's function XIRR allows you to input sums at any date. The function provides the IRR between any two dates given a starting value, cash flows at various dates in between (with additions given as negative numbers, and withdrawals as positive values), and a final value on the closing date.

eXcel APPLICATIONS: Simple Investment Account

An investment account starts with an initial contribution of $10,000 dollars. Over a time period of 2 years, the account experiences both inflows and outflows of cash in the form of additional contributions, withdrawals, and dividends (not reinvested). Using Excel's XIRR function, this spreadsheet shows the dollar-weighted average return of the account.

Excel Questions

1. What would happen to the rate of return if, instead of withdrawing security S2 in October, the investor holds onto it? Explain the difference in returns.
2. How much would the investor need to contribute to the account in 03/12 to bring the dollar-weighted average return up to a value of zero?

	A	B	C	D	E	F	G	H	I	J
1				Data						
2	LEGEND:			Initial contribution	$ 10,000.00					
3	Given data					Initial Price	# of Shares	Monthly Dividend	Total Dividend	End Price
4	Value calculated			(S1) Dividend-yielding security	$ 5,000.00	$ 9.50	526	$ 0.30	$ 157.89	$ 13.00
5	See comment			(S2) Non-dividend-yielding security	$ 5,000.00	$ 9.50	526	0	0	$ 11.00
6				(S3) Non-div yielding security (03/12)	$ 8,000.00	$ 8.80	909	0	0	$ 13.50
7										
8										
9		1-Mar-11	$ (10,000.00)							
10		1-Apr-11	$ 157.89							
11		2-May-11	$ 157.89							
12		1-Jun-11	$ 157.89							
13		1-Jul-11	$ 157.89							
14		1-Aug-11	$ 157.89							
15		1-Sep-11	$ 157.89							
16		3-Oct-11	$ (5,631.58)							
17		1-Nov-11	$ 157.89							

In 10/11, investor withdraws all of security S2 from the account at the current price.

are grouped into one "universe," growth stock equity funds are grouped into another universe, and so on. Then the (usually time-weighted) average returns of each fund within the universe are ordered, and each portfolio manager receives a percentile ranking depending on relative performance with the **comparison universe.** For example, the manager with the ninth-best performance in a universe of 100 funds would be the 90th percentile manager: Her performance was better than 90% of all competing funds over the evaluation period.[3] The nearby box reports on Vanguard's recent revamp of its benchmark indexes for several asset classes.

These relative rankings are usually displayed in a chart such as that in Figure 24.1. The chart summarizes performance rankings over four periods: 1 quarter, 1 year, 3 years, and 5 years. The top and bottom lines of each box are drawn at the rate of return of the 95th and 5th percentile managers. The three dashed lines correspond to the rates of return of the 75th, 50th (median), and 25th percentile managers. The diamond is drawn at the average return of a particular fund and the square is drawn at the return of a benchmark index such as the S&P 500. The placement of the diamond within the box is an easy-to-read representation of the performance of the fund relative to the comparison universe.

This comparison of performance with other managers of similar investment style is a useful first step in evaluating performance. However, such rankings can be misleading. Within a particular universe, some managers may concentrate on particular subgroups, so that

[3] In previous chapters (particularly in Chapter 11 on the efficient market hypothesis), we have examined whether actively managed portfolios can outperform a passive index. For this purpose we looked at the distribution of alpha values for samples of mutual funds. We noted that any conclusion from such samples was subject to error due to survivorship bias if funds that failed during the sample period were excluded from the sample. In this chapter, we are interested in how to assess the performance of individual funds (or other portfolios) of interest. When a particular portfolio is chosen today for inspection of its returns going forward, survivorship bias is not an issue. However, comparison groups must be free of survivorship bias. A sample comprised only of surviving funds will bias the return of the benchmark group upward and the *relative* performance of any particular fund downward.

Vanguard to Change Target Benchmarks for 22 Index Funds

Vanguard plans to transition six international stock index funds to FTSE benchmarks and 16 U.S. stock and balanced index funds to new benchmarks developed by the University of Chicago's Center for Research in Security Prices (CRSP). The transition from the current MSCI benchmarks for the 22 funds is expected to result in considerable savings for the funds' shareholders over time.

"The indexes from FTSE and CRSP are well constructed, offer comprehensive coverage of their respective markets, and meet Vanguard's 'best practice' standards for market benchmarks," said Vanguard Chief Investment Officer Gus Sauter. "Equally important, and with our clients' best interests in mind, we negotiated licensing agreements for these benchmarks that we expect will enable us to deliver significant value to our index fund and ETF shareholders and lower expense ratios over time." In an environment in which index licensing fees, in general, have represented a growing portion of the expenses that investors pay to own index funds and ETFs, Mr. Sauter noted that the long-term agreements with FTSE and CRSP will provide cost certainty going forward with these two index providers.

In 2009, CRSP engaged with Vanguard to create a new series of investable indexes, the CRSP Indexes. Vanguard will be the first investment management firm to track CRSP's broadly diversified benchmarks that cover the broad U.S. market, market capitalization segments, and styles. CRSP's capitalization-weighted methodology introduces the unique concept of "packeting," which cushions the movement of stocks between adjacent indexes and allows holdings to be shared between two indexes of the same family. This approach maximizes style purity while minimizing index turnover.

Sixteen Vanguard stock and balanced index funds, with aggregate assets of $367 billion, will track CRSP benchmarks, including Vanguard's largest index fund, the $197 billion Vanguard Total Stock Market Index Fund. The fund and its ETF Shares (ticker: VTI) will transition from the MSCI U.S. Broad Market Index to the CRSP US Total Market Index.

The benchmark changes will encompass all share classes of the 22 funds, including ETFs. The transitions will be staggered and are expected to occur collectively over a number of months. No changes are planned for Vanguard U.S. stock index funds seeking to track Russell and Standard & Poor's benchmarks, or the 11 Vanguard sector equity funds currently seeking to track MSCI benchmarks.

Source: October 2, 2012 © The Vanguard Group, Inc., used with permission.

portfolio characteristics are not truly comparable. For example, within the equity universe, one manager may concentrate on high-beta or aggressive growth stocks. Similarly, within fixed-income universes, durations can vary across managers. These considerations suggest that a more precise means for risk adjustment is desirable.

Methods of risk-adjusted performance evaluation using mean-variance criteria came on stage simultaneously with the capital asset pricing model. Jack Treynor,[4] William Sharpe,[5] and Michael Jensen[6] recognized immediately the implications of the CAPM for rating the performance of managers. Within a short time, academicians were in command of a battery of performance measures, and a bounty of scholarly investigation of mutual fund performance was pouring from ivory towers. Shortly thereafter, agents emerged who were willing to supply rating services to portfolio managers and their clients.

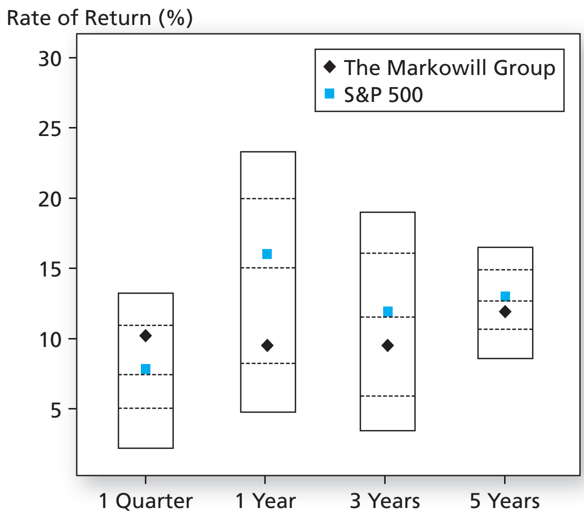


Figure 24.1 Universe comparison, periods ending December 31, 2010

[4]Jack L. Treynor, "How to Rate Management Investment Funds," *Harvard Business Review* 43 (January–February 1966).

[5]William F. Sharpe, "Mutual Fund Performance," *Journal of Business* 39 (January 1966).

[6]Michael C. Jensen, "The Performance of Mutual Funds in the Period 1945–1964," *Journal of Finance,* May 1968; and "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios," *Journal of Business,* April 1969.

But while widely used, risk-adjusted performance measures each have their own limitations. Moreover, their reliability requires quite a long history of consistent management with a steady level of performance and a representative sample of investment environments: bull as well as bear markets.

We start by cataloging some possible risk-adjusted performance measures for a portfolio, *P*, and examine the circumstances in which each measure might be most relevant.

1. *Sharpe ratio:* $(\bar{r}_P - \bar{r}_f)/\sigma_P$

   Sharpe's ratio divides average portfolio excess return over the sample period by the standard deviation of returns over that period. It measures the reward to (total) volatility trade-off.[7]

2. *Treynor measure:* $(\bar{r}_P - \bar{r}_f)/\beta_P$

   Like the Sharpe ratio, **Treynor's measure** gives excess return per unit of risk, but it uses systematic risk instead of total risk.

3. *Jensen's alpha:* $\alpha_P = \bar{r}_P - [\bar{r}_f + \beta_P(\bar{r}_M - \bar{r}_f)]$

   Jensen's alpha is the average return on the portfolio over and above that predicted by the CAPM, given the portfolio's beta and the average market return.[8]

4. *Information ratio:* $\alpha_P/\sigma(e_P)$

   The **information ratio** divides the alpha of the portfolio by the nonsystematic risk of the portfolio, called "tracking error" in the industry. It measures abnormal return per unit of risk that in principle could be diversified away by holding a market index portfolio.

5. *Morningstar risk-adjusted return:* $MRAR(\gamma) = \left[\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1 + r_t}{1 + r_{ft}}\right)^{-\gamma}\right]^{\frac{12}{\gamma}} - 1$

   The Morningstar rating is a sort of harmonic average of excess returns, where $t = 1, \ldots, T$ are monthly observations,[9] and γ measures risk aversion. Higher γ means greater punishment for risk. For mutual funds, Morningstar uses $\gamma = 2$, which is considered a reasonable coefficient for an average retail client.[10] MRAR can be interpreted as the risk-free equivalent excess return of the portfolio for an investor with risk aversion measured by γ.

Each performance measure has some appeal. But each does not necessarily provide consistent assessments of performance, because the risk measures used to adjust returns differ substantially.

[7]We place bars over r_f as well as r_P to denote the fact that because the risk-free rate may not be constant over the measurement period, we are taking a sample average, just as we do for r_P. Equivalently, we may simply compute sample average *excess* returns.

[8]In many cases performance evaluation assumes a multifactor market. For example, when the Fama-French 3-factor model is used, Jensen's alpha will be: $\alpha_P = \overline{r_P - r_f} - \beta_{PM}(\overline{r_M - r_f}) - s_P\bar{r}_{SMB} - h_P\bar{r}_{HML}$ where s_P is the loading on the SMB portfolio and h_P is the loading on the HML portfolio. A multifactor version of the Treynor measure also exists. See footnote 13.

[9]The fraction $(1 + r_t)/(1 + r_{ft})$ is well approximated by 1 plus the excess return, R_t.

[10]The MRAR measure is the *certainty-equivalent geometric average excess return* derived from a more sophisticated utility function than the mean-variance function we used in Chapter 6. The utility function is called *constant relative risk aversion (CRRA).* When investors have CRRA, their capital allocation (the fraction of the portfolio placed in risk-free versus risky assets) does not change with wealth. The coefficient of risk aversion is: $A = 1 + \gamma$. When $\gamma = 0$ (equivalently, $A = 1$), the utility function is just the geometric average of gross excess returns:

$$MRAR(0) = \left[\prod_{t=1}^{T}(1 + R_t)\right]^{\frac{12}{T}} - 1$$

CONCEPT CHECK 24.2

Consider the following data for a particular sample period:

	Portfolio *P*	Market *M*
Average return	35%	28%
Beta	1.20	1.00
Standard deviation	42%	30%
Tracking error (nonsystematic risk), $\sigma(e)$	18%	0

Calculate the following performance measures for portfolio *P* and the market: Sharpe, Jensen (alpha), Treynor, information ratio. The T-bill rate during the period was 6%. By which measures did portfolio *P* outperform the market?

The M^2 Measure of Performance

While the Sharpe ratio can be used to rank portfolio performance, its numerical value is not easy to interpret. Comparing the ratios for portfolios *M* and *P* in Concept Check 2, you should have found that $S_P = .69$ and $S_M = .73$. This suggests that portfolio *P* underperformed the market index. But is a difference of .04 in the Sharpe ratio economically meaningful? We often compare rates of return, but these pure numbers are difficult to interpret.

An equivalent representation of Sharpe's ratio was proposed by Graham and Harvey, and later popularized by Leah Modigliani of Morgan Stanley and her grandfather Franco Modigliani, past winner of the Nobel Prize in Economics.[11] Their approach has been dubbed the M^2 measure (for Modigliani-squared). Like the Sharpe ratio, M^2 focuses on total volatility as a measure of risk, but its risk adjustment leads to an easy-to-interpret differential return relative to the benchmark index.

To compute M^2, we imagine that a managed portfolio, *P*, is mixed with a position in T-bills so that the complete, or "adjusted," portfolio matches the volatility of a market index such as the S&P 500. If the managed portfolio has 1.5 times the standard deviation of the index, the adjusted portfolio would be two-thirds invested in the managed portfolio and one-third in bills. The adjusted portfolio, which we call *P**, would then have the same standard deviation as the index. (If the managed portfolio had *lower* standard deviation than the index, it would be leveraged by borrowing money and investing the proceeds in the portfolio.) Because the market index and portfolio *P** have the same standard deviation, we may compare their performance simply by comparing returns. This is the M^2 measure for portfolio *P*:

$$M_P^2 = r_{P^*} - r_M \tag{24.1}$$

[11] John R. Graham and Campbell R. Harvey, "Market Timing Ability and Volatility Implied in Investment Advisors' Asset Allocation Recommendations," National Bureau of Economic Research Working Paper 4890, October 1994. The part of this paper dealing with volatility-adjusted returns was ultimately published as "Grading the Performance of Market Timing Newsletters," *Financial Analysts Journal* 53 (November/December 1997), pp. 54–66. Franco Modigliani and Leah Modigliani, "Risk-Adjusted Performance," *Journal of Portfolio Management,* Winter 1997, pp. 45–54.

Example 24.1 M^2 Measure

Using the data of Concept Check 2, P has a standard deviation of 42% versus a market standard deviation of 30%. Therefore, the adjusted portfolio P^* would be formed by mixing bills and portfolio P with weights 30/42 = .714 in P and 1 − .714 = .286 in bills. The return on this portfolio would be (.286 × 6%) + (.714 × 35%) = 26.7%, which is 1.3% less than the market return. Thus portfolio P has an M_P^2 measure of −1.3%.

A graphical representation of M^2 appears in Figure 24.2. We move down the capital allocation line corresponding to portfolio P (by mixing P with T-bills) until we reduce the standard deviation of the adjusted portfolio to match that of the market index. M_P^2 is then the vertical distance (the difference in expected returns) between portfolios P^* and M. You can see from Figure 24.2 that P will have a negative M^2 when its capital allocation line is less steep than the capital market line, that is, when its Sharpe ratio is less than that of the market index.[12]

Sharpe's Ratio Is the Criterion for Overall Portfolios

Suppose that Jane Close constructs a portfolio and holds it for a considerable period of time. She makes no changes in portfolio composition during the period. In addition, suppose that the daily rates of return on all securities have constant means, variances, and covariances. These assumptions are unrealistic, and the need for them highlights the shortcoming of conventional applications of performance measurement.

Now we want to evaluate the performance of Jane's portfolio. Has she made a good choice of securities? This is really a three-pronged question. First, "good choice" compared with what alternatives? Second, in choosing between two distinct alternative portfolios, what are the appropriate criteria to evaluate performance? Finally, the performance criteria having been identified, is there a rule that will separate basic ability from the random luck of the draw?

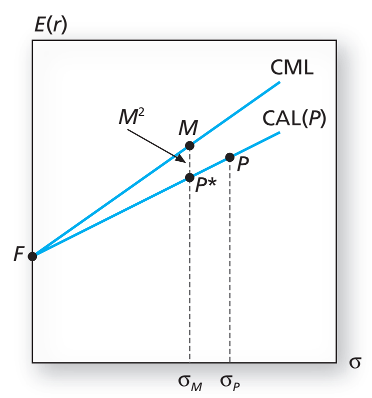


Figure 24.2 M^2 of portfolio P

Earlier chapters of this text help to determine portfolio choice criteria. If investor preferences can be summarized by a mean-variance utility function such as that introduced in Chapter 6, we can arrive at a relatively simple criterion. The particular utility function that we used is

$$U = E(r_P) - \tfrac{1}{2}A\sigma_P^2$$

where A is the coefficient of risk aversion. With mean-variance preferences, Jane wants to maximize the Sharpe ratio $[E(r_P) - r_f]/\sigma_P$. Recall that this criterion led to the selection of the tangency portfolio in Chapter 7. Jane's problem reduces to the search for the portfolio with the highest possible Sharpe ratio.

[12] M^2 is positive when the portfolio's Sharpe ratio exceeds the market's. Letting R denote excess returns and S denote Sharpe measures, the geometry of Figure 24.2 implies that $R_{P^*} = S_P\sigma_M$, and therefore that

$$M^2 = r_{p^*} - r_M = R_{p^*} - R_M = S_P\sigma_M - S_M\sigma_M = (S_P - S_M)\sigma_M$$

M^2 and the Sharpe ratio will therefore rank order portfolios identically.

Appropriate Performance Measures in Two Scenarios

To evaluate Jane's portfolio choice, we first ask whether this portfolio is her exclusive investment vehicle. If the answer is no, we need to know her "complementary" portfolio. The appropriate measure of portfolio performance depends critically on whether the portfolio is the entire investment fund or only a portion of the investor's overall wealth.

Jane's Portfolio Represents Her Entire Risky Investment Fund In this simplest case we need to ascertain only whether Jane's portfolio has the highest Sharpe measure. We can proceed in three steps:

1. Assume that past security performance is representative of expected performance, meaning that realized security returns over Jane's holding period exhibit averages and covariances similar to those that Jane had anticipated.
2. Determine the benchmark (alternative) portfolio that Jane would have held if she had chosen a passive strategy, such as the S&P 500.
3. Compare Jane's Sharpe measure or M^2 to that of the best portfolio.

In sum, when Jane's portfolio represents her entire investment fund, the benchmark is the market index or another specific portfolio. The performance criterion is the Sharpe measure of the actual portfolio versus the benchmark.

Jane's Choice Portfolio Is One of Many Portfolios Combined into a Large Investment Fund This case might describe a situation where Jane, as a corporate financial officer, manages the corporate pension fund. She parcels out the entire fund to a number of portfolio managers. Then she evaluates the performance of individual managers to reallocate the fund to improve future performance. What is the correct performance measure?

The Sharpe ratio is based on average excess return (the reward) against total SD (total portfolio risk). It measures the slope of the CAL. However, when Jane employs a number of managers, nonsystematic risk will be largely diversified away, so systematic risk becomes the relevant measure of risk. The appropriate performance metric is now Treynor's, which takes the ratio of average excess return to beta (because systematic SD = $\beta \times$ market SD).

Consider portfolios P and Q in Table 24.1 and the graph in Figure 24.3. We plot P and Q in the expected return–beta (rather than the expected return–standard deviation) plane, because we assume that P and Q are two of many subportfolios in the fund, and thus that

Table 24.1

Portfolio performance

	Portfolio P	Portfolio Q	Market
Beta	.90	1.60	1.0
Excess return ($\bar{r} - \bar{r}_f$)	11%	19%	10%
Alpha*	2%	3%	0

*Alpha = Excess return − (Beta × Market excess return)
$= (\bar{r} - \bar{r}_f) - \beta(\bar{r}_M - \bar{r}_f) = \bar{r} - [\bar{r}_f + \beta(\bar{r}_M - \bar{r}_f)]$

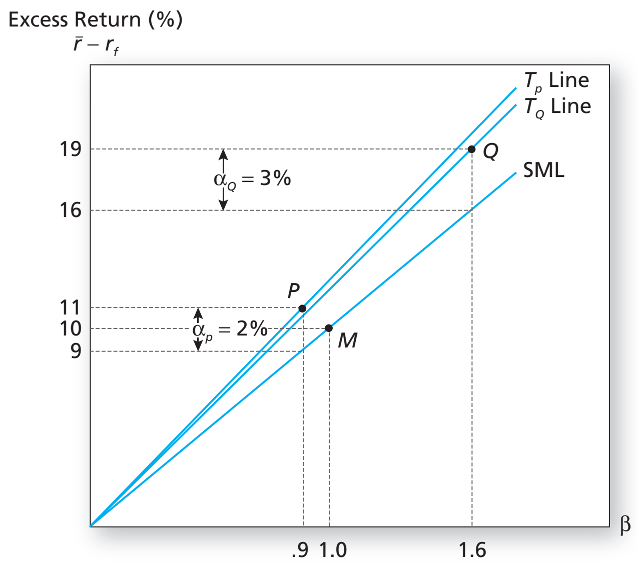


Figure 24.3 Treynor's measure

nonsystematic risk will be largely diversified away. The security market line (SML) shows the value of α_P and α_Q as the distance of P and Q above the SML.

If we invest w_Q in Q and $w_F = 1 - w_Q$ in T-bills, the resulting portfolio, Q^*, will have alpha and beta values proportional to Q's alpha and beta scaled down by w_Q:

$$\alpha_{Q^*} = w_Q \alpha_Q$$
$$\beta_{Q^*} = w_Q \beta_Q$$

Thus all portfolios such as Q^*, generated by mixing Q with T-bills, plot on a straight line from the origin through Q. We call it the T-line for the Treynor measure, which is the slope of this line.

Figure 24.3 shows the T-line for portfolio P as well. P has a steeper T-line; despite its lower alpha, P is a better portfolio after all. For any *given* beta, a mixture of P with T-bills will give a better alpha than a mixture of Q with T-bills.

Example 24.2 Equalizing Beta

Suppose we choose to mix Q with T-bills to create a portfolio Q^* with a beta equal to that of P. We find the necessary proportion by solving for w_Q:

$$\beta_{Q^*} = w_Q \beta_Q = 1.6 w_Q = \beta_P = .9$$
$$w_Q = 9/16$$

Portfolio Q^* therefore has an alpha of

$$\alpha_{Q^*} = 9/16 \times 3\% = 1.69\%$$

which is less than that of P.

The slope of the T-line, giving the trade-off between excess return and beta, is the appropriate performance criterion in this case. The slope for P, denoted by T_P, is given by

$$T_P = \frac{\bar{r}_P - \bar{r}_f}{\beta_P}$$

Like M^2, Treynor's measure is a percentage. If you subtract the market excess return from Treynor's measure, you will obtain the difference between the return on the T_P line in Figure 24.3 and the SML, at the point where $\beta = 1$. We might dub this difference T^2, analogous to M^2. Be aware though that M^2 and T^2 are as different as Sharpe's measure is from Treynor's measure. They may well rank portfolios differently.

The Role of Alpha in Performance Measures

With some algebra we can derive the relationship between the three performance measures discussed so far. The following table shows these relationships.

eXcel APPLICATIONS: Performance Measurement

The following performance measurement spreadsheet computes all the performance measures discussed in this section. You can see how relative ranking differs according to the criterion selected. This Excel model is available at the Online Learning Center (**www.mhhe.com/bkm**).

Excel Questions

1. Examine the performance measures of the funds included in the spreadsheet. Rank performance and determine whether the rankings are consistent using each measure. What explains these results?
2. Which fund would you choose if you were considering investing the entire risky portion of your portfolio? What if you were considering adding a small position in one of these funds to a portfolio currently invested in the market index?

	A	B	C	D	E	F	G	H	I	J	K
1	**Performance Measurement**							**LEGEND**			
2								**Enter data**			
3								**Value calculated**			
4								**See comment**			
5											
6					**Non-**						
7		**Average**	**Standard**	**Beta**	**systematic**	**Sharpe's**	**Treynor's**	**Jensen's**	**M2**	**T2**	**Appraisal**
8	**Fund**	**Return**	**Deviation**	**Coefficient**	**Risk**	**Measure**	**Measure**	**Measure**	**Measure**	**Measure**	**Ratio**
9	Alpha	28.00%	27.00%	1.7000	5.00%	0.8148	0.1294	−0.0180	−0.0015	−0.0106	−0.3600
10	Omega	31.00%	26.00%	1.6200	6.00%	0.9615	0.1543	0.0232	0.0235	0.0143	0.3867
11	Omicron	22.00%	21.00%	0.8500	2.00%	0.7619	0.1882	0.0410	−0.0105	0.0482	2.0500
12	Millennium	40.00%	33.00%	2.5000	27.00%	1.0303	0.1360	−0.0100	0.0352	−0.0040	−0.0370
13	Big Value	15.00%	13.00%	0.9000	3.00%	0.6923	0.1000	−0.0360	−0.0223	−0.0400	−1.2000
14	Momentum Watcher	29.00%	24.00%	1.4000	16.00%	0.9583	0.1643	0.0340	0.0229	0.0243	0.2125
15	Big Potential	15.00%	11.00%	0.5500	1.50%	0.8182	0.1636	0.0130	−0.0009	0.0236	0.8667
16	S & P Index Return	20.00%	17.00%	1.0000	0.00%	0.8235	0.1400	0.0000	0.0000	0.0000	0.0000
17	T-Bill Return	6.00%		0.0000							
18											
19	**Ranking By Sharpe's Measure**				**Non-**						
20		**Average**	**Standard**	**Beta**	**systematic**	**Sharpe's**	**Treynor's**	**Jensen's**	**M2**	**T2**	**Appraisal**
21	**Fund**	**Return**	**Deviation**	**Coefficient**	**Risk**	**Measure**	**Measure**	**Measure**	**Measure**	**Measure**	**Ratio**

	Treynor (T_p)	**Sharpe* (S_p)**
Relation to alpha	$\dfrac{E(r_p) - r_f}{\beta_p} = \dfrac{\alpha_p}{\beta_p} + T_M$	$\dfrac{E(r_p) - r_f}{\sigma_p} = \dfrac{\alpha_p}{\sigma_p} + \rho S_M$
Deviation from market performance	$T_p^2 = T_p - T_M = \dfrac{\alpha_p}{\beta_p}$	$S_p - S_M = \dfrac{\alpha_p}{\sigma_p} - (1 - \rho)S_M$

*ρ denotes the correlation coefficient between portfolio *P* and the market, and is less than 1.

All of these measures are consistent in that superior performance requires a positive alpha. Hence, alpha is the most widely used performance measure. However, positive alpha alone cannot guarantee a better Sharpe ratio for a portfolio. Taking advantage of mispricing means departing from full diversification, which entails a cost in terms of nonsystematic risk. A mutual fund can achieve a positive alpha, yet, at the same time, increase its SD enough that its Sharpe ratio will actually fall.[13]

[13]With a multifactor model, alpha must be adjusted for the additional factors. When you have K factors, $k = 1, \ldots, K$ (the first of which, $k = 1$, is the market index M), a portfolio P's average realized excess return is given by: $\overline{R}_P = \alpha_P + \sum_{k=1}^{K} \beta_{Pk}\overline{R}_k$, where $\overline{R}_k$ is the average return on the zero-investment factor portfolio, or the average excess rate when the direct factor growth rate is used. Hence, the generalization of Jensen's alpha is $\alpha_P^K = \alpha_P - \sum_{k=2}^{K} \beta_{Pk}\overline{R}_k$. The generalized Treynor measure that accounts for all K factors is given by: $GT_P = \alpha_P^K \dfrac{\sum_k \beta_{kM}\overline{R}_k}{\sum_k \beta_{Pk}\overline{R}_k}$, where β_{kM} is the beta of factor k on the index M, and β_{Pk} is the beta of P on factor k. [This measure was developed by Georges Hubner (HEC School of Management, yet unpublished]). Notice that with just one factor, the alpha reduces to the original Jensen's alpha and GT to the single-index Treynor measure.

Actual Performance Measurement: An Example

Now that we have examined possible criteria for performance evaluation, we need to deal with a statistical issue: Can we assess the quality of ex ante decisions using ex post data? Before we plunge into a discussion of this problem, let us look at the rate of return on Jane's portfolio over the last 12 months. Table 24.2 shows the excess return recorded each month for Jane's portfolio *P*, one of her alternative portfolios *Q*, and the benchmark index portfolio *M*. The last rows in Table 24.2 give sample average and standard deviations. From these, and regressions of *P* and *Q* on *M*, we obtain the necessary performance statistics.

The performance statistics in Table 24.3 show that portfolio *Q* is more aggressive than *P*, in the sense that its beta is significantly higher (1.40 vs. .70). At the same time, from its residual standard deviation, *P* is better diversified (2.02% vs. 9.81%). Both portfolios outperformed the benchmark market index, as is evident from their larger Sharpe ratios (and thus positive M^2), their positive alphas, and better Morningstar RAR.

Table 24.2

Excess returns for portfolios *P* and *Q* and the benchmark *M* over 12 months

Month	Jane's Portfolio *P*	Alternative *Q*	Benchmark *M*
1	3.58%	2.81%	2.20%
2	−4.91	−1.15	−8.41
3	6.51	2.53	3.27
4	11.13	37.09	14.41
5	8.78	12.88	7.71
6	9.38	39.08	14.36
7	−3.66	−8.84	−6.15
8	5.56	0.83	2.74
9	−7.72	0.85	−15.27
10	7.76	12.09	6.49
11	−4.01	−5.68	−3.13
12	0.78	−1.77	1.41
Average	2.77	7.56	1.64
Standard deviation	6.45	15.55	8.84

Table 24.3

Performance statistics

	Portfolio *P*	Portfolio *Q*	Portfolio *M*
Sharpe ratio	0.43	0.49	0.19
M^2	2.16	2.66	0.00
Morningstar RAR	0.30	0.80	0.07
SCL regression statistics			
Alpha	1.63	5.26	0.00
Beta	0.70	1.40	1.00
Treynor	3.97	5.38	1.64
T^2	2.34	3.74	0.00
$\sigma(e)$	2.02	9.81	0.00
Information ratio	0.81	0.54	0.00
R-SQR	0.91	0.64	1.00

Which portfolio is more attractive based on reported performance? If P or Q represents the entire investment fund, Q would be preferable on the basis of its higher Sharpe measure (.49 vs. .43) and better M^2 (2.66% vs. 2.16%). For the second scenario, where P and Q are competing for a role as one of a number of subportfolios, Q also dominates because its Treynor measure is higher (5.38 vs. 3.97). However, as an active portfolio to be mixed with the index portfolio, P is preferred because its information ratio (IR $= \alpha/\sigma(e)$) is larger (.81 vs. .54), as discussed in Chapter 8 and restated in the next section. Thus, the example illustrates that the right way to evaluate a portfolio depends in large part on how the portfolio fits into the investor's overall wealth.

This analysis is based on 12 months of data only, a period too short to lend statistical significance to the conclusions. Even longer observation intervals may not be enough to make the decision clear-cut, which represents a further problem. A model that calculates these performance measures is available on the Online Learning Center **(www.mhhe.com/bkm).**

Performance Manipulation and the Morningstar Risk-Adjusted Rating

Performance evaluation so far has been based on this assumption: Rates of return in each period are independent and drawn from the same distribution; in statistical jargon, returns are independent and identically distributed. This assumption can crumble in an insidious way when managers, whose compensation depends on performance, try to game the system. They may employ strategies designed to improve *measured* performance even if they harm investors. Managers' compensation may then lose its anchor to beneficial performance.

Managers can affect performance measures over a given evaluation period because they observe how returns unfold over the course of the period and can adjust portfolios accordingly. Once they do so, rates of return in the later part of the evaluation period come to depend on rates in the beginning of the period.

Ingersoll, Spiegel, Goetzmann, and Welch[14] show how all but one of the performance measures covered in this chapter can be manipulated. The sole exception is the Morningstar RAR, which is in fact a manipulation-proof performance measure (MPPM). While the details of their model are challenging, the logic is straightforward, as we now illustrate using the Sharpe ratio.

As we saw when analyzing capital allocation (Chapter 6), investment in the risk-free asset (lending or borrowing) will not affect the Sharpe ratio of the portfolio. Put differently, the Sharpe ratio is invariant to the fraction y in the risky portfolio (leverage occurs when $y > 1$). The reason is that excess returns are proportional to y and therefore so are both the risk premium and SD, leaving the Sharpe ratio unchanged. But what if y is changed during a period? If the decision to change leverage in mid-stream is made before any performance is observed, then again, the Sharpe measure will not be affected because rates in the two portions of the period will still be uncorrelated.

But imagine a manager already partway into an evaluation period. While realized excess returns (average return, SD, and Sharpe ratio) are now known for the first part of the evaluation period, the distribution of the remaining future rates is still the same as before. The overall Sharpe ratio will be some (complicated) average of the known Sharpe ratio in the first leg and the yet unknown ratio in the second leg of the evaluation period. Increasing leverage during the second leg will increase the weight of that performance in the average

[14]Jonathan Ingersoll, Matthew Spiegel, William Goetzmann, and Ivo Welch, "Portfolio Performance Manipulation and Manipulation Proof Performance Measures," *Review of Financial Studies* 20 (2007).

because leverage will amplify returns, both good and bad. Therefore, managers will wish to increase leverage in the latter part of the period if early returns are poor.[15] Conversely, good first-part performance calls for deleveraging to increase the weight on the initial period. With an extremely good first leg, a manager will shift almost the entire portfolio to the risk-free asset. This strategy induces a (negative) correlation between returns in the first and second legs of the evaluation period.

Investors lose, on average, from this strategy. Arbitrary variation in leverage (and therefore risk) is utility-reducing. It benefits managers only because it allows them to adjust the weighting scheme of the two subperiods over the full evaluation/compensation period after observing their initial performance.[16] Hence investors would like to prohibit or at least eliminate the incentive to pursue this strategy. Unfortunately, only one performance measure is impossible to manipulate.

A manipulation-proof performance measure (MPPM) must fulfill four requirements:

1. The measure should produce a single-value score to rank a portfolio.
2. The score should not depend on the dollar value of the portfolio.
3. An uninformed investor should not expect to improve the expected score by deviating from the benchmark portfolio.
4. The measure should be consistent with standard financial market equilibrium conditions.

Ingersoll et al. prove that the Morningstar RAR fulfills these requirements and is in fact a manipulation-proof performance measure (MPPM). Interestingly, Morningstar was not aiming at an MPPM when it developed the MRAR—it was simply attempting to accommodate investors with constant relative risk aversion.

Panel A of Figure 24.4 shows a scatter of Sharpe ratios vs. MRAR of 100 portfolios based on statistical simulation. Thirty-six excess returns were randomly generated for each portfolio, all with an annual expected return of 7% and SDs varying from 10% to 30%. Thus the true Sharpe ratios of these simulated "mutual funds" are in the range of 0.7 to 0.23, with a mean of .39. Because of sampling variation, the actual 100 Sharpe ratios in the simulation differ quite a bit from these population parameters; they range from −1.02 to 2.46 and average .32. The 100 MRARs range from −28% to 37% and average 0.7%. The correlation between the measures was .94, suggesting that Sharpe ratios track MRAR quite well. Indeed the scatter is pretty tight along a line with a slope of 0.19.

Panel B of Figure 24.4 (drawn on the same scale as panel A) illustrates the effect of manipulation when one leverage change is allowed after initial performance is observed, specifically in the middle of the 36-month evaluation period.[17] The effect of manipulation is evident from the extreme-value portfolios. For high-positive initial MRARs, the switch toward risk-free investments preserves the first-half high Sharpe ratios that might otherwise be diluted or possibly even reversed in the second half. For the high-negative initial MRARs, when leverage ratios are increased, we see two effects. First, MRARs look worse because of cases where the high leverage backfired and worsened the MRARs compared to panel A (points move to the left). In contrast, Sharpe ratios look better than in panel A

[15]Managers who are precluded from increasing leverage will instead shift to high-beta stocks. If this is a widespread phenomenon, it could help explain why high-beta stocks appear, on average, to be overpriced relative to low-beta ones.

[16]One way to reduce the potency of manipulation is to evaluate performance more frequently. This will reduce the statistical precision of the measure, however.

[17]To keep the exercise realistic, leverage ratios were capped at 2 (a debt-to-equity ratio of 1.0).

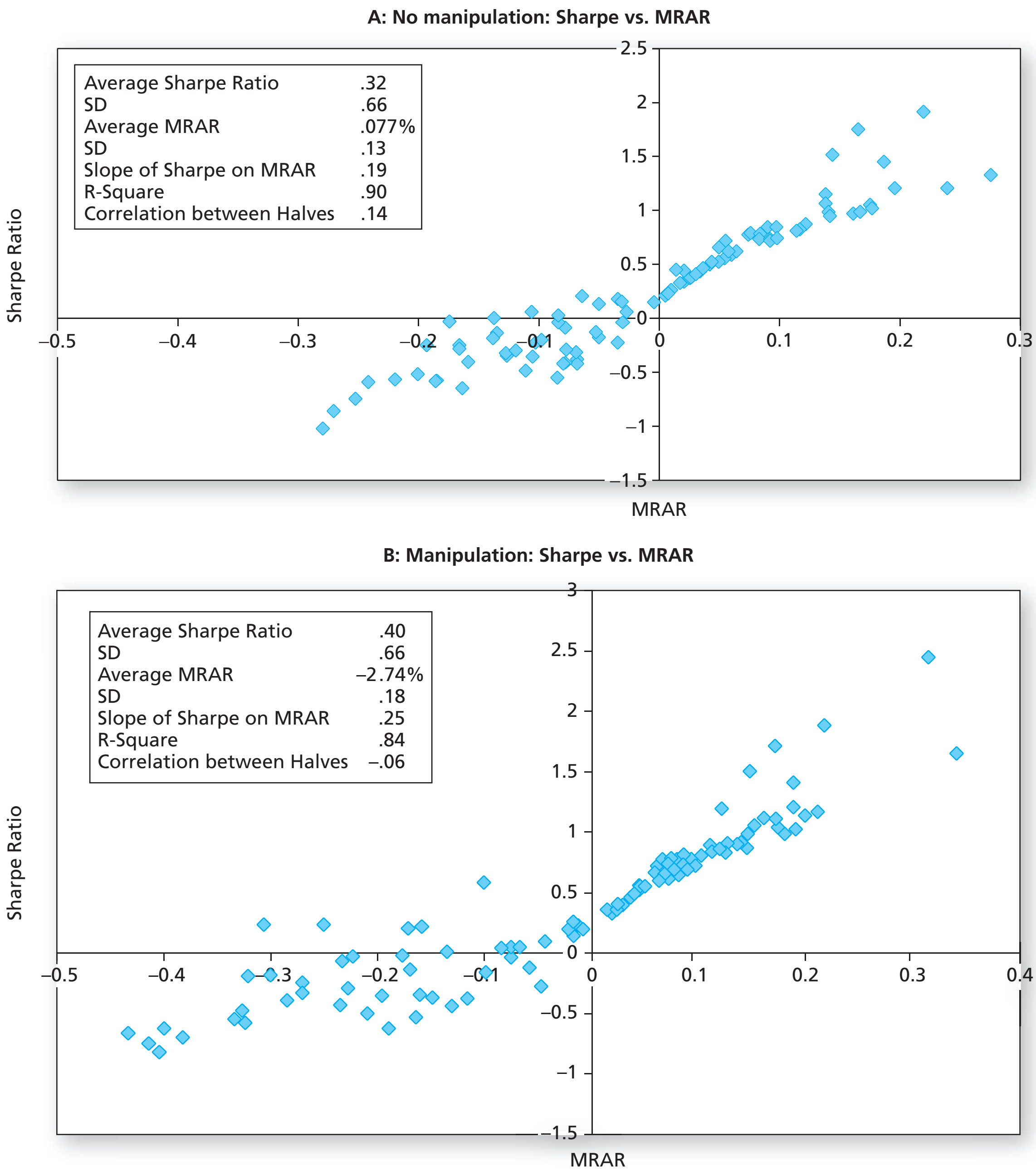


Figure 24.4 MRAR scores and Sharpe ratios with and without manipulation

(they move upward). Some Sharpe ratios move from negative to positive values, while others do not look worse (because the increased SD in the second period reduced the absolute value of the negative Sharpe ratios).

The statistics in the box of panel B quantify the improvement of measured Sharpe ratios; in contrast, MRARs clearly deteriorated from a slight positive value to a certainty-equivalent of −2.74% per year! As predicted, the correlation between average returns in the first and second legs of the period changes from positive to negative. All this happened because of an average increase in leverage from 1.0 to 1.39.[18]

Morningstar introduced the MRAR in 2002. It is particularly relevant to hedge funds, where managers have great latitude and incentive to manipulate. See Chapter 26 for further discussion. Given its immunity to manipulation, we would expect the MRAR measure to become a standard performance statistic sometime in the future, required especially of managers who have the most discretion over investment policy.

Realized Returns versus Expected Returns

When evaluating a portfolio, the evaluator knows neither the portfolio manager's original expectations nor whether those expectations made sense. One can only observe performance after the fact and hope that random results are neither taken for, nor hide, true underlying ability. But risky asset returns are "noisy," which complicates the inference problem. To avoid making mistakes, we have to determine the "significance level" of a performance measure to know whether it reliably indicates ability.

Consider Joe Dart, a portfolio manager. Suppose that his portfolio has an alpha of 20 basis points per month, which makes for a hefty 2.4% per year before compounding. Let us assume that the return distribution of Joe's portfolio has constant mean, beta, and alpha, a heroic assumption, but one that is in line with the usual treatment of performance measurement. Suppose that for the measurement period Joe's portfolio beta is 1.2 and the monthly standard deviation of the residual (nonsystematic risk) is .02 (2%). With a market index standard deviation of 6.5% per month (22.5% per year), Joe's portfolio systematic variance is

$$\beta^2\sigma_M^2 = 1.2^2 \times 6.5^2 = 60.84$$

and hence the correlation coefficient between his portfolio and the market index is

$$\rho = \left[\frac{\beta^2\sigma_M^2}{\beta^2\sigma_M^2 + \sigma^2(e)}\right]^{1/2} = \left[\frac{60.84}{60.84 + 4}\right]^{1/2} = .97$$

which shows that his portfolio is quite well diversified.

To estimate Joe's portfolio alpha from the security characteristic line (SCL), we regress the portfolio excess returns on the market index. Suppose that we are in luck and the regression estimates yield precisely the true parameters. That means that our SCL estimates for the N months are

$$\hat{\alpha} = .2\%, \quad \hat{\beta} = 1.2, \quad \hat{\sigma}(e) = 2\%$$

The evaluator who runs such a regression, however, does not know the true values, and hence must compute the t-statistic of the alpha estimate to determine whether to reject the hypothesis that Joe's alpha is zero, that is, that he has no superior ability.

The standard error of the alpha estimate in the SCL regression is approximately

$$\hat{\sigma}(\alpha) = \frac{\hat{\sigma}(e)}{\sqrt{N}}$$

where N is the number of observations and $\hat{\sigma}(e)$ is the sample estimate of nonsystematic risk. The t-statistic for the alpha estimate is then

$$t(\hat{\alpha}) = \frac{\hat{\alpha}}{\hat{\sigma}(\alpha)} = \frac{\hat{\alpha}\sqrt{N}}{\hat{\sigma}(e)} \tag{24.2}$$

[18]Out of 100 funds, the leverage ratio was decreased in 38 portfolios, was increased to less than 2 in 14 portfolios, and was increased to 2 (and would have been increased even more absent the cap) in 48 portfolios.

Suppose that we require a significance level of 5%. This requires a $t(\hat{\alpha})$ value of 1.96 if N is large. With $\hat{\alpha} = .2$ and $\hat{\sigma}(e) = 2$, we solve Equation 24.2 for N and find that

$$1.96 = \frac{.2\sqrt{N}}{2}$$
$$N = 384 \text{ months}$$

or 32 years!

What have we shown? Here is an analyst who has very substantial ability. The example is biased in his favor in the sense that we have assumed away statistical complications. Nothing changes in the parameters over a long period of time. Furthermore, the sample period "behaves" perfectly. Regression estimates are all perfect. Still, it will take Joe's entire working career to get to the point where statistics will confirm his true ability. We have to conclude that the problem of statistical inference makes performance evaluation extremely difficult in practice.

Now add to the imprecision of performance estimates the fact that the average tenure of a fund manager is only about 4.5 years. By the time you are lucky enough to find a fund whose historic superior performance you are confident of, its manager is likely to be about to move, or has already moved elsewhere. The nearby box explores this topic further.

CONCEPT CHECK 24.3

Suppose an analyst has a measured alpha of .2% with a standard error of 2%, as in our example. What is the probability that the positive alpha is due to luck of the draw and that true ability is zero?

24.2 Performance Measurement for Hedge Funds

In describing Jane's portfolio performance evaluation we left out one scenario that may well be relevant.

Suppose Jane has been satisfied with her well-diversified mutual fund, but now she stumbles upon information on hedge funds. Hedge funds are rarely designed as candidates for an investor's overall portfolio. Rather than focusing on Sharpe ratios, which would entail establishing an attractive trade-off between expected return and overall volatility, these funds tend to concentrate on opportunities offered by temporarily mispriced securities, and show far less concern for broad diversification. In other words, these funds are *alpha driven,* and best thought of as possible *additions* to core positions in more traditional portfolios established with concerns of diversification in mind.

In Chapter 8, we considered precisely the question of how best to mix an actively managed portfolio with a broadly diversified core position. We saw that the key statistic for this mixture is the information ratio of the actively managed portfolio; this ratio, therefore, becomes the active fund's appropriate performance measure.

To briefly review, call the active portfolio established by the hedge fund H, and the investor's baseline passive portfolio M. Then the optimal position of H in the overall portfolio, denoted P^*, would be

$$w_H = \frac{w_H^0}{1 + (1 - \beta_H)w_H^0};\ w_H^0 = \frac{\dfrac{\alpha_H}{\sigma^2(e_H)}}{\dfrac{E(R_M)}{\sigma_M^2}} \tag{24.3}$$

WORDS FROM THE STREET

Should You Follow Your Fund Manager?

The whole idea of investing in a mutual fund is to leave the stock and bond picking to the professionals. But frequently, events don't turn out quite as expected—the manager resigns, gets transferred or dies. A big part of the investor's decision to buy a managed fund is based on the manager's record, so changes like these can come as an unsettling surprise.

There are no rules about what happens in the wake of a manager's departure. It turns out, however, that there is strong evidence to suggest that the managers' real contribution to fund performance is highly overrated. For example, research company Morningstar compared funds that experienced management changes between 1990 and 1995 with those that kept the same managers. In the five years ending in June 2000, the top-performing funds of the previous five years tended to keep beating their peers—despite losing any fund managers. Those funds that performed badly in the first half of the 1990s continued to do badly, regardless of management changes. While mutual fund management companies will undoubtedly continue to create star managers and tout their past records, investors should stay focused on fund performance.

Funds are promoted on their managers' track records, which normally span a three- to five-year period. But performance data that goes back only a few years is hardly a valid measure of talent. To be statistically sound, evidence of a manager's track record needs to span, at a minimum, 10 years or more.

The mutual fund industry may look like a merry-go-round of managers, but that shouldn't worry most investors. Many mutual funds are designed to go through little or no change when a manager leaves. That is because, according to a strategy designed to reduce volatility and succession worries, mutual funds are managed by teams of stock pickers, who each run a portion of the assets, rather than by a solo manager with co-captains. Meanwhile, even so-called star managers are nearly always surrounded by researchers and analysts, who can play as much of a role in performance as the manager who gets the headlines.

Don't forget that if a manager does leave, the investment is still there. The holdings in the fund haven't changed. It is not the same as a chief executive leaving a company whose share price subsequently falls. The best thing to do is to monitor the fund more closely to be on top of any changes that hurt its fundamental investment qualities.

In addition, don't underestimate the breadth and depth of a fund company's "managerial bench." The larger, established investment companies generally have a large pool of talent to draw on. They are also well aware that investors are prone to depart from a fund when a managerial change occurs.

Lastly, for investors who worry about management changes, there is a solution: index funds. These mutual funds buy stocks and bonds that track a benchmark index like the S&P 500 rather than relying on star managers to actively pick securities. In this case, it doesn't really matter if the manager leaves. At the same time, index investors don't have to pay tax bills that come from switching out of funds when managers leave. Most importantly, index fund investors are not charged the steep fees that are needed to pay star management salaries.

Source: Shauna Carther, "Should You Follow Your Fund Manager?" *Investopedia.com,* March 3, 2010. Provided by *Forbes.*

As we saw in Chapter 8, when the hedge fund is optimally combined with the baseline portfolio using Equation 24.3, the improvement in the Sharpe measure will be determined by its information ratio $\alpha_H/\sigma(e_H)$, according to

$$S_{P*}^2 = S_M^2 + \left[\frac{\alpha_H}{\sigma(e_H)}\right]^2 \tag{24.4}$$

Equation 24.4 tells us that the appropriate performance measure for the hedge fund is its information ratio (IR).

Looking back at Table 24.3, we can calculate the IR of portfolios P and Q as

$$\text{IR}_P = \frac{\alpha_P}{\sigma(e_P)} = \frac{1.63}{2.02} = .81; \ \text{IR}_Q = \frac{5.38}{9.81} = .54 \tag{24.5}$$

If we were to interpret P and Q as hedge funds, the low beta of P, .70, could result from short positions the fund holds in some assets. The relatively high beta of Q, 1.40, might result from leverage that would also increase the firm-specific risk of the fund, $\sigma(e_Q)$. Using these calculations, Jane would favor hedge fund P with the higher information ratio.

Sharpe Point: Risk Gauge is Misused

WORDS FROM THE STREET

William F. Sharpe was probably the biggest expert in the room when economists from around the world gathered to hash out a pressing problem: How to gauge hedge-fund risk. About 40 years ago, Dr. Sharpe created a simple calculation for measuring the return that investors should expect for the level of volatility they are accepting. In other words: How much money do they stand to make compared with the size of the up-and-down swings they will lose sleep over?

The so-called Sharpe ratio became a cornerstone of modern finance, as investors used it to help select money managers and mutual funds. But the use of the ratio has been criticized by many prominent academics—including Dr. Sharpe himself.

The ratio is commonly used—"misused," Dr. Sharpe says—for promotional purposes by hedge funds. Hedge funds, loosely regulated private investment pools, often use complex strategies that are vulnerable to surprise events and elude any simple formula for measuring risk. "Past average experience may be a terrible predictor of future performance," Dr. Sharpe says.

Dr. Sharpe designed the ratio to evaluate portfolios of stocks, bonds, and mutual funds. The higher the Sharpe ratio, the better a fund is expected to perform over the long term. However, at a time when smaller investors and pension funds are pouring money into hedge funds, the ratio can foster a false sense of security.

Dr. Sharpe says the ratio doesn't foreshadow hedge-fund woes because "no number can." The formula can't predict such troubles as the inability to sell off investments quickly if they start to head south, nor can it account for extreme unexpected events. Long-Term Capital Management, a huge hedge fund in Connecticut, had a glowing Sharpe ratio before it abruptly collapsed in 1998 when Russia devalued its currency and defaulted on debt. Plus, hedge funds are generally secretive about their strategies, making it difficult for investors to get an accurate picture of risk.

Another problem with the Sharpe ratio is that it is designed to evaluate the risk-reward profile of an investor's entire portfolio, not small pieces of it. This shortcoming is particularly telling for hedge funds.

Source: Ianthe Jeanne Dugan, "Sharpe Point: Risk Gauge is Misused," *The Wall Street Journal,* August 31, 2005, p. C1.

In practice, evaluating hedge funds poses considerable practical challenges. We will discuss many of these in Chapter 26, which is devoted to these funds. But for now we can briefly mention a few of the difficulties:

1. The risk profile of hedge funds (both total volatility and exposure to relevant systematic factors) may change rapidly. Hedge funds have far greater leeway than mutual funds to change investment strategy opportunistically. This instability makes it hard to measure exposure at any given time.
2. Hedge funds tend to invest in illiquid assets. We therefore must disentangle liquidity premiums from true alpha to properly assess their performance. Moreover, it can be difficult to accurately price inactively traded assets, and correspondingly difficult to measure rates of return.
3. Many hedge funds pursue strategies that may provide apparent profits over long periods of time, but expose the fund to infrequent but severe losses. Therefore, very long time periods may be required to formulate a realistic picture of their true risk–return trade-off.
4. Hedge funds have ample latitude to change their risk profiles and therefore considerable ability to manipulate conventional performance measures. Only the MRAR is manipulation-proof, and investors should urge these funds to use them.
5. When hedge funds are evaluated as a group, survivorship bias can be a major consideration, because turnover in this industry is far higher than for investment companies such as mutual funds.

The nearby box discusses some of the misuses of conventional performance measures in evaluating hedge funds.

24.3 Performance Measurement with Changing Portfolio Composition

We have seen already that the volatility of stock returns requires a very long observation period to determine performance levels with any precision, even if portfolio returns are distributed with constant mean and variance. Imagine how this problem is compounded when portfolio return distributions are constantly changing.

It is acceptable to assume that the return distributions of passive strategies have constant mean and variance when the measurement interval is not too long. However, under an active strategy return distributions change by design, as the portfolio manager updates the portfolio in accordance with the dictates of financial analysis. In such a case, estimating various statistics from a sample period assuming a constant mean and variance may lead to substantial errors. Let us look at an example.

Example 24.3 Changing Portfolio Risk

Suppose that the Sharpe measure of the market index is .4. Over an initial period of 52 weeks, the portfolio manager executes a low-risk strategy with an annualized mean excess return of 1% and standard deviation of 2%. This makes for a Sharpe measure of .5, which beats the passive strategy. Over the next 52-week period this manager finds that a *high*-risk strategy is optimal, with an annual mean excess return of 9% and standard deviation of 18%. Here, again, the Sharpe measure is .5. Over the 2-year period our manager maintains a better-than-passive Sharpe measure.

Figure 24.5 shows a pattern of (annualized) quarterly returns that are consistent with our description of the manager's strategy of 2 years. In the first four quarters the excess returns are −1%, 3%, −1%, and 3%, making for an average of 1% and standard deviation of 2%. In the next four quarters the returns are −9%, 27%, −9%, 27%, making for an average of 9% and standard deviation of 18%. Thus *both* years exhibit a Sharpe measure of .5. However, over the eight-quarter sequence the mean and standard deviation are 5% and 13.42%, respectively, making for a Sharpe measure of only .37, apparently inferior to the passive strategy!

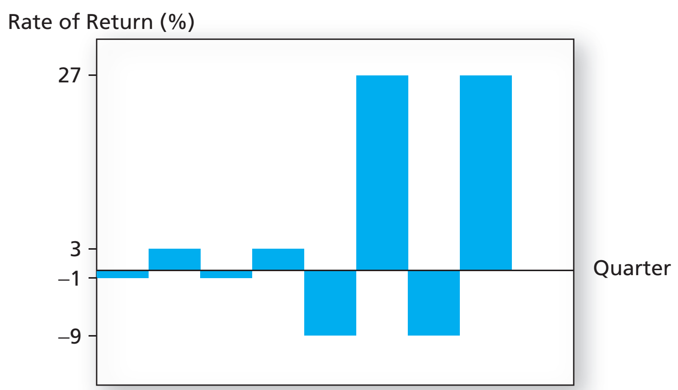


Figure 24.5 Portfolio returns. Returns in last four quarters are more variable than in the first four.

What happened in Example 24.3? The shift of the mean from the first four quarters to the next was not recognized as a shift in strategy. Instead, the difference in mean returns in the 2 years added to the *appearance* of volatility in portfolio returns. The active strategy with shifting means appears riskier than it really is and biases the estimate of the Sharpe measure downward. We conclude that for actively managed portfolios it is helpful to keep track of portfolio composition and changes in portfolio mean and risk. We will see another example of this problem in the next section, which deals with market timing.

24.4 Market Timing

In its pure form, market timing involves shifting funds between a market-index portfolio and a safe asset, depending on whether the market index is expected to outperform the safe asset. In practice, most managers do not shift fully between T-bills and the market. How can we account for partial shifts into the market when it is expected to perform well?

To simplify, suppose that an investor holds only the market-index portfolio and T-bills. If the weight of the market were constant, say, .6, then portfolio beta would also be constant, and the SCL would plot as a straight line with slope .6, as in Figure 24.6, panel A. If, however, the investor could correctly time the market and shift funds into it in periods when the market does well, the SCL would plot as in Figure 24.6, panel B. If bull and bear markets can be predicted, the investor will shift more into the market when the market is about to go up. The portfolio beta and the slope of the SCL will be higher when r_M is higher, resulting in the curved line that appears in Figure 24.6, panel B.

Treynor and Mazuy were the first to propose estimating such a line by adding a squared term to the usual linear index model:[19]

$$r_P - r_f = a + b(r_M - r_f) + c(r_M - r_f)^2 + e_P$$

where r_P is the portfolio return, and *a, b,* and *c* are estimated by regression analysis. If *c* turns out to be positive, we have evidence of timing ability, because this last term will make the characteristic line steeper as $r_M - r_f$ is larger. Treynor and Mazuy estimated this equation for a number of mutual funds, but found little evidence of timing ability.

A similar but simpler methodology was proposed by Henriksson and Merton.[20] These authors suggested that the beta of the portfolio take only two values: a large value if the market is expected to do well and a small value otherwise. Under this scheme the portfolio characteristic line appears as Figure 24.6, panel C. Such a line appears in regression form as

$$r_P - r_f = a + b(r_M - r_f) + c(r_M - r_f)D + e_P$$

where D is a dummy variable that equals 1 for $r_M > r_f$ and zero otherwise. Hence the beta of the portfolio is b in bear markets and $b + c$ in bull markets. Again, a positive value of c implies market timing ability.

Henriksson[21] estimated this equation for 116 mutual funds. He found that the average value of c for the funds was *negative,* and equal to −.07. In sum, the results showed little evidence of market timing ability. Perhaps this should be expected; given the tremendous values to be reaped by a successful market timer, it would be surprising in nearly efficient markets to uncover clear-cut evidence of such skills.

To illustrate a test for market timing, return to Table 24.2. Regressing the excess returns of portfolios P and Q on the excess returns of M and the square of these returns,

$$r_P - r_f = a_P + b_P(r_M - r_f) + c_P(r_M - r_f)^2 + e_P$$
$$r_Q - r_f = a_Q + b_Q(r_M - r_f) + c_Q(r_M - r_f)^2 + e_Q$$

[19]Jack L. Treynor and Kay Mazuy, "Can Mutual Funds Outguess the Market?" *Harvard Business Review* 43 (July–August 1966).

[20]Roy D. Henriksson and R. C. Merton, "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecast Skills," *Journal of Business* 54 (October 1981).

[21]Roy D. Henriksson, "Market Timing and Mutual Fund Performance: An Empirical Investigation," *Journal of Business* 57 (January 1984).

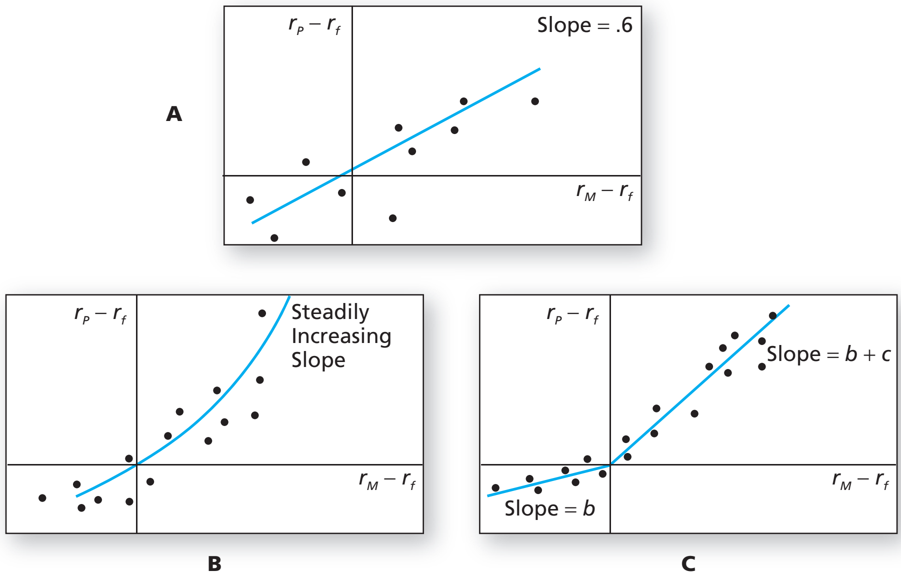


Figure 24.6 Characteristic lines. ***Panel A:*** No market timing, beta is constant. ***Panel B:*** Market timing, beta increases with expected market excess return. ***Panel C:*** Market timing with only two values of beta.

we derive the following statistics:

	Portfolio	
Estimate	***P***	***Q***
Alpha (*a*)	1.77 (1.63)	–2.29 (5.28)
Beta (*b*)	0.70 (0.69)	1.10 (1.40)
Timing (*c*)	0.00	0.10
R-SQR	0.91 (0.91)	0.98 (0.64)

The numbers in parentheses are the regression estimates from the single variable regression reported in Table 24.3. The results reveal that portfolio *P* shows no timing. It is not clear whether this is a result of Jane's making no attempt at timing or that the effort to time the market was in vain and served only to increase portfolio variance unnecessarily.

The results for portfolio *Q,* however, reveal that timing has, in all likelihood, successfully been attempted. The timing coefficient, *c,* is estimated at .10. The evidence thus suggests successful timing (positive *c*) offset by unsuccessful stock selection (negative *a*). Note that the alpha estimate, *a,* is now −2.29% as opposed to the 5.28% estimate derived from the regression equation that did not allow for the possibility of timing activity.

This example illustrates the inadequacy of conventional performance evaluation techniques that assume constant mean returns and constant risk. The market timer constantly shifts beta and mean return, moving into and out of the market. Whereas the expanded regression captures this phenomenon, the simple SCL does not. The relative desirability of portfolios P and Q remains unclear in the sense that the value of the timing success and selectivity failure of Q compared with P has yet to be evaluated. The important point for performance evaluation, however, is that expanded regressions can capture many of the effects of portfolio composition change that would confound the more conventional mean-variance measures.

The Potential Value of Market Timing

Suppose we define perfect market timing as the ability to tell (with certainty) at the beginning of each year whether the S&P 500 portfolio will outperform the strategy of rolling over 1-month T-bills throughout the year. Accordingly, at the beginning of each year, the market timer shifts all funds into either cash equivalents (T-bills) or equities (the all U.S. stock portfolio), whichever is predicted to do better. Beginning with $1 on January 1, 1927, how would the perfect timer end an 86-year experiment on December 31, 2012, in comparison with investors who kept their funds in either equity or T-bills for the entire period?

Table 24.4, columns 1–3, presents summary statistics for each of the three passive strategies, computed from the historical annual returns of bills and equities. From the returns on stocks and bills, we calculate wealth indexes of the all-bills and all-equity investments and show terminal values for these investors at the end of 2012. The return for the perfect timer in each year is the *maximum* of the return on stocks and the return on bills.

The first row in Table 24.4 tells all. The terminal value of investing $1 in bills over the 86 years (1927–2012) is $20, while the terminal value of the same initial investment in equities is about $2,652. We saw a similar pattern for a 25-year investment in Chapter 5; the much larger terminal values (and difference between them) when extending the horizon from 25 to 86 years is just another manifestation of the power of compounding. We argued in Chapter 5 that as impressive as the difference in terminal values is, it is best interpreted as no more than fair compensation for the risk borne by equity investors. Notice that the standard deviation of the all-equity investor was a hefty 20.39%. This is also why the geometric average of stocks for the period is "only" 9.60%, compared with the arithmetic average of 11.63%. (The difference between the two averages increases with volatility.)

Strategy	Bills	Equities	Perfect Timer	Imperfect Timer*
Terminal Value	20	2,652	352,796	8,859
Arithmetic Average	3.59	11.63	16.75	11.98
Standard Deviation	3.12	20.39	13.49	14.36
Geometric Average	3.54	9.60	16.01	11.09
LPSD (relative to bills)	0	21.18	0	17.15
Minimum†	−0.04	−44.00	−0.02	−27.09
Maximum	14.72	57.42	57.42	57.42
Skew	0.99	−0.42	0.72	0.71
Kurtosis	0.98	0.02	−0.13	1.50

Table 24.4

Performance of bills, equities, and (annual) timers—perfect and imperfect

*The imperfect timer has $P_1 = .7$, and $P_2 = .7$. Therefore, $P_1 + P_2 - 1 = .4$.
†A negative rate on "bills" was observed in 1940. The Treasury security used in the data series in these early years was actually not a T-bill but a T-bond with 30 days to maturity.

Now observe that the terminal value of the perfect timer is about \$353,000, a 133-fold increase over the already large terminal value of the all-equity strategy! In fact, this result is even better than it looks, because the return to the market timer is truly risk-free. This is the classic case where a large standard deviation (13.49%) has nothing to do with risk. Because the timer never delivers a return below the risk-free rate, the standard deviation is a measure of *good* surprises only. The positive skew of the distribution (compared with the negative skew of equities) is a manifestation of the fact that the extreme values are all positive. Another indication of this stellar performance is the minimum and maximum returns—the minimum return equals the minimum return on bills (in 1940) and the maximum return is that of equities (in 1933)—so that all negative returns on equities (as low as −44% in 1931) were avoided by the timer. Finally, the best indication of the performance of the timer is a lower partial standard deviation, LPSD.[22] The LPSD of the all-equity portfolio is only slightly greater than the conventional standard deviation, but it is necessarily zero for the perfect timer.

If we interpret the terminal value of the all-equity portfolio in excess of the value of the T-bill portfolio entirely as a risk premium commensurate with investment risk, we must conclude that the risk-adjusted equivalent value of the all-equity terminal value is the same as that of the T-bill portfolio, \$20.[23] In contrast, the perfect timer's portfolio has no risk, and so receives no discount for risk. Hence, it is fair to say that the forecasting ability of the perfect timer converts a \$20 final value to a value of \$352,796.

Valuing Market Timing as a Call Option

The key to valuing market timing ability is to recognize that perfect foresight is equivalent to holding a call option on the equity portfolio. The perfect timer invests 100% in either the safe asset or the equity portfolio, whichever will provide the higher return. The rate of return is *at least* the risk-free rate. This is shown in Figure 24.7.

To see the value of information as an option, suppose that the market index currently is at S_0 and that a call option on the index has an exercise price of $X = S_0(1 + r_f)$. If the market outperforms bills over the coming period, S_T will exceed X, whereas it will be less than X otherwise. Now look at the payoff to a portfolio consisting of this option and S_0 dollars invested in bills:

	$S_T < X$	$S_T \geq X$
Bills	$S_0(1 + r_f)$	$S_0(1 + r_f)$
Call	0	$S_T - X$
Total	$S_0(1 + r_f)$	S_T

The portfolio pays the risk-free return when the market is bearish (i.e., the market return is less than the risk-free rate), and it pays the market return when the market is bullish and beats bills. Such a portfolio is a perfect market timer.[24]

[22]The conventional LPSD is based on the average squared deviation below the mean. Because the threshold performance in this application is the risk-free rate, we modify the LPSD for this discussion by taking squared deviations from that rate and the observations are conditional on being below that threshold. It ignores the number of such events.

[23]It may seem hard to attribute such a big difference in final outcome solely to risk aversion. But think of it this way: the final value of the equity position is 133 times that of the bills position (\$2,652 versus \$20). Over 86 years, this implies a reasonable annualized risk premium of 5.85%: $133^{1/86} = 1.0585$.

[24]The analogy between market timing and call options, and the valuation formulas that follow from it, were developed in Robert C. Merton, "On Market Timing and Investment Performance: An Equilibrium Theory of Value for Market Forecasts," *Journal of Business,* July 1981.

Because the ability to predict the better-performing investment is equivalent to holding a call option on the market, we can use option-pricing models to assign a dollar value to perfect timing ability. This value would constitute the fair fee that a perfect timer could charge investors for its services. Placing a value on perfect timing also enables us to assign value to less-than-perfect timers.

The exercise price of the perfect-timer call option on \$1 of the equity portfolio is the final value of the T-bill investment. Using continuous compounding, this is $\$1 \times e^{rT}$. When you use this exercise price in the Black-Scholes formula for the value of the call option, the formula simplifies considerably to[25]

$$\text{MV(Perfect timer per \$ of assets)} = C = 2N(\tfrac{1}{2}\,\sigma_M\sqrt{T}) - 1 \qquad (24.6)$$

We have so far assumed annual forecasts, that is, $T = 1$ year. Using $T = 1$, and the standard deviation of stocks from Table 24.4, 20.39%, we compute the value of this call option as 8.12 cents, or 8.12% of the value of the equity portfolio. This is less than the historical-average return of perfect timing shown in Table 24.5, reflecting the fact that actual timing value is sensitive to fat tails in the distribution of returns, whereas Black-Scholes presumes a log-normal distribution.

Equation 24.6 tells us that perfect market timing would be equivalent to enhancing the annual equity return by .0812 (or 8.12% per year). Since the average equity return over the last 86 years has been 11.63%, this would be similar in value to enjoying an annual return of $1.1162 \times 1.0812 - 1 = .2069$, or 20.69%.

If a timer could make the correct choice every month instead of every year, the value of the forecasts would dramatically increase. Of course, making perfect forecasts more frequently requires even better powers of prediction. As the frequency of such perfect predictions increases without bound, the value of the services will increase without bound as well.

Suppose the perfect timer could make perfect forecasts every month. In this case, each forecast would be for a shorter interval, and the value of each individual forecast would be lower, but there would be 12 times as many forecasts, each of which could be valued as another call option. The net result is a big increase in total value. With monthly predictions, the value of the call will be $2N(\tfrac{1}{2} \times .2039 \times \sqrt{1/12}) - 1 = .0235$. Using a monthly T-bill rate of 3.6%/12, the present value of a 1-year string of such monthly calls, each worth \$.0235, is \$.28. Thus, the annual value of the monthly perfect timer is 28 cents on the dollar, compared to 8.12 cents for an annual timer. For an investment period of 86 years, the forecast future value of a \$1 investment would be a far greater $[(1 + .28)(1 + .1163)]^{86} = \$2.1 \times 10^{13}$. This value suggests the otherworldly power of these forecasts.

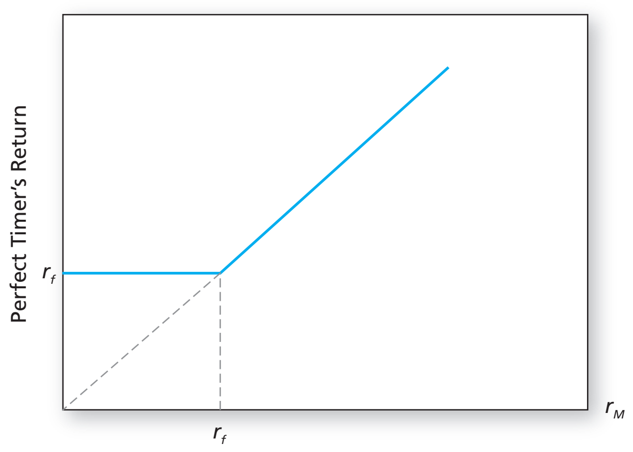


Figure 24.7 Rate of return of a perfect market timer as a function of the rate of return on the market index.

[25]Substitute $S_0 = \$1$ for the current value of the equity portfolio and $X = \$1 \times e^{rT}$ in Equation 21.1 of Chapter 21, and you will obtain Equation 24.6.

The Value of Imperfect Forecasting

A weather forecaster in Tucson, Arizona, who *always* predicts no rain may be right 90% of the time. But a high success rate for a "stopped-clock" strategy is not evidence of forecasting ability. Similarly, the appropriate measure of market forecasting ability is not the overall proportion of correct forecasts. If the market is up 2 days out of 3 and a forecaster always predicts market advance, the two-thirds success rate is not a measure of forecasting ability. We need to examine the proportion of bull markets ($r_M < r_f$) correctly forecast *and* the proportion of bear markets ($r_M > r_f$) correctly forecast.

If we call P_1 the proportion of the correct forecasts of bull markets and P_2 the proportion for bear markets, then $P_1 + P_2 - 1$ is the correct measure of timing ability. For example, a forecaster who always guesses correctly will have $P_1 = P_2 = 1$, and will show ability of $P_1 + P_2 - 1 = 1$ (100%). An analyst who always bets on a bear market will mispredict all bull markets ($P_1 = 0$), will correctly "predict" all bear markets ($P_2 = 1$), and will end up with timing ability of $P_1 + P_2 - 1 = 0$.

CONCEPT CHECK 24.4

What is the market timing score of someone who flips a fair coin to predict the market?

When timing is imperfect, Merton shows that if we measure overall accuracy by the statistic $P_1 + P_2 - 1$, the market value of the services of an imperfect timer is simply

$$\text{MV(Imperfect timer)} = (P_1 + P_2 - 1) \times C = (P_1 + P_2 - 1)\left[2N(\tfrac{1}{2}\sigma_M\sqrt{T}) - 1\right] \quad (24.7)$$

The last column in Table 24.4 provides an assessment of the imperfect market-timer. To simulate the performance of an imperfect timer, we drew random numbers to capture the possibility that the timer will sometimes issue an incorrect forecast (we assumed here both P_1 and $P_2 = .7$) and compiled results for the 86 years of history.[26] The statistics of this exercise resulted in an average terminal value for the imperfect timer of "only" $8,859, compared with the perfect timer's $352,796, but still considerably superior to the $2,562 for the all-equity investments.[27]

A further variation on the valuation of market timing is a case in which the timer does not shift fully from one asset to the other. In particular, if the timer knows her forecasts are imperfect, one would not expect her to shift fully between markets. She presumably would moderate her positions. Suppose that she shifts a fraction ω of the portfolio between T-bills and equities. In that case, Equation 24.7 can be generalized as follows:

$$\text{MV(Imperfect timer)} = \omega(P_1 + P_2 - 1)\left[2N(\sigma_M\sqrt{T}) - 1\right]$$

If the shift is $\omega = .50$ (50% of the portfolio), the timer's value will be one-half of the value we would obtain for full shifting, for which $\omega = 1.0$.

[26]In each year, we started with the correct forecast, but then used a random number generator to occasionally change the timer's forecast to an incorrect prediction. We set the probability that the timer's forecast would be correct equal to .70 for both up and down markets.

[27]Notice that Equation 24.7 implies that an investor with a value of $P = 0$ who attempts to time the market would add zero value. The shifts across markets would be no better than a random decision concerning asset allocation.

24.5 Style Analysis

Style analysis was introduced by Nobel laureate William Sharpe.[28] The popularity of the concept was aided by a well-known study[29] concluding that 91.5% of the variation in returns of 82 mutual funds could be explained by the funds' asset allocation to bills, bonds, and stocks. Later studies that considered asset allocation across a broader range of asset classes found that as much as 97% of fund returns can be explained by asset allocation alone.

Sharpe's idea was to regress fund returns on indexes representing a range of asset classes. The regression coefficient on each index would then measure the fund's implicit allocation to that "style." Because funds are barred from short positions, the regression coefficients are constrained to be either zero or positive and to sum to 100%, so as to represent a complete asset allocation. The *R*-square of the regression would then measure the percentage of return variability attributable to style or asset allocation, while the remainder of return variability would be attributable either to security selection or to market timing by periodic changes in the asset-class weights.

To illustrate Sharpe's approach, we use monthly returns on Fidelity Magellan's Fund during the famous manager Peter Lynch's tenure between October 1986 and September 1991, with results shown in Table 24.5. While seven asset classes are included in this analysis (of which six are represented by stock indexes and one is the T-bill alternative), the regression coefficients are positive for only three, namely, large capitalization stocks, medium cap stocks, and high P/E (growth) stocks. These portfolios alone explain 97.5% of the variance of Magellan's returns. In other words, a tracking portfolio made up of the three style portfolios, with weights as given in Table 24.5, would explain the vast majority of Magellan's variation in monthly performance. We conclude that the fund returns are well represented by three style portfolios.

The proportion of return variability *not* explained by asset allocation can be attributed to security selection within asset classes, as well as timing that shows up as periodic changes in allocation. For Magellan, residual variability was 100 − 97.5 = 2.5%. This sort

Table 24.5

Style analysis for Fidelity's Magellan Fund

Style Portfolio	Regression Coefficient
T-Bill	0
Small Cap	0
Medium Cap	35
Large Cap	61
High P/E (growth)	5
Medium P/E	0
Low P/E (value)	0
Total	100
R-square	97.5

Source: Authors' calculations. Return data for Magellan obtained from **finance.yahoo.com/funds** and return data for style portfolios obtained from the Web page of Professor Kenneth French: **mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html**.

[28] William F. Sharpe, "Asset Allocation: Management Style and Performance Evaluation," *Journal of Portfolio Management,* Winter 1992, pp. 7–19.

[29] Gary Brinson, Brian Singer, and Gilbert Beebower, "Determinants of Portfolio Performance," *Financial Analysts Journal,* May/June 1991.

of result is commonly used to play down the importance of security selection and timing in fund performance, but such a conclusion misses the important role of the intercept in this regression. (The *R*-square of the regression can be 100%, and yet the intercept can be nonzero due to a superior risk-adjusted abnormal return.) For Magellan, the intercept was 32 basis points per month, resulting in a cumulative abnormal return over the 5-year period of 19.19%. The superior performance of Magellan is displayed in Figure 24.8, which plots the cumulative impact of the intercept plus monthly residuals relative to the tracking portfolio composed of the style portfolios. Except for the period surrounding the crash of October 1987, Magellan's return consistently increased relative to the benchmark portfolio.

Style analysis provides an alternative to performance evaluation based on the security market line (SML) of the CAPM. The SML uses only one comparison portfolio, the broad market index, whereas style analysis more freely constructs a tracking portfolio from a number of specialized indexes. To compare the two approaches, the security characteristic line (SCL) of Magellan was estimated by regressing its excess return on the excess return of a market index composed of all NYSE, Amex, and NASDAQ stocks. The beta estimate of Magellan was 1.11 and the *R*-square of the regression was .99. The alpha value (intercept) of this regression was "only" 25 basis points per month, reflected in a cumulative abnormal return of 15.19% for the period.

How can we explain the higher *R*-square of the regression with only one factor (the market index) relative to the style regression, which deploys six stock indexes? The answer is that style analysis imposes extra constraints on the regression coefficients: It forces them to be positive and to sum to 1.0. This "neat" representation may not be consistent with actual portfolio weights that are constantly changing over time. So which representation better gauges Magellan's performance over the period? There is no clear-cut answer. The SML benchmark is a better representation of performance relative to the theoretically prescribed passive portfolio, that is, the broadest market index available. On the other hand,

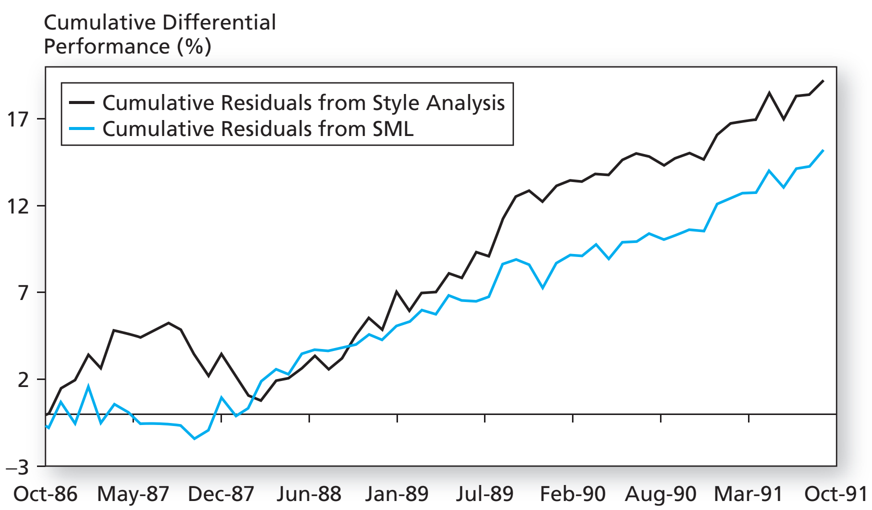


Figure 24.8 Fidelity Magellan Fund cumulative return difference: Fund versus style benchmark and fund versus SML benchmark

Source: Authors' calculations.

style analysis reveals the strategy that most closely tracks the fund's activity and measures performance relative to this strategy. If the strategy revealed by the style analysis method is consistent with the one stated in the fund prospectus, then the performance relative to this strategy is the correct measure of the fund's success.

Figure 24.9 shows the frequency distribution of average residuals across 636 mutual funds from Sharpe's style analysis. The distribution has the familiar bell shape with a slightly negative mean of −.074% per month. This should remind you of Figure 11.7, where we presented the frequency distribution of CAPM alphas for a large sample of mutual funds. As in Sharpe's study, these risk-adjusted returns plot as a bell-shaped curve with slightly negative mean.

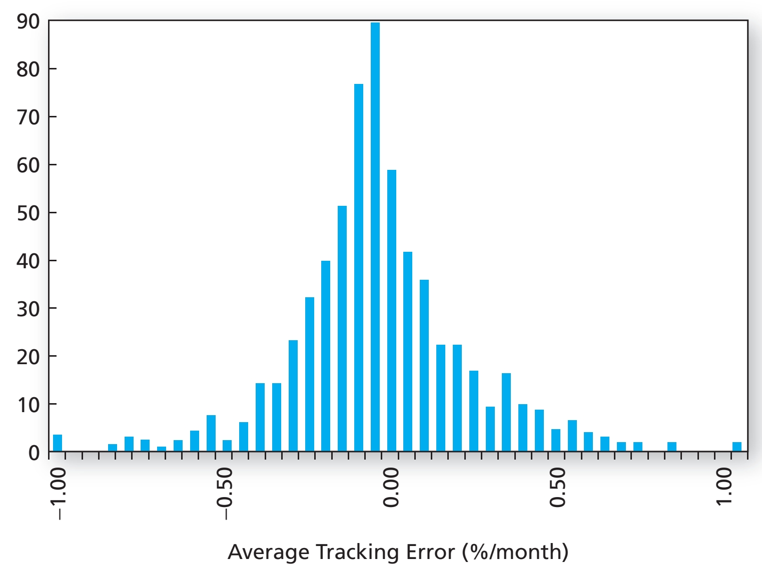


Figure 24.9 Average tracking error for 636 mutual funds, 1985–1989

Source: William F. Sharpe, "Asset Allocation: Management Style and Performance Evaluation," *Journal of Portfolio Management,* Winter 1992, pp. 7–19. Used with permission of Institutional Investor, Inc., **www.iijournals.com.** All Rights Reserved.

Style Analysis and Multifactor Benchmarks

Style analysis raises an interesting question for performance evaluation. Suppose a growth-index portfolio exhibited superior performance relative to a mutual fund benchmark such as the S&P 500 over some measurement period. Including this growth index in a style analysis would eliminate this superior performance from the portfolio's estimated alpha value. Is this proper? Quite plausibly, the fund's analysts predicted that an active portfolio of growth stocks was underpriced and tilted the portfolio to take advantage of it. Clearly, the contribution of this decision to an alpha value relative to the benchmark is a legitimate part of the overall alpha value of the fund, and should not be eliminated by style analysis. This brings up a related question.

Chapter 11 pointed out that the conventional performance benchmark today is a four-factor model, which employs the three Fama-French factors (the return on the market index, and returns to portfolios based on size and book-to-market ratio) augmented by a momentum factor (a portfolio constructed based on prior-year stock return). Alphas estimated from these four factor portfolios control for a wide range of style choices that may affect average returns. But using alpha values from a multifactor model presupposes that a passive strategy would include the aforementioned factor portfolios. When is this reasonable?

Use of any benchmark other than the fund's single-index benchmark is legitimate only if we assume that the factor portfolios in question are part of the fund's alternative passive strategy. This assumption may be unrealistic in many cases where a single-index benchmark is used for performance evaluation even if research shows a multifactor model better explains asset returns. In Section 24.8 on performance attribution we show how portfolio managers attempt to uncover which decisions contributed to superior performance. This performance attribution procedure starts with benchmark allocations to various indexes and attributes performance to asset allocation on the basis of deviation of actual from benchmark allocations. The performance benchmark may be and often is specified in advance without regard to any particular style portfolio.

Style Analysis in Excel

Style analysis has become very popular in the investment management industry and has spawned quite a few variations on Sharpe's methodology. Many portfolio managers utilize Web sites that help investors identify their style and stock selection performance.

You can do style analysis with Excel's Solver. The strategy is to regress a fund's rate of return on those of a number of style portfolios (as in Table 24.5). The style portfolios are passive (index) funds that represent a style alternative to asset allocation. Suppose you choose three style portfolios, labeled 1–3. Then the coefficients in your style regression are alpha (the intercept that measures abnormal performance) and three slope coefficients, one for each style index. The slope coefficients reveal how sensitively the performance of the fund follows the return of each passive style portfolio. The residuals from this regression, $e(t)$, represent "noise," that is, fund performance at each date, t, that is independent of any of the style portfolios. We cannot use a standard regression package in this analysis, however, because we wish to constrain each coefficient to be nonnegative and sum to 1.0, representing a portfolio of styles.

To do style analysis using Solver, start with arbitrary coefficients (e.g., you can set $\alpha = 0$ and set each $\beta = 1/3$). Use these to compute the time series of residuals from the style regression according to

$$e(t) = R(t) - [\alpha + \beta_1 R_1(t) + \beta_2 R_2(t) + \beta_3 R_3(t)] \tag{24.8}$$

where

$R(t)$ = Excess return on the measured fund for date t

$R_i(t)$ = Excess return on the ith style portfolio ($i = 1, 2, 3$)

α = Abnormal performance of the fund over the sample period

β_i = Beta of the fund on the ith style portfolio

Equation 24.8 yields the time series of residuals from your "regression equation" with those arbitrary coefficients. Now square each residual and sum the squares. At this point, you call on the Solver to minimize the sum of squares by changing the value of the four coefficients. You will use the "by changing variables" command. You also add four constraints to the optimization: three that force the betas to be nonnegative and one that forces them to sum to 1.0.

Solver's output will give you the three style coefficients, as well as the estimate of the fund's unique, abnormal performance as measured by the intercept. The sum of squares also allows you to calculate the R-square of the regression and p-values as explained in Chapter 8.

24.6 Performance Attribution Procedures

Rather than focus on risk-adjusted returns, practitioners often want simply to ascertain which decisions resulted in superior or inferior performance. Superior investment performance depends on an ability to be in the "right" securities at the right time. Such timing and selection ability may be considered broadly, such as being in equities as opposed to fixed-income securities when the stock market is performing well. Or it may be defined at a more detailed level, such as choosing the relatively better-performing stocks within a particular industry.

Portfolio managers constantly make broad-brush asset allocation decisions as well as more detailed sector and security allocation decisions within asset classes. Performance attribution studies attempt to decompose overall performance into discrete components that may be identified with a particular level of the portfolio selection process.

Attribution studies start from the broadest asset allocation choices and progressively focus on ever-finer details of portfolio choice. The difference between a managed portfolio's performance and that of a benchmark portfolio then may be expressed as the sum of the contributions to performance of a series of decisions made at the various levels of the portfolio construction process. For example, one common attribution system decomposes performance into three components: (1) broad asset market allocation choices across equity, fixed-income, and money markets; (2) industry (sector) choice within each market; and (3) security choice within each sector.

The attribution method explains the difference in returns between a managed portfolio, P, and a selected benchmark portfolio, B, called the **bogey.** Suppose that the universe of assets for P and B includes n asset classes such as equities, bonds, and bills. For each asset class, a benchmark index portfolio is determined. For example, the S&P 500 may be chosen as a benchmark for equities. The bogey portfolio is set to have fixed weights in each asset class, and its rate of return is given by

$$r_B = \sum_{i=1}^{n} w_{Bi} r_{Bi}$$

where w_{Bi} is the weight of the bogey in asset class i, and r_{Bi} is the return on the benchmark portfolio of that class over the evaluation period. The portfolio managers choose weights in each class, w_{Pi}, based on their capital market expectations, and they choose a portfolio of the securities within each class based on their security analysis, which earns r_{Pi} over the evaluation period. Thus the return of the managed portfolio will be

$$r_P = \sum_{i=1}^{n} w_{Pi} r_{Pi}$$

The difference between the two rates of return, therefore, is

$$r_P - r_B = \sum_{i=1}^{n} w_{Pi} r_{Pi} - \sum_{i=1}^{n} w_{Bi} r_{Bi} = \sum_{i=1}^{n} (w_{Pi} r_{Pi} - w_{Bi} r_{Bi}) \tag{24.9}$$

Each term in the summation of Equation 24.9 can be rewritten in a way that shows how asset allocation decisions versus security selection decisions for each asset class contributed to overall performance. We decompose each term of the summation into a sum of two terms as follows. Note that the two terms we label as contribution from asset allocation and contribution from security selection in the following decomposition do in fact sum to the total contribution of each asset class to overall performance.

Contribution from asset allocation	$(w_{Pi} - w_{Bi}) r_{Bi}$
+ Contribution from security selection	$w_{Pi}(r_{Pi} - r_{Bi})$
= Total contribution from asset class i	$w_{Pi} r_{Pi} - w_{Bi} r_{Bi}$

The first term of the sum measures the impact of asset allocation because it shows how deviations of the actual weight from the benchmark weight for that asset class multiplied

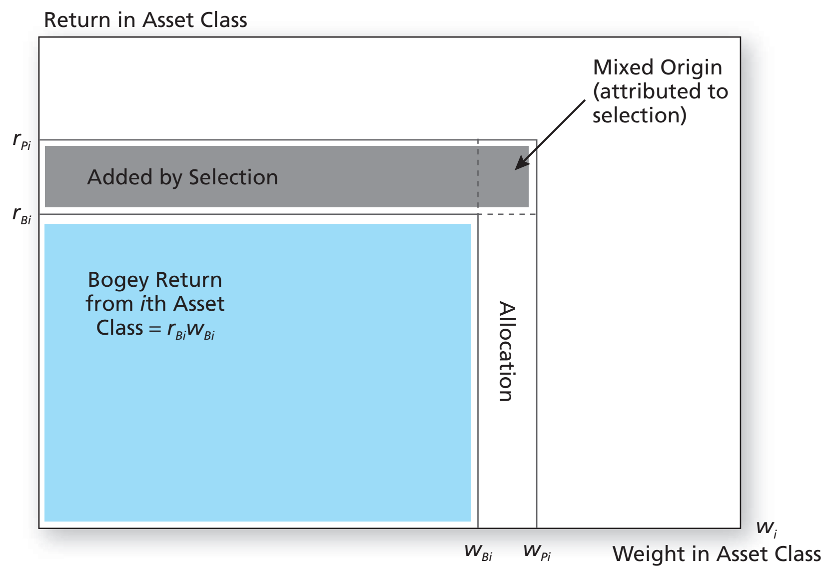


Figure 24.10 Performance attribution of ith asset class. Enclosed area indicates total rate of return.

by the index return for the asset class added to or subtracted from total performance. The second term of the sum measures the impact of security selection because it shows how the manager's excess return *within* the asset class compared to the benchmark return for that class multiplied by the portfolio weight for that class added to or subtracted from total performance. Figure 24.10 presents a graphical interpretation of the attribution of overall performance into security selection versus asset allocation.

To illustrate this method, consider the attribution results for a hypothetical portfolio. The portfolio invests in stocks, bonds, and money market securities. An attribution analysis appears in Tables 24.6 through 24.9. The portfolio return over the month is 5.34%.

The first step is to establish a benchmark level of performance against which performance ought to be compared. This benchmark, again, is called the bogey. It is designed to measure the returns the portfolio manager would earn if he or she were to follow a completely passive strategy. "Passive" in this context has two attributes. First, it means that the allocation of funds across broad asset classes is set in accord with a notion of "usual," or neutral, allocation across sectors. This would be considered a passive asset-market allocation. Second, it means that *within* each asset class, the portfolio manager holds an indexed portfolio such as the S&P 500 index for the equity sector. In such a manner, the passive strategy used as a performance benchmark rules out asset allocation as well as security selection decisions. Any departure of the manager's return from the passive benchmark must be due to either asset allocation bets (departures from the neutral allocation across markets) or security selection bets (departures from the passive index within asset classes).

While we have already discussed in earlier chapters the justification for indexing within sectors, it is worth briefly explaining the determination of the neutral allocation of funds across the broad asset classes. Weights that are designated as "neutral" will depend on the risk tolerance of the investor and must be determined in consultation with the client. For example, risk-tolerant clients may place a large fraction of their portfolio in the equity market, perhaps directing the fund manager to set neutral weights of 75% equity, 15% bonds, and 10% cash equivalents. Any deviation from these weights must be justified by a belief that one or another market will either

Table 24.6

Performance of the managed portfolio

Component	Benchmark Weight (Bogey Performance and Excess Return)	Return of Index during Month (%)
Equity (S&P 500)	.60	5.81
Bonds (Barclays Aggregate Index)	.30	1.45
Cash (money market)	.10	0.48
Bogey = (.60 × 5.81) + (.30 × 1.45) + (.10 × 0.48) = 3.97%		
Return of managed portfolio	5.34%	
− Return of bogey portfolio	3.97	
Excess return of managed portfolio	1.37%	

over- or underperform its usual risk–return profile. In contrast, more risk-averse clients may set neutral weights of 45%/35%/20% for the three markets. Therefore, their portfolios in normal circumstances will be exposed to less risk than that of the risk-tolerant client. Only intentional bets on market performance will result in departures from this profile.

In Table 24.6, the neutral weights have been set at 60% equity, 30% fixed income, and 10% cash (money market securities). The bogey portfolio, comprised of investments in each index with the 60/30/10 weights, returned 3.97%. The managed portfolio's measure of performance is positive and equal to its actual return less the return of the bogey: 5.34 − 3.97 = 1.37%. The next step is to allocate the 1.37% excess return to the separate decisions that contributed to it.

Asset Allocation Decisions

Our hypothetical managed portfolio is invested in the equity, fixed-income, and money markets with weights of 70%, 7%, and 23%, respectively. The portfolio's performance could have to do with the departure of this weighting scheme from the benchmark 60/30/10 weights and/or to superior or inferior results *within* each of the three broad markets.

To isolate the effect of the manager's asset allocation choice, we measure the performance of a hypothetical portfolio that would have been invested in the indexes for each market with weights 70/7/23. This return measures the effect of the shift away from the benchmark 60/30/10 weights without allowing for any effects attributable to active management of the securities selected within each market.

Superior performance relative to the bogey is achieved by overweighting investments in markets that turn out to perform well and by underweighting those in poorly performing markets. The contribution of asset allocation to superior performance equals the sum over all markets of the excess weight (sometimes called the *active weight* in the industry) in each market times the return of the market index.

Panel A of Table 24.7 demonstrates that asset allocation contributed 31 basis points to the portfolio's overall excess return of 137 basis points. The major factor contributing to superior performance in this month is the heavy weighting of the equity market in a month when the equity market has an excellent return of 5.81%.

Table 24.7

Performance attribution

A. Contribution of Asset Allocation to Performance

Market	(1) Actual Weight in Market	(2) Benchmark Weight in Market	(3) Active or Excess Weight	(4) Market Return (%)	(5) = (3) × (4) Contribution to Performance (%)
Equity	.70	.60	.10	5.81	.5810
Fixed-income	.07	.30	−.23	1.45	−.3335
Cash	.23	.10	.13	.48	.0624
Contribution of asset allocation					.3099

B. Contribution of Selection to Total Performance

Market	(1) Portfolio Performance (%)	(2) Index Performance (%)	(3) Excess Performance (%)	(4) Portfolio Weight	(5) = (3) × (4) Contribution (%)
Equity	7.28	5.81	1.47	.70	1.03
Fixed-income	1.89	1.45	0.44	.07	0.03
Contribution of selection within markets					1.06

Sector and Security Selection Decisions

If .31% of the excess performance (Table 24.7, panel A) can be attributed to advantageous asset allocation across markets, the remaining 1.06% then must be attributable to sector selection and security selection within each market. Table 24.7, panel B, details the contribution of the managed portfolio's sector and security selection to total performance.

Panel B shows that the equity component of the managed portfolio has a return of 7.28% versus a return of 5.81% for the S&P 500. The fixed-income return is 1.89% versus 1.45% for the Barclays Aggregate Bond Index. The superior performance in both equity and fixed-income markets weighted by the portfolio proportions invested in each market sums to the 1.06% contribution to performance attributable to sector and security selection.

Table 24.8 documents the sources of the equity market performance by each sector within the market. The first three columns detail the allocation of funds within the equity market compared to their representation in the S&P 500. Column (4) shows the rate of return of each sector. The contribution of each sector's allocation presented in column (5) equals the product of the difference in the sector weight and the sector's performance.

Note that good performance (a positive contribution) derives from overweighting well-performing sectors such as consumer noncyclicals, as well as underweighting poorly performing sectors such as technology. The excess return of the equity component of the portfolio attributable to sector allocation alone is 1.29%. Table 24.7, panel B, column (3), shows that the equity component of the portfolio outperformed the S&P 500 by 1.47%. We conclude that the effect of security selection *within* sectors must have contributed an additional 1.47% − 1.29%, or .18%, to the performance of the equity component of the portfolio.

A similar sector analysis can be applied to the fixed-income portion of the portfolio, but we do not show those results here.

eXcel APPLICATIONS: Performance Attribution

The performance attribution spreadsheet develops the attribution analysis that is presented in this section. Additional data can be used in the analysis of performance for other sets of portfolios. The model can be used to analyze performance of mutual funds and other managed portfolios.

You can find this Excel model on the Online Learning Center at **www.mhhe.com/bkm**.

Excel Questions

1. What would happen to the contribution of asset allocation to overall performance if the actual weights had been 75/12/13 instead of 70/7/23? Explain your result.
2. What would happen to the contribution of security selection to overall performance if the actual return on the equity portfolio had been 6.81% instead of 5.81% and the return on the bond portfolio had been 0.45% instead of 1.45%? Explain your result.

	A	B	C	D	E	F
1	**Performance Attribution**					
2						
3						
4	**Bogey**					
5	**Portfolio**		**Benchmark**	**Return on**	**Portfolio**	
6	**Component**	**Index**	**Weight**	**Index**	**Return**	
7	Equity	S&P 500	0.60	5.8100%	3.4860%	
8	Bonds	Barclays Index	0.30	1.4500%	0.4350%	
9	Cash	Money Market	0.10	0.4800%	0.0480%	
10			Return on Bogey		3.9690%	
11						
12		**Managed**				
13		**Portfolio**	**Portfolio**	**Actual**	**Portfolio**	
14		**Component**	**Weight**	**Return**	**Return**	
15		Equity	0.70	5.8100%	5.0960%	
16		Bonds	0.07	1.4500%	0.1323%	
17		Cash	0.23	0.4800%	0.1104%	
18			Return on Managed		5.3387%	
19			Excess Return		1.3697%	

Table 24.8

Sector selection within the equity market

	(1)	(2)	(3)	(4)	(5) = (3) × (4)
	Beginning of Month Weights (%)		Active Weights (%)	Sector Return (%)	Sector Allocation Contribution
Sector	**Portfolio**	**S&P 500**			
Basic materials	1.96	8.3	−6.34	6.9	−0.4375
Business services	7.84	4.1	3.74	7.0	0.2618
Capital goods	1.87	7.8	−5.93	4.1	−0.2431
Consumer cyclical	8.47	12.5	−4.03	8.8	0.3546
Consumer noncyclical	40.37	20.4	19.97	10.0	1.9970
Credit sensitive	24.01	21.8	2.21	5.0	0.1105
Energy	13.53	14.2	−0.67	2.6	−0.0174
Technology	1.95	10.9	−8.95	0.3	−0.0269
TOTAL					1.2898

Summing Up Component Contributions

In this particular month, all facets of the portfolio selection process were successful. Table 24.9 details the contribution of each aspect of performance. Asset allocation across the major security markets contributes 31 basis points. Sector and security allocation within those markets contributes 106 basis points, for total excess portfolio performance of 137 basis points.

The sector and security allocation of 106 basis points can be partitioned further. Sector allocation within the equity market results in excess performance of 129 basis points, and security selection within sectors contributes 18 basis points. (The total equity excess performance of 147 basis points is multiplied by the 70% weight in equity to obtain contribution to portfolio performance.) Similar partitioning could be done for the fixed-income sector.

CONCEPT CHECK 24.5

a. Suppose the benchmark weights in Table 24.7 had been set at 70% equity, 25% fixed-income, and 5% cash equivalents. What would have been the contributions of the manager's asset allocation choices?

b. Suppose the S&P 500 return is 5%. Compute the new value of the manager's security selection choices.

Table 24.9

Portfolio attribution: summary

		Contribution (basis points)
1. Asset allocation		31
2. Selection		
a. Equity excess return (basis points)		
i. Sector allocation	129	
ii. Security selection	18	
	147 × .70 (portfolio weight) =	102.9
b. Fixed-income excess return	44 × .07 (portfolio weight) =	3.1
Total excess return of portfolio		137.0

SUMMARY

1. The appropriate performance measure depends on the role of the portfolio to be evaluated. Appropriate performance measures are as follows:
   a. Sharpe: when the portfolio represents the entire investment fund.
   b. Information ratio: when the portfolio represents the active portfolio to be optimally mixed with the passive portfolio.
   c. Treynor or Jensen: when the portfolio represents one subportfolio of many.
2. Many observations are required to eliminate the effect of the "luck of the draw" from the evaluation process because portfolio returns commonly are very "noisy."
3. Hedge funds or other active positions meant to be mixed with a passive indexed portfolio should be evaluated based on their information ratio.

4. The shifting mean and variance of actively managed portfolios make it even harder to assess performance. A typical example is the attempt of portfolio managers to time the market, resulting in ever-changing portfolio betas.
5. A simple way to measure timing and selection success simultaneously is to estimate an expanded security characteristic line, with a quadratic term added to the usual index model. Another way to evaluate timers is based on the implicit call option embedded in their performance.
6. Style analysis uses a multiple regression model where the factors are category (style) portfolios such as bills, bonds, and stocks. A regression of fund returns on the style portfolio returns generates residuals that represent the value added of stock selection in each period. These residuals can be used to gauge fund performance relative to similar-style funds.
7. The Morningstar Star Rating method compares each fund to a peer group represented by a style portfolio within four asset classes. Risk-adjusted ratings (RAR) are based on fund returns relative to the peer group and used to award each fund one to five stars based on the rank of its RAR. The MRAR is the only manipulation-proof performance measure.
8. Common attribution procedures partition performance improvements to asset allocation, sector selection, and security selection. Performance is assessed by calculating departures of portfolio composition from a benchmark or neutral portfolio.

Related Web sites for this chapter are available at **www.mhhe.com/bkm**

KEY TERMS

time-weighted average
dollar-weighted rate of return
comparison universe
Sharpe's ratio
Treynor's measure
Jensen's alpha
information ratio
bogey

KEY EQUATIONS

Sharpe ratio: $S = \dfrac{r_P - r_f}{\sigma}$

M^2 of portfolio P relative to its Sharpe ratio: $M^2 = \sigma_M(S_P - S_M)$

Treynor measure: $T = \dfrac{r_P - r_f}{\beta}$

Jensen's alpha: $\alpha_P = \bar{r}_P - [\bar{r}_f + \beta_P(\bar{r}_M - \bar{r}_f)]$

Information ratio: $\dfrac{\alpha_P}{\sigma(e_P)}$

Morningstar risk-adjusted return: $MRAR(\gamma) = \left[\dfrac{1}{T}\sum_{t=1}^{T}\left(\dfrac{1 + r_t}{1 + r_{ft}}\right)^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1$

PROBLEM SETS

Basic

1. A household (HH) savings-account spreadsheet shows the following entries:

Date	Additions	Withdrawals	Value
1/1/10			148,000
1/3/10	2,500		
3/20/10	4,000		
7/5/10	1,500		
12/2/10	13,460		
3/10/11		23,000	
4/7/11	3,000		
5/3/11			198,000

Calculate the dollar-weighted average return on the HH savings account between the first and final dates.

2. Is it possible that a positive alpha will be associated with inferior performance? Explain.
3. We know that the geometric average (time-weighted return) on a risky investment is always lower than the corresponding arithmetic average. Can the IRR (the dollar-weighted return) similarly be ranked relative to these other two averages?
4. We have seen that market timing has tremendous potential value. Would it therefore be wise to shift resources to timing at the expense of security selection?

Intermediate

5. Consider the rate of return of stocks ABC and XYZ.

Year	r_{ABC}	r_{XYZ}
1	20%	30%
2	12	12
3	14	18
4	3	0
5	1	−10

a. Calculate the arithmetic average return on these stocks over the sample period.
b. Which stock has greater dispersion around the mean?
c. Calculate the geometric average returns of each stock. What do you conclude?
d. If you were equally likely to earn a return of 20%, 12%, 14%, 3%, or 1%, in each year (these are the five annual returns for stock ABC), what would be your expected rate of return? What if the five possible outcomes were those of stock XYZ?

6. XYZ stock price and dividend history are as follows:

Year	Beginning-of-Year Price	Dividend Paid at Year-End
2013	$100	$4
2014	120	4
2015	90	4
2016	100	4

An investor buys three shares of XYZ at the beginning of 2013, buys another two shares at the beginning of 2014, sells one share at the beginning of 2015, and sells all four remaining shares at the beginning of 2016.

a. What are the arithmetic and geometric average time-weighted rates of return for the investor?
b. What is the dollar-weighted rate of return? (*Hint:* Carefully prepare a chart of cash flows for the *four* dates corresponding to the turns of the year for January 1, 2013, to January 1, 2016. If your calculator cannot calculate internal rate of return, you will have to use trial and error.)

7. A manager buys three shares of stock today, and then sells one of those shares each year for the next 3 years. His actions and the price history of the stock are summarized below. The stock pays no dividends.

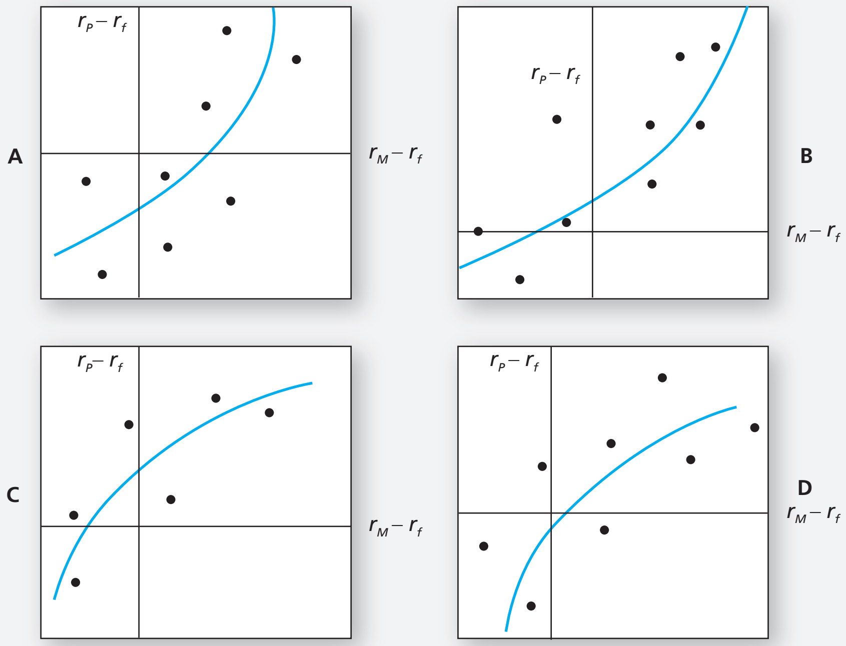
Time	Price	Action
0	$ 90	Buy 3 shares
1	100	Sell 1 share
2	100	Sell 1 share
3	100	Sell 1 share

a. Calculate the time-weighted geometric average return on this "portfolio."
b. Calculate the time-weighted arithmetic average return on this portfolio.
c. Calculate the dollar-weighted average return on this portfolio.

8. Based on current dividend yields and expected capital gains, the expected rates of return on portfolios *A* and *B* are 12% and 16%, respectively. The beta of *A* is .7, while that of *B* is 1.4. The T-bill rate is currently 5%, whereas the expected rate of return of the S&P 500 index is 13%. The standard deviation of portfolio *A* is 12% annually, that of *B* is 31%, and that of the S&P 500 index is 18%.
   a. If you currently hold a market-index portfolio, would you choose to add either of these portfolios to your holdings? Explain.
   b. If instead you could invest *only* in T-bills and *one* of these portfolios, which would you choose?
9. Consider the two (excess return) index-model regression results for stocks *A* and *B*. The risk-free rate over the period was 6%, and the market's average return was 14%. Performance is measured using an index model regression on excess returns.

	Stock *A*	Stock *B*
Index model regression estimates	$1\% + 1.2(r_M - r_f)$	$2\% + .8(r_M - r_f)$
R-square	.576	.436
Residual standard deviation, $\sigma(e)$	10.3%	19.1%
Standard deviation of excess returns	21.6%	24.9%

   a. Calculate the following statistics for each stock:
      i. Alpha
      ii. Information ratio
      iii. Sharpe ratio
      iv. Treynor measure
   b. Which stock is the best choice under the following circumstances?
      i. This is the only risky asset to be held by the investor.
      ii. This stock will be mixed with the rest of the investor's portfolio, currently composed solely of holdings in the market-index fund.
      iii. This is one of many stocks that the investor is analyzing to form an actively managed stock portfolio.
10. Evaluate the market timing and security selection abilities of four managers whose performances are plotted in the accompanying diagrams.



11. Consider the following information regarding the performance of a money manager in a recent month. The table represents the actual return of each sector of the manager's portfolio in column 1, the fraction of the portfolio allocated to each sector in column 2, the benchmark or neutral sector allocations in column 3, and the returns of sector indices in column 4.

	Actual Return	Actual Weight	Benchmark Weight	Index Return
Equity	2%	.70	.60	2.5% (S&P 500)
Bonds	1	.20	.30	1.2 (Salomon Index)
Cash	0.5	.10	.10	0.5

a. What was the manager's return in the month? What was her overperformance or underperformance?
b. What was the contribution of security selection to relative performance?
c. What was the contribution of asset allocation to relative performance? Confirm that the sum of selection and allocation contributions equals her total "excess" return relative to the bogey.

12. A global equity manager is assigned to select stocks from a universe of large stocks throughout the world. The manager will be evaluated by comparing her returns to the return on the MSCI World Market Portfolio, but she is free to hold stocks from various countries in whatever proportions she finds desirable. Results for a given month are contained in the following table:

Country	Weight In MSCI Index	Manager's Weight	Manager's Return in Country	Return of Stock Index for That Country
U.K.	.15	.30	20%	12%
Japan	.30	.10	15	15
U.S.	.45	.40	10	14
Germany	.10	.20	5	12

a. Calculate the total value added of all the manager's decisions this period.
b. Calculate the value added (or subtracted) by her *country* allocation decisions.
c. Calculate the value added from her stock selection ability within countries. Confirm that the sum of the contributions to value added from her country allocation plus security selection decisions equals total over- or underperformance.

13. Conventional wisdom says that one should measure a manager's investment performance over an entire market cycle. What arguments support this convention? What arguments contradict it?

14. Does the use of universes of managers with similar investment styles to evaluate relative investment performance overcome the statistical problems associated with instability of beta or total variability?

15. During a particular year, the T-bill rate was 6%, the market return was 14%, and a portfolio manager with beta of .5 realized a return of 10%.

a. Evaluate the manager based on the portfolio alpha.
b. Reconsider your answer to part (*a*) in view of the Black-Jensen-Scholes finding that the security market line is too flat. Now how do you assess the manager's performance?

16. Bill Smith is evaluating the performance of four large-cap equity portfolios: Funds *A*, *B*, *C*, and *D*. As part of his analysis, Smith computed the Sharpe ratio and the Treynor measure for all four funds. Based on his finding, the ranks assigned to the four funds are as follows:

KAPLAN SCHWESER

Fund	Treynor Measure Rank	Sharpe Ratio Rank
A	1	4
B	2	3
C	3	2
D	4	1

The difference in rankings for Funds *A* and *D* is most likely due to:

a. A lack of diversification in Fund *A* as compared to Fund *D*.
b. Different benchmarks used to evaluate each fund's performance.
c. A difference in risk premiums.

Use the following information to answer Problems 17–20: Primo Management Co. is looking at how best to evaluate the performance of its managers. Primo has been hearing more and more about benchmark portfolios and is interested in trying this approach. As such, the company hired Sally Jones, CFA, as a consultant to educate the managers on the best methods for constructing a benchmark portfolio, how best to choose a benchmark, whether the style of the fund under management matters, and what they should do with their global funds in terms of benchmarking.

For the sake of discussion, Jones put together some comparative 2-year performance numbers that relate to Primo's current domestic funds under management and a potential benchmark.

	Weight		Return	
Style Category	Primo	Benchmark	Primo	Benchmark
Large-cap growth	.60	.50	17%	16%
Mid-cap growth	.15	.40	24	26
Small-cap growth	.25	.10	20	18

As part of her analysis, Jones also takes a look at one of Primo's global funds. In this particular portfolio, Primo is invested 75% in Dutch stocks and 25% in British stocks. The benchmark invested 50% in each—Dutch and British stocks. On average, the British stocks outperformed the Dutch stocks. The euro appreciated 6% versus the U.S. dollar over the holding period while the pound depreciated 2% versus the dollar. In terms of the local return, Primo outperformed the benchmark with the Dutch investments, but underperformed the index with respect to the British stocks.

17. What is the within-sector selection effect for each individual sector?
18. Calculate the amount by which the Primo portfolio out- (or under-)performed the market over the period, as well as the contribution to performance of the pure sector allocation and security selection decisions.
19. If Primo decides to use return-based style analysis, will the R^2 of the regression equation of a passively managed fund be higher or lower than that of an actively managed fund?
20. Which of the following statements about Primo's global fund is most correct? Primo appears to have a positive currency allocation effect as well as

    a. A negative market allocation effect and a positive security allocation effect.
    b. A negative market allocation effect and a negative security allocation effect.
    c. A positive market allocation effect and a negative security allocation effect.

21. Kelli Blakely is a portfolio manager for the Miranda Fund (Miranda), a core large-cap equity fund. The market proxy and benchmark for performance measurement purposes is the S&P 500. Although the Miranda portfolio generally mirrors the asset class and sector weightings of the S&P, Blakely is allowed a significant amount of leeway in managing the fund. Her portfolio holds only stocks found in the S&P 500 and cash.

    Blakely was able to produce exceptional returns last year (as outlined in the table below) through her market timing and security selection skills. At the outset of the year, she became extremely concerned that the combination of a weak economy and geopolitical uncertainties would negatively impact the market. Taking a bold step, she changed her market allocation. For the entire year her asset class exposures averaged 50% in stocks and 50% in cash. The S&P's allocation between stocks and cash during the period was a constant 97% and 3%, respectively. The risk-free rate of return was 2%.

One-Year Trailing Returns

	Miranda Fund	S&P 500
Return	10.2%	–22.5%
Standard deviation	37%	44%
Beta	1.10	1.00

a. What are the Sharpe ratios for the Miranda Fund and the S&P 500?
b. What are the M^2 measures for Miranda and the S&P 500?
c. What is the Treynor measure for the Miranda Fund and the S&P 500?
d. What is the Jensen measure for the Miranda Fund?

Challenge

excel
Please visit us at
www.mhhe.com/bkm

22. Go to Kenneth French's data library site at **http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html**. Select two industry portfolios of your choice and download 36 months of data. Download other data from the site as needed to perform the following tasks.

a. Compare the portfolio's performance to that of the market index on the basis of the various performance measures discussed in the chapter. Plot the monthly values of alpha plus residual return.
b. Now use the Fama-French three-factor model as the return benchmark. Compute plots of alpha plus residual return using the FF model. How does performance change using this benchmark instead of the market index?

CFA® PROBLEMS

1. You and a prospective client are considering the measurement of investment performance, particularly with respect to international portfolios for the past 5 years. The data you discussed are presented in the following table:

International Manager or Index	Total Return	Country and Security Return	Currency Return
Manager A	−6.0%	2.0%	−8.0%
Manager B	−2.0	−1.0	−1.0
International Index	−5.0	0.2	−5.2

a. Assume that the data for manager A and manager B accurately reflect their investment skills and that both managers actively manage currency exposure. Briefly describe one strength and one weakness for each manager.
b. Recommend and justify a strategy that would enable your fund to take advantage of the strengths of each of the two managers while minimizing their weaknesses.

2. Carl Karl, a portfolio manager for the Alpine Trust Company, has been responsible since 2015 for the City of Alpine's Employee Retirement Plan, a municipal pension fund. Alpine is a growing community, and city services and employee payrolls have expanded in each of the past 10 years. Contributions to the plan in fiscal 2020 exceeded benefit payments by a three-to-one ratio.

The plan board of trustees directed Karl 5 years ago to invest for total return over the long term. However, as trustees of this highly visible public fund, they cautioned him that volatile or erratic results could cause them embarrassment. They also noted a state statute that mandated that not more than 25% of the plan's assets (at cost) be invested in common stocks.

At the annual meeting of the trustees in November 2020, Karl presented the following portfolio and performance report to the board:

Alpine Employee Retirement Plan

Asset Mix as of 9/30/20	At Cost (millions)		At Market (millions)	
Fixed-income assets:				
Short-term securities	$ 4.5	11.0%	$ 4.5	11.4%
Long-term bonds and mortgages	26.5	64.7	23.5	59.5
Common stocks	10.0	24.3	11.5	29.1
	$41.0	100.0%	$39.5	100.0%

Investment Performance

	Annual Rates of Return for Periods Ending 9/30/20	
	5 Years	1 Year
Total Alpine Fund:		
Time-weighted	8.2%	5.2%
Dollar-weighted (internal)	7.7%	4.8%
Assumed actuarial return	6.0%	6.0%
U.S. Treasury bills	7.5%	11.3%
Large sample of pension funds (average 60% equities, 40% fixed income)	10.1%	14.3%
Common stocks—Alpine Fund	13.3%	14.3%
Alpine portfolio beta coefficient	0.90	0.89
Standard & Poor's 500 stock index	13.8%	21.1%
Fixed-income securities—Alpine Fund	6.7%	1.0%
Salomon Brothers' bond index	4.0%	−11.4%

Karl was proud of his performance and was chagrined when a trustee made the following critical observations:

a. "Our 1-year results were terrible, and it's what you've done for us lately that counts most."
b. "Our total fund performance was clearly inferior compared to the large sample of other pension funds for the last 5 years. What else could this reflect except poor management judgment?"
c. "Our common stock performance was especially poor for the 5-year period."
d. "Why bother to compare your returns to the return from Treasury bills and the actuarial assumption rate? What your competition could have earned for us or how we would have fared if invested in a passive index (which doesn't charge a fee) are the only relevant measures of performance."
e. "Who cares about time-weighted return? If it can't pay pensions, what good is it!"

Appraise the merits of each of these statements and give counterarguments that Mr. Karl can use.

3. The Retired Fund is an open-ended mutual fund composed of $500 million in U.S. bonds and U.S. Treasury bills. This fund has had a portfolio duration (including T-bills) of between 3 and 9 years. Retired has shown first-quartile performance over the past 5 years, as measured by an independent fixed-income measurement service. However, the directors of the fund would like to

measure the market timing skill of the fund's sole bond investor manager. An external consulting firm has suggested the following three methods:

a. Method I examines the value of the bond portfolio at the beginning of every year, then calculates the return that would have been achieved had that same portfolio been held throughout the year. This return would then be compared with the return actually obtained by the fund.

b. Method II calculates the average weighting of the portfolio in bonds and T-bills for each year. Instead of using the actual bond portfolio, the return on a long-bond market index and T-bill index would be used. For example, if the portfolio on average was 65% in bonds and 35% in T-bills, the annual return on a portfolio invested 65% in a long-bond index and 35% in T-bills would be calculated. This return is compared with the annual return that would have been generated using the indexes and the manager's actual bond/T-bill weighting for each quarter of the year.

c. Method III examines the net bond purchase activity (market value of purchases less sales) for each quarter of the year. If net purchases were positive (negative) in any quarter, the performance of the bonds would be evaluated until the net purchase activity became negative (positive). Positive (negative) net purchases would be viewed as a bullish (bearish) view taken by the manager. The correctness of this view would be measured.

Critique *each* method with regard to market timing measurement problems.

Use the following data to solve CFA Problems 4–6: The administrator of a large pension fund wants to evaluate the performance of four portfolio managers. Each portfolio manager invests only in U.S. common stocks. Assume that during the most recent 5-year period, the average annual total rate of return including dividends on the S&P 500 was 14%, and the average nominal rate of return on government Treasury bills was 8%. The following table shows risk and return measures for each portfolio:

Portfolio	Average Annual Rate of Return	Standard Deviation	Beta
P	17%	20%	1.1
Q	24	18	2.1
R	11	10	0.5
S	16	14	1.5
S&P 500	14	12	1.0

4. What is the Treynor performance measure for portfolio *P*?
5. What is the Sharpe performance measure for portfolio *Q*?
6. An analyst wants to evaluate portfolio *X*, consisting entirely of U.S. common stocks, using both the Treynor and Sharpe measures of portfolio performance. The following table provides the average annual rate of return for portfolio *X*, the market portfolio (as measured by the S&P 500), and U.S. Treasury bills during the past 8 years:

	Average Annual Rate of Return	Standard Deviation of Return	Beta
Portfolio *X*	10%	18%	0.60
S&P 500	12	13	1.00
T-bills	6	N/A	N/A

a. Calculate the Treynor and Sharpe measures for both portfolio *X* and the S&P 500. Briefly explain whether portfolio *X* underperformed, equaled, or outperformed the S&P 500 on a risk-adjusted basis using both the Treynor measure and the Sharpe ratio.

b. On the basis of the performance of portfolio *X* relative to the S&P 500 calculated in part (*a*), briefly explain the reason for the conflicting results when using the Treynor measure versus the Sharpe ratio.

7. Assume you invested in an asset for 2 years. The first year you earned a 15% return, and the second year you earned a negative 10% return. What was your annual geometric return?

8. A portfolio of stocks generates a −9% return in 2013, a 23% return in 2014, and a 17% return in 2015. What is the annualized return (geometric mean) for the entire period?
9. A 2-year investment of $2,000 results in a cash flow of $150 at the end of the first year and another cash flow of $150 at the end of the second year, in addition to the return of the original investment. What is the internal rate of return on the investment?
10. In measuring the performance of a portfolio, the time-weighted rate of return is superior to the dollar-weighted rate of return because:

    a. When the rate of return varies, the time-weighted return is higher.
    b. The dollar-weighted return assumes all portfolio deposits are made on day 1.
    c. The dollar-weighted return can only be estimated.
    d. The time-weighted return is unaffected by the timing of portfolio contributions and withdrawals.

11. A pension fund portfolio begins with $500,000 and earns 15% the first year and 10% the second year. At the beginning of the second year, the sponsor contributes another $500,000. What were the time-weighted and dollar-weighted rates of return?
12. During the annual review of Acme's pension plan, several trustees questioned their investment consultant about various aspects of performance measurement and risk assessment.

    a. Comment on the appropriateness of using each of the following benchmarks for performance evaluation:
    - Market index.
    - Benchmark normal portfolio.
    - Median of the manager universe.

    b. Distinguish among the following performance measures:
    - The Sharpe ratio.
    - The Treynor measure.
    - Jensen's alpha.

    i. Describe how each of the three performance measures is calculated.
    ii. State whether each measure assumes that the relevant risk is systematic, unsystematic, or total. Explain how each measure relates excess return and the relevant risk.

13. Trustees of the Pallor Corp. pension plan ask consultant Donald Millip to comment on the following statements. What should his response be?

    a. Median manager benchmarks are statistically unbiased measures of performance over long periods of time.
    b. Median manager benchmarks are unambiguous and are therefore easily replicated by managers wishing to adopt a passive/indexed approach.
    c. Median manager benchmarks are not appropriate in all circumstances because the median manager universe encompasses many investment styles.

14. James Chan is reviewing the performance of the global equity managers of the Jarvis University endowment fund. Williamson Capital is currently the endowment fund's only large-capitalization global equity manager. Performance data for Williamson Capital are shown in Table 24A.

    Chan also presents the endowment fund's investment committee with performance information for Joyner Asset Management, which is another large-capitalization global equity manager. Performance data for Joyner Asset Management are shown in Table 24B. Performance data for the relevant risk-free asset and market index are shown in Table 24C.

    a. Calculate the Sharpe ratio and Treynor measure for both Williamson Capital and Joyner Asset Management.
    b. The Investment Committee notices that using the Sharpe ratio versus the Treynor measure produces different performance rankings of Williamson and Joyner. Explain why these criteria may result in different rankings.

Average annual rate of return	22.1%
Beta	1.2
Standard deviation of returns	16.8%

Table 24A

Williamson capital performance data, 1999–2010

Average annual rate of return	24.2%
Beta	0.8
Standard deviation of returns	20.2%

Table 24B

Joyner asset management performance data, 1999–2010

Risk-Free Asset	
Average annual rate of return	5.0%
Market Index	
Average annual rate of return	18.9%
Standard deviation of returns	13.8%

Table 24C

Relevant risk-free asset and market index performance data, 1999–2010

E-INVESTMENTS EXERCISES

Several popular finance-related Web sites offer mutual fund screeners. Go to **moneycentral.msn.com** and click on the *Investing* link on the top menu. Choose *Funds* from the submenu, then look for the *Easy Screener* link on the left-side menu. Before you start to specify your preferences using the drop-down boxes, look for the *Show More Options* link toward the bottom of the page and select it. When all of the options are shown, devise a screen for funds that meet the following criteria: 5-star Morningstar Overall Rating, a Minimum Initial Investment as low as possible, Low Morningstar Risk, No Load, Manager Tenure of at least 5 years, Morningstar Overall Return high, 12b-1 fees as low as possible, and Expense Ratio as low as possible. Click on the *Find Funds* link to run the screen.

When you get the list of results, you can sort them according to any one criterion that interests you by clicking on its column heading. Are there any funds you would rule out based on what you see? If you want to rerun the screen with different choices click on the *Change Criteria* link toward the top of the page and make the changes. Click on *Find Funds* again to run the new screen. You can click on any fund symbol to get more information about it.

Are any of these funds of interest to you? How might your screening choices differ if you were choosing funds for various clients?

SOLUTIONS TO CONCEPT CHECKS

1.

Time	Action	Cash Flow
0	Buy two shares	−40
1	Collect dividends; then sell one of the shares	4 + 22
2	Collect dividend on remaining share, then sell it	2 + 19

a. Dollar-weighted return:

$$-40 + \frac{26}{1+r} + \frac{21}{(1+r)^2} = 0$$

$$r = .1191, \text{ or } 11.91\%$$

b. Time-weighted return:
The rates of return on the stock in the 2 years were:

$$r_1 = \frac{2 + (22 - 20)}{20} = .20$$
$$r_2 = \frac{2 + (19 - 22)}{22} = -.045$$
$$(r_1 + r_2)/2 = .077, \text{ or } 7.7\%$$

2. Sharpe: $(\bar{r} - \bar{r}_f)/\sigma$

$$S_P = (35 - 6)/42 = .69$$
$$S_M = (28 - 6)/30 = .733$$

Alpha: $\bar{r} - [\bar{r}_f + \beta(\bar{r}_M - \bar{r}_f)]$

$$\alpha_P = 35 - [6 + 1.2(28 - 6)] = 2.6$$
$$\alpha_M = 0$$

Treynor: $(\bar{r} - \bar{r}_f)/\beta$

$$T_P = (35 - 6)/1.2 = 24.2$$
$$T_M = (28 - 6)/1.0 = 22$$

Information ratio: $\alpha/\sigma(e)$

$$I_P = 2.6/18 = .144$$
$$I_M = 0$$

3. The alpha exceeds zero by .2/2 = .1 standard deviations. A table of the normal distribution (or, somewhat more appropriately, the distribution of the *t*-statistic) indicates that the probability of such an event, if the analyst actually has no skill, is approximately 46%.

4. The timer will guess bear or bull markets completely randomly. One-half of all bull markets will be preceded by a correct forecast, and similarly for bear markets. Hence $P_1 + P_2 - 1 = ½ + ½ - 1 = 0$.

5. First compute the new bogey performance as $(.70 \times 5.81) + (.25 \times 1.45) + (.05 \times .48) = 4.45$.

a. Contribution of asset allocation to performance:

Market	(1) Actual Weight in Market	(2) Benchmark Weight in Market	(3) Active or Excess Weight	(4) Market Return (%)	(5) = (3) × (4) Contribution to Performance (%)
Equity	.70	.70	.00	5.81	.00
Fixed-income	.07	.25	−.18	1.45	−.26
Cash	.23	.05	.18	0.48	.09
Contribution of asset allocation					−.17

b. Contribution of selection to total performance:

Market	(1) Portfolio Performance (%)	(2) Index Performance (%)	(3) Excess Performance (%)	(4) Portfolio Weight	(5) = (3) × (4) Contribution (%)
Equity	7.28	5.00	2.28	.70	1.60
Fixed-income	1.89	1.45	0.44	.07	0.03
Contribution of selection within markets					1.63