CHAPTER 1

Portfolio Selection

FRANK J. FABOZZI, PhD, CFA, CPA
Professor in the Practice of Finance, Yale School of Management

HARRY M. MARKOWITZ, PhD
Consultant

FRANCIS GUPTA, PhD
Director, Research, Dow Jones Indexes

Abstract: The goal of portfolio selection is the construction of portfolios that maximize expected returns consistent with individually acceptable levels of risk. Using both historical data and investor expectations of future returns, portfolio selection uses modeling techniques to quantify "expected portfolio returns" and "acceptable levels of portfolio risk," and provides methods to select an optimal portfolio. It would not be an overstatement to say that modern portfolio theory has revolutionized the world of investment management. Allowing managers to quantify the investment risk and expected return of a portfolio has provided the scientific and objective complement to the subjective art of investment management. More importantly, whereas at one time the focus of portfolio management used to be the risk of individual assets, the theory of portfolio selection has shifted the focus to the risk of the entire portfolio. This theory shows that it is possible to combine risky assets and produce a portfolio whose expected return reflects its components, but with considerably lower risk. In other words, it is possible to construct a portfolio whose risk is smaller than the sum of all its individual parts.

Keywords: portfolio selection, modern portfolio theory, mean-variance analysis, utility function, efficient portfolio, optimal portfolio, covariance, correlation, portfolio diversification, beta, portfolio variance, feasible portfolio

In this chapter, we present the theory of portfolio selection as formulated by Markowitz (1952). This theory is also referred to as mean-variance portfolio analysis or simply *mean-variance analysis*.

SOME BASIC CONCEPTS

Portfolio theory draws on concepts from two fields: financial economic theory and probability and statistical theory. This section presents the concepts from financial economic theory used in portfolio theory. While many of the concepts presented here have a more technical or rigorous definition, the purpose is to keep the explanations simple and intuitive so the reader can appreciate the importance and contribution of these concepts to the development of modern portfolio theory.

Utility Function and Indifference Curves

In life there are many situations where entities (that is, individuals and firms) face two or more choices. The economic "theory of choice" uses the concept of a utility function developed by von Neuman and Morgenstern (1944), to describe the way entities make decisions when faced with a set of choices. A *utility function* assigns a (numeric) value to all possible choices faced by the entity. The higher the value of a particular choice, the greater the utility derived from that choice. The choice that is selected is the one that results in the maximum utility given a set of (budget) constraints faced by the entity.

In portfolio theory too, entities are faced with a set of choices. Different portfolios have different levels of expected return and risk. Also, the higher the level of expected return, the larger the risk. Entities are faced with the decision of choosing a portfolio from the set of all possible risk/return combinations: where return is a desirable which increases the level of utility, and risk is an undesirable which decreases the level of utility. Therefore, entities obtain different levels of utility from different risk/return combinations. The utility obtained from any possible risk/return combination is expressed by the utility function. Put simply, the utility function expresses the preferences of entities over perceived risk and expected return combinations.

A utility function can be expressed in graphical form by a set of indifference curves. Figure 1.1 shows indifference curves labeled u_1, u_2, and u_3. By convention, the horizontal axis measures risk and the vertical axis measures expected return. Each curve represents a set of portfolios with different combinations of risk and return. All the points on a given indifference curve indicate combinations of risk and expected return that will give the same level of utility to a given investor. For example, on utility curve u_1, there are two points u and u', with u having a higher expected return than u', but also having a higher risk. Because the two points lie on the same indifference curve, the investor has an equal preference for (or is indifferent to) the two points, or, for that matter, any point on the curve. The (positive) slope of an indifference curve reflects the fact that, to obtain the same level of utility, the investor requires a higher expected return in order to accept higher risk.

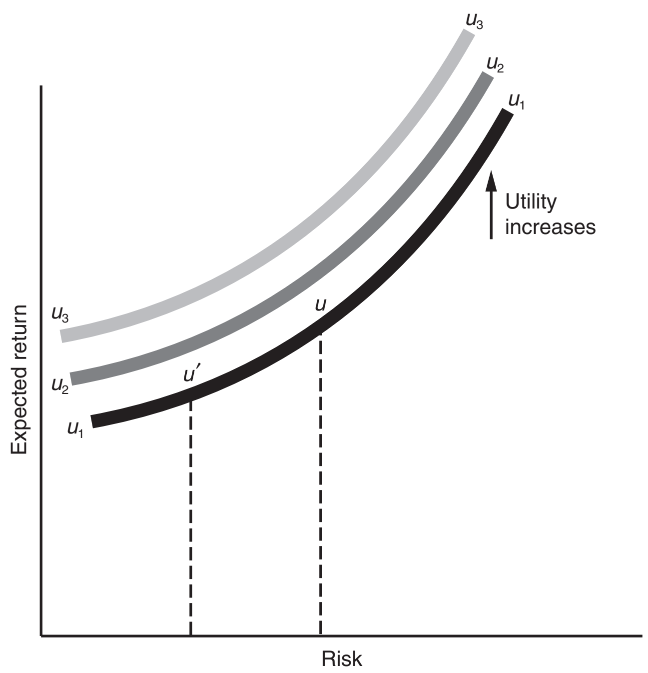

Figure 1.1 Indifference Curves

For the three indifference curves shown in Figure 1.1, the utility the investor receives is greater the further the indifference curve is from the horizontal axis, because that curve represents a higher level of return at every level of risk. Thus, for the three indifference curves shown in the exhibit, u_3 has the highest utility and u_1 the lowest.

The Set of Efficient Portfolios and the Optimal Portfolio

Portfolios that provide the largest possible expected return for given levels of risk are called *efficient portfolios*. To construct an efficient portfolio, it is necessary to make some assumption about how investors behave when making investment decisions. One reasonable assumption is that investors are risk averse. A risk-averse investor is an investor who, when faced with choosing between two investments with the same expected return but two different risks, prefers the one with the lower risk.

In selecting portfolios, an investor seeks to maximize the expected portfolio return given his tolerance for risk. Alternatively stated, an investor seeks to minimize the risk that he is exposed to given some target expected return. Given a choice from the set of efficient portfolios, an *optimal portfolio* is the one that is most preferred by the investor.

Risky Assets versus Risk-Free Assets

A risky asset is one for which the return that will be realized in the future is uncertain. For example, an investor who purchases the stock of Pfizer Corporation today with the intention of holding it for some finite time does not know what return will be realized at the end of the holding period. The return will depend on the price of Pfizer's stock at the time of sale and on the dividends that the company pays during the holding period. Thus, Pfizer stock, and indeed the stock of all companies, is a risky asset.

Securities issued by the U.S. government are also risky. For example, an investor who purchases a U.S. government bond that matures in 30 years does not know the return that will be realized if this bond is held for only one year. This is because changes in interest rates in that year will affect the price of the bond one year from now and that will impact the return on the bond over that year.

There are assets, however, for which the return that will be realized in the future is known with certainty today. Such assets are referred to as risk-free or riskless assets. The risk-free asset is commonly defined as a short-term obligation of the U.S. government. For example, if an investor buys a U.S. government security that matures in one year and plans to hold that security for one year, then there is no uncertainty about the return that will be realized. (Note: Here "return" refers to the nominal return. The "real" return, which adjusts for inflation, is uncertain.) The investor knows that in one year, the maturity date of the security, the government will pay a specific amount to retire the debt. Notice how this situation differs for the U.S. government security that matures in 30 years. While the 1-year and the 30-year securities are obligations of the U.S. government, the former matures in one year so that there is no uncertainty about the return that will be realized. In contrast, while the investor knows what the government will pay at the end of 30 years for the 30-year bond, he does not know what the price of the bond will be one year from now.

MEASURING A PORTFOLIO'S EXPECTED RETURN

We are now ready to define the actual and expected return of a risky asset and a portfolio of risky assets.

Measuring Single-Period Portfolio Return

The actual return on a portfolio of assets over some specific time period is straightforward to calculate using the following:

$$R_p = w_1 R_1 + w_2 R_2 + \ldots + w_G R_G \tag{1.1}$$

where

R_p = rate of return on the portfolio over the period
R_g = rate of return on asset g over the period
w_g = weight of asset g in the portfolio (that is, market value of asset g as a proportion of the market value of the total portfolio) at the beginning of the period
G = number of assets in the portfolio

In shorthand notation, equation (1.1) can be expressed as follows:

$$R_p = \sum_{g=1}^{G} w_g R_g \tag{1.2}$$

Equation (1.2) states that the return on a portfolio (R_p) of G assets is equal to the sum over all individual assets' weights in the portfolio times their respective return. The portfolio return R_p is sometimes called the holding period return or the ex post return.

For example, consider the following portfolio consisting of three assets:

Asset	Market Value at the Beginning of Holding Period	Rate of Return Over Holding Period
1	\$6 million	12%
2	8 million	10%
3	11 million	5%

The portfolio's total market value at the beginning of the holding period is \$25 million. Therefore,

w_1 = \$6 million/\$25 million = 0.24, or 24% and R_1 = 12%
w_2 = \$8 million/\$25 million = 0.32, or 32% and R_2 = 10%
w_3 = \$11 million/\$25 million = 0.44, or 44% and R_3 = 5%

Notice that the sum of the weights is equal to 1. Substi- tuting into equation (1.1), we get the holding period portfolio return,

$$R_p = 0.24(12\%) + 0.32(10\%) + 0.44(5\%) = 8.28\%$$

Note that since the holding period portfolio return is 8.28%, the growth in the portfolio's value over the holding period is given by (\$25 million) × 0.0828 = \$2.07 million.

The Expected Return of a Portfolio of Risky Assets

Equation (1.1) shows how to calculate the actual return of a portfolio over some specific time period. In portfolio management, the investor also wants to know the expected (or anticipated) return from a portfolio of risky assets. The expected portfolio return is the weighted average of the expected return of each asset in the portfolio. The weight assigned to the expected return of each asset is the percentage of the market value of the asset to the total market value of the portfolio. That is,

$$E(R_p) = w_1 E(R_1) + w_2 E(R_2) + \ldots + w_G E(R_G) \tag{1.3}$$

The $E(\,)$ signifies expectations, and $E(R_P)$ is sometimes called the *ex ante* return, or the expected portfolio return over some specific time period.

The expected return, $E(R_i)$, on a risky asset i is calculated as follows. First, a probability distribution for the possible rates of return that can be realized must be specified. A probability distribution is a function that assigns a probability of occurrence to all possible outcomes for a random variable. Given the probability distribution, the expected

Table 1.1 Probability Distribution for the Rate of Return for Stock XYZ

n	Rate of Return	Probability of Occurrence
1	12%	0.18
2	10	0.24
3	8	0.29
4	4	0.16
5	−4	0.13
Total		1.00

value of a random variable is simply the weighted average of the possible outcomes, where the weight is the probability associated with the possible outcome.

In our case, the random variable is the uncertain return of asset *i*. Having specified a probability distribution for the possible rates of return, the expected value of the rate of return for asset *i* is the weighted average of the possible outcomes. Finally, rather than use the term "expected value of the return of an asset," we simply use the term "expected return." Mathematically, the expected return of asset *i* is expressed as:

$$E(R_i) = p_1 R_1 + p_2 R_2 + \cdots + p_N R_N \tag{1.4}$$

where,

R_n = the nth possible rate of return for asset i
p_n = the probability of attaining the rate of return n for asset i
N = the number of possible outcomes for the rate of return

How do we specify the probability distribution of returns for an asset? We shall see later on in this chapter that in most cases the probability distribution of returns is based on historical returns. Probabilities assigned to different return outcomes that are based on the past performance of an uncertain investment act as a good estimate of the probability distribution. However, for purpose of illustration, assume that an investor is considering an investment, stock XYZ, which has a probability distribution for the rate of return for some time period as given in Table 1.1. The stock has five possible rates of return and the probability distribution specifies the likelihood of occurrence (in a probabilistic sense) for each of the possible outcomes.

Substituting into equation (1.4) we get

$$\begin{aligned} E(R_{XYZ}) &= 0.18(12\%) + 0.24(10\%) + 0.29(8\%) \\ &\quad + 0.16(4\%) + 0.13(-4\%) \\ &= 7\% \end{aligned}$$

Thus, 7% is the expected return or mean of the probability distribution for the rate of return on stock XYZ.

MEASURING PORTFOLIO RISK

The dictionary defines risk as "hazard, peril, exposure to loss or injury." With respect to investments, investors have used a variety of definitions to describe risk. Markowitz (1952, 1959) quantified the concept of risk using the well-known statistical measures of variances and covariances. He defined the risk of a portfolio as the sum of the variances of the investments and covariances among the investments. The notion of introducing the covariances among returns of the investments in the portfolio to measure the risk of a portfolio forever changed how the investment community thought about the concept of risk.

Variance and Standard Deviation as a Measure of Risk

The variance of a random variable is a measure of the dispersion or variability of the possible outcomes around the expected value (mean). In the case of an asset's return, the variance is a measure of the dispersion of the possible rate of return outcomes around the expected return.

The equation for the variance of the expected return for asset *i*, denoted var(R_i), is

$$\begin{aligned} \text{var}(R_i) &= p_1[r_1 - E(R_i)]^2 + p_2[r_2 - E(R_i)]^2 \\ &\quad + \cdots + p_N[r_N - E(R_i)]^2 \end{aligned}$$

or,

$$\text{var}(R_i) = \sum_{i=1}^{N} p_n[r_n - E(R_i)]^2 \tag{1.5}$$

Using the probability distribution of the return for stock XYZ, we can illustrate the calculation of the variance:

$$\begin{aligned} \text{var}(R_{XYZ}) &= 0.18(12\% - 7\%)^2 + 0.24(10\% - 7\%)^2 \\ &\quad + 0.29(8\% - 7\%)^2 + 0.16(4\% - 7\%)^2 \\ &\quad + 0.13(-4\% - 7\%)^2 \\ &= 24.1\% \end{aligned}$$

The variance associated with a distribution of returns measures the compactness with which the distribution is clustered around the mean or expected return. Markowitz argued that this compactness or variance is equivalent to the uncertainty or riskiness of the investment. If an asset is riskless, it has an expected return dispersion of zero. In other words, the return (which is also the expected return in this case) is certain, or guaranteed.

Standard Deviation

Since the variance is squared units, it is common to see the variance converted to the standard deviation by taking the positive square root of the variance:

$$\text{SD}(R_i) = \sqrt{\text{var}(R_i)}$$

For stock XYZ, then, the standard deviation is

$$\text{SD}(R_{\text{XYZ}}) = \sqrt{24.1\%} = 4.9\%$$

The variance and standard deviation are conceptually equivalent; that is, the larger the variance or standard deviation, the greater the investment risk.

There are two criticisms of the use of the variance as a measure of risk. The first criticism is that since the variance measures the dispersion of an asset's return around its expected return, it treats both the returns above and below the expected return identically. There has been research in the area of behavioral finance to suggest that investors do not view return outcomes above the expected return in the same way as they view returns below the expected return. Whereas returns above the expected return are considered favorable, outcomes below the expected return are disliked. Because of this, some researchers have argued that measures of risk should not consider the possible return outcomes above the expected return.

Markowitz recognized this limitation and, in fact, suggested a measure of downside risk—the risk of realizing an outcome below the expected return—called the semivariance. The semivariance is similar to the variance except that in the calculation no consideration is given to returns above the expected return. However, because of the computational problems with using the semivariance and the limited resources available to him at the time, he used the variance in developing portfolio theory.

Today, practitioners use various measures of downside risk. However, regardless of the measure used, the basic principles of portfolio theory developed by Markowitz and set forth in this chapter are still applicable. That is, the choice of the measure of risk may affect the calculation but doesn't invalidate the theory.

The second criticism is that the variance is only one measure of how the returns vary around the expected return. When a probability distribution is not symmetrical around its expected return, then a statistical measure of the skewness of a distribution should be used in addition to the variance (see Ortobelli et al., 2005).

Because expected return and variance are the only two parameters that investors are assumed to consider in making investment decisions, the Markowitz formulation of portfolio theory is often referred to as a "two-parameter model." It is also referred to as "mean-variance analysis."

Measuring the Portfolio Risk of a Two-Asset Portfolio

Equation (1.5) gives the variance for an individual asset's return. The variance of a portfolio consisting of two assets is a little more difficult to calculate. It depends not only on the variance of the two assets, but also upon how closely the returns of one asset track those of the other asset. The formula is:

$$\text{var}(R_p) = \text{var}(R_i) + \text{var}(R_j) + 2w_i w_j \,\text{cov}(R_i, R_j) \tag{1.6}$$

where

$\text{cov}(R_i, R_j)$ = covariance between the return for assets i and j

In words, equation (1.6) states that the variance of the portfolio return is the sum of the squared weighted variances of the two assets plus two times the weighted covariance between the two assets. We will see that this equation can be generalized to the case where there are more than two assets in the portfolio.

Covariance

Like the variance, the covariance has a precise mathematical translation. Its practical meaning is the degree to which the returns on two assets co-vary or change together. In fact, the covariance is just a generalized concept of the variance applied to multiple assets. A positive covariance between two assets means that the returns on two assets tend to move or change in the same direction, while a negative covariance means the returns tend to move in opposite directions. The covariance between any two assets i and j is computed using the following formula:

$$\begin{aligned}\text{cov}(R_i, R_j) = {} & p_1[r_{i1} - E(R_i)][r_{j1} - E(R_j)] \\ & + p_2[r_{i2} - E(R_i)][r_{j2} - E(R_j)] \\ & + \cdots + p_N[r_{iN} - E(R_i)][r_{jN} - E(R_j)]\end{aligned} \tag{1.7}$$

where

r_{in} = the nth possible rate of return for asset i
r_{jn} = the nth possible rate of return for asset j
p_n = the probability of attaining the rate of return n for assets i and j
N = the number of possible outcomes for the rate of return.

The covariance between asset i and i is just the variance of asset i.

To illustrate the calculation of the covariance between two assets, we use the two stocks in Table 1.2. The first is stock XYZ from Table 1.1 that we used earlier to illustrate the calculation of the expected return and the standard deviation. The other hypothetical stock is ABC whose data

Table 1.2 Probability Distribution for the Rate of Return for Asset XYZ and Asset ABC

n	Rate of Return for Asset XYZ	Rate of Return for Asset ABC	Probability of Occurrence
1	12%	21%	0.18
2	10	14	0.24
3	8	9	0.29
4	4	4	0.16
5	−4	−3	0.13
Total			1.00
Expected return	7.0%	10.0%	
Variance	24.1%	53.6%	
Standard deviation	4.9%	7.3%	

are shown in Table 1.2. Substituting the data for the two stocks from Table 1.2 in equation (1.7), the covariance between stocks XYZ and ABC is calculated as follows:

$$\begin{aligned}\mathrm{cov}(R_{XYZ}, R_{ABC}) &= 0.18(12\% - 7\%)(21\% - 10\%)\\ &\quad + 0.24(10\% - 7\%)(14\% - 10\%)\\ &\quad + 0.29(8\% - 7\%)(9\% - 10\%)\\ &\quad + 0.16(4\% - 7\%)(4\% - 10\%)\\ &\quad + 0.13(-4\% - 7\%)(-3\% - 10\%)\\ &= 34\end{aligned}$$

Relationship between Covariance and Correlation

The *correlation* is analogous to the *covariance* between the expected returns for two assets. Specifically, the correlation between the returns for assets i and j is defined as the covariance of the two assets divided by the product of their standard deviations:

$$\mathrm{cor}(R_i, R_j) = \mathrm{cov}(R_i, R_j)/[\mathrm{SD}(R_i)\mathrm{SD}(R_j)] \tag{1.8}$$

The correlation and the covariance are conceptually equivalent terms. Dividing the covariance between the returns of two assets by the product of their standard deviations results in the correlation between the returns of the two assets. Because the correlation is a standardized number (that is, it has been corrected for differences in the standard deviation of the returns), the correlation is comparable across different assets. The correlation between the returns for stock XYZ and stock ABC is

$$\mathrm{cor}(R_{XYZ}, R_{ABC}) = 34/(4.9 \times 7.3) = 0.94$$

The correlation coefficient can have values ranging from +1.0, denoting perfect co-movement in the same direction, to –1.0, denoting perfect co-movement in the opposite direction. Also note that because the standard deviations are always positive, the correlation can be negative only if the covariance is a negative number. A correlation of zero implies that the returns are uncorrelated. Finally, though causality implies correlation, correlation does not imply causality.

Measuring the Risk of a Portfolio Comprised of More than Two Assets

So far, we have defined the risk of a portfolio consisting of two assets. The extension to three assets—i, j, and k—is as follows:

$$\begin{aligned}\mathrm{var}(R_p) &= w_i^2\,\mathrm{var}(R_i) + w_j^2\,\mathrm{var}(R_j) + w_k^2\,\mathrm{var}(R_k)\\ &\quad + 2w_iw_j\,\mathrm{cov}(R_i, R_j) + 2w_iw_k\,\mathrm{cov}(R_i, R_k)\\ &\quad + 2w_jw_k\,\mathrm{cov}(R_j, R_k)\end{aligned} \tag{1.9}$$

In words, equation (1.9) states that the variance of the portfolio return is the sum of the squared weighted variances of the individual assets plus two times the sum of the weighted pair-wise covariances of the assets. In general, for a portfolio with G assets, the portfolio variance is given by,

$$\mathrm{var}(R_p) = \sum_{g=1}^{G}\sum_{h=1}^{G} w_g w_h\,\mathrm{cov}(R_g, R_h) \tag{1.10}$$

In (1.10), the terms for which $h = g$ results in the variances of the G assets, and the terms for which $h \neq g$ results in all possible pair-wise covariances amongst the G assets. Therefore, (1.10) is a shorthand notation for the sum of all G variances and the possible covariances amongst the G assets.

PORTFOLIO DIVERSIFICATION

Often, one hears investors talking about diversifying their portfolio. By this an investor means constructing a portfolio in such a way as to reduce portfolio risk without sacrificing return. This is certainly a goal that investors should seek. However, the question is how to do this in practice.

Some investors would say that including assets across all asset classes could diversify a portfolio. For example, a investor might argue that a portfolio should be diversified by investing in stocks, bonds, and real estate. While that might be reasonable, two questions must be addressed in order to construct a diversified portfolio. First, how much should be invested in each asset class? Should 40% of the portfolio be in stocks, 50% in bonds, and 10% in real estate, or is some other allocation more appropriate? Second, given the allocation, which specific stocks, bonds, and real estate should the investor select?

Some investors who focus only on one asset class such as common stock argue that such portfolios should also be diversified. By this they mean that an investor should not place all funds in the stock of one corporation, but rather should include stocks of many corporations. Here, too, several questions must be answered in order to construct a diversified portfolio. First, which corporations should be represented in the portfolio? Second, how much of the portfolio should be allocated to the stocks of each corporation?

Prior to the development of portfolio theory, while investors often talked about diversification in these general terms, they did not possess the analytical tools by which to answer the questions posed above. For example, in 1945, D. H. Leavens (1945) wrote:

> An examination of some fifty books and articles on investment that have appeared during the last quarter of a century shows that most of them refer to the desirability of diversification. The majority, however, discuss it in general terms and do not clearly indicate why it is desirable.

Leavens illustrated the benefits of diversification on the assumption that risks are independent. However, in the last paragraph of his article, he cautioned:

> The assumption, mentioned earlier, that each security is acted upon by independent causes, is important, although it cannot always be fully met in practice. Diversification among companies in one industry cannot protect against unfavorable factors that may affect the whole industry; additional diversification among industries is needed for that purpose. Nor can diversification among industries protect against cyclical factors that may depress all industries at the same time.

A major contribution of the theory of portfolio selection is that using the concepts discussed above, a quantitative measure of the diversification of a portfolio is possible, and it is this measure that can be used to achieve the maximum diversification benefits.

The Markowitz diversification strategy is primarily concerned with the degree of covariance between asset returns in a portfolio. Indeed a key contribution of Markowitz diversification is the formulation of an asset's risk in terms of a portfolio of assets, rather than in isolation. Markowitz diversification seeks to combine assets in a portfolio with returns that are less than perfectly positively correlated, in an effort to lower portfolio risk (variance) without sacrificing return. It is the concern for maintaining return while lowering risk through an analysis of the covariance between asset returns, that separates Markowitz diversification from a naive approach to diversification and makes it more effective.

Markowitz diversification and the importance of asset correlations can be illustrated with a simple two-asset portfolio example. To do this, we will first show the general relationship between the risk of a two-asset portfolio and the correlation of returns of the component assets. Then we will look at the effects on portfolio risk of combining assets with different correlations.

Portfolio Risk and Correlation

In our two-asset portfolio, assume that asset C and asset D are available with expected returns and standard deviations as shown:

Asset	E(R)	SD(R)
Asset C	12%	30%
Asset D	18%	40%

If an equal 50% weighting is assigned to both stocks C and D, the expected portfolio return can be calculated as shown:

$$E(R_p) = 0.50(12\%) + 0.50(18\%) = 15\%$$

The variance of the return on the two-stock portfolio from equation (1.6) is:

$$\begin{aligned}\text{var}(R_p) &= w_C^2\,\text{var}(R_C) + w_D^2 w_D\, 2\,\text{var}(R_D) \\ &\quad + 2w_C w_D\,\text{cov}(R_C, R_D) \\ &= (0.5)^2(30\%)^2 + (0.5)^2(40\%)^2 \\ &\quad + 2(0.5)(0.5)\,\text{cov}(R_C, R_D)\end{aligned}$$

From equation (1.8),

$$\text{cor}(R_C, R_D) = \text{cov}(R_C, R_D)/[\text{SD}(R_C)\text{SD}(R_D)]$$

so

$$\text{cov}(R_C, R_D) = \text{SD}(R_C)\,\text{SD}(R_D)\,\text{cor}(R_C, R_D)$$

Since $\text{SD}(R_C) = 30\%$ and $\text{SD}(R_D) = 40\%$, then

$$\text{cov}(R_C, R_D) = (30\%)(40\%)\,\text{cor}(R_C, R_D)$$

Substituting into the expression for $\text{var}(R_p)$, we get

$$\begin{aligned}\text{var}(R_p) &= (0.5)^2(30\%)^2 + (0.5)^2(40\%)^2 \\ &\quad + 2(0.5)(0.5)(30\%)(40\%)\,\text{cor}(R_C, R_D)\end{aligned}$$

Taking the square root of the variance gives

$$\begin{aligned}&\text{SD}(R_p) \\ &= \sqrt{(0.5)^2(30\%)^2 + (0.5)^2(40\%)^2 + 2(0.5)(0.5)(30\%)(40\%)\,\text{cor}(R_C, R_D)} \\ &= \sqrt{625 + (600)\,\text{cor}(R_C, R_D)}\end{aligned} \tag{1.11}$$

The Effect of the Correlation of Asset Returns on Portfolio Risk

How would the risk change for our two-asset portfolio with different correlations between the returns of the component stocks? Let's consider the following three cases for $\text{cor}(R_C,R_D)$: +1.0, 0, and –1.0. Substituting into equation (1.11) for these three cases of $\text{cor}(R_C,R_D)$, we get

cor	E	SD
+1.0	15%	35%
0.0	15	25
−1.0	15	5

As the correlation between the expected returns on stocks C and D decreases from +1.0 to 0.0 to –1.0, the standard deviation of the expected portfolio return also decreases from 35% to 5%. However, the expected portfolio return remains 15% for each case.

This example clearly illustrates the effect of Markowitz diversification. The principle of Markowitz diversification states that as the correlation (covariance) between the returns for assets that are combined in a portfolio decreases, so does the variance (hence the standard deviation) of the return for the portfolio. This is due to the degree of correlation between the expected asset returns.

The good news is that investors can maintain expected portfolio return and lower portfolio risk by combining assets with lower (and preferably negative) correlations. However, the bad news is that very few assets have small to negative correlations with other assets! The problem, then, becomes one of searching among large numbers of assets in an effort to discover the portfolio with the minimum risk at a given level of expected return or, equivalently, the highest expected return at a given level of risk.

The stage is now set for a discussion of efficient portfolios and their construction.

CHOOSING A PORTFOLIO OF RISKY ASSETS

Diversification in the manner suggested by Professor Markowitz leads to the construction of portfolios that have the highest expected return at a given level of risk. Such portfolios are called efficient portfolios. In order to construct efficient portfolios, the theory makes some basic assumptions about asset selection behavior by the entities. The assumptions are as follows:

Assumption 1: The only two parameters that affect an investor's decision are the expected return and the variance. (That is, investors make decisions using the two-parameter model formulated by Markowitz.)

Assumption 2: Investors are risk averse. (That is, when faced with two investments with the same expected return but two different risks, investors will prefer the one with the lower risk.)

Assumption 3: All investors seek to achieve the highest expected return at a given level of risk.

Table 1.3 Portfolio Expected Returns and Standard Deviations for Five Mixes of Assets C and D

Portfolio	Proportion of Asset C	Proportion of Asset D	$E(R_p)$	$SD(R_p)$
1	100%	0%	12.0%	30.0%
2	75	25	13.5%	19.5%
3	50	50	15.0%	18.0%
4	25	75	16.5%	27.0%
5	0	100	18.0%	40.0%

Asset C: $E(R_C) = 12\%$, $SD(R_C) = 30\%$
Asset D: $E(R_D) = 18\%$, and $SD(R_D) = 40\%$
Correlation between Asset C. and D = $cor(R_C,R_D) = -0.5$

Assumption 4: All investors have the same expectations regarding expected return, variance, and covariances for all risky assets. (This is referred to as the homogeneous expectations assumption.)
Assumption 5: All investors have a common one-period investment horizon.

Constructing Efficient Portfolios

The technique of constructing efficient portfolios from large groups of stocks requires a massive number of calculations. In a portfolio of G securities, there are $(G^2 - G)/2$ unique covariances to calculate. Hence, for a portfolio of just 50 securities, there are 1,224 covariances that must be calculated. For 100 securities, there are 4,950. Furthermore, in order to solve for the portfolio that minimizes risk for each level of return, a mathematical technique called quadratic programming must be used. A discussion of this technique is beyond the scope of this chapter. However, it is possible to illustrate the general idea of the construction of efficient portfolios by referring again to the simple two-asset portfolio consisting of assets C and D.

Recall that for two assets, C and D, $E(R_C) = 12\%$, $SD(R_C) = 30\%$, $E(R_D) = 18\%$, and $SD(R_D) = 40\%$. We now further assume that $cor(R_C,R_D) = -0.5$. Table 1.3 presents the expected portfolio return and standard deviation for five different portfolios made up of varying proportions of C and D.

Feasible and Efficient Portfolios

A *feasible portfolio* is any portfolio that an investor can construct given the assets available. The five portfolios presented in Table 1.3 are all feasible portfolios. The collection of all feasible portfolios is called the feasible set of portfolios. With only two assets, the feasible set of portfolios is graphed as a curve that represents those combinations of risk and expected return that are attainable by constructing portfolios from all possible combinations of the two assets.

Figure 1.2 presents the feasible set of portfolios for all combinations of assets C and D. As mentioned earlier, the portfolio mixes listed in Table 1.3 belong to this set and are shown by the points 1 through 5, respectively. Starting from 1 and proceeding to 5, asset C goes from 100% to 0%, while asset D goes from 0% to 100%—therefore, all

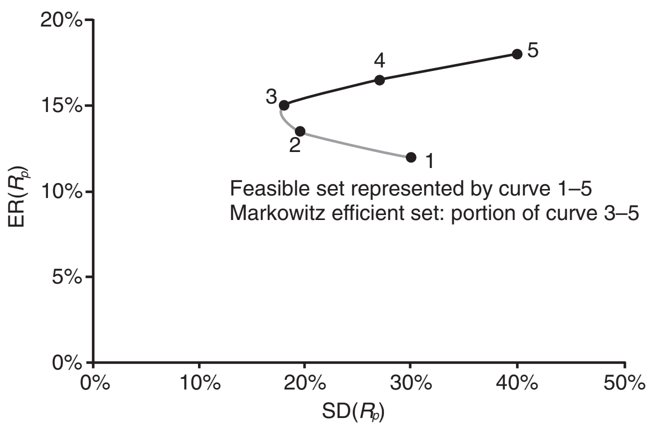

Figure 1.2 Feasible and Efficient Portfolios for Assets C and D

possible combinations of C and D lie between portfolios 1 and 5, or on the curve labeled 1–5. In the case of two assets, any risk/return combination not lying on this curve is not attainable, since there is no mix of assets C and D that will result in that risk/return combination. Consequently, the curve 1–5 can also be thought of as the feasible set.

In contrast to a feasible portfolio, an *efficient portfolio* is one that gives the highest expected return of all feasible portfolios with the same risk. An efficient portfolio is also said to be a mean-variance efficient portfolio. Thus, for each level of risk there is an efficient portfolio. The collection of all efficient portfolios is called the efficient set.

The efficient set for the feasible set presented in Figure 1.2 is differentiated by the bold curve section 3–5. Efficient portfolios are the combinations of assets C and D that result in the risk/return combinations on the bold section of the curve. These portfolios offer the highest expected return at a given level of risk. Notice that two of our five portfolio mixes—portfolio 1 with $E(R_p) = 12\%$ and $SD(R_p) = 20\%$ and portfolio 2 with $E(R_p) = 13.5\%$ and $SD(R_p) = 19.5\%$—are not included in the efficient set. This is because there is at least one portfolio in the efficient set (for example, portfolio 3) that has a higher expected return and lower risk than both of them. We can also see that portfolio 4 has a higher expected return and lower risk than portfolio 1. In fact, the whole curve section 1–3 is not efficient. For any given risk/return combination on this curve section, there is a combination (on the curve section 3–5) that has the same risk and a higher return, or the same return and a lower risk, or both. In other words, for any portfolio that results in the return/risk combination on the curve section 1–3 (excluding portfolio 3), there exists a portfolio that dominates it by having the same return and lower risk, or the same risk and a higher return, or a lower risk and a higher return. For example, portfolio 4 dominates portfolio 1, and portfolio 3 dominates both portfolios 1 and 2.

Figure 1.3 shows the feasible and efficient sets when there are more than two assets. In this case, the feasible set is not a line, but an area. This is because, unlike the two-asset case, it is possible to create asset portfolios that result in risk/return combinations that not only result in

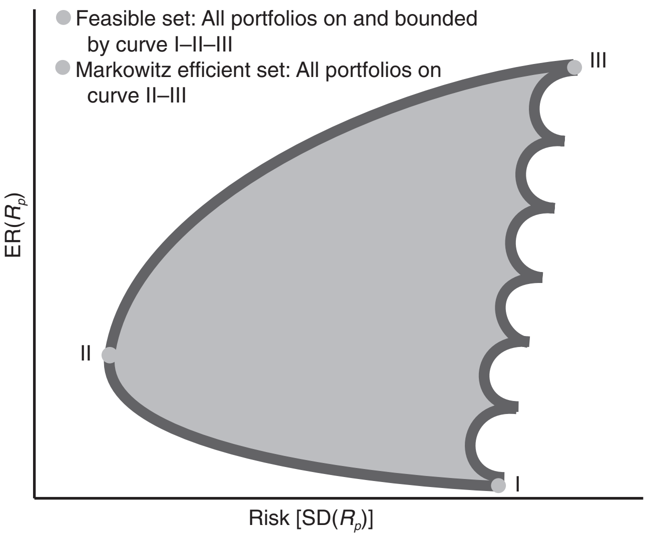

Figure 1.3 Feasible and Efficient Portfolios with More Than Two Assets*

* The picture is for illustrative purposes only. The actual shape of the feasible region depends on the returns and risks of the assets chosen and the correlation among them.

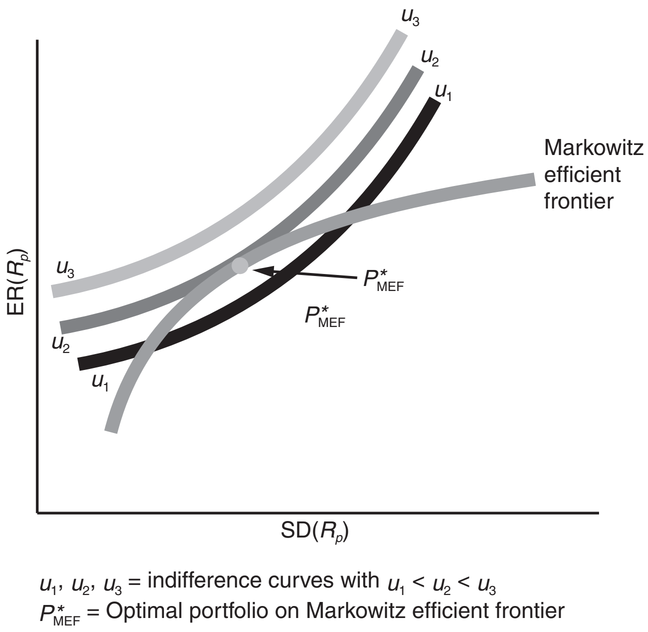

Figure 1.4 Selection of the Optimal Portfolio

combinations that lie on the curve I–II–III, but all combinations that lie in the shaded area. However, the efficient set is given by the curve II–III. It is easily seen that all the portfolios on the efficient set dominate the portfolios in the shaded area.

The efficient set of portfolios is sometimes called the efficient frontier, because graphically all the efficient portfolios lie on the boundary of the set of feasible portfolios that have the maximum return for a given level of risk. Any risk/return combination above the efficient frontier cannot be achieved, while risk/return combinations of the portfolios that make up the efficient frontier dominate those that lie below the efficient frontier.

Choosing the Optimal Portfolio in the Efficient Set

Now that we have constructed the efficient set of portfolios, the next step is to determine the optimal portfolio.

Since all portfolios on the efficient frontier provide the greatest possible return at their level of risk, an investor or entity will want to hold one of the portfolios on the efficient frontier. Notice that the portfolios on the efficient frontier represent trade-offs in terms of risk and return. Moving from left to right on the efficient frontier, the risk increases, but so does the expected return. The question is which one of those portfolios should an investor hold? The best portfolio to hold of all those on the efficient frontier is the optimal portfolio.

Intuitively, the optimal portfolio should depend on the investor's preference over different risk/return trade-offs. As explained earlier, this preference can be expressed in terms of a utility function.

In Figure 1.4, three indifference curves representing a utility function and the efficient frontier are drawn on the same diagram. An indifference curve indicates the combinations of risk and expected return that give the same level of utility. Moreover, the farther the indifference curve from the horizontal axis, the higher the utility.

From Figure 1.4, it is possible to determine the optimal portfolio for the investor with the indifference curves shown. Remember that the investor wants to get to the highest indifference curve achievable given the efficient frontier. Given that requirement, the optimal portfolio is represented by the point where an indifference curve is tangent to the efficient frontier. In Figure 1.4, that is the portfolio P^*_{MEF}. For example, suppose that P^*_{MEF} corresponds to portfolio 4 in Figure 1.2. We know from Table 1.3 that this portfolio is made up of 25% of asset C and 75% of asset D, with an $E(R_p) = 16.5\%$ and $\text{SD}(R_p) = 27.0\%$.

Consequently, for the investor's preferences over risk and return as determined by the shape of the indifference curves represented in Figure 1.4, and expectations for asset C and D inputs (returns and variance-covariance) represented in Table 1.2, portfolio 4 is the optimal portfolio because it maximizes the investor's utility. If this investor had a different preference for expected risk and return, there would have been a different optimal portfolio. For example, Figure 1.5 shows the same efficient frontier but three other indifference curves. In this case, the optimal portfolio is P^{**}_{MEF}, which has a lower expected return and risk than P^*_{MEF} in Figure 1.4. Similarly, if the investor had a different set of input expectations, the optimal portfolio would be different.

At this point in our discussion, a natural question is how to estimate an investor's utility function so that the indifference curves can be determined. Unfortunately, there is little guidance about how to construct one. In general,

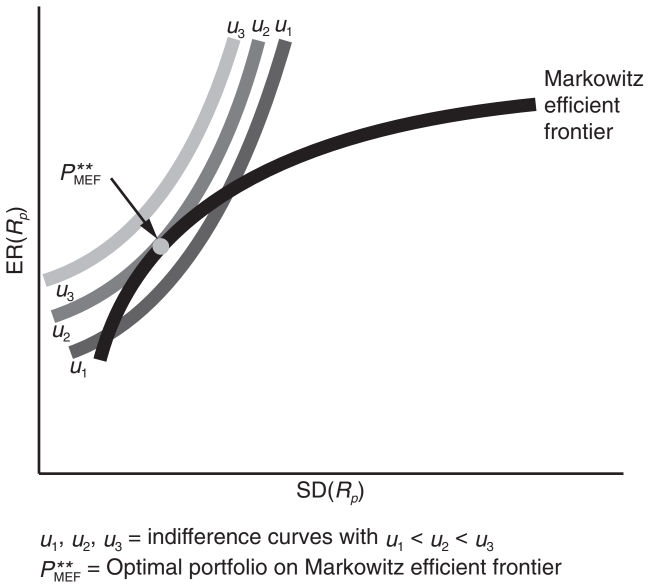

Figure 1.5 Selection of Optimal Portfolio with Different Indifference Curves (Utility Function)

economists have not been successful in estimating utility functions.

The inability to estimate utility functions does not mean that the theory is flawed. What it does mean is that once an investor constructs the efficient frontier, the investor will subjectively determine which efficient portfolio is appropriate given his or her tolerance to risk.

INDEX MODEL'S APPROXIMATIONS TO THE COVARIANCE STRUCTURE

The inputs to mean-variance analysis include expected returns, variance of returns, and either covariance or correlation of returns between each pair of securities. For example, an analysis that allows 200 securities as possible candidates for portfolio selection requires 200 expected returns, 200 variances of return, and 19,900 correlations or covariances. An investment team tracking 200 securities may reasonably be expected to summarize their analyses in terms of 200 means and variances, but it is clearly unreasonable for them to produce 19,900 carefully considered correlations or covariances.

It was clear to Markowitz that some kind of model of covariance structure was needed for the practical application of normative analysis to large portfolios. He did little more than point out the problem and suggest some possible models of covariance for research.

One model Markowitz proposed to explain the correlation structure among security returns assumed that the return on the ith security depends on an "underlying factor, the general prosperity of the market as expressed by some index." Mathematically, the relationship is expressed as follows (Markowitz, 1959, pp. 96–101):

$$r_i = \alpha_i + \beta_i F + u_i \tag{1.12}$$

where

r_i = the return on security i
F = value of some index
u_i = error term

(Note that Markowitz (1959) used the notation I in proposing the model given by equation (1.12).) The expected value of u_i is zero and u_i is uncorrelated with F and every other u_j.

The parameters α_i and β_i are parameters to be estimated. When measured using regression analysis, β_i is the ratio of the covariance of asset i's return and F to the variance of F.

Markowitz further suggested that the relationship need not be linear and that there could be several underlying factors.

Single-Index Market Model

Sharpe (1963) tested equation (1.12) as an explanation of how security returns tend to go up and down together with general market index, F. For the index in the market model he used a market index for F. Specifically, Sharpe estimated using regression analysis the following model:

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_{it} \tag{1.13}$$

where

r_{it} = return on asset i over the period t
r_{mt} = return on the market index over the period t
α_i = a term that represents the nonmarket component of the return on asset i
β_i = the ratio of the covariance of the return of asset i and the return of the market index to the variance of the return of the market index
u_{it} = a zero mean random error term

The model given by equation (1.13) is called the single index market model or simply the market model. It is important to note that when Markowitz discussed the possibility of using equation (1.12) to estimate the covariance structure, the index he suggested was not required to be a market index.

Suppose that the Dow Jones Wilshire 5000 is used to represent the market index, then for a portfolio of G assets regression analysis is used to estimate the values of the βs and αs. The *beta* of the portfolio (β_p), is simply a weighted average of the computed betas of the individual assets (β_i), where the weights are the percentage of the market value of the individual assets relative to the total market value of the portfolio. That is,

$$\beta_p = \sum_{i=1}^{G} w_i \beta_i$$

So, for example, the beta for a portfolio comprised of the following:

Company	Weight	β
General Electric	20%	1.24
McGraw-Hill	25%	0.86
IBM	15%	1.22
General Motors	10%	1.11
Xerox	30%	1.27

would have the following beta:

$$\begin{aligned}\text{Portfolio beta} &= 20\%(1.24) + 25\%(0.86) + 15\%(1.22)\\ &\quad + 10\%(1.11) + 30\%(1.27)\\ &= 1.14\end{aligned}$$

Multi-Index Market Models

Sharpe concluded that equation (1.12) was as complex a covariance as seemed to be needed. This conclusion was supported by research of Cohen and Pogue (1967). King (1966) found strong evidence for industry factors in addition to the marketwide factor. Rosenberg (1974) found other sources of risk beyond a marketwide factor and industry factor.

One alternative approach to full mean-variance analysis is the use of these multi-index or factor models to obtain the covariance structure.

SUMMARY

In this chapter we have introduced portfolio theory. Developed by Markowitz, this theory explains how investors should construct efficient portfolios and select the best or optimal portfolio from among all efficient portfolios. The theory differs from previous approaches to portfolio selection in that he demonstrated how the key parameters should be measured. These parameters include the risk and the expected return for an individual asset and a portfolio of assets. Moreover, the concept of diversifying a portfolio, the goal of which is to reduce a portfolio's risk without sacrificing expected return, can be cast in terms of these key parameters plus the covariance or correlation between assets. All these parameters are estimated from historical data and other sources of information and draw from concepts in probability and statistical theory.

A portfolio's expected return is simply a weighted average of the expected return of each asset in the portfolio. The weight assigned to each asset is the market value of the asset in the portfolio relative to the total market value of the portfolio. The variance or the standard deviation of an asset's returns measures its risk. Unlike the portfolio's expected return, a portfolio's risk is not a simple weighting of the standard deviation of the individual assets in the portfolio. Rather, the covariance or correlation between the assets that comprise the portfolio affects the portfolio risk. The lower the correlation, the smaller the risk of the portfolio.

Markowitz has set forth the theory for the construction of an efficient portfolio, which has come to be called a efficient portfolio—a portfolio that has the highest expected return of all feasible portfolios with the same level of risk. The collection of all efficient portfolios is called the efficient set of portfolios or the efficient frontier.

The optimal portfolio is the one that maximizes an investor's preferences with respect to return and risk. An investor's preference is described by a utility function, which can be represented graphically by a set of indifference curves. The utility function shows how much an investor is willing to trade off between expected return and risk. The optimal portfolio is the one where an indifference curve is tangent to the efficient frontier.

REFERENCES

Rosenberg, B. (1974). Extra-market components of covariance in security returns. *Journal of Financial and Quantitative Analysis* 19, 2: 23–274.

Cohen, K. J., and Pogue, G. A. (1967). An empirical evaluation of alternative portfolio selection models. *Journal of Business* 40, 2: 166–193.

King, B. F. (1966). Market and industry factors in stock price behavior. *Journal of Business* 39, 1 (Part 2: Supplement on Security Prices): 139–190.

Leavens, D. H. (1945). Diversification of investments. *Trusts and Estates* 80: 469–473.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance* 7, 1: 77–91

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Cowles Foundation Monograph 16, New York: John Wiley & Sons.

Ortobelli, S., Rachev, S. T., Stoyanov, S., Fabozzi, F. J., and Biglova, A. (2005). The proper use of risk measures in portfolio theory. *International Journal of Theoretical and Applied Finance* 8, 8: 1–27.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science* 9, 2: 277–293.

von Neumann, J., and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.