Chapter 2

Mean-Variance Analysis

The preceding chapter studied an investor's choice between a risk-free asset and a single risky asset. This chapter adds realism by giving the investor the opportunity to choose among multiple risky assets. As a University of Chicago graduate student, Harry Markowitz, wrote a path-breaking article on this topic (Markowitz 1952).[1] Markowitz's insight was to recognize that, in allocating wealth among various risky assets, a risk-averse investor should focus on the expectation and the risk of her combined portfolio's return, a return that is affected by the individual assets' diversification possibilities. Because of diversification, the attractiveness of a particular asset when held in a portfolio can differ from its appeal when it is the sole asset held by an investor.

Markowitz proxied the risk of a portfolio's return by the variance of its return. Of course, the variance of an investor's total portfolio return depends on the return variances of the individual assets included in the portfolio. But portfolio return variance also depends on the covariances of the individual assets'

[1] His work on portfolio theory, of which this article was the beginning, won him a share of the Nobel prize in economics in 1990. Initially, the importance of his work was not widely-recognized. Milton Friedman, a member of Markowitz's doctoral dissertation committee and who also became a Nobel laureate, questioned whether the work met the requirements for an economics Ph.D. See Bernstein (Bernstein 1992).

returns. Hence, in selecting an optimal portfolio, the investor needs to consider how the co-movement of individual assets' returns affects diversification possibilities.

A rational investor would want to choose a portfolio of assets that efficiently trades off higher expected return for lower variance of return. Interestingly, not all portfolios that an investor can create are efficient in this sense. Given the expected returns and covariances of returns on individual assets, Markowitz solved the investor's problem of constructing an efficient portfolio. His work has had an enormous impact on the theory and practice of portfolio management and asset pricing.

Intuitively, it makes sense that investors would want their wealth to earn a high average return with as little variance as possible. However, in general, an individual who maximizes expected utility may care about moments of the distribution of wealth in addition to its mean and variance.[2] Though Markowitz's mean - variance analysis fails to consider the effects of these other moments, in later chapters of this book we will see that his model's insights can be generalized to more complicated settings.

The next section outlines the assumptions on investor preferences and the distribution of asset returns that would allow us to simplify the investor's portfolio choice problem to one that considers only the mean and variance of portfolio returns. We then analyze a risk-averse investor's preferences by showing that he has indifference curves that imply a trade-off of expected return for variance.

[2] For example, expected utility can depend on the skewness (the third moment) of the return on wealth. The observation that some people purchase lottery tickets, even though these investments have a negative expected rate of return, suggests that their utility is enhanced by positive skewness. Alan Kraus and Robert Litzenberger (Kraus and Litzenberger 1976) developed a single-period portfolio selection and asset pricing model that extends Markowitz's analysis to consider investors who have a preference for skewness. Their results generalize Markowitz's model, but his fundamental insights are unchanged. For simplicity, this chapter focuses on the orginal Markowitz framework. Recent empirical work by Campbell Harvey and Akhtar Siddique (Harvey and Siddique 2000) examines the effect of skewness on asset pricing.

Subsequently, we show how a portfolio can be allocated among a given set of risky assets in a mean-variance efficient manner. We solve for the *efficient frontier*, defined as the set of portfolios that maximize expected returns for a given variance of returns, and show that any two frontier portfolios can be combined to create a third. In addition, we show that a fundamental simplification to the investor's portfolio choice problem results when one of the assets included in the investor's choice set is a risk-free asset. The final section of this chapter applies mean-variance analysis to a problem of selecting securities to hedge the risk of commodity prices. This application is an example of how modern portfolio analysis has influenced the practice of risk management.

2.1 Assumptions on Preferences and Asset Returns

Suppose an expected utility maximizing individual invests her beginning-of-period wealth, W_0, in a particular portfolio of assets. Let $\widetilde{R}_p$ be the random return on this portfolio, so that the individual's end-of-period wealth is $\tilde{W} = W_0\widetilde{R}_p$.[3] Denote this individual's end-of-period utility by $U(\tilde{W})$. Given W_0, for notational simplicity we write $U(\tilde{W}) = U\left(W_0\widetilde{R}_p\right)$ as just $U(\widetilde{R}_p)$, because $\tilde{W}$ is completely determined by $\widetilde{R}_p$.

Let us express $U(\widetilde{R}_p)$ by expanding it in a Taylor series around the mean of $\widetilde{R}_p$, denoted as $E[\widetilde{R}_p]$. Let $U'(\cdot)$, $U''(\cdot)$, and $U^{(n)}(\cdot)$ denote the first, second, and n^{th} derivatives of the utility function:

[3] Thus, $\widetilde{R}_p$ is defined as one plus the rate of return on the portfolio.

$$\begin{aligned} U(\widetilde{R}_p) &= U\left(E[\widetilde{R}_p]\right) + \left(\widetilde{R}_p - E[\widetilde{R}_p]\right) U'\left(E[\widetilde{R}_p]\right) \\ &\quad + \tfrac{1}{2}\left(\widetilde{R}_p - E[\widetilde{R}_p]\right)^2 U''\left(E[\widetilde{R}_p]\right) + ... \\ &\quad + \tfrac{1}{n!}\left(\widetilde{R}_p - E[\widetilde{R}_p]\right)^n U^{(n)}\left(E[\widetilde{R}_p]\right) + ... \end{aligned} \tag{2.1}$$

Now let us investigate the conditions that would make this individual's expected utility depend only on the mean and variance of the portfolio return. We first analyze the restrictions on the form of utility, and then the restrictions on the distribution of asset returns, that would produce this result.

Note that if the utility function is quadratic, so that all derivatives of order 3 and higher are equal to zero ($U^{(n)} = 0, \forall\ n \geq 3$), then the individual's expected utility is

$$\begin{aligned} E\left[U(\widetilde{R}_p)\right] &= U\left(E[\widetilde{R}_p]\right) + \tfrac{1}{2}E\left[\left(\widetilde{R}_p - E[\widetilde{R}_p]\right)^2\right] U''\left(E[\widetilde{R}_p]\right) \\ &= U\left(E[\widetilde{R}_p]\right) + \tfrac{1}{2}V[\widetilde{R}_p]U''\left(E[\widetilde{R}_p]\right) \end{aligned} \tag{2.2}$$

where $V[\widetilde{R}_p]$ is the variance of the return on the portfolio.[4] Therefore, for any probability distribution of the portfolio return, $\widetilde{R}_p$, quadratic utility leads to expected utility that depends only on the mean and variance of $\widetilde{R}_p$.

Next, suppose that utility is not quadratic, but any general increasing, concave form. Are there particular probability distributions for portfolio returns that make expected utility, again, depend only on the portfolio return's mean

[4] The expected value of the second term in the Taylor series, $E\left[\left(\widetilde{R}_p - E[\widetilde{R}_p]\right) U'\left(E[\widetilde{R}_p]\right)\right]$, equals zero.

and variance? Such distributions would need to be fully determined by their means and variances, that is, they must be two-parameter distributions whereby higher order moments could be expressed in terms of the first two moments (mean and variance). Many distributions, such as the gamma, normal, and lognormal, satisfy this criterion. But in the context of an investor's portfolio selection problem, such distributions need to satisfy another condition. Since an individual is able to choose which assets to combine into a portfolio, all portfolios created from a combination of individual assets or other portfolios must have distributions that continue to be determined by their means and variances. In other words, we need a distribution where if the individual assets' return distributions depend on just mean and variance, then the return on a combination (portfolio) of these assets has a distribution that depends on just mean and variance. The only distributions that satisfy this "additivity" restriction is the stable family of distributions, and, among this family, the only distribution that has finite variance is the normal (Gaussian) distribution. A portfolio (sum) of assets whose returns are multivariate normally distributed also has a return that is normally distributed.

To verify that expected utility depends only on the portfolio return's mean and variance when this return is normally distributed, note that the third, fourth, and all higher central moments of the normal distribution are either zero or a function of the variance: $E\left[\left(\widetilde{R}_p - E[\widetilde{R}_p]\right)^n\right] = 0$, for n odd, and $E\left[\left(\widetilde{R}_p - E[\widetilde{R}_p]\right)^n\right] = \frac{n!}{(n/2)!}\left(\frac{1}{2}V[\widetilde{R}_p]\right)^{n/2}$, for n even. Therefore, for this case the individual's expected utility equals

$$\begin{aligned} E\left[U(\widetilde{R}_p)\right] &= U\left(E[\widetilde{R}_p]\right) + \tfrac{1}{2}V[\widetilde{R}_p]U''\left(E[\widetilde{R}_p]\right) + 0 + \frac{1}{8}\left(V[\widetilde{R}_p]\right)^2 U''''\left(E[\widetilde{R}_p]\right) \\ &\quad +0 + ... + \frac{1}{(n/2)!}\left(\frac{1}{2}V[\widetilde{R}_p]\right)^{n/2} U^{(n)}\left(E[\widetilde{R}_p]\right) + ... \end{aligned} \tag{2.3}$$

which depends only on the mean and variance of the portfolio return.

In summary, restricting utility to be quadratic or restricting the distribution of asset returns to be normal allows us to write $E\left[U(\widetilde{R}_p)\right]$ as a function of only the mean, $E[\widetilde{R}_p]$, and the variance, $V[\widetilde{R}_p]$, of the portfolio return. Are either of these assumptions realistic? If not, it may be unjustified to suppose that only the first two moments of the portfolio return distribution matter to the individual investor.

The assumption of quadratic utility clearly is problematic. As mentioned earlier, quadratic utility displays negative marginal utility for levels of wealth greater than the "bliss point," and it has the unattractive characteristic of increasing absolute risk aversion. There are also difficulties with the assumption of normally distributed asset returns. When asset returns measured over any finite time period are normally distributed, there exists the possibility that their end of period values could be negative since realizations from the normal distribution have no lower (or upper) bound. This is an unrealistic description of returns for many assets such as stocks and bonds because, being limited-liability assets, their minimum value is non-negative.[5]

As will be demonstrated in Chapter 12, the assumption of normal returns can be modified if we generalize the model to have multiple periods and assume that asset rates of return follow continuous-time stochastic processes. In that context, one can assume that assets' rates of return are *instantaneously* normally distributed, which implies that if their means and variances are constant over infinitesimal intervals, then over any finite interval asset values are lognormally distributed. This turns out to be a better way of modeling limited liability assets because the lognormal distribution bounds these assets' values to be no less than zero. When we later study continuous-time, multi-period models, we shall

[5] A related problem is that many standard utility functions, such as constant relative risk aversion, are not defined for negative values of portfolio wealth.

see that the results derived here assuming a single-period, discrete-time model continue to hold, under particular conditions, in the more realistic multi-period context. Moreover, in more complex multiperiod models that permit assets to have time-varying return distributions, we will show that optimal portfolio choices are straightforward generalizations of the mean-variance results derived in this chapter.

2.2 Investor Indifference Relations

Therefore, let us proceed by assuming that the individual's utility function, U, is a general concave utility function and that individual asset returns are normally distributed. Hence, a portfolio of these assets have a return $\tilde{R}_p$ that is normally distributed with probability density function $f(R; \bar{R}_p, \sigma_p^2)$, where we use the short-hand notation $\bar{R}_p \equiv E[\tilde{R}_p]$ and $\sigma_p^2 \equiv V[\tilde{R}_p]$. In this section we analyze an investor's "tastes," that is, the investor's risk - expected return preferences when utility depends on the mean (expected return) and variance (risk) of the return on wealth. The following section analyzes investment "technologies" represented by the combinations of portfolio risk and expected return that can be created from different portfolios of individual assets. Historically, mean - variance analysis has been illustrated graphically, and we will follow that convention while also providing analytic results.

Note that an investor's expected utility can then be written

$$E\left[U\left(\widetilde{R}_p\right)\right] = \int_{-\infty}^{\infty} U(R) f(R; \bar{R}_p, \sigma_p^2) dR \tag{2.4}$$

To gain insight regarding this investor's preferences over portfolio risk and expected return, we wish to determine the characteristics of this individual's indifference curves in portfolio mean-variance space. An indifference curve represents

the combinations of portfolio mean and variance that would give the individual the same level of expected utility.[6] To understand this relation, let us begin by defining $\widetilde{x} \equiv \frac{\tilde{R}_p - \bar{R}_p}{\sigma_p}$. Then

$$E\left[U\left(\widetilde{R}_p\right)\right] = \int_{-\infty}^{\infty} U(\bar{R}_p + x\sigma_p)n(x)dx \tag{2.5}$$

where $n(x) \equiv f(x; 0, 1)$ is the standardized normal probability density function, that is, the normal density having a zero mean and unit variance. Now consider how expected utility varies with changes in the mean and variance of the return on wealth. Taking the partial derivative with respect to $\bar{R}_p$:

$$\frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \bar{R}_p} = \int_{-\infty}^{\infty} U'n(x)dx > 0 \tag{2.6}$$

since U' is always greater than zero. Next, take the partial derivative of (2.5) with respect to σ_p^2:

$$\frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \sigma_p^2} = \frac{1}{2\sigma_p}\frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \sigma_p} = \frac{1}{2\sigma_p}\int_{-\infty}^{\infty} U'xn(x)dx \tag{2.7}$$

While U' is always positive, x ranges between $-\infty$ and $+\infty$. Because x has a standard normal distribution, which is symmetric, for each positive realization there is a corresponding negative realization with the same probability density. For example, take the positive and negative pair $+x_i$ and $-x_i$. Then $n(+x_i) = n(-x_i)$. Comparing the integrand of equation (2.7) for equal absolute realizations of x, we can show

[6] Indifference curves are used in microeconomics to analyze an individual's choice of consuming different quantities of multiple goods. For example, if utility, $u(x, y)$, derives from consuming two goods, with x being the quantity of good X consumed and y being the quantity of good Y consumed, then an indifference curve is the locus of points in X, Y space that gives a constant level of utilitiy. That is, combinations of goods X and Y for which the individual would be indifferent between consuming. Mathematically, these combinations are represented as the points (x, y) such that $u(x, y) = \overline{U}$, a constant. In this section, we employ a similar concept but where expected utility depends on the mean and variance of the return on wealth.

$$
\begin{aligned}
& U'(\bar{R}_p + x_i\sigma_p)x_i n(x_i) + U'(\bar{R}_p - x_i\sigma_p)(-x_i)n(-x_i) \qquad (2.8)\\
= \;& U'(\bar{R}_p + x_i\sigma_p)x_i n(x_i) - U'(\bar{R}_p - x_i\sigma_p)x_i n(x_i)\\
= \;& x_i n(x_i)\left[U'(\bar{R}_p + x_i\sigma_p) - U'(\bar{R}_p - x_i\sigma_p)\right] < 0
\end{aligned}
$$

because

$$U'(\bar{R} + x_i\sigma_p) < U'(\bar{R} - x_i\sigma_p) \qquad (2.9)$$

due to the assumed concavity of U, that is, the individual is risk averse so that $U'' < 0$. Thus, comparing $U' x_i n(x_i)$ for each positive and negative pair, we conclude that

$$\frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \sigma_p^2} = \frac{1}{2\sigma_p}\int_{-\infty}^{\infty} U' x n(x)dx < 0 \qquad (2.10)$$

which is the intuitive result that higher portfolio variance, without higher portfolio expected return, reduces a risk-averse individual's expected utility.

Finally, an indifference curve is the combinations of portfolio mean and variance that leaves expected utility unchanged. In other words, it is combinations of $\left(\bar{R}_p, \sigma_p^2\right)$ that satisfy the equation $E\left[U\left(\widetilde{R}_p\right)\right] = \overline{U}$, a constant. Higher levels of $\overline{U}$ denote different indifference curves providing a greater level of utility. If we totally differentiate this equation, we obtain:

$$dE\left[U\left(\widetilde{R}_p\right)\right] = \frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \sigma_p^2} d\sigma_p^2 + \frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \bar{R}_p} d\bar{R}_p = 0 \qquad (2.11)$$

which, based on our previous results, tells us that each indifference curve is positively sloped in $\left(\bar{R}_p, \sigma_p^2\right)$ space:

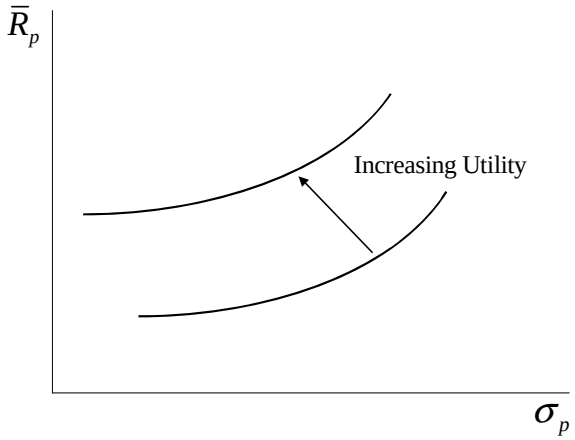


Figure 2.1: Indifference Curves

$$\frac{d\bar{R}_p}{d\sigma_p^2} = -\frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \sigma_p^2} / \frac{\partial E\left[U\left(\widetilde{R}_p\right)\right]}{\partial \bar{R}_p} > 0 \tag{2.12}$$

Thus, the indifference curve's slope in (2.12) quantifies the extent to which the individual requires a higher portfolio mean for accepting a higher portfolio variance.

Indifference curves are typically drawn in mean - standard deviation space, rather than mean - variance space, because standard deviations of returns are in the same unit of measurement as returns or interest rates (rather than squared returns). Figure 2.1 illustrates such a graph, where the arrow indicates an increase in the utility level, $\overline{U}$.[7] It is left as an end-of-chapter exercise to show that the curves are convex due to the assumed concavity of the utility function.

Having analyzed an investor's preferences over different combinations of

[7] Clearly, these indifference curves cannot "cross" (intersect) because we showed that utility is always increasing in expected portfolio return for a given level of portfolio standard deviation.

portfolio means and standard deviations (or variances), let us consider next what portfolio means and standard deviations are possible given the available distributions of returns for individual assets.

2.3 The Efficient Frontier

The individual's optimal choice of portfolio mean and variance is determined by the point where one of these indifference curves is tangent to the set of means and standard deviations for all feasible portfolios, what we might describe as the "risk versus expected return investment opportunity set." This set represents all possible ways of combining various *individual* assets to generate alternative combinations of *portfolio* mean and variance (or standard deviation). This set includes inefficient portfolios (those in the interior of the opportunity set) as well as efficient portfolios (those on the "frontier" of the set). Efficient portfolios are those that make best use of the benefits of diversification. As we shall later prove, efficient portfolios have the attractive characteristic that any two efficient portfolios can be used to create any other efficient portfolio.

2.3.1 A Simple Example

To illustrate the effects of diversification, consider the following simple example. Suppose there are two assets, assets A and B, that have expected returns $\bar{R}_A$ and $\bar{R}_B$ and variances of σ_A^2 and σ_B^2, respectively. Further the correlation between their returns is given by ρ. Let us assume that $\bar{R}_A < \bar{R}_B$ but $\sigma_A^2 < \sigma_B^2$. Now form a portfolio with a proportion ω invested in asset A and a proportion $1-\omega$ invested in asset B.[8] The expected return on this portfolio is

[8] It is assumed that ω can be any real number. A $\omega < 0$ indicates a short position in asset A while $\omega > 1$ indicates a short position in asset B.

$$\overline{R}_p = \omega \bar{R}_A + (1-\omega)\bar{R}_B \tag{2.13}$$

The expected return of a portfolio is a simple weighted average of the expected returns of the individual financial assets. Expected returns are not fundamentally transformed by combining individual assets into a portfolio. The standard deviation of the return on the portfolio is

$$\sigma_p = \left[\omega^2\sigma_A^2 + 2\omega(1-\omega)\sigma_A\sigma_B\rho + (1-\omega)^2\sigma_B^2\right]^{\frac{1}{2}} \tag{2.14}$$

In general, portfolio risk, as measured by the portfolio's return standard deviation, is a nonlinear function of the individual assets' variances and covariances. Thus, risk is altered in a relatively complex way when individual assets are combined in a portfolio.

Let us consider some special cases regarding the correlation between the two assets. Suppose $\rho = 1$, so that the two assets are perfectly positively correlated. Then the portfolio standard deviation equals

$$\begin{aligned} \sigma_p &= \left[\omega^2\sigma_A^2 + 2\omega(1-\omega)\sigma_A\sigma_B + (1-\omega)^2\sigma_B^2\right]^{\frac{1}{2}} \\ &= |\omega\sigma_A + (1-\omega)\sigma_B| \end{aligned} \tag{2.15}$$

which is a simple weighted average of the individual assets' standard deviations. Solving (2.15) for asset A's portfolio proportion gives $\omega = (\sigma_B \pm \sigma_p)/(\sigma_B - \sigma_A)$. Then, by substituting for ω in (2.13), we obtain

$$\begin{aligned} \bar{R}_p &= \bar{R}_B + \left[\frac{\pm\sigma_p - \sigma_B}{\sigma_B - \sigma_A}\right](\bar{R}_B - \bar{R}_A) \\ &= \frac{\sigma_B\bar{R}_A - \sigma_A\bar{R}_B}{\sigma_B - \sigma_A} \pm \frac{\bar{R}_B - \bar{R}_A}{\sigma_B - \sigma_A}\sigma_p \end{aligned} \tag{2.16}$$

Thus, the relationship between portfolio risk and expected return are two straight lines in σ_p, $\bar{R}_p$ space. They have the same intercept of $\left(\sigma_B\bar{R}_A - \sigma_A\bar{R}_B\right)/\left(\sigma_B - \sigma_A\right)$ and have slopes of the same magnitude but opposite signs. The positively sloped line goes through the points $(\sigma_A, \bar{R}_A)$ and $(\sigma_B, \bar{R}_B)$ when $\omega = 1$ and $\omega = 0$, respectively. When $\omega = \sigma_B/\left(\sigma_B - \sigma_A\right) > 1$, indicating a short position in asset B, we see from (2.15) that all portfolio risk is eliminated ($\sigma_p = 0$). Figure 2.2 provides a graphical illustration of these relationships.

Next, suppose $\rho = -1$, so that the assets are perfectly negatively correlated. Then

$$\begin{aligned} \sigma_p &= \left[(\omega\sigma_A - (1-\omega)\sigma_B)^2\right]^{\frac{1}{2}} \\ &= |\omega\sigma_A - (1-\omega)\sigma_B| \end{aligned} \tag{2.17}$$

In a manner similar to the previous case, we can show that

$$\bar{R}_p = \frac{\sigma_A\bar{R}_B + \sigma_B\bar{R}_A}{\sigma_A + \sigma_B} \pm \frac{\bar{R}_B - \bar{R}_A}{\sigma_A + \sigma_B}\sigma_p \tag{2.18}$$

which, again, represents two straight lines in σ_p, $\bar{R}_p$ space. The intercept at $\sigma_p = 0$ is given by $\omega = \sigma_B/\left(\sigma_A + \sigma_B\right)$, so that all portfolio risk is eliminated with positive amounts invested in each asset. Furthermore, the negatively sloped line goes through the point $(\sigma_A, \bar{R}_A)$ when $\omega = 1$, while the positively sloped line goes through the point $(\sigma_B, \bar{R}_B)$ when $\omega = 0$. Figure 2.2 summarizes these

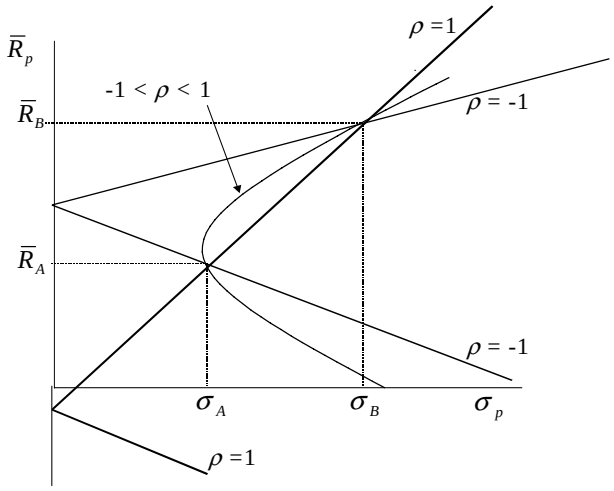


Figure 2.2: Efficient Frontier with Two Risky Assets

risk - expected return constraints.

For either the $\rho = 1$ or $\rho = -1$ case, an investor would always choose a portfolio represented by the positively sloped lines because they give the highest average portfolio return for any given level of portfolio risk. These lines represent the so-called *efficient portfolio frontier*. The exact portfolio chosen by the individual would be where her indifference curve is tangent to the frontier.

When correlation between the assets is imperfect ($-1 < \rho < 1$), the relationship between portfolio expected return and standard deviation is not linear, but, as illustrated in Figure 2.2, is hyperbolic. In this case, it is no longer possible to create a riskless portfolio, so that the portfolio having minimum standard deviation is one where $\sigma_p > 0$. We now set out to prove these assertions for the general case of n assets.

2.3.2 Mathematics of the Efficient Frontier

Robert C. Merton (Merton 1972) provided an analytical solution to the following portfolio choice problem: Given the expected returns and the matrix of covariances of returns for n individual assets, find the set of portfolio weights that minimizes the variance of the portfolio for each feasible portfolio expected return. The locus of these points in portfolio expected return - variance space is the portfolio frontier. This section presents the derivation of Merton's solution. We begin by specifying the problem's notation and assumptions.

Let $\bar{R} = (\bar{R}_1 \bar{R}_2 ... \bar{R}_n)'$ be an $n \times 1$ vector of the expected returns of the n assets. Also let V be the $n \times n$ covariance matrix of the returns on the n assets. V is assumed to be of full rank.[9] Since it is a covariance matrix, it is also symmetric and positive definite. Next, let $\omega = (\omega_1 \ \omega_2 ... \omega_n)'$ be an $n \times 1$ vector of portfolio proportions, such that ω_i is the proportion of total portfolio wealth invested in the i^{th} asset. It follows that the expected return on the portfolio is given by

$$\bar{R}_p = \omega' \bar{R} \tag{2.19}$$

and the variance of the portfolio return is given by

$$\sigma_p^2 = \omega' V \omega \tag{2.20}$$

The constraint that the portfolio proportions must sum to 1 can be written as $\omega' e = 1$ where e is defined to be an $n \times 1$ vector of ones.

The problem of finding the portfolio frontier now can be stated as a quadratic optimization exercise: minimize the portfolio's variance subject to the con-

[9] This implies that there are no redundant assets among the n assets. An asset would be redundant if its return was an exact linear combination of the the returns on other assets. If such an asset exists, it can be ignored, since its availability does not affect the efficient portfolio frontier.

straints that the portfolio's expected return equals $\overline{R}_p$ and the portfolio's weights sum to one.[10]

$$\min_{\omega} \tfrac{1}{2}\omega' V \omega + \lambda \left[\overline{R}_p - \omega' \bar{R}\right] + \gamma[1 - \omega' e] \tag{2.21}$$

The first order conditions with respect to ω and the two Lagrange multipliers, λ and γ, are:

$$V\omega - \lambda \bar{R} - \gamma e = 0 \tag{2.22}$$

$$\overline{R}_p - \omega' \bar{R} = 0 \tag{2.23}$$

$$1 - \omega' e = 0 \tag{2.24}$$

Solving (2.22), the optimal portfolio weights satisfy

$$\omega^* = \lambda V^{-1} \bar{R} + \gamma V^{-1} e \tag{2.25}$$

Pre-multiplying equation (2.25) by $\bar{R}'$, we have

$$\overline{R}_p = \bar{R}' \omega^* = \lambda \bar{R}' V^{-1} \bar{R} + \gamma \bar{R}' V^{-1} e \tag{2.26}$$

Pre-multiplying equation (2.25) by e' we have

$$1 = e' \omega^* = \lambda e' V^{-1} \bar{R} + \gamma e' V^{-1} e \tag{2.27}$$

Equations (2.26) and (2.27) are two linear equations in two unknowns, λ and γ.

[10] In (2.21), the problem actually minimizes one-half the portfolio variance to avoid carrying an extra "2" in the first order condition (2.22). The solution is the same as minimizing the total variance and only changes the scale of the Lagrange multipliers.

The solution is

$$\lambda = \frac{\delta \overline{R}_p - \alpha}{\varsigma\delta - \alpha^2} \tag{2.28}$$

$$\gamma = \frac{\varsigma - \alpha \overline{R}_p}{\varsigma\delta - \alpha^2} \tag{2.29}$$

where $\alpha \equiv \bar{R}' V^{-1} e = e' V^{-1} \bar{R}$, $\varsigma \equiv \bar{R}' V^{-1} \bar{R}$, and $\delta \equiv e' V^{-1} e$ are scalars. Note that the denominators of λ and γ, given by $\varsigma\delta - \alpha^2$, are guaranteed to be positive when V is of full rank.[11] Substituting for λ and γ in equation (2.25), we have

$$\omega^* = \frac{\delta \overline{R}_p - \alpha}{\varsigma\delta - \alpha^2} V^{-1} \bar{R} + \frac{\varsigma - \alpha \overline{R}_p}{\varsigma\delta - \alpha^2} V^{-1} e \tag{2.30}$$

Collecting terms in $\bar{R}_p$ gives

$$\omega^* = a + b \overline{R}_p \tag{2.31}$$

where $a \equiv \dfrac{\varsigma V^{-1} e - \alpha V^{-1} \bar{R}}{\varsigma\delta - \alpha^2}$ and $b \equiv \dfrac{\delta V^{-1} \bar{R} - \alpha V^{-1} e}{\varsigma\delta - \alpha^2}$.

Equation (2.31) is both a necessary and sufficient condition for a frontier portfolio. Given $\overline{R}_p$, a portfolio must have weights satisfying (2.31) to minimize its return variance.

Having found the optimal portfolio weights for a given $\overline{R}_p$, the variance of the frontier portfolio is

[11] To see this, note that since V is positive definite, so is V^{-1}. Therefore the quadratic form $\left(\alpha\overline{R} - \varsigma e\right) V^{-1} \left(\alpha\overline{R} - \varsigma e\right) = \alpha^2\varsigma - 2\alpha^2\varsigma + \varsigma^2\delta = \varsigma\left(\varsigma\delta - \alpha^2\right)$ is positive. But since $\varsigma \equiv \overline{R} V^{-1} \overline{R}$ is a positive quadratic form, then $\left(\varsigma\delta - \alpha^2\right)$ must also be positive.

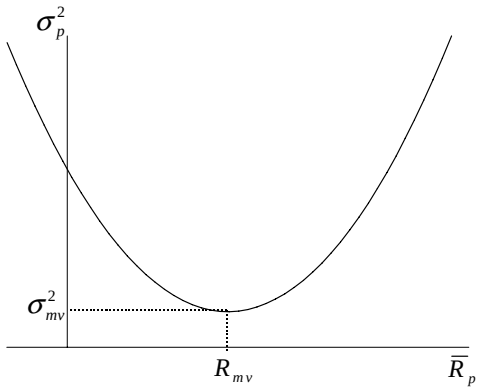


Figure 2.3: Frontier Portfolios

$$\begin{aligned} \sigma_p^2 &= \omega^{*\prime} V \omega^* = (a + b\overline{R}_p)' V (a + b\overline{R}_p) \\ &= \frac{\delta \overline{R}_p^2 - 2\alpha \overline{R}_p + \varsigma}{\varsigma\delta - \alpha^2} \\ &= \frac{1}{\delta} + \frac{\delta \left(\overline{R}_p - \frac{\alpha}{\delta}\right)^2}{\varsigma\delta - \alpha^2} \end{aligned} \tag{2.32}$$

where the second line in equation (2.32) results from substituting in the definitions of a and b and simplifying the resulting expression. Equation (2.32) is a parabola in σ_p^2, $\overline{R}_p$ space and is graphed in Figure 2.3. From the third line in equation (2.32), it is obvious that the unique minimum is at the point $\overline{R}_p = R_{mv} \equiv \frac{\alpha}{\delta}$, which corresponds to a global minimum variance of $\sigma_{mv}^2 \equiv \frac{1}{\delta}$. Substituting $\overline{R}_p = \frac{\alpha}{\delta}$ into 2.30 shows that this minimum variance portfolio has weights $\omega_{mv} = \frac{1}{\delta} V^{-1} e$.

Each point on the parabola in Figure 2.3 represents an investor's lowest possible portfolio variance given some target level of expected return, $\overline{R}_p$. However,

an investor whose utility is increasing in expected portfolio return and is decreasing in portfolio variance would never choose a portfolio having $\overline{R}_p < R_{mv}$, that is, points on the parabola to the left of R_{mv}. This is because the frontier portfolio's variance actually increases as the target expected return falls when $\overline{R}_p < R_{mv}$, making this target expected return region irrelevant to an optimizing investor. Hence, the *efficient* portfolio frontier is represented only by the region $\overline{R}_p \geq R_{mv}$.

Traditionally, portfolios satisfying (2.32) are graphed in σ_p, $\overline{R}_p$ space. Taking the square root of both sides of equation (2.32), σ_p becomes a hyperbolic function of $\overline{R}_p$. When this is graphed as in Figure 2.4 with $\overline{R}_p$ on the vertical axis and σ_p on the horizontal one, only the upper arc of the hyperbola is relevant because, as stated above, investors would not choose target levels of $\overline{R}_p < R_{mv}$. Differentiating (2.32), we can also see that the hyperbola's slope equals

$$\frac{\partial \overline{R}_p}{\partial \sigma_p} = \frac{\varsigma\delta - \alpha^2}{\delta\left(\overline{R}_p - \frac{\alpha}{\delta}\right)}\sigma_p \tag{2.33}$$

The upper arc asymptotes to the straight line $\overline{R}_p = R_{mv} + \sqrt{\frac{\varsigma\delta-\alpha^2}{\delta}}\sigma_p$, while the lower arc, representing inefficient frontier portfolios, asymptotes to the straight line $\overline{R}_p = R_{mv} - \sqrt{\frac{\varsigma\delta-\alpha^2}{\delta}}\sigma_p$.[12]

2.3.3 Portfolio Separation

We now state and prove a fundamental result:

> Every portfolio on the mean-variance frontier can be replicated by a combination of any two frontier portfolios; and an individual will be indifferent between choosing among the n financial assets, or

[12] To see that the slope of the hyperbola asymptotes to a magnitude of $\sqrt{(\varsigma\delta - \alpha^2)/\delta}$, use (2.32) to substitute for $\left(\overline{R}_p - \frac{\alpha}{\delta}\right)$ in (2.33) to obtain $\partial\overline{R}_p/\partial\sigma_p = \pm\sqrt{(\varsigma\delta - \alpha^2)}/\sqrt{\delta - 1/\sigma_p^2}$. Taking the limit of this expression as $\sigma_p \to \infty$ gives the desired result.

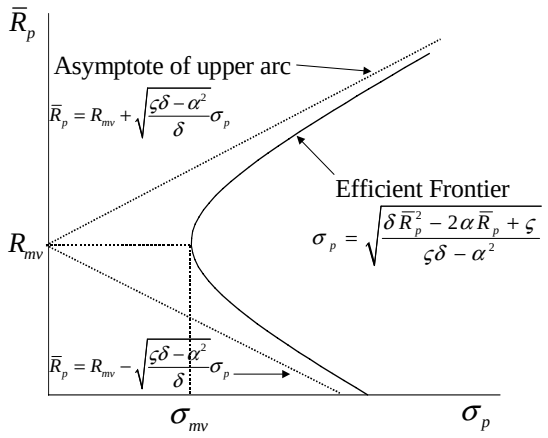


Figure 2.4: Efficient Frontier

choosing a combination of just two frontier portfolios.

This remarkable finding has an immediate practical implication. If all investors have the same beliefs regarding the distribution of asset returns, namely, returns are distributed $N\left(\overline{R}, V\right)$ and, therefore, the frontier is (2.32), then they can form their individually preferred frontier portfolios by trading in as little as two frontier portfolios. For example, if a security market offered two mutual funds, each invested in a different frontier portfolio, any mean-variance investor could replicate his optimal portfolio by appropriately dividing his wealth between only these two mutual funds.[13]

The proof of this separation result is as follows. Let $\bar{R}_{1p}$ and $\bar{R}_{2p}$ be the expected returns on any two distinct frontier portfolios. Let $\bar{R}_{3p}$ be the expected

[13] To form his preferred frontier portfolio, an investor may require a short position in one of the frontier mutual funds. Since short positions are not possible with typical open ended mutual funds, the better analogy would be that these funds are exchange-traded funds (ETFs) which do permit short positions.

return on a third frontier portfolio. Now consider investing a proportion of wealth, x, in the first frontier portfolio and the remainder, $(1-x)$, in the second frontier portfolio. Clearly, a value for x can be found that makes the expected return on this "composite" portfolio equal to that of the third frontier portfolio.[14]

$$\bar{R}_{3p} = x\bar{R}_{1p} + (1-x)\bar{R}_{2p} \tag{2.34}$$

In addition, because portfolios 1 and 2 are frontier portfolios, we can write their portfolio proportions as a linear function of their expected returns. Specifically, we have $\omega^1 = a + b\bar{R}_{1p}$ and $\omega^2 = a + b\bar{R}_{2p}$ where ω^i is the $n \times 1$ vector of optimal portfolio weights for frontier portfolio i. Now create a new portfolio with an $n \times 1$ vector of portfolio weights given by $x\omega^1 + (1-x)\omega^2$. The portfolio proportions of this new portfolio can be written as

$$\begin{aligned} x\omega^1 + (1-x)\omega^2 &= x(a + b\bar{R}_{1p}) + (1-x)(a + b\bar{R}_{2p}) \\ &= a + b(x\bar{R}_{1p} + (1-x)\bar{R}_{2p}) \\ &= a + b\bar{R}_{3p} = \omega^3 \end{aligned} \tag{2.35}$$

where, in the last line of (2.35) we have substituted in equation (2.34). Based on the portfolio weights of the composite portfolio, $x\omega^1 + (1-x)\omega^2$,equalling $a + b\bar{R}_{3p}$, which is the portfolio weights of the third frontier portfolio, ω^3, this composite portfolio equals the third frontier portfolio. Hence, any given efficient portfolio can be replicated by two frontier portfolios.

Portfolios on the mean-variance frontier have an additional property that will prove useful to the next section's analysis of portfolio choice when a riskless asset

[14] x may be any positive or negative number.

exists and also to understanding equilibrium asset pricing in Chapter 3. Except for the global minimum variance portfolio, ω_{mv}, for each frontier portfolio one can find another frontier portfolio with which its returns have zero covariance. That is, one can find pairs of frontier portfolios whose returns are orthogonal. To show this, note that the covariance between two frontier portfolios, w^1 and w^2, is

$$\begin{aligned} \omega^1 V \omega^2 &= (a + b\overline{R}_{1p})' V (a + b\overline{R}_{2p}) \\ &= \frac{1}{\delta} + \frac{\delta}{\varsigma\delta - \alpha^2}\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)\left(\overline{R}_{2p} - \frac{\alpha}{\delta}\right) \end{aligned} \tag{2.36}$$

Setting this equal to zero and solving for $\overline{R}_{2p}$, the expected return on the portfolio that has zero covariance with portfolio ω^1 is

$$\begin{aligned} \overline{R}_{2p} &= \frac{\alpha}{\delta} - \frac{\varsigma\delta - \alpha^2}{\delta^2\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)} \\ &= R_{mv} - \frac{\varsigma\delta - \alpha^2}{\delta^2\left(\overline{R}_{1p} - R_{mv}\right)} \end{aligned} \tag{2.37}$$

Note that if $\left(\overline{R}_{1p} - R_{mv}\right) > 0$ so that frontier portfolio ω^1 is efficient, then equation (2.37) indicates that $\overline{R}_{2p} < R_{mv}$, implying that frontier portfolio ω^2 must be inefficient. We can determine the relative locations of these zero covariance portfolios by noting that in σ_p, $\overline{R}_p$ space, a line tangent to the frontier at the point $\left(\sigma_{1p}, \overline{R}_{1p}\right)$ is of the form

$$\overline{R}_p = \overline{R}_0 + \frac{\partial \overline{R}_p}{\partial \sigma_p}\Big|_{\sigma_p = \sigma_{1p}} \sigma_p \tag{2.38}$$

where $\frac{\partial \overline{R}_p}{\partial \sigma_p}\Big|_{\sigma_p = \sigma_{1p}}$ denotes the slope of the hyperbola at point $\left(\sigma_{1p}, \overline{R}_{1p}\right)$ and $\overline{R}_0$ denotes the tangent line's intercept at $\sigma_p = 0$. Using (2.33) and (2.32), we

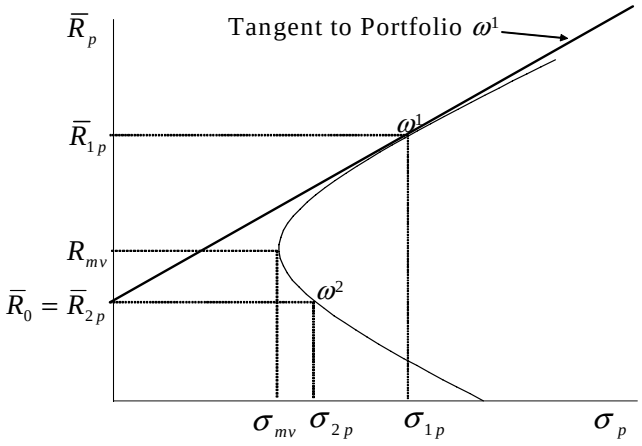


Figure 2.5: Frontier Portfolios with Zero Covariance

can solve for $\overline{R}_0$ by evaluating (2.38) at the point $\left(\sigma_{1p}, \overline{R}_{1p}\right)$:

$$\begin{aligned}
\overline{R}_0 &= \overline{R}_{1p} - \frac{\partial \overline{R}_p}{\partial \sigma_p}\Big|_{\sigma_p=\sigma_{1p}} \sigma_{1p} = \overline{R}_{1p} - \frac{\varsigma\delta - \alpha^2}{\delta\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)}\sigma_{1p}\sigma_{1p} \qquad (2.39)\\
&= \overline{R}_{1p} - \frac{\varsigma\delta - \alpha^2}{\delta\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)}\left[\frac{1}{\delta} + \frac{\delta\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)^2}{\varsigma\delta - \alpha^2}\right]\\
&= \frac{\alpha}{\delta} - \frac{\varsigma\delta - \alpha^2}{\delta^2\left(\overline{R}_{1p} - \frac{\alpha}{\delta}\right)}\\
&= \overline{R}_{2p}
\end{aligned}$$

Hence, as shown in Figure 2.5, the intercept of the line tangent to frontier portfolio ω^1 equals the expected return of its zero covariance counterpart, frontier portfolio ω^2.

2.4 The Efficient Frontier with a Riskless Asset

Thus far, we have assumed that investors can hold only risky assets. An implication of our analysis was that while all investors would choose efficient portfolios of risky assets, these portfolios would differ based on the particular investor's level of risk aversion. However, as we shall now see, introducing a riskless asset can simplify the investor's portfolio choice problem. This augmented portfolio choice problem, whose solution was first derived by James Tobin (Tobin 1958), is one that we now consider.[15]

Assume that there is a riskless asset with return R_f. Let ω continue to be the $n \times 1$ vector of portfolio proportions invested in the risky assets. Now, however, the constraint $\omega' e = 1$ does not apply, because $1 - \omega' e$ is the portfolio proportion invested in the riskless asset. We can impose the restriction that the portfolio weights for all $n + 1$ assets sum to one by writing the expected return on the portfolio as

$$\bar{R}_p = R_f + \omega'(\bar{R} - R_f e) \tag{2.40}$$

The variance of the return on the portfolio continues to be given by $\omega' V \omega$. Thus, the individual's optimization problem is changed to:

$$\min_{\omega} \tfrac{1}{2}\omega' V \omega + \lambda \left\{ \overline{R}_p - \left[R_f + \omega'(\bar{R} - R_f e) \right] \right\} \tag{2.41}$$

In a manner similar to the previous derivation, the first order conditions lead to the solution

$$\omega^* = \lambda V^{-1}(\bar{R} - R_f e) \tag{2.42}$$

[15] Tobin's work on portfolio selection was one of his contributions cited by the selection committee that awarded him the Nobel prize in economics in 1981.

where $\lambda \equiv \dfrac{\overline{R}_p - R_f}{\left(\bar{R} - R_f e\right)' V^{-1}(\bar{R} - R_f e)} = \dfrac{\overline{R}_p - R_f}{\varsigma - 2\alpha R_f + \delta R_f^2}$. Since V^{-1} is positive definite, λ is non-negative when $\overline{R}_p \geq R_f$, the region of the efficient frontier where investors' expected portfolio return is at least as great as the risk-free return. Given (2.42), the amount optimally invested in the riskless asset is determined by $1 - e'w^*$. Note that since λ is linear in $\overline{R}_p$, so is ω^*, similar to the previous case of no riskless asset. The variance of the portfolio now takes the form

$$\sigma_p^2 = \omega^{*\prime} V \omega^* = \frac{(\overline{R}_p - R_f)^2}{\varsigma - 2\alpha R_f + \delta R_f^2} \tag{2.43}$$

Taking the square root of each side of (2.43) and re-arranging

$$\overline{R}_p = R_f \pm \left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}} \sigma_p \tag{2.44}$$

which indicates that the frontier is *linear* in σ_p, $\overline{R}_p$ space. Corresponding to the hyperbola for the no riskless asset case, the frontier when a riskless asset is included becomes two straight lines, each with an intercept of R_f but one having a positive slope of $\left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}}$, the other having a negative slope of $-\left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}}$. Of course, only the positively sloped line is the efficient portion of the frontier.

Since $\omega^*$is linear in $\overline{R}_p$, the previous section's separation result continues to hold: any portfolio on the frontier can be replicated by two other frontier portfolios. However, when $R_f \neq R_{mv} \equiv \frac{\alpha}{\delta}$ holds, an even stronger separation principle obtains.[16] In this case, any portfolio on the linear efficient frontier can be replicated by two particular portfolios: one portfolio that is located on the "risky asset only" frontier and another portfolio that holds only the riskless

[16] We continue to let R_{mv} denote the expected return on the minimum variance portfolio that holds only risky assets. Of course with a riskless asset, the minimum variance portfolio would be one that holds only the riskless asset.

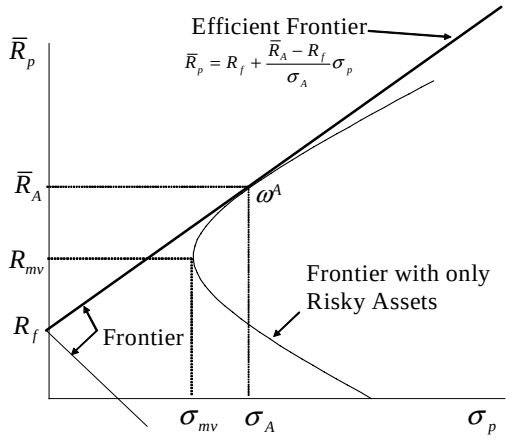


Figure 2.6: Efficient Frontier with a Riskless Asset

asset.

Let us start by proving this result for the situation where $R_f < R_{mv}$. We assert that the efficient frontier given by the line $\overline{R}_p = R_f + \left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}} \sigma_p$ can be replicated by a portfolio consisting of only the riskless asset and a portfolio on the risky asset only frontier that is determined by a straight line tangent to this frontier whose intercept is R_f. This situation is illustrated in Figure 2.6 where ω^A denotes the portfolio of risky assets determined by the tangent line having intercept R_f. If we can show that the slope of this tangent line equals $\left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}}$, then our assertion is proved.[17] Let $\overline{R}_A$ and σ_A be the expected return and standard deviation of return, respectively, on this tangency portfolio. Then the results of (2.37) and (2.39) allow us to write the slope of

[17] Note that if a proportion x is invested in any risky asset portfolio having expected return and standard deviation of $\overline{R}_A$and σ_A, respectively, and a proportion $1 - x$ is invested in the riskless asset having certain return R_f, then the combined portfolio has an expected return and standard deviation of $\overline{R}_p = R_f + x\left(\overline{R}_A - R_f\right)$ and $\sigma_p = x\sigma_A$, respectively. When graphed in $\overline{R}_p, \sigma_p$ space, we can substitute for x to show that these combination portfolios are represented by the straight line $\overline{R}_p = R_f + \frac{R_A - R_f}{\sigma_A}\sigma_p$ whose intercept is R_f.

the tangent as

$$\begin{aligned}\frac{\overline{R}_A - R_f}{\sigma_A} &= \left[\frac{\alpha}{\delta} - \frac{\varsigma\delta - \alpha^2}{\delta^2\left(R_f - \frac{\alpha}{\delta}\right)} - R_f\right]/\sigma_A \\ &= \left[\frac{2\alpha R_f - \varsigma - \delta R_f^2}{\delta\left(R_f - \frac{\alpha}{\delta}\right)}\right]/\sigma_A\end{aligned} \tag{2.45}$$

Furthermore, we can use (2.32) and (2.37) to write

$$\begin{aligned}\sigma_A^2 &= \frac{1}{\delta} + \frac{\delta\left(\overline{R}_A - \frac{\alpha}{\delta}\right)^2}{\varsigma\delta - \alpha^2} \\ &= \frac{1}{\delta} + \frac{\varsigma\delta - \alpha^2}{\delta^3\left(R_f - \frac{\alpha}{\delta}\right)^2} \\ &= \frac{\delta R_f^2 - 2\alpha R_f + \varsigma}{\delta^2\left(R_f - \frac{\alpha}{\delta}\right)^2}\end{aligned} \tag{2.46}$$

Substituting the square root of (2.46) into (2.45) gives[18]

$$\begin{aligned}\frac{\overline{R}_A - R_f}{\sigma_A} &= \left[\frac{2\alpha R_f - \varsigma - \delta R_f^2}{\delta\left(R_f - \frac{\alpha}{\delta}\right)}\right]\frac{-\delta\left(R_f - \frac{\alpha}{\delta}\right)}{\left(\delta R_f^2 - 2\alpha R_f + \varsigma\right)^{\frac{1}{2}}} \\ &= \left(\delta R_f^2 - 2\alpha R_f + \varsigma\right)^{\frac{1}{2}}\end{aligned} \tag{2.47}$$

which is the desired result.

This result is an important simplification. If all investors agree on the distribution of asset returns (returns are distributed $N\left(\overline{R}, V\right)$), then they all consider the linear efficient frontier to be $\overline{R}_p = R_f + \left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}}\sigma_p$ and all will choose to hold risky assets in the same relative proportions given by the tangency portfolio ω^A. Investors differ only in the amount of wealth

[18] Because it is assumed that $R_f < \frac{\alpha}{\delta}$, the square root of (2.46) has an opposite sign in order for σ_A to be positive.

they choose to allocate to this portfolio of risky assets versus the risk-free asset. Along the efficient frontier depicted in Figure 2.7, the proportion of an investor's total wealth held in the tangency portfolio, $e\acute{\omega}^*$, increases as one moves to the right. At point $\left(\sigma_p, \overline{R}_p\right) = (0, R_f)$, $e\acute{\omega}^* = 0$ and all wealth is invested in the risk-free asset. In between points $(0, R_f)$ and $\left(\sigma_A, \overline{R}_A\right)$, which would be the case if, say, investor 1 had an indifference curve tangent to the efficient frontier at point $\left(\sigma_1, \overline{R}_{p1}\right)$, then $0 < e\acute{\omega}^* < 1$ and positive proportions of wealth are invested in the risk-free asset and the tangency portfolio of risky assets. At point $\left(\sigma_A, \overline{R}_A\right)$, $e\acute{\omega}^* = 1$ and all wealth is invested in risky assets and none in the risk-free asset. Finally, to the right of this point, which would be the case if, say, investor 2 had an indifference curve tangent to the efficient frontier at point $\left(\sigma_2, \overline{R}_{p2}\right)$, then $e\acute{\omega}^* > 1$. This implies a negative proportion of wealth in the risk-free asset. The interpretation is that investor 2 borrows at the risk-free rate to invest more than 100 percent of her wealth in the tangency portfolio of risky assets. In practical terms, such an investor could be viewed as buying risky assets "on margin," that is leveraging her asset purchases with borrowed money.

It will later be argued that $R_f < R_{mv}$, the situation depicted in Figures 2.6 and 2.7, is required for asset market equilibrium. However, we briefly describe the implications of other parametric cases. When $R_f > R_{mv}$, the efficient frontier of $\overline{R}_p = R_f + \left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}} \sigma_p$ is always above the risky-asset-only frontier. Along this efficient frontier, the investor short-sells the tangency portfolio of risky assets. This portfolio is located on the inefficient portion of the risky-asset-only frontier at the point where the line $\overline{R}_p = R_f - \left(\varsigma - 2\alpha R_f + \delta R_f^2\right)^{\frac{1}{2}} \sigma_p$ becomes tangent. The proceeds from this short selling are then wholly invested in the risk-free asset. Lastly, when $R_f = R_{mv}$ the portfolio frontier is given by the asymptotes illustrated in Figure 2.4. It is

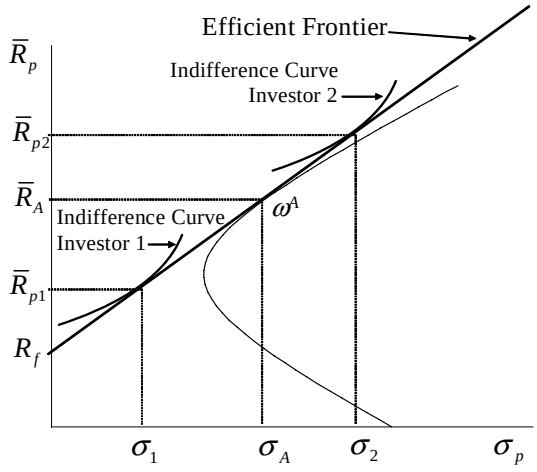


Figure 2.7: Investor Portfolio Choice

straightforward to show that $e\acute{\omega}^* = 0$ for this case, so that total wealth is invested in the risk-free asset. However, the investor also holds a risky, but zero net wealth, position in risky assets. In other words, the proceeds from short-selling particular risky assets are used to finance long positions in other risky assets.

2.4.1 An Example with Negative Exponential Utility

To illustrate our results, let us specify a form for an individual's utility function. This enables us to determine the individual's preferred efficient portfolio, that is, the point of tangency between the individual's highest indifference curve and the efficient frontier. Given a specific utility function and normally distributed asset returns, we show how the individual's optimal portfolio weights can be derived directly by maximizing expected utility.

As before, let $\tilde{W}$ be the individual's end-of-period wealth and assume that

she maximizes expected negative exponential utility.

$$U(\tilde{W}) = -e^{-b\tilde{W}} \tag{2.48}$$

where b is the individual's coefficient of absolute risk aversion. Now define $b_r \equiv bW_0$, which is the individual's coefficient of relative risk aversion at initial wealth W_0. Equation (2.48) can be re-written:

$$U(\tilde{W}) = -e^{-b_r\tilde{W}/W_0} = -e^{-b_r\tilde{R}_p} \tag{2.49}$$

where $\tilde{R}_p$ is the total return (one plus the rate of return) on the portfolio.

In this problem, we assume that initial wealth can be invested in a riskless asset and n risky assets. As before, denote the return on the riskless asset as R_f and the returns on the n risky assets as the $n \times 1$ vector $\tilde{R}$. Also as before, let $\omega = (\omega_1 \ldots \omega_n)'$ be the vector of portfolio weights for the n risky assets. The risky assets' returns are assumed to have a joint normal distribution where $\bar{R}$ is the $n \times 1$ vector of expected returns on the n risky assets and V is the $n \times n$ covariance matrix of returns. Thus, the expected return on the portfolio can be written $\bar{R}_p \equiv R_f + \omega'(\bar{R} - R_f e)$ and the variance of the return on the portfolio is $\sigma_p^2 \equiv \omega' V \omega$.

Now recall the properties of the lognormal distribution. If $\tilde{x}$ is a normally distributed random variable, for example, $\tilde{x} \sim N(\mu, \sigma^2)$, then $\tilde{z} = e^{\tilde{x}}$ is lognormally distributed. The expected value of $\tilde{z}$ is

$$E[\tilde{z}] = e^{\mu + \frac{1}{2}\sigma^2} \tag{2.50}$$

From (2.49), we see that if $\tilde{R}_p = R_f + \omega'(\tilde{R} - R_f e)$ is normally distributed, then $U\left(\tilde{W}\right)$ is lognormally distributed. Using equation (2.50), we have

$$E\left[U\left(\widetilde{W}\right)\right] = -e^{-b_r\left[R_f+\omega'(\bar{R}-R_f e)\right]+\frac{1}{2}b_r^2\omega' V\omega} \tag{2.51}$$

The individual chooses portfolio weights by maximizing expected utility:

$$\max_{\omega} E\left[U\left(\widetilde{W}\right)\right] = \max_{\omega} - e^{-b_r\left[R_f+\omega'(\bar{R}-R_f e)\right]+\frac{1}{2}b_r^2\omega' V\omega} \tag{2.52}$$

Because the expected utility function is monotonic in its exponent, the maximization problem in (2.52) is equivalent to

$$\max_{\omega} \omega'(\bar{R} - R_f e) - \tfrac{1}{2}b_r\omega' V\omega \tag{2.53}$$

The n first order conditions are

$$\bar{R} - R_f e - b_r V\omega = 0 \tag{2.54}$$

Solving for ω, we obtain

$$\omega^* = \frac{1}{b_r}V^{-1}(\bar{R} - R_f e) \tag{2.55}$$

Thus, we see that the individual's optimal portfolio choice depends on b_r, her coefficient of relative risk aversion, and the expected returns and covariances of the assets. Comparing (2.55) to (2.42), note that

$$\frac{1}{b_r} = \lambda \equiv \frac{\overline{R}_p - R_f}{\left(\bar{R} - R_f e\right)' V^{-1}(\bar{R} - R_f e)} \tag{2.56}$$

so that the greater the investor's relative risk aversion, b_r, the smaller is her target mean portfolio return, $\overline{R}_p$, and the smaller is the proportion of wealth invested in the risky assets. In fact, multiplying both sides of (2.55) by W_0, we see that the absolute amount of wealth invested in the risky assets is

$$W_0\omega^* = \frac{1}{b}V^{-1}(\bar{R} - R_f e) \tag{2.57}$$

Therefore, the individual with constant absolute risk aversion, b, invests a fixed dollar amount in the risky assets, independent of her initial wealth. As wealth increases, each additional dollar is invested in the risk-free asset. Recall that this same result was derived at the end of Chapter 1 for the special case of a single risky asset.

As in this example, constant absolute risk aversion's property of making risky asset choice independent of wealth often allows for simple solutions to portfolio choice problems when asset returns are assumed to be normally distributed. However, the unrealistic implication that both wealthy and poor investors invest the same dollar amount in risky assets limits the empirical applications of using this form of utility. As we shall see in later chapters of this book, models where utility displays constant relative risk aversion is more typical.

2.5 An Application to Cross-Hedging

The following application of mean-variance analysis is based on Anderson and Danthine (Anderson and Danthine 1981). Consider a one-period model of an individual or institution that is required to buy or sell a commodity in the future and would like to hedge the risk of such a transaction by taking positions in futures (or other financial securities) markets. Assume that this financial operator is committed at the beginning of the period, date 0, to buy y units of a risky commodity at the end of the period, date 1, at the then-prevailing spot price p_1. For example, a commitment to buy could arise if the commodity is a necessary input in the operator's production process.[19] Conversely, $y < 0$

[19] An example of this case would be a utility that generates electricity from oil.

represents a commitment to sell $-y$ units of a commodity, which could be due to the operator producing a commodity that is non-storable.[20] What is important is that, as of date 0, y is deterministic while p_1 is stochastic.

There are n financial securities (for example, futures contracts) in the economy. Denote the date 0 price of the i^{th} financial security as p^s_{i0}. Its date 1 price is p^s_{i1}, which is uncertain as of date 0. Let s_i denote the amount of the i^{th} security purchased at date 0. Thus, $s_i < 0$ indicates a short position in the security.

Define the $n \times 1$ quantity and price vectors $s \equiv [s_1 \dots s_n]'$, $p^s_0 \equiv [p^s_{10} \dots p^s_{n0}]'$, and $p^s_1 \equiv [p^s_{11} \dots p^s_{n1}]'$. Also define $p^s \equiv p^s_1 - p^s_0$ as the $n \times 1$ vector of security price changes. This is the profit at date 1 from having taken unit long positions in each of the securities (futures contracts) at date 0, so that the operator's profit from its security position is $p^{s\prime}s$. Also define the first and second moments of the date 1 prices of the spot commodity and the financial securities:

$E[p_1] = \bar{p}_1$, $Var[p_1] = \sigma_{00}$, $E[p^s_1] = \bar{p}^s_1$, $E[p^s] = \bar{p}^s$, $Cov[p^s_{i1}, p^s_{j1}] = \sigma_{ij}$, $Cov[p_1, p^s_{i1}] = \sigma_{0i}$, and the $(n+1) \times (n+1)$ covariance matrix of the spot commodity and financial securities is

$$\Sigma = \begin{bmatrix} \sigma_{00} & \Sigma_{01} \\ \Sigma'_{01} & \Sigma_{11} \end{bmatrix} \tag{2.58}$$

where Σ_{11} is an $n \times n$ matrix whose i, j^{th} element is σ_{ij}, and Σ_{01} is an $1 \times n$ vector whose i^{th} element is σ_{0i}.

For simplicity, let us assume that y is fixed and, therefore, is not a decision variable at date 0. Then the end-of-period profit (wealth) of the financial operator, W, is given by

[20] For example, the operator could be a producer of an agricultural good, such as corn, wheat, or soybeans.

$$W = p^{s\prime}s - p_1 y \tag{2.59}$$

What the operator must decide is the date 0 positions in the financial securities. We assume that the operator chooses s in order to maximize the following objective function that depends linearly on the mean and variance of profit:

$$\max_s E[W] - \tfrac{1}{2}\alpha Var[W] \tag{2.60}$$

As was shown in the previous section's (2.53), this objective function results from maximizing expected utility of wealth when portfolio returns are normally distributed and utility displays constant absolute risk aversion.[21] Substituting in for the operator's profit, we have

$$\max_s \bar{p}^{s\prime}s - \bar{p}_1 y - \tfrac{1}{2}\alpha\left[y^2\sigma_{00} + s'\Sigma_{11}s - 2y\Sigma_{01}s\right] \tag{2.61}$$

The first order conditions are

$$\bar{p}^s - \alpha\left[\Sigma_{11}s - y\Sigma'_{01}\right] = 0 \tag{2.62}$$

Thus, the optimal positions in financial securities are

$$\begin{aligned} s &= \frac{1}{\alpha}\Sigma_{11}^{-1}\bar{p}^s + y\Sigma_{11}^{-1}\Sigma'_{01} \\ &= \frac{1}{\alpha}\Sigma_{11}^{-1}\left(\overline{p}_1^s - p_0^s\right) + y\Sigma_{11}^{-1}\Sigma'_{01} \end{aligned} \tag{2.63}$$

[21] Similar to the previous derivation, the objective function (2.60) can be derived from an expected utility function of the form $E\left[U\left(W\right)\right] = -\exp\left[-\alpha W\right]$ where α is the operator's coefficient of absolute risk aversion. Unlike the previous example, here the objective function is written in terms of total profit (wealth) not portfolio returns per unit wealth. Also, risky asset holdings, s, are in terms of absolute amounts purchased, not portfolio proportions. Hence, α is the coefficient of absolute risk aversion, not relative risk aversion.

Let us first consider the case of $y = 0$. This can be viewed as the situation faced by a pure *speculator*, by which we mean a trader who has no requirement to hedge. If $n = 1$ and $\bar{p}_1^s > p_0^s$, the speculator takes a long position in (purchases) the security, while if $\bar{p}_1^s < p_0^s$, the speculator takes a short position in (sells) the security. The magnitude of the position is tempered by the volatility of the security ($\Sigma_{11}^{-1} = 1/\sigma_{11}$), and the speculator's level of risk aversion, α. However, for the general case of $n > 1$, an expected price decline or rise is not sufficient to determine whether a speculator takes a long or short position in a particular security. All of the elements in Σ_{11}^{-1} need to be considered, since a position in a given security may have particular diversification benefits.

For the general case of $y \neq 0$, the situation faced by a *hedger*, the demand for financial securities is similar to that of a pure speculator in that it also depends on price expectations. In addition, there are hedging components to the demand for financial assets, call them s^h :

$$s^h \equiv y\Sigma_{11}^{-1}\Sigma_{01}' \tag{2.64}$$

This is the solution to the problem $\min_s \ Var(W)$. Thus, even for a hedger, it is never optimal to minimize volatility (risk) unless risk aversion is infinitely large. Even a risk-averse, expected utility maximizing hedger should behave somewhat like a speculator in that securities' expected returns matter. From definition (2.64), note that when $n = 1$ the pure hedging demand per unit of the commodity purchased, s^h/y, simplifies to[22]

$$\frac{s^h}{y} = \frac{Cov(p_1, p_1^s)}{Var(p_1^s)} \tag{2.65}$$

[22] Note that if the correlation between the commodity price and the financial security return were equal to 1, so that a perfect hedge exists, then (2.65) beomces $s^h/y = \sqrt{\sigma_{00}}/\sqrt{\sigma_{11}}$, that is, the hedge ratio equals the ratio of the commodity price's standard deviation to that of the security price.

For the general case, $n > 1$, the elements of the vector $\Sigma_{11}^{-1}\Sigma_{01}'$ equal the coefficients $\beta_1, ..., \beta_n$ in the multiple regression model

$$\Delta p_1 = \beta_0 + \beta_1 \Delta p_1^s + \beta_2 \Delta p_2^s + ... + \beta_n \Delta p_n^s + \varepsilon \tag{2.66}$$

where $\Delta p_1 \equiv p_1 - p_0$, $\Delta p_i^s \equiv p_{1i}^s - p_{0i}^s$, and ε is a mean-zero error term. An implication of (2.66) is that an operator might estimate the *hedge ratios*, s^h/y, by performing a statistical regression using an historical times series of the $n \times 1$ vector of security price changes. In fact, this is a standard way that practitioners use to calculate hedge ratios.

2.6 Summary

When the returns on individual assets are multivariate normally distributed, a risk-averse investor optimally chooses among a set of mean-variance efficient portfolios. Such portfolios make best use of the benefits of diversification by providing the highest mean portfolio return for a given portfolio variance. The particular efficient portfolio chosen by a given investor depends on her level of risk-aversion. However, the ability to trade in only two efficient portfolios is sufficient to satisfy all investors, because any efficient portfolio can be created from any other two. When a riskless asset exists, the set of efficient portfolios has the characteristic that the portfolios' mean returns are linear in their portfolio variances. In such a case, a more risk averse investor optimally holds a positive amount of the riskless asset and a positive amount of a particular risky asset portfolio, while a less risk averse investor optimally borrows at the riskless rate to purchase the same risky asset portfolio in an amount exceeding his wealth.

This chapter provided insights on how individuals should optimally allocate

their wealth among various assets. Taking the distribution of returns for all available assets as given, we determined any individual's portfolio demands for these assets. Having now derived a theory of investor asset demands, the next chapter will consider the equilibrium asset pricing implications of this investor behavior.

2.7 Exercises

1. Prove that the indifference curves graphed in Figure 2.1 are convex if the utility function is concave. Hint: Suppose there are two portfolios, portfolios 1 and 2, that lie on the same indifference curve, where this indifference curve has expected utility of $\overline{U}$. Let the mean returns on portfolios 1 and 2 be $\bar{R}_{1p}$ and $\bar{R}_{2p}$, respectively, and let the standard deviations of returns on portfolios 1 and 2 be σ_{1p} and σ_{2p}, respectively. Consider a third portfolio located in $(\overline{R}_p, \sigma_p)$ space that happens to be on a straight line between portfolios 1 and 2, that is, a portfolio having a mean and standard deviation satisfying $\bar{R}_{3p} = x\bar{R}_{1p} + (1-x)\bar{R}_{2p}$ and $\sigma_{3p} = x\sigma_{1p} + (1-x)\sigma_{2p}$ where $0 < x < 1$. Prove that the indifference curve is convex by showing that the expected utility of portfolio 3 exceeds $\overline{U}$. Do this by showing that the utility of portfolio 3 exceeds the convex combination of utilities for portfolios 1 and 2 for each standardized normal realization. Then, integrate over all realizations to show this inequality holds for expected utilities.

2. Show that the covariance between the return on the minimum variance portfolio and the return on *any* other portfolio equals the variance of the return on the minimum variance portfolio. Hint: Write down the variance of a portfolio that consists of a proportion x invested in the minimum