CORE CURRICULUM

HARVARD BUSINESS PUBLISHING

Finance

Mihir A. Desai, Series Editor

READING + INTERACTIVE ILLUSTRATIONS

Risk and Return 2:

Portfolio Theory

TIMOTHY A. LUEHRMAN

8603 | Published: April 11, 2017

Table of Contents

This reading contains links to online interactive illustrations, denoted by the icon above. To access these exercises, you will need a broadband Internet connection. Verify that your browser meets the minimum technical requirements by visiting http://hbsp.harvard.edu/tech-specs.

Former Harvard Business School professor Timothy A. Luehrman developed this Core Reading with the assistance of writer Barbara Wall Lobosco, HBS MBA 1995.

1 INTRODUCTION

This is the second of two readings on risk and return. In the first, *Core Reading: Risk and Return 1: Stock Returns and Diversification* (HBP No. 5220), we defined risk and expected return from the perspective of a risk-averse investor and introduced statistical methods for measuring each. We examined probability distributions for stock returns and their depiction in μ-σ space, which allowed us to see how returns on one asset might dominate another, for example, by having the same expected return and lower risk, or the same risk but higher expected return. We looked at statistical measures of the tendency for different stocks to move together, or co-vary. This led us to compute risk and expected return parameters for portfolios composed of two or more stocks and to an understanding of diversification—the reduction in portfolio risk that results from less-than-perfect positive correlation of individual stocks.

This reading presumes an understanding of the concepts and statistical calculations presented in *Risk and Return 1*. We will now use those insights to tackle some big questions. How is risk priced? In other words, what is the relationship between risk and expected return in a well-functioning capital market? How do investors' perceptions of risk and return affect corporations and corporate activities such as mergers and acquisitions, divestitures, and capital allocation generally?

To address such topics, we consider whether and under what assumptions we might say some portfolio of risky assets is "best." Further, if an investor already owns the "best" portfolio (or even a pretty good one) and a new investment opportunity comes along, how should he or she decide whether to add it to the existing portfolio? That is, how can an investor decide if a new investment will improve his or her portfolio? To simplify the presentation of key ideas, we assume investors' preferences are such that a portfolio's expected returns and standard deviation are the only parameters the investor needs to consider (this assumption is analogous to our assumption of normally distributed returns in the first reading). From there we will consider the possibility of a general equilibrium: if all investors optimize *individually*, what is implied *collectively* for asset prices and returns? This will lead us to the famous capital asset pricing model (CAPM), one of the cornerstones of modern finance. Understanding the CAPM will facilitate subsequent study in many other branches of financial economics, including corporate finance, investment finance, public finance, and personal finance, to name a few.

2 ESSENTIAL READING

2.1 Diversification and Portfolio Returns

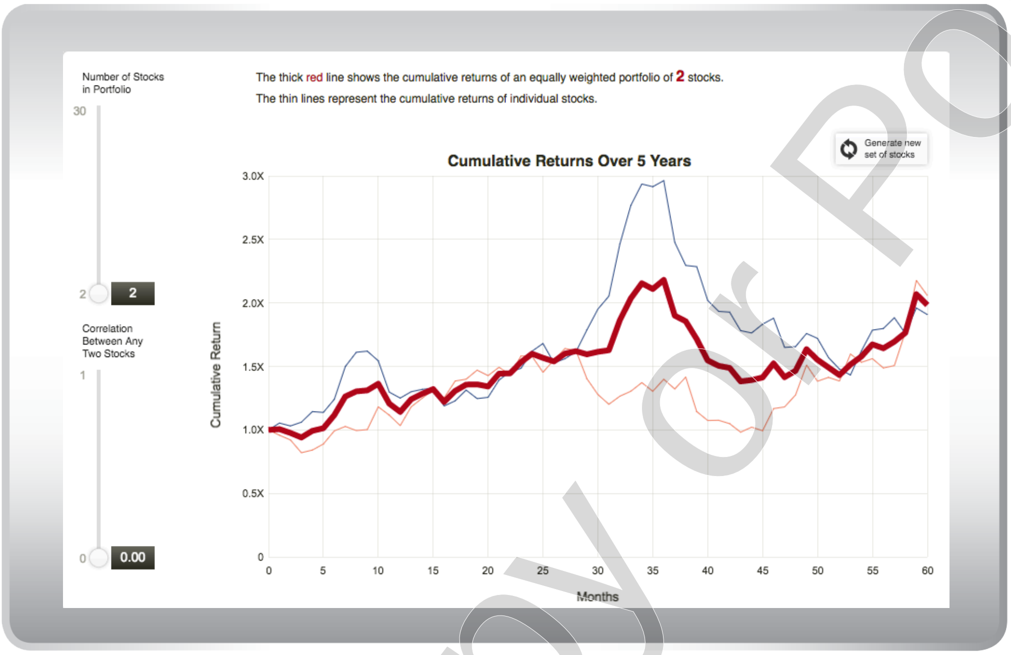
In *Risk and Return 1*, we saw that by combining two or more stocks together, it is possible to reduce portfolio risk without sacrificing expected return. The term ***diversification*** refers to the reduction of return volatility (risk) that occurs when combining assets that are less than perfectly positively correlated. How much risk can we eliminate through diversification? To consider this, we will compare graphs of accumulated returns over time for portfolios that differ in their degree of diversification. Such graphs are similar to the time series of stock returns we examined *Risk and Return 1*, only now we'll look at *cumulative* returns.

Interactive Illustration 1 lets you construct equal-weighted portfolios of up to 30 stocks and observe the effect on cumulative portfolio returns as the number of stocks in the portfolio increases. The graph shows cumulative returns for each stock as well as the cumulative return for the portfolio. Both are expressed in relation to the initial investment at t = 0. For example, a point at 2.0X indicates that the value of the stock (or portfolio) that month is double the amount invested at t = 0.

Stock returns shown in the illustration are randomly generated, using a normal distribution with the same mean and standard deviation for all 30 individual stocks. To begin, the illustration shows cumulative returns for two stocks (represented by the lighter lines) and the cumulative returns of an equal-weighted portfolio formed from the two stocks (shown by the bold red line). Now raise the slider from two to three stocks. The illustration adds another stock (shown by another lighter line of a different color) and the portfolio returns are recalculated, now composed of all three stocks in equal amounts. Move the slider to add a fourth stock, then a fifth, and so forth, all the way up to 30 stocks.

INTERACTIVE ILLUSTRATION 1 Diversification and Portfolio Returns

Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2rdPf1H



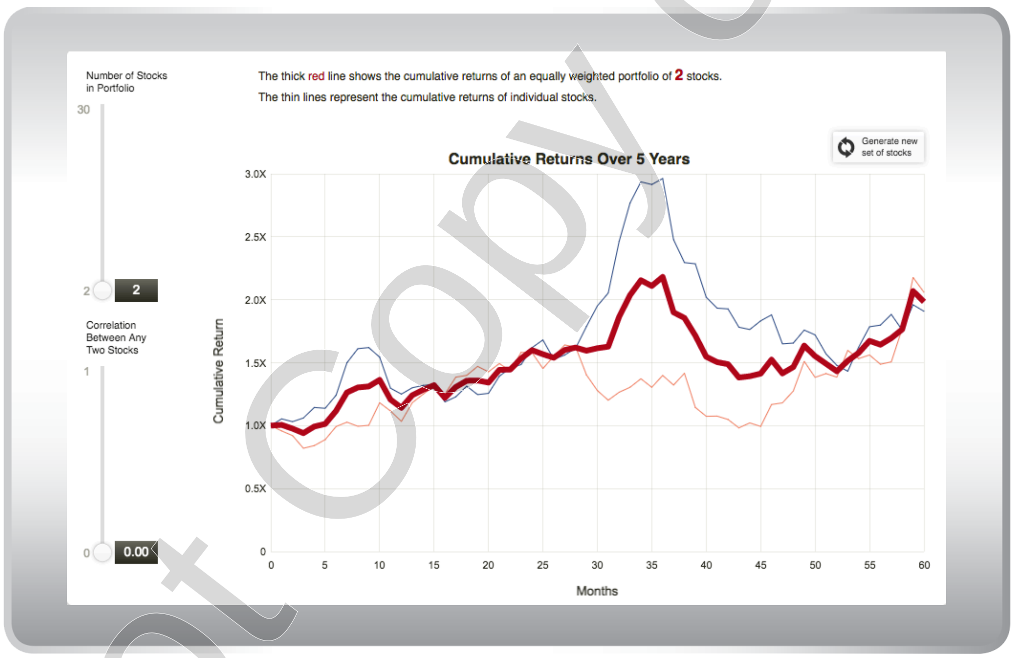
What happens to the portfolio returns? What began as a fairly choppy red line becomes smoother and smoother as you add additional stocks to the portfolio. The "choppiness" in cumulative returns is volatility—portfolio risk—and it nearly disappears as the number of stocks approaches 30. Here we see the risk-reducing effect of diversification in the increasing smoothness of the red line, whereas in *Risk and Return 1* we observed it in a lower computed σ for portfolio returns. Now reduce the number of stocks back to two, click the reset button to generate a new set of stocks, and repeat the experiment. Once again, the choppiness in portfolio returns diminishes as the number of stocks in the portfolio increases. Run as many scenarios as you need to convince yourself that regardless of the specific pattern of individual stock returns, the volatility of the overall portfolio decreases. What does this suggest about the degree to which the stocks in the illustration are correlated?

If you guessed that correlations between stocks must be low, you were right: They were set to equal zero. Now let's repeat the same experiment with non-zero correlations. **Interactive Illustration 2** is the same as the previous one, except that it includes a slider for correlation, which can vary from 0.0 to 1.0. If you would like to repeat the previous experiment, set the correlation to 0.0 and begin adding stocks to the portfolio. Once again, the cumulative returns graph for the portfolio becomes very smooth as volatility diminishes with the addition of more stocks. Now increase the correlation to 0.5 and repeat the experiment.[a] Portfolio returns are still less choppy than the individual stocks—volatility has been reduced—but not by as much as when the stocks were uncorrelated. Now boost the correlation to 0.75 and you will observe still less risk reduction. Finally, perform the experiment with $\rho = 1.0$—all 30 stocks are perfectly positively correlated—and you will see *no* risk reduction from diversification. When all stocks are perfectly positively correlated, their returns are identical, so the cumulative returns of the portfolio are identical to those of the individual stocks.

INTERACTIVE ILLUSTRATION 2

Diversification and Porftolio Returns with Correlation

Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2IRP2bD



[a] To perform the necessary calculations, the illustration has embedded within it correlation matrices, containing a correlation coefficient for all possible pairs of stocks as the number of stocks increases from 1 to 30. All elements of every correlation matrix are set equal to the value chosen by the user.

The Mathematics of Correlation and Portfolio Risk

To cement our understanding of the role of diversification in reducing portfolio risk, let's return briefly to the mathematical expression for portfolio risk, σ_P, that we developed in *Risk and Return 1*,

$$\sigma_P = \sqrt{w^2\sigma_A^2 + (1-w)^2\sigma_B^2 + 2w(1-w)\rho_{A,B}\sigma_A\sigma_B}$$

where *w* denotes the weight of stock A in the portfolio and $0 \le w \le 1$; σ_A, σ_B, and σ_P denote the standard deviations of returns for stock A, stock B, and the portfolio P composed of A and B, respectively; and $\rho_{A,B}$ denotes the correlation between returns of A and B. With a bit of algebraic manipulation, we can rearrange the terms in the expression above to state the standard deviation of portfolio returns as

$$\sigma_P = \sqrt{\left(w\sigma_A + (1-w)\sigma_B\right)^2 - \left[2w(1-w)(1-\rho_{A,B})\sigma_A\sigma_B\right]}$$

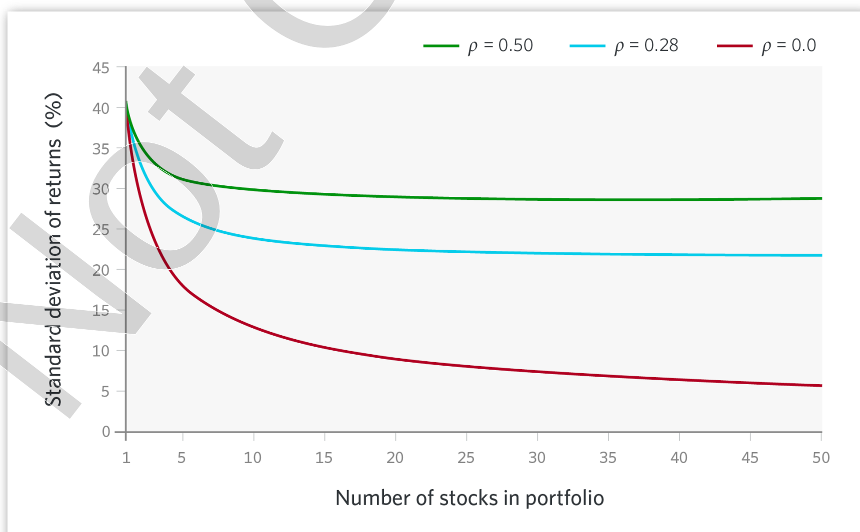
Let's examine this expression. The left term under the radical (in parentheses) is clearly positive. The right term (in square brackets) is the only place that the correlation between stocks A and B ($\rho_{A,B}$) appears. Notice that whenever $\rho_{A,B} < 1$, the term inside the square brackets is positive: It reduces portfolio risk (because it is subtracted). Moreover, the lower $\rho_{A,B}$ is, the lower σ_P will be. Indeed, when the correlation is zero or negative, portfolio risk may even be lower than the risk of either asset alone. But even when two stocks are positively correlated (as most are), so long as $\rho_{A,B} < 1$, the riskiness of a portfolio combining them is reduced by diversification. Finally, when A and B are perfectly positively correlated ($\rho_{A,B} = 1$), the standard deviation simplifies to

$$\sigma_P = w\sigma_A + (1-w)\sigma_B$$

and there is no benefit from diversification. The risk of the portfolio is simply the weighted average of the securities' individual risks.

In the interactive illustrations we demonstrated the effect of diversification on portfolio risk using artificially constructed stocks, all with the same expected returns and standard deviations, but with varying degrees of correlation. Though our stocks were artificial, the result was not; real-world stocks behave the same way. They tend to be positively correlated, and, as a result, some portfolio risk remains no matter how many stocks we put in the portfolio. The risk that remains goes by various names; it is sometimes called *systematic* or *market* or *nondiversifiable* risk. It is risk that no amount of diversification can eliminate. We see this in **Exhibit 1**, which graphs portfolio volatility as a function of the number of stocks in the portfolio. Here again, the underlying stock returns that produced the graph are artificial—each is constructed with an annual standard deviation of 40% (about average for large company stocks in the real world). If we set the correlation between any pair of stocks at 28% (again, the real-world average), the blue line in Exhibit 1 shows portfolio risk falling as we add stocks to the portfolio—the same effect we saw in Interactive Illustrations 1 and 2. We also see now that though portfolio volatility falls dramatically at first, the marginal risk reduction from additional diversification diminishes fairly quickly. That is, the risk of a portfolio with 50 stocks is not substantially lower than for one with only 15. More and more diversification will not bring portfolio risk below 20%, which is about the historical average for the stock market as a whole.

EXHIBIT 1 Portfolio Volatility as a Function of Portfolio Size and Correlation Between Stocks



The green and red lines in Exhibit 1 show the same phenomenon but use different correlation coefficients. The red line shows what happens when $\rho = 0$: risk reduction is even more dramatic at first and the marginal benefit diminishes less rapidly. This portfolio's risk will approach zero if we continue adding uncorrelated stocks. In contrast, the green line shows the opposite effect when correlation is higher ($\rho = 0.5$). The green portfolio reflects much less risk reduction due to diversification, and most of the benefit is realized very early, with as few as 10 stocks.

In *Risk and Return 1* we observed that risk-averse investors will demand to be compensated for bearing risk. In the trade-off between risk and expected return we expect higher risk (σ) to be associated with higher expected returns (μ). Now that we understand diversification and its effect on portfolio risk, we can say a bit more about the nature of this trade-off: specifically, that the trade-off must be between ***systematic* (*nondiversifiable*) *risk*** and expected return. Since diversification is "free"—it reduces risk without sacrificing expected return—investors should not expect to earn higher returns for bearing risk they can diversify away. They get higher expected returns only for bearing *non*diversifiable risk. This also suggests that investors' optimal portfolios will be diversified. An investor who does not diversify will be bearing risk that he or she doesn't have to bear, and for which he or she is not being offered a higher expected return. A risk-averse investor simply would not do this, but, rather, would diversify.[b]

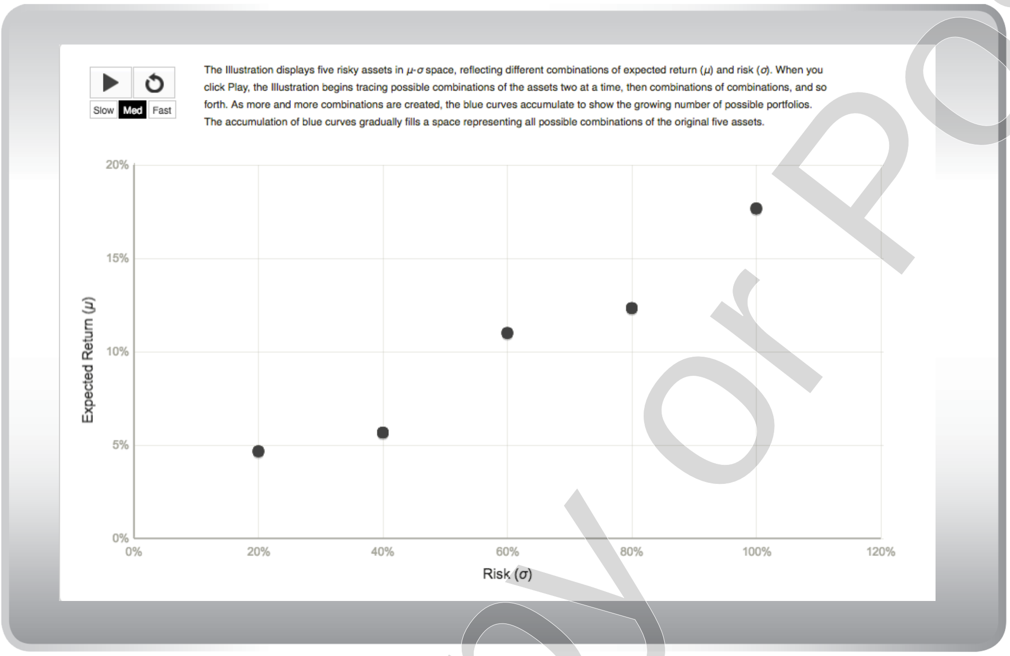
2.2 Efficient Portfolios

So investors will diversify, but how? There are many ways to diversify. Are some investments better than others? To consider this problem, we return to μ-σ space and consider investors' possibilities. In *Risk and Return 1* we examined all possible combinations of two risky stocks and saw that the set of possible portfolios described a convex curve in μ-σ space as long as the correlation between the stocks' returns was less than 1.0. What if there are many such stocks, not just two? What does the set of possible portfolios look like in μ-σ space? **Interactive Illustration 3** shows the set of possible portfolios created from multiple stocks.

[b] These are the insights of Harry Markowitz, whose work on modern portfolio theory revolutionized finance following its introduction in a 1952 essay, and for which he was awarded the Nobel Prize in 1990.

INTERACTIVE ILLUSTRATION 3 Region of Feasible Portfolios

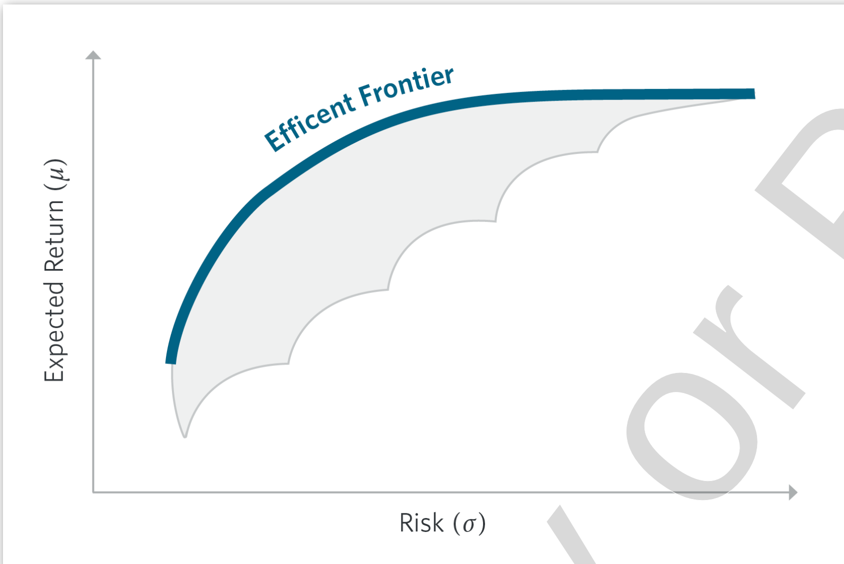
Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2DXEPXv



Interactive Illustration 3 shows five risky assets, each with its own expected return and standard deviation and represented by a unique point in μ-σ space. When you click Play, the illustration begins identifying possible portfolios—combinations of any two of the original assets—and plotting their risk and expected return parameters as blue curves. Then it identifies combinations of combinations and so forth, continuing to plot risk and return parameters as still more blue curves. As more and more curves are plotted and saved, they begin to fill a space. Click "Stop" when you get the idea or get tired of watching. The collection of all feasible portfolios eventually fills a space shaped like a broken eggshell, with a smooth curve on the northwest side and a more jagged edge on the southeast. No matter how many assets we might add to the original five, the result will always be qualitatively similar to the one derived in the illustration owing to the statistical properties of risk and expected returns for portfolios.

Exhibit 2 presents a stylized picture of the set of all possible portfolios composed of risky assets in μ-σ space. It is the final state of Interactive Illustration 3. We don't need to know μ and σ for all the individual risky assets to know that the set of all possible combinations of them has this characteristic shape in μ-σ space.

EXHIBIT 2 Portfolio Combinations and the Efficient Frontier



Given this set of possible portfolios, which one(s) would an investor prefer? We cannot say yet which, if any, portfolio is best, but it is easy to see that some are better than others. A risk-averse investor will always prefer higher expected returns for a given amount of risk. If we draw a vertical line through the filled-in portion of the curve in Exhibit 2, every portfolio on the line has the same risk (x-coordinate). But the uppermost portfolio on the line dominates all the others in the same way we defined dominance in *Risk and Return 1*. We say this portfolio is "efficient": it has the highest expected return (y-coordinate) for a given amount of risk. No risk-averse investor would invest in any of the dominated portfolios; to do so would be to accept a lower expected return for the same risk. If we consider in a similar fashion the entire range of risk (x-coordinates) for the space, we can easily identify a set of efficient portfolios. They all lie on the edge of the space, beginning on the left at the portfolio with the lowest risk and continuing right along the upper edge of the space to the portfolio with the maximum expected return. This set of portfolios is called the ***efficient frontier*** and is depicted in blue in Exhibit 2. Every portfolio on the efficient frontier represents the highest expected return for a given amount of risk or, alternatively, the least risk for a given expected return. Regardless of an individual investor's other preferences, if he or she is risk-averse, he or she will prefer at least one portfolio on the efficient frontier to *any* portfolio not on the efficient frontier.

Let's pause a moment and consider adding another asset to the set of five in Interactive Illustration 3—specifically, an asset with μ and σ such that it plots *inside* the broken eggshell shape and *below* the efficient frontier. If the new asset is not on the efficient frontier, no one would want it, right? Wrong. We can say *no one would* *want it all by itself,* that is, in lieu of all other possible investments. But we cannot say no one would want it in a *portfolio*. This is because our new asset might be uncorrelated with the other five assets, for example, and therefore create valuable risk reduction opportunities: New combinations of it with the other assets might produce a "better" efficient frontier—one that plots north or west (or both) of the existing frontier. In short, a new asset won't change the basic shape of the set of possibilities, but it might change the *location* of this set and of the efficient frontier. Whether it does so depends on its correlation with the other assets.

2.3 Introducing a Risk-Free Asset

In the preceding section we saw that there were many possible portfolios created from an underlying set of risky assets, but that a particular subset of them—those on the efficient frontier—collectively dominated all the others. A risk-averse investor has no interest in nonefficient portfolios. There was no one portfolio that we could deem best for *all* investors, however; each investor would choose to invest in a portfolio on the efficient frontier that best suits his or her tolerance for risk.

Now we add an additional feature to the investment problem: the possibility of investing in a *risk-free* asset in addition to a risky asset (or a portfolio of risky assets). What is a risk-free asset? One with a guaranteed return over the holding period—in other words, a return with no risk.[c] *No risk* in our framework means no uncertainty: the standard deviation of riskless returns equals zero, and they are uncorrelated with risky returns. Therefore, in μ-σ space a riskless asset plots on the vertical axis at $\sigma = 0$ and $\mu = r_f =$ the risk-free rate of interest. Now we can consider portfolios that combine a risk-free asset (call it asset F) with some risky asset A (or A could as well be a portfolio of risky assets).

Interactive Illustration 4 shows the risk and expected-return parameters for all combinations of the risk-free asset with a risky asset. Point F represents the risk and expected return of a portfolio invested 100% in the risk-free asset F. Likewise, point A represents a 100% investment in risky asset A. Check the box at the top of the illustration to activate the slider that lets you choose portfolio weights for A and F. Now move the slider to vary the portfolio weights between 0% and 100%, and watch portfolio P_{FA} move along the line connecting F and A. Each point on this line

[c] For investments in US dollars, US Treasury securities are commonly considered to be risk-free assets, even though they are not explicitly guaranteed as such. See the Supplemental Reading for further discussion of the risk-free rate of return.

represents the risk and expected return of a particular combination of F and A. Why is this set of possible combinations a straight line and not a convex curve? Each portfolio's expected return is a weighted average of the returns of A and F, as always

$$[\,E(r_P) = wE(r_A) + (1-w)E(r_F)\,]$$

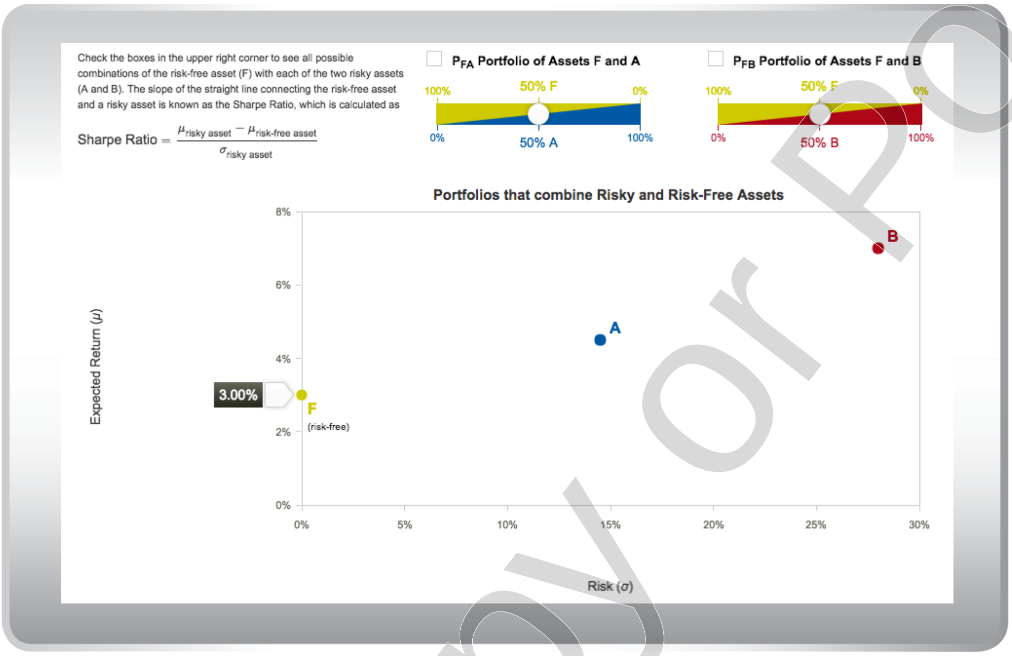
But because F is risk-free, the portfolio's standard deviation is simply the weight of A times the standard deviation of A, since the risk-free asset has a σ of zero and the assets A and F have zero covariance. To see this, substitute $\sigma_F = 0$ and $\text{cov}(r_A, r_F) = 0$ in the expression for portfolio risk:

$$\sigma_P = \sqrt{w^2\sigma_A^2 + (1-w)^2\,\sigma_F^2 + 2w(1-w)\text{cov}(r_A, r_F)}$$
$$\sigma_P = \sqrt{w^2\sigma_A^2 + (1-w)^2\,(0) + 2w(1-w)(0)}$$
$$\sigma_P = \sqrt{w^2\sigma_A^2} = w\sigma_A$$

The value of σ_P equals σ_A multiplied by the weight of A in the portfolio. An equal-weighted portfolio of assets A and F (50% A and 50% F) will have an expected return of 3.75% and a standard deviation of 7.25% (these figures use μ of 4.5% and 3.0%, and σ of 14.5% and 0% for A and F, respectively). Now move northeast along the line, which represents an increase in the portfolio's weight of A. As asset A's weight in the portfolio increases, so do the portfolio's expected return and risk. Conversely, if we increase the weight of F (move southwest), the risk and expected return of the portfolio decrease. The important point is that *all portfolios composed of a risky and a risk-free asset lie on a straight line in μ-σ space.*

INTERACTIVE ILLUSTRATION 4 Combining Risky and Risk-Free Assets

Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2pILTTB



2.3.1 The Sharpe Ratio

The slope of the line combining risky asset A and risk-free asset F describes a trade-off between risk and expected return that characterizes *all* combinations of a risk-free asset with a particular risky asset or portfolio. The slope is also known as the ***Sharpe Ratio***,[d] and it measures how much an investor in any such portfolio is compensated for bearing additional risk. The slope of this line is

$$\text{Sharpe Ratio} = \text{Slope} = \left(\frac{E(r_A) - r_f}{\sigma_A}\right) = \left(\frac{\mu_A - \mu_F}{\sigma_A}\right)$$

In Interactive Illustration 4, the Sharpe Ratio for all combinations of assets A and F was 0.10.

Interactive Illustration 4 also allows you to examine portfolios containing a second risky asset, B, and shows the straight line representing all possible combinations of B and F. Asset B has both higher risk and higher expected return than asset A. Check the box that allows you to see portfolios combining risky asset B with risk-free asset F (leave the box for A checked as well).

[d] The ratio is named for its inventor, the Nobel Prize–winning economist William Sharpe.

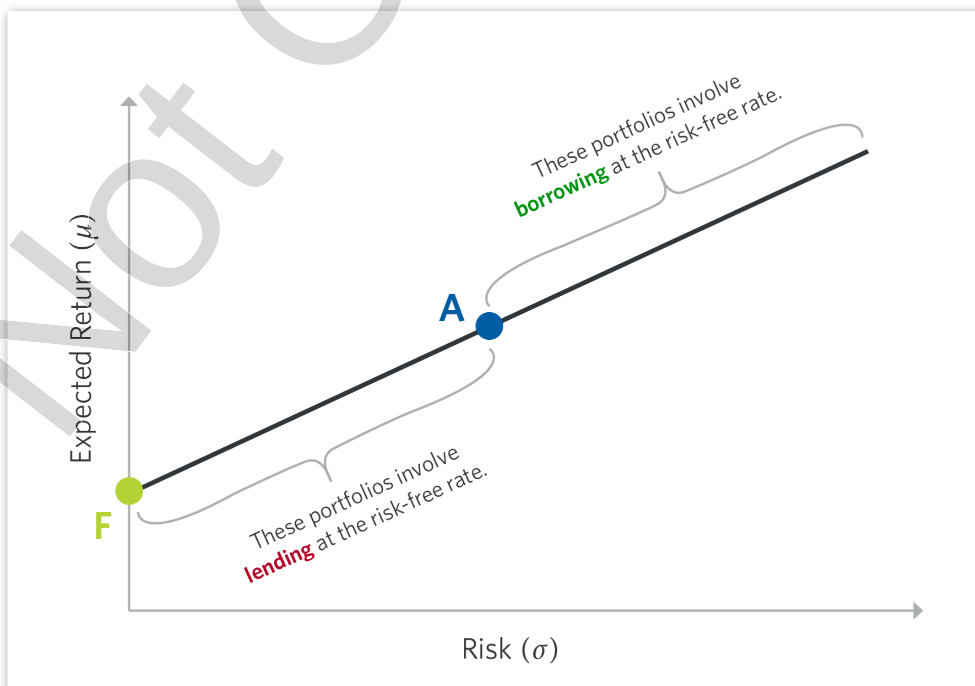
What do you observe about the slope and Sharpe Ratio of portfolios combining B and F? Given a choice between P_{FA} and P_{FB}, which one would you choose? The Sharpe Ratio is the key to answering this question. A risk-averse investor will always prefer a higher Sharpe Ratio to a lower one—that is, greater expected return for a given amount of risk. Given a risk-free rate of 3%, all possible combinations of B and F have a higher Sharpe Ratio (0.14) than all combinations of A and F (0.10). Put more simply, we say that B's Sharpe Ratio is higher than (and preferable to) A's.

Now suppose that the risk-free rate were lower. Does that change your view of which portfolio is superior? Changing the risk-free rate changes the Sharpe Ratio for both A and B. Depending on how it changes, it might reverse your preference; that is, A's Sharpe Ratio might become higher than B's. To see how that happens, try changing the slider next to F in Interactive Illustration 4.

2.3.2 Borrowing Versus Lending the Risk-Free Asset

We have seen that points in μ-σ space on the line between a risky asset and the risk-free asset represent combinations of the two, and that all such portfolios have the same Sharpe Ratio. **Exhibit 3** is a simplified version of the illustration we just examined—it shows combinations of asset A and the risk-free asset F.

EXHIBIT 3 Combinations of a Risky Asset with the Risk-Free Asset, with Borrowing and Lending



Point F is a portfolio composed of 100% F; Point A is 100% A. Points in between contain A and F in amounts between 0% and 100%. What about points on the line beyond A (i.e., to the upper right, or northeast)? Are these feasible portfolios? Yes. They involve investing *more than 100%* of the portfolio in asset A. How is this possible? By placing a *negative* weight on the risk-free asset F. That is, portfolio weights still must sum to 1.0 (100%), but the weight on the risk-free asset may be less than zero. A negative weight signifies that the investor is *borrowing* rather than *lending* at the risk-free rate of interest.

To illustrate, suppose the risk-free rate is 3% per year (asset F has μ and σ of 3% and 0%, respectively), and we have risky asset A with μ and σ of 4.5% and 14.5%, respectively. As we have seen, an equal-weighted portfolio of F and A has an expected return of 3.75% (= (0.50 × 3%) + (0.50 × 4.5%)) and risk of 7.25% (= $w\sigma_A$ = 0.50 × 14.5%). We say the owner of this portfolio is *lending* at the risk-free rate because he or she has invested in a security that will pay a return of 3% per year—he or she is *receiving* a guaranteed 3% return on however much has been invested in asset F. But if he or she agrees to *pay* someone else 3% per year guaranteed, then he or she is a *borrower* rather than a lender.

Let's work a numerical example involving borrowing rather than lending in a portfolio (also known as "leverage" or "levering" the portfolio). Suppose you have $1,000 to invest, and risky asset A offers an expected return (μ_A) of 4.5% per year and risk (σ_A) of 14.5% per year, as above. You would like to achieve an expected return *higher* than 4.5%. How is this possible, given that A's expected return is only 4.5% and F's is even lower—only 3%? Here's how: you borrow $500 at an interest rate of 3% per year. Then you invest $1,500 in asset A—your original $1,000 plus the $500 you borrowed. The expected return on this portfolio is 5.25% (= (1.5 × 4.5%) + (0.5 × (−3.0%))). Your expected return is indeed greater than 4.5%, even though you have to pay 3% interest on the funds you borrowed. But the higher expected return isn't "free"—it entails more risk. The standard deviation for this portfolio is 21.75% per year (= $w\sigma_A$ = 1.5 × 14.5%). The new portfolio plots in μ-σ space with coordinates (21.75%, 5.25%), which puts it on the same line through F and A as the equal-weighted portfolio we considered, but it lies beyond (northeast of) point A.

Both portfolios—the equal-weighted portfolio and the leveraged portfolio—have the same Sharpe Ratio:

$$\text{Sharpe Ratio} = \left(\frac{3.75\% - 3.0\%}{7.25\%}\right) = \left(\frac{5.25\% - 3.0\%}{21.75\%}\right) = 0.1035$$

Indeed, both portfolios have the same Sharpe Ratio as asset A alone ([(4.5% – 3.0%)/14.5%] = 0.1035). All three assets/portfolios lie on the same line through F and A. The important point is that *all* possible portfolios combining a risky asset and the risk-free asset, including points *beyond* the risky asset, lie on a straight line connecting the two points in μ-σ space. As Exhibit 3 shows, points beyond (northeast of) the risky asset are reached by borrowing at the risk-free rate of interest (assigning a negative weight to the risk-free asset) and investing the proceeds in the risky asset (assigning it a portfolio weight greater than 1.0).[e]

Now that we've seen how risky portfolios may be combined with a risk-free asset—lending or borrowing at the risk-free rate—we return to the problem of choosing among the portfolios on the efficient frontier. As a set of possible portfolios, the efficient frontier dominates all other combinations of risky assets, but they *all* are efficient—we need a way to choose *among* efficient portfolios, and combining a risk-free asset with portfolios on the efficient frontier provides us with a way of doing so.

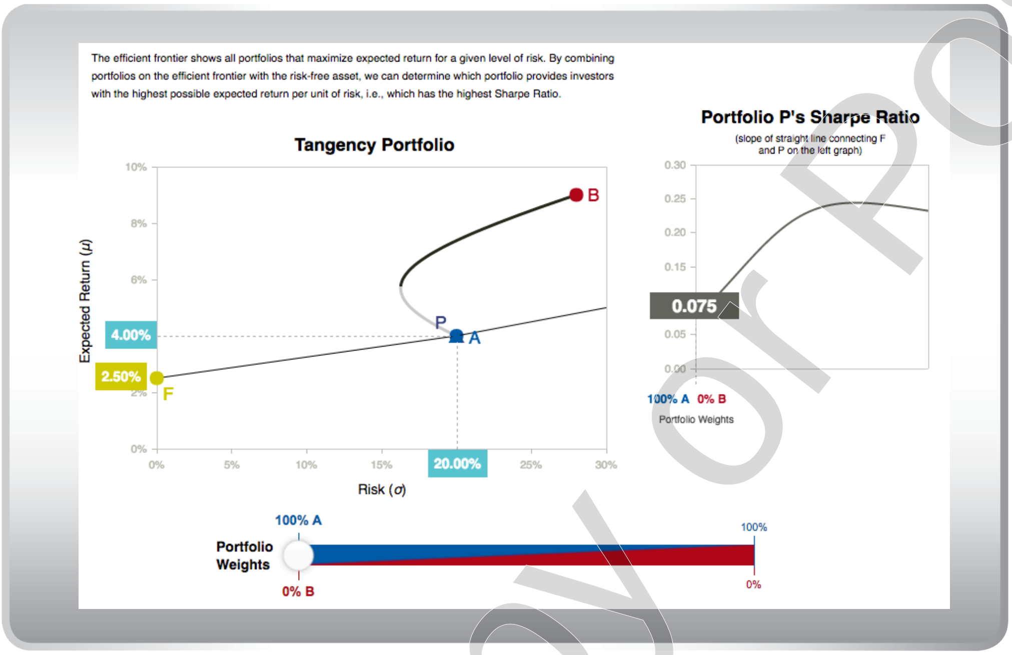
2.3.3 The Tangency Portfolio

Interactive Illustration 5 allows you to examine possible combinations of each portfolio on the efficient frontier with a risk-free asset. We'll work with a simple world in which there are only two risky assets, A and B. The illustration initially shows the convex curve representing all possible combinations of A and B. Assets A and B are correlated, but not perfectly: $\rho_{A,B}$ = 0.40. The efficient frontier lies on this curve (it is the portion of the curve above the minimum-σ portfolio, shown in black, and would look the same, qualitatively, even if there were many risky assets). The illustration also shows the risk-free asset F and all possible combinations of F and portfolio P, which comprises some combination of assets A and B. Initially, P is 100% A. As you move the slider for portfolio weights in P to the right, the illustration shows the corresponding combinations of F and P. Every such set of combinations consists of a straight line between the risk-free asset and a given efficient portfolio. Now notice that some portfolios on the efficient frontier produce better—more efficient—combinations with the risk-free asset than others. That is, the slope of the line between them and F is steeper—they have higher Sharpe Ratios—which means they represent superior combinations of expected return and risk. In effect, by allowing combinations of efficient portfolios with the risk-free asset, we create a new efficient frontier or, we might say, an even-*more*-efficient frontier.

[e] Some readers may object that in the real world it is not possible to borrow at the risk-free rate. In fact, large institutional investors can and do borrow large sums for short periods at rates practically equal to the risk-free rate. In any case, we could introduce a "spread" between rates at which one can borrow and lend; it would not change the fundamental insight in the text—that one can boost expected returns (and risk) by borrowing—while unnecessarily complicating the mathematics.

INTERACTIVE ILLUSTRATION 5 Tangency Portfolio

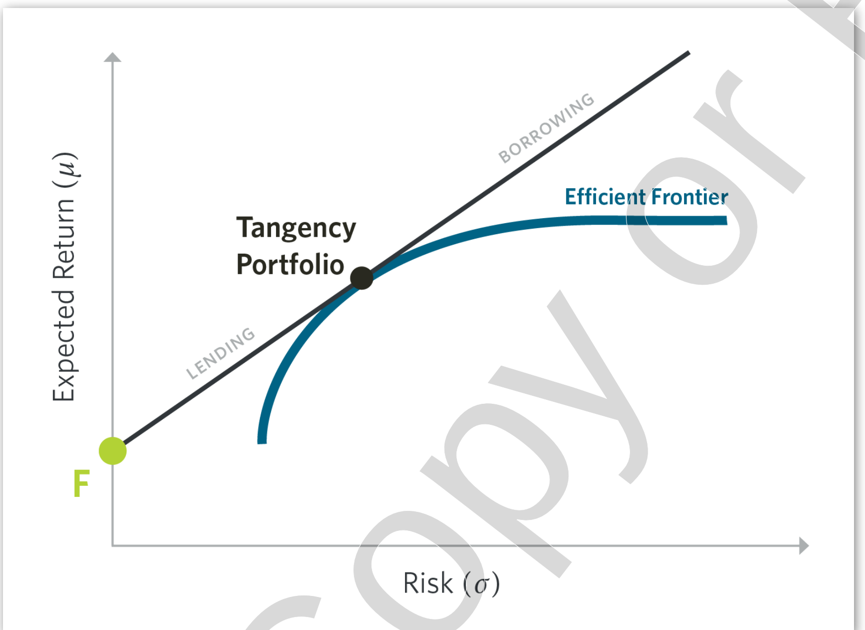
Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2pKuSYU



As the slope of the straight line connecting F and P increases, the Sharpe Ratio also increases, as shown in the box on the right of the illustration. Use the slider to examine different combinations of F and the efficient frontier until you find the one with the highest Sharpe Ratio. This is the point at which the slope of the line is maximized. The graph of the Sharpe Ratio on the right will confirm your choice. Where on the efficient frontier is the best portfolio located? At the point where a line connecting the risk-free asset to the efficient frontier is *tangent* to the efficient frontier. This portfolio is sometimes called the ***tangency portfolio***. In what sense is it best? In the same sense we used above: combinations of it with the risk-free asset produce the highest possible expected return for a given amount of risk. Put simply, it maximizes the Sharpe Ratio.

Let's summarize. In μ-σ space the efficient frontier represents the set of the best of all possible combinations of all risky assets (i.e., the highest μ for a given σ). Any of these (risky) efficient portfolios may be combined with a risk-free asset, and such combinations lie on the straight line connecting the risky portfolio and the risk-free asset. Only one efficient portfolio is the "best of the best": It is the tangency portfolio; it lies on a line running through the risk-free asset and tangent to the efficient frontier. Combinations of the tangency portfolio and the risk-free asset maximize the Sharpe Ratio. Offered this set of investment opportunities, risk-averse investors will choose a portfolio somewhere on this line connecting the risk-free asset and the tangency portfolio. Portfolios between those two points entail lending at the risk-free rate; points to the right of the tangency portfolio involve borrowing at the risk-free rate and investing the proceeds in the tangency portfolio. **Exhibit 4** shows the efficient frontier, the tangency portfolio, the line between them, and regions of borrowing and lending.

EXHIBIT 4 The Efficient Frontier and the Tangency Portfolio



More generally, we can now see the significance of the Sharpe Ratio in understanding the trade-off between risk and expected return. By construction, the Sharpe Ratio measures the incremental expected return associated with an additional unit of risk. Risk-averse investors prefer a higher Sharpe Ratio to a lower one. Another way to state such investors' objective is that they want to choose an investment portfolio that *maximizes* their Sharpe Ratio. What if all investors behave this way? Returning to Exhibit 4, we can say that, given the opportunity set depicted, *everyone will invest in some combination of the tangency portfolio and the risk-free asset*. All combinations of the tangency portfolio and the risk-free asset offer the same Sharpe Ratio—the line has the same slope everywhere—and, further, it represents the maximum achievable Sharpe Ratio, given the set of available risky assets. Individual preferences will lead different investors to choose different portfolios on the tangency line—different degrees of borrowing or lending—but the tangency portfolio will be the risky-asset portfolio of choice for everyone.

2.4 The Portfolio Improvement Rule

The framework developed so far is robust enough to address a practical question. Let's suppose that a risk-averse investor already holds a portfolio of risky assets and also that she can borrow or lend at the risk-free rate. A new investment opportunity comes along (it, too, is risky). Should she add a bit of the new risky asset (call it asset X) to the existing portfolio?

We have a straightforward rule to guide the investor's decision. The new asset should be added to the portfolio only if its addition increases the portfolio's Sharpe Ratio. An increase in the Sharpe Ratio implies an increase in the portfolio's return, per unit of risk. This rule is known as the ***portfolio improvement rule***.

How can we implement the portfolio improvement rule? The first step is to determine the investor's ***required expected return*** on X. The required expected return on X is the return that would leave the Sharpe Ratio of the portfolio *unchanged*. (Any return higher than this would boost the portfolio's Sharpe Ratio; any lower return would reduce it.) In other words, it is the minimum expected return on the new investment sufficient to offset the incremental effect of X on the portfolio's risk. Then the decision rule will be to invest only if the *actual* expected return on the investment, denoted as $E(r_x)$ or μ_X, is greater than the investor's *required* expected return (to be calculated).

Before we perform the analysis, let's prime our intuition with some qualitative insights drawn from the material above. There are two main issues. One is how much ***excess return*** per unit of risk the existing portfolio already earns. ("Excess return" is simply the return over and above the risk-free rate of return: $\mu_X - r_f$, the numerator of the Sharpe Ratio.) In other words, what is the existing portfolio's Sharpe Ratio? The second issue is how the new asset X will affect the portfolio's risk. As we have seen, this depends not only on the risk of X itself but also on the extent to which X's returns are correlated with those of the existing portfolio. If we can measure the effect of the new asset X on portfolio risk, we can find the expected return required to leave the Sharpe Ratio unchanged.

Interactive Illustration 6 is designed to apply the portfolio improvement rule using trial and error. The illustration examines a hypothetical situation in which an investor is considering adding a small amount of asset X to an existing portfolio P, which has an expected return $E(r_P)$ of 6.0% and risk σ_P of 19.3%. The inclusion of X would create a new portfolio N comprising 98% P and 2% X. Why might the investor be interested in adding a bit of X to the portfolio? The risk of asset X, σ_X, is known to be 15.3%, which is less than σ_P (19.3%). Further, asset X's returns are positively correlated with those of P, but not perfectly so: $\rho_{X,P}$ is 40%. So adding X to the portfolio should decrease portfolio risk. Indeed, we have enough information to calculate the risk σ_N of the new portfolio N. The question is what expected return on X, μ_X, is required to give the new portfolio N the same Sharpe Ratio as the existing portfolio P?

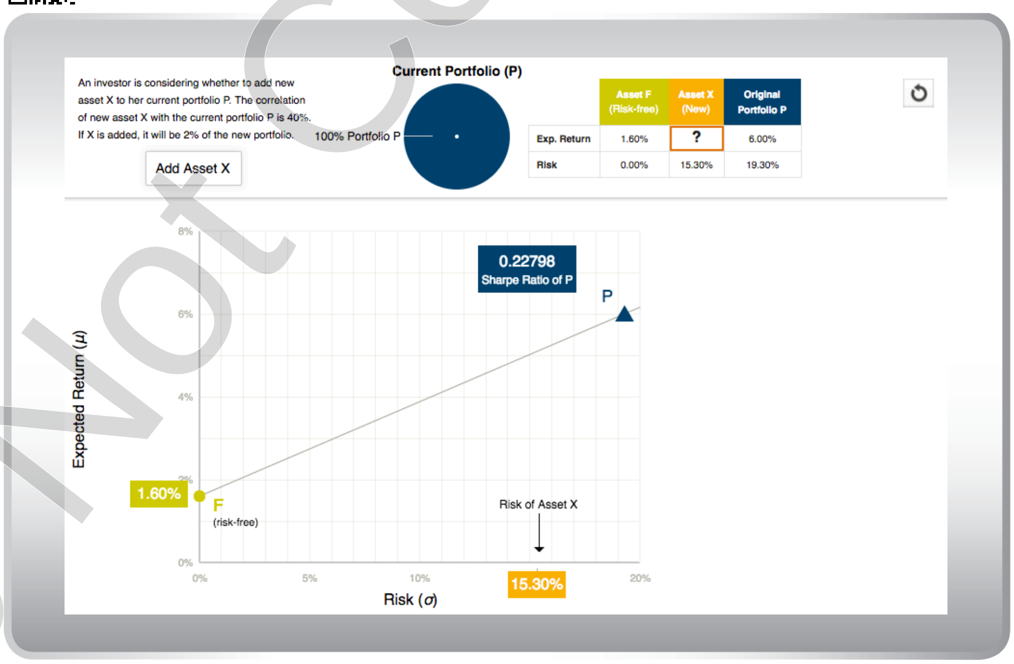
The existing portfolio's Sharpe Ratio is

$$\frac{\left(\mathrm{E}(r_P)-r_f\right)}{\sigma_P}=\frac{(6.0\%-1.6\%)}{19.3\%}=0.22798$$

as displayed in the illustration. The table at the upper right shows expected returns and risks for the risk-free asset F, asset X, and portfolio P. Click "Add Asset X" to see expected return and risk calculations for the new portfolio N. The dark gray curve represents all possible combinations of X and P, from 100% X to 100% P. Note that the new portfolio N is on that line with weights of 2% X and 98% P (also shown in the pie chart at the top).

INTERACTIVE ILLUSTRATION 6 Portfolio Improvement Rule

Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2GgEssO



At first, the illustration adopts an arbitrary expected return for asset X of 2.1%. Using this return for X, the illustration calculates expected return and risk for N, and then N's Sharpe Ratio. When the expected return for X is 2.1%, N's Sharpe Ratio is

0.22701, which is lower than P's Sharpe Ratio. What does this tell us? If asset X's expected return and risk are μ_X = 2.1% and σ_X = 15.3%, adding X would not be beneficial—it would lower the portfolio's Sharpe Ratio. While it's true that X lowers portfolio risk (σ_N = 19.04% < σ_P = 19.30%), an expected return on X of 2.1% is not high enough to keep the Sharpe Ratio constant. How much higher does X's expected return need to be before it makes sense to add it? Click on the magnifying icon to open a zoom view of the space around portfolio P and get a closer look at the new portfolio N relative to P. To raise the expected return on X, use the slider on the vertical axis of the graph. (Alternatively, you can type a value for X's return in the box to the left of the slider.) As you do so, the illustration shows the required expected return you selected in the table at the top, as well as the Sharpe Ratio for new portfolio N. Continue moving the slider until the new portfolio's Sharpe Ratio equals that of the existing portfolio. The expected return on X that equalizes the Sharpe Ratios for N and P, μ_X = 3.02%, is the required expected return for asset X. It is the minimum expected return that exactly compensates for X's incremental effect on portfolio risk. In the zoom box, you can confirm that the Sharpe Ratios of N and P are the same because N lies on the same line that connects F and P.

Continue increasing the expected return by moving the slider or typing in higher values for the return on X. Suppose the investor is told that the actual expected return on asset X is 3.50%. Should she make the investment? Yes. As the illustration shows, with a return of 3.5% (or, for that matter, any return greater than 3.02%), the new investment will improve the portfolio's Sharpe Ratio. This is the essence of the portfolio improvement rule: accept investments that raise the existing portfolio's Sharpe Ratio.

What would happen to the required expected return on asset X if its correlation with the existing portfolio were greater than 40%? We know that the *less* correlated X is, the more diversification (risk reduction) is achieved when we add it to the portfolio; the opposite is true when X is *more* correlated with the existing portfolio—it increases the risk of the portfolio, all else being equal.[f] That suggests that the required expected return on X should be higher for higher correlations between X and P. Let's check that hypothesis with some calculations.

Exhibit 5 maintains all the existing parameters for asset X and portfolio P, but it lets the correlation between the two vary from 0.20 to 0.70 and solves for the required

[f] Should you wish to review this point, see *Risk and Return 1*, especially Interactive Illustration 7.

expected return on X (denoted required E(r_X)) that equates the Sharpe Ratios of portfolios N and P.[g] As expected, as correlation increases, so does the required E(r_X).

EXHIBIT 5 Required Expected Return as a Function of Correlation

$\rho_{X,P}$	σ_X	σ_P	$\beta_X = \rho_{X,P}\dfrac{\sigma_X}{\sigma_P}$	$E(r_P)$	Required $E(r_X)$
0.20	0.1530	0.1930	0.1585	6.000%	2.298%
0.30	0.1530	0.1930	0.2378	6.000%	2.646%
0.40	0.1530	0.1930	0.3171	6.000%	2.995%
0.50	0.1530	0.1930	0.3964	6.000%	3.344%
0.60	0.1530	0.1930	0.4756	6.000%	3.693%
0.70	0.1530	0.1930	0.5549	6.000%	4.042%

Exhibit 5 also shows the ***beta*** of asset X, (β_X), measured with respect to the existing portfolio. We'll return to the relevance of beta shortly; for now, simply note that beta also increases as correlation increases, because $\beta_X = \rho_{X,P}(\sigma_X/\sigma_P)$.

A robust way to compare expected returns on risky assets is to express them in terms of the implied ***risk premium***. The risk premium is the difference between an asset's expected return and the risk-free rate (i.e., its expected excess return); it is the numerator of the Sharpe Ratio and reflects only that part of the expected return that depends on risk. In the example above, the expected return on portfolio P is 6.0% and the risk-free rate is 1.6% (as in Interactive Illustration 6), so the portfolio's risk premium is 4.4%. Similarly, we can express asset X's required expected return as a *required expected risk premium* by subtracting the risk-free rate. In Interactive Illustration 6 we found a required expected return for X of 3.02% (the corresponding figure in Exhibit 5 is 2.995%, in the row where correlation equals 0.40; the small difference is explained in footnote g). The implied required expected risk premium for X is 1.395% (= 2.995% – 1.6%). Note that this required risk premium for X is significantly smaller than the risk premium for portfolio P—about one-third the value. Let's examine relative risk premiums, that is,

[g] The values in Exhibits 5 and 6 are calculated on a basis similar to that used in Interactive Illustration 6, but the results differ slightly. The illustration assumes a portfolio weight for asset X of 2%; here we assume the weight of X approaches zero. The resulting difference is extremely small. For example, when ρ = 0.40, the result for the required E(r_x) using the interactive is 3.02%, and in Exhibit 5 it is 2.995%.

$$\frac{\left(\mathrm{E}\left(r_X\right)-r_f\right)}{\left(\mathrm{E}\left(r_P\right)-r_f\right)},$$

as we change the correlation between X and P.

Exhibit 6 reproduces the figures from Exhibit 5, but now includes the risk premiums for X and P (holding the risk-free rate constant at 1.6%) and the ratios between premiums in the last column.

EXHIBIT 6 Ratios of Risk Premiums for Asset X and Portfolio P

$\rho_{X,P}$	σ_X	σ_P	$\beta_X=\rho_{X,P}\frac{\sigma_X}{\sigma_P}$	$\mathrm{E}\left(r_P\right)$	Required $\mathrm{E}\left(r_X\right)$	Risk Premium, P $\mathrm{E}\left(r_P\right)-r_f$	Required Risk Premium, X $\mathrm{E}\left(r_X\right)-r_f$	$\frac{\left(\mathrm{E}\left(r_X\right)-r_f\right)}{\left(\mathrm{E}\left(r_P\right)-r_f\right)}$
0.20	0.1530	0.1930	0.1585	6.000%	2.298%	4.400%	0.698%	0.1585
0.30	0.1530	0.1930	0.2378	6.000%	2.646%	4.400%	1.046%	0.2378
0.40	0.1530	0.1930	0.3171	6.000%	2.995%	4.400%	1.395%	0.3171
0.50	0.1530	0.1930	0.3964	6.000%	3.344%	4.400%	1.744%	0.3964
0.60	0.1530	0.1930	0.4756	6.000%	3.693%	4.400%	2.093%	0.4756
0.70	0.1530	0.1930	0.5549	6.000%	4.042%	4.400%	2.442%	0.5549

Now notice that in every row the ratio of the risk premiums equals the beta. That is, asset X's beta (measured with respect to the existing portfolio P) equals X's required risk premium divided by the existing portfolio's risk premium:

$$\beta_X=\frac{\left(\mathrm{E}\left(r_X\right)-r_f\right)}{\left(\mathrm{E}\left(r_P\right)-r_f\right)}$$

Rearranging terms, we have

$$\left(\mathrm{E}\left(r_X\right)-r_f\right)=\beta_X\left(\mathrm{E}\left(r_P\right)-r_f\right)$$

This is an important result. The required risk premium for a proposed asset (one that will leave the Sharpe Ratio of a portfolio unchanged) equals the proposed asset's beta (measured with respect to the portfolio) times the portfolio's risk premium. Using this result, we can state the portfolio improvement rule in terms of an inequality:

Invest if

$$\left(\mathrm{E}\left(r_X\right)-r_f\right)>\beta_X\left(\mathrm{E}\left(r_P\right)-r_f\right)$$

Alternatively, we can state the rule in terms of the Sharpe Ratios for the proposed investment on the one hand and the existing portfolio on the other. Substituting $\rho_{X,P}\,(\sigma_X/\sigma_P)$ for β_X:

$$\left(\mathrm{E}\left(r_X\right)-r_f\right)>\rho_{X,P}\frac{\sigma_X}{\sigma_P}\left(\mathrm{E}\left(r_P\right)-r_f\right)$$

Rearranging:

$$\frac{\left(\mathrm{E}\left(r_X\right)-r_f\right)}{\sigma_X}>\rho_{X,P}\frac{\left(\mathrm{E}\left(r_P\right)-r_f\right)}{\sigma_P}$$

Stating the rule this way shows that the Sharpe Ratio of the proposed investment (the left side) must exceed the product of the investment's correlation with the existing portfolio and the Sharpe Ratio of the existing portfolio (the right side). Put yet another way, the ratio of the Sharpe Ratios must exceed the correlation coefficient.

Now that we have an equation (an inequality) for the Portfolio Improvement Rule, we can use it to find the required expected return on a new asset X, $\mathrm{E}(r_x)$, directly, without using the cumbersome trial-and-error method shown in Interactive Illustration 6. The required expected return on asset X that leaves a portfolio's Sharpe Ratio unchanged is given by

$$\mathrm{E}\left(r_X\right)=r_f+\beta_X\left(\mathrm{E}\left(r_P\right)-r_f\right)$$

Notice that the right side of this equation contains only one parameter specific to asset X: its beta. In effect, β_X measures the incremental risk that asset X will bring to the portfolio and for which the investor must be compensated, by means of the expected return on asset X.

To summarize, the Sharpe Ratio led us to a prescription for individual investors trying to decide whether to add a small amount of a new stock to their portfolios: the portfolio improvement rule. Using beta, we stated the portfolio improvement rule as an equation that gives us the required expected return on a given stock as a function of its beta:

$$\mathrm{E}\left(r_X\right)=r_f+\beta_X\left(\mathrm{E}\left(r_P\right)-r_f\right)$$

To broaden this result, we need to extend the rule to cover all investors and all risky assets. In other words, what happens if everyone behaves like the individual investor we just considered? Under certain assumptions, this leads to a general market equilibrium described by the capital asset pricing model. Not surprisingly, the CAPM equation that gives the required expected return for a risky asset looks very much like the portfolio improvement rule.

2.5 Capital Asset Pricing Model

The ***capital asset pricing model*** **(*CAPM*)** is one of the most famous theories in modern finance.[h] It addresses the relationship between risk and expected return when the market for risky assets is in equilibrium—a state in which all risky assets are priced so that risk-averse investors collectively are willing to hold them. Put differently, equilibrium prices are prices that are "just right": no assets are overpriced (in which case no one would want to own them) and no assets are underpriced (everyone would want to own only them). CAPM's fundamental result is that not all risks affect equilibrium expected returns—only those risks that cannot be diversified away. If this sounds similar to the portfolio improvement rule, it is.

The portfolio improvement rule gives us the required expected return on a risky asset that leaves the Sharpe Ratio of a portfolio unchanged. From there it is straightforward to derive the capital asset pricing model.[1] Four key assumptions are required:

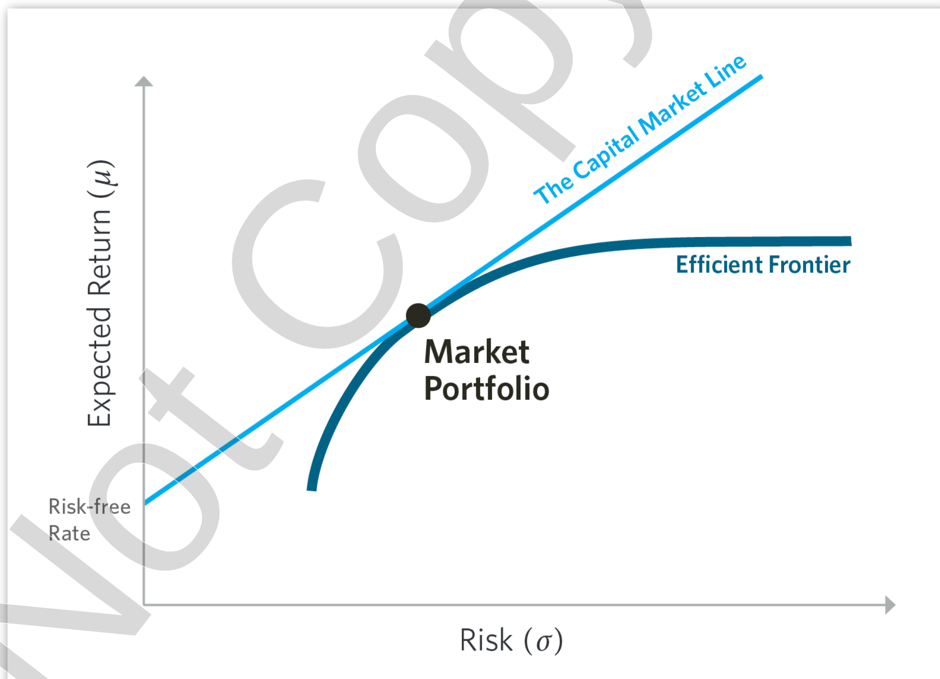
- Investors are risk-averse and evaluate their investment portfolios solely on the basis of expected returns and standard deviation of returns measured over the same holding period.
- Capital markets are perfect; in other words, all assets are infinitely divisible, there are no transaction costs, information is costless and available to everyone, and all investors can borrow and lend at a single risk-free rate.
- All investors share the same set of investment opportunities.

[h] CAPM was developed in the early 1960s by William Sharpe, Jack Treynor, John Lintner, and Jan Mossin. William Sharpe, Harry Markowitz, and Merton Miller jointly received a Nobel Prize in 1990 for their groundbreaking work in portfolio theory.

- Investors' expectations are homogeneous—they all have the same estimates of individual assets' expected returns, standard deviations of returns, and correlations among asset returns.

Under these assumptions, given prevailing prices, all investors will choose the *same* portfolio of risky assets—the one with the highest Sharpe Ratio—a result we suggested above. They may assign different portfolio weights to the optimal risky portfolio and the risk-free asset (e.g., some may borrow while others lend), but all will select a portfolio lying on the same line in μ-σ space. This line is shown in **Exhibit 7** and is called the ***capital market line*** (***CML***). Put another way, investors may have different risk preferences (we have assumed that they all are risk-averse, but not that they have identical degrees of risk aversion) and hence prefer different portfolios on the CML. But every portfolio on the CML represents some combination of the risk-free asset and the unique tangency portfolio.

EXHIBIT 7 The CAPM and the Capital Market Line



In equilibrium every asset is priced so that investors collectively will hold exactly the supply of all available assets—the demand for risky assets must equal the supply. Further, all investors will hold risky assets in the same relative proportions[i]—that is,

[i] Some readers may find a comparison to well-known market index funds helpful in thinking about the "equal proportions" aspect of the CAPM. Owners of an index fund all hold the same *percentage* of a

they all hold the tangency portfolio—some hold more and others less, but portfolio weights (for individual securities) are the same for everyone. Since all investors collectively hold all risky assets, and in the same proportions, and since they all hold the portfolio with the highest attainable Sharpe Ratio, in equilibrium that portfolio of risky assets with the highest Sharpe Ratio—the tangency portfolio—must be *the market portfolio*. Think of the market portfolio simply as a giant portfolio that contains all risky assets. How can everyone hold it? Since assets are infinitely divisible, we can create as many "mini" market portfolios—each with a pro rata share of all risky assets—as necessary. We've already determined that the tangency portfolio is the optimal (and only) portfolio investors will hold; therefore, in equilibrium the tangency portfolio must also be the market portfolio.

Because the market portfolio has the highest attainable Sharpe Ratio, there is no way to improve its risk-adjusted return by holding more or less of a given asset. Therefore, just as it did above, the portfolio improvement rule itself must give the equilibrium risk premium for each asset:

$$\left(\mathrm{E}\left(r_X\right)-r_f\right)=\beta_X\left(\mathrm{E}\left(r_P\right)-r_f\right).$$

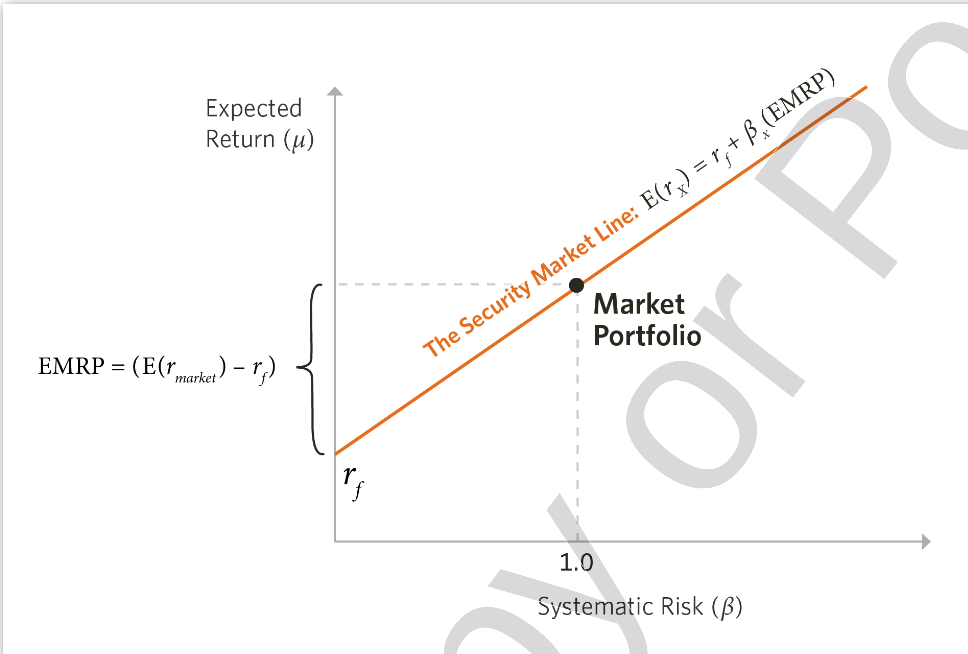
In other words, we can now substitute the *market* portfolio for the investor's existing portfolio in our expression for required expected return. (And now beta for any given security is measured with respect to returns on the market portfolio.) This gives us the fundamental equation of the CAPM, which states that, in equilibrium, the expected return of a given asset, $\mathrm{E}(r_x)$, is denoted by

$$\mathrm{E}\left(r_X\right)=r_f+\beta_X\left(\mathrm{E}\left(r_{market}\right)-r_f\right).$$

In equilibrium all assets are priced so that the expected return on any given asset equals the risk-free return plus a risk premium consisting of the asset's beta multiplied by the risk premium for the market as a whole. Another way to say this is that in the CAPM's equilibrium, the expected return for every security equals its *required* expected return. If this were not the case, investors would be able to further improve their portfolios and prices would adjust until actual $\mathrm{E}(r_x)$ = required $\mathrm{E}(r_x)$. The CAPM equation is also called the ***security market line*** **(*SML*)**. The SML is graphed in **Exhibit 8**, which shows expected return as a function of beta.

given stock in their portfolio—that is, all investors will own the same percentage of Apple, Inc., stock—even though the dollar amount of their investments in the fund differs.

EXHIBIT 8 The CAPM and the Security Market Line



Note that the SML is graphed in μ-β space rather than μ-σ space, and the slope of the line equals the market risk premium, $(E(r_{market}) - r_f)$. The beta of the market portfolio must be 1.0—it is the beta of the market measured with respect to the market. At every point on the SML, the *expected* return of a security is equal to its *required* return ($E(r)$ = required $E(r)$); that is, all securities on the SML are fairly priced. In the CAPM equilibrium, *all* traded assets fall on the SML; if one were to fall above or below the line, prices would adjust until its expected return equaled its required expected return—in other words, until all assets were once again on the SML.

We can now summarize a few implications of the CAPM and the underlying relationships:

- The expected return on an individual asset depends on β rather than σ. Beta measures the stock's nondiversifiable, or systematic, risk.

- A portfolio's β equals the value-weighted average of the betas of its individual securities. In other words, betas combine *linearly* (as we have seen, this is not true for σ).

- Standard deviation, σ, is the pertinent measure of risk for diversified portfolios, but β is the pertinent measure for individual securities. This is not to say that an individual security's σ does not measure risk; rather, σ for an individual stock measures *both* diversifiable and nondiversifiable (both nonsystematic and systematic) risk.

These points deserve some amplification. We discovered in *Risk and Return 1* that β measures an individual stock's co-movement with the market. It can be expressed in terms of either covariance or correlation with the market, and it is estimated as the slope of a linear regression of the stock's returns on the market's returns. But this is a statistical observation, not an economic one. As we have discovered in this reading, the economic significance of beta is that it measures an asset's systematic risk. As such, beta is the key to relating risk and expected return. How do we know that β measures systematic risk? We established that investors are compensated only for systematic (i.e., nondiversifiable) risk. And the portfolio improvement rule, generalized to the CAPM's SML, states that a security's required expected return depends on its beta. If investors are compensated only for systematic risk, and expected returns (over and above the risk-free rate) depend on beta, then beta must be a proxy for systematic risk.

In short, the logical measure of risk for a diversified portfolio is the standard deviation (σ) of portfolio returns. The expected return on a diversified portfolio compensates investors for the portfolio's standard deviation. But for an *individual asset*, the expected return depends not on the standard deviation of the asset's returns, but rather on the *incremental* risk it contributes to a diversified portfolio. According to the CAPM, the optimal diversified portfolio is the market portfolio, and an individual asset's beta measures the risk it contributes to the market portfolio. Beta does this in relative, not absolute, terms: It measures an individual asset's incremental systematic risk relative to the market's systematic risk.

3 SUPPLEMENTAL READING

3.1 Using the CAPM

The CAPM is widely used in the real world, primarily as a way to estimate discount rates used in discounted cash flow (DCF) valuations of risky securities, such as equities, but also risky assets in general, such as business assets and even whole business entities. In a DCF valuation, the value of some risky asset X is computed as

$$PV_X = \sum_{t=1}^{T} \frac{E\left(Cf_{X,t}\right)}{\left(1+E\left(r_X\right)\right)^t}$$

where $E(Cf_{X,t})$ denotes the expected cash flow for asset X at each future date t, T denotes the number of periods in the calculation, and $E(r_X)$ denotes the required expected return on asset X. Now we can use the CAPM equation (the SML) to estimate $E(r_X)$:[j]

$$E\left(r_X\right) = r_f + \beta_X\left(E\left(r_{market}\right) - r_f\right)$$

In words, the required expected return on asset X equals the risk-free rate plus a risk premium; the appropriate risk premium equals the beta of asset X times the risk premium for the market as a whole, $(E(r_{market}) - r_f)$. (Sometimes the equity market risk premium is abbreviated as EMRP or MRP.) Both the risk-free rate and the EMRP are market parameters—they are not specific to asset X. The only asset-specific component of the SML is β_X—the beta of asset X measured with respect to the market.

3.2 Estimating Betas

Stock betas are widely published and available in periodicals, in databases, and on websites. Alternatively, they can be calculated using the tools presented in *Risk and Return 1* (e.g., by regressing a stock's returns on corresponding market returns; the slope of this regression is the stock's beta). It is important to know that if you obtain a beta for a given stock from two different sources, they may differ. Why should this be so? Remember that beta is a statistic and depends on the data from which it is calculated. Different sources may have used different data—the sample periods may differ, for example, or the frequency of observation (say, weekly versus monthly). Some sources use simple linear regressions; others make adjustments to the regression intended to overcome the effects of known or possible anomalies in the data. Understanding differences in published betas requires some research to determine how they were estimated.

What if the business being evaluated is not publicly traded? This is a common problem, but the CAPM is still widely relied on. Common practice is to identify publicly traded businesses that are similar to the one being evaluated and use *their* betas (or an average of them) in the SML to estimate the required expected return on

[j] For a more complete exposition of how this is done with real data, see *Core Reading*: *Cost of Capital* (HBP No. 8293).

the nontraded business. The idea is that similar businesses have similar amounts of systematic risk and should therefore have similar risk premiums. Establishing similarity is not always easy, but the idea is sound and widely used.

3.3 Risk-Free Rate

The risk-free rate is not asset-specific; it is the same regardless of the particular asset or business we are evaluating. In theory, the risk-free rate is easily defined: it is the return an investor earns on a risk-free investment over a given holding period. In reality, the problem is more complex. First, in the real world there is no perfectly risk-free investment. For investments in US dollars, the return on US government securities is commonly used as the risk-free rate; US Treasury obligations are widely regarded as the closest thing to a genuinely risk-free investment. Second, the time horizon matters because the term structure of US Treasury rates may not be flat. That is, on any given day, the quoted rates on 1-, 5-, 10-, and 20-year Treasury obligations are likely to be different from one another. Which one(s) should we use? In theory, we may need all of them—the whole term structure. That is, the expected cash flows for year 1 should be discounted at a rate in which the 1-year risk-free rate is embedded, expected cash flows for year 2, at a rate in which the 2-year risk-free rate is embedded, and so forth. Some investments are indeed evaluated this way (e.g., carefully constructed hedging portfolios), whereas other analyses ignore this complication and simply pick one rate that corresponds as nearly as possible to the longest relevant horizon. (Many corporate applications, such as evaluating capital budgeting proposals, use this approach.)

3.4 Equity Market Risk Premium

The ***equity market risk premium***, or ***EMRP*** (also known as the market risk premium), $(E(r_{market}) - r_f)$, is simply the amount by which the return on the portfolio of all risky assets is expected to exceed the risk-free return over a given holding period. Simple as it sounds, experts disagree about how best to estimate it. One common approach is to use an average of historical returns on the market over and above corresponding historical risk-free returns. This approach yields estimates in the range of about 4% to 7.5%, depending on the time period selected and the particular data set used. Another approach is to survey market participants, in effect asking them what they expect the market risk premium to be over the next N years.

Conclusions of such surveys depend on who is surveyed (finance professionals, finance professors, finance textbooks, or corporate executives, for example), where they reside, and how the questions are asked. In general, however, survey results have exhibited two patterns. First, they report somewhat lower values for the EMRP than do historical returns data, and, second, the mean EMRP reported in surveys has fallen over the past 30 years or so—the same types of respondents are using lower estimates of the EMRP today than their counterparts used in the 1980s. Finally, another method for estimating the EMRP is to infer it not from historical time series but from aggregated data on corporate earnings growth, dividends, and the level of the stock market. This so-called implied risk premium for the US stock market has averaged about 4% over the past 40 years and has exceeded 6% only once.

While there is disagreement about what value to assign the EMRP in applications of the CAPM, the practical reality for many analysts is that they needn't decide themselves. In most large firms—corporations, banks, investment banks, consulting practices, appraisal firms, and so on—a single value is chosen and applied throughout the firm. The value is reviewed periodically in light of market conditions and observations about what value other professionals are using, but the estimate tends to be "sticky"—it doesn't often change within a given firm.

3.5 A Simple Application of the CAPM

Suppose at the end of 2015 we wished to estimate the expected return on the stock of McDonald's Corporation during the upcoming calendar year, 2016. Perhaps we own the stock itself (we certainly do if we own a piece of the market portfolio). Or maybe we are analyzing some other business that we think has systematic risk similar to McDonald's in 2016. For either purpose, we can use the CAPM to estimate a required expected return. Recall the CAPM (SML) equation:

$$E(r_X) = r_f + \beta_X \left(E(r_{market}) - r_f\right)$$

We'll use 5% as our estimate of the EMRP (i.e., $(E(r_{market}) - r_f) = 5\%$). We need to look up a risk-free rate and a beta for McDonald's as of year-end 2015. Our holding period is one year, so we look up the one-year return on US Treasury obligations. On December 31, 2015, it was 0.65%.[2] Next we look up a beta for McDonald's and find it is 0.79.[3] This estimate implies that McDonald's stock is less risky than the stock market as a whole—a 1% change in the market tends to be accompanied by a change of about 0.79% in the price of McDonald's. The rest of the exercise is simple arithmetic:

$$E(r_{McDonald's}) = 0.65\% + 0.79(5\%) = 4.60\%$$

So our estimate of the expected return on McDonald's stock during 2016 is 4.60%, according to the CAPM. What does this mean? McDonald's stock closed at a price of \$118.14 per share at the end of 2015, and it paid annual dividends of \$3.56 per share (note, though, that dividends are subject to change). The dividend yield equals the dividend divided by the stock price, or 3.01% (= \$3.56 / \$118.14). If the dividend remains the same in 2016, then we'll receive about 3% of our expected 4.6% return in the form of dividends. The remaining 1.6% must come in the form of an expected capital gain. To offer a capital gain of 1.6%, the expected stock price at the end of 2016 must be \$120.03 (= \$118.14 × 1.016). Of course, there is no guarantee this will happen—the stock is risky—this is simply an estimate of 2016's return derived from a model—the CAPM. Is it realistic? In terms of the portfolio theory we have developed, yes. Someone who owns McDonald's shares deserves a higher expected return than the risk-free rate. The latter is only 0.65%, so 4.6% is considerably higher. And statistical data (β, specifically) shows that McDonald's has less systematic risk than the stock market as a whole, so our theory implies that the expected return on McDonald's should be lower than the expected return on the market. How much lower? The expected return on the market equals 5.65% (= 0.65% + 1.0(5%)) compared to McDonald's 4.6%, so about 1% lower. But remember that both of these returns—the market's and McDonald's—may well turn out to be different when 2016 is over.

3.6 Criticisms of the CAPM

The capital asset pricing model is one of the most important achievements of modern financial economics. It has been widely praised and, perhaps inevitably, widely criticized. Just as some of the praise has been lavish, some of the criticism has been harsh. In this section we'll mention some of the common types of criticisms and, where possible, assess their validity.

The first serious critiques of the CAPM emerged from early attempts to test it empirically using data drawn primarily from real-world stock returns. Researchers wanted to know whether stock returns would plot on a security market line as described by the CAPM. And, if they did not, were the apparent differences purely random, or could some of them be explained by other economic variables not contained in the CAPM? The answers generally were that stock returns were not entirely explained by the CAPM, and that some of the apparent errors of the CAPM could in fact be explained by other factors. This raised further questions. Perhaps the basic idea was right, but the model's specifications needed to be refined to accommodate one or more market imperfections. For example, what would the model look like if there were a spread between the rates at which one could lend and borrow? What if there was no truly risk-free asset? What if there were different countries and currencies? What if different investors faced different taxes or regulations? Or possessed different information? And so forth.

Some questions were of a more fundamental nature. What if investors are not risk-averse or, even more fundamentally, are simply not rational in the way the model presumed? Or what if stock returns do not obey normal distributions? The significance of these issues depends in large part on the answers to follow-up questions: How do investors actually behave? Are departures from "rationality" random or somewhat predictable? Or, how do asset returns actually behave? If they are not normal, what other probability distribution describes them? Can we specify and give it parameters? And so forth. Some of the research spawned by these questions led to more elaborate models and still more empirical testing.

In 1977 the financial economist Richard Roll published a famous critique of the (empirical) critiques.[4] Roll observed that the CAPM in fact has only one testable implication, namely, that the market portfolio is efficient in the μ-σ sense described earlier—in effect, that the market portfolio lies on the efficient frontier. He went on to observe that because the true market portfolio contains *everything* one could invest in—including illiquid, intangible, nontraditional, and nontraded assets of all sorts—the theory is practically untestable. In effect, tests of the CAPM based on stock market data are joint tests of both the CAPM equilibrium and whether broad indices such as the S&P 500 are good proxies for the true market portfolio. We can't tell whether a "failed" test means that the CAPM failed or the proxy failed (or both).

Research in the fields of psychology and behavioral finance provides compelling evidence that investors' behavior is not as simple as presumed by the CAPM (and many other economic theories)—that humans are not simply risk-averse wealth maximizers who care only about their portfolios' expected returns and standard deviations. There appear to be important ways in which people depart from so-called rational behavior in their decision making, and the departures are neither rare nor random. Certainly, we observe many investors who deliberately choose not to be fully diversified, some dramatically so. Having said that, these fields of research have not so far produced a general equilibrium model of asset prices under alternative investor behavior.

Some of the harshest criticism of the CAPM has come from practitioners—investment professionals in fields such as venture capital, real estate development, and private equity—who claim that its implications are simply and starkly different from economic reality as they know it. They claim that required investment returns in venture capital, for example, are simply too high to be predicted or explained by the CAPM. Put differently, asset prices implied by CAPM-predicted returns grossly overvalue such assets. More often than not, such criticisms are unfair or at least misplaced. A common problem is that the critics are applying CAPM-generated returns as discount rates to future values (or cash flows) that are statistically inappropriate. As a simple example, suppose that a risky venture capital investment will be worth \$1,000 one year from now, assuming an underlying key technology succeeds in the meantime. An investment adviser tells the venture capitalist that, based on the project's beta, CAPM says the required expected return on the investment is 10%. The VC investor explodes with indignation. "That implies a value today of more than \$900 (= \$1,000/(1 + 10%) = \$909)! I *know* that if we held an auction *no one* would even pay \$750 for this thing! It requires a return of at least 35%, not 10%." And sure enough, \$1,000/(1 + 35%) = \$740. But these calculations ignore the possibility of failure. Suppose there is an 80% chance of success and a 20% chance of failure—in which case the future value of the project is \$0. Then the expected value one year from now is not \$1,000 but \$800 (= \$1,000 × 0.80). Discounting \$800 at the CAPM rate of 10% gives a present value of \$727, actually below the VC investor's \$740. The investor is right (pretty much) about value, but wrong (almost entirely) about what the CAPM means. He or she is applying a CAPM-derived discount rate not to the *expected* future value (or cash flow) but to a *conditional expected* future value—conditional on technological success. The CAPM-derived discount rate applies to the *un*conditional expected future value: \$800, not \$1,000. It is true that the CAPM is "wrong" for discounting \$1,000 in this case. But that is not the fault of the model.

The misunderstanding of the VC investor is not uncommon, and part of the fault lies with CAPM's proponents and teachers. The CAPM's content is commonly summarized with a phrase such as "CAPM shows that only systematic (non-diversifiable) risk matters." Such a summary is often followed by an example or two intended to illuminate the difference between diversifiable and nondiversifiable risk. Moreover, "technological risk" is likely to be dismissed as obviously diversifiable (nonsystematic). In other words, the "right" discount rate is 10%, not 35%, because technological risk clearly is diversifiable and doesn't matter. To say that technological risk doesn't matter is silly—of course it matters. But in our framework it matters for *cash flows, not discount rates*. If we ignore the technological uncertainty in *both* cash flows (we discount \$1,000 instead of \$800) *and* discount rates (we use 10% instead of 35%), we will indeed get the wrong value. This is not CAPM's fault—it's ours.

Finally, we want to point out that the basic CAPM equation—the SML—actually can be derived more compactly than we did in this pair of *Risk and Return* readings.

Why did we take the long way? For two main reasons. First, to make sure readers followed the progression from prices to returns to points in μ-σ space, and then from combinations of two assets to many assets and to the "magic" of diversification in portfolios. Second, to avoid the most restrictive assumptions of the CAPM for as long as possible. We were able to state the portfolio improvement rule without the final assumptions needed for the CAPM's general equilibrium. The portfolio improvement rule looks a lot like the CAPM's SML, but it works whether a particular investor—to say nothing of all investors—is well diversified or not. It does not suppose that all investors have the same information and expectations, or imply that the market portfolio is efficient. Instead, it justifies and implements this simple advice: maximize your Sharpe Ratio. This is quite sensible for many investors, even those who are reluctant to accept the full list of CAPM assumptions. In short, modern portfolio theory has a lot to offer even to CAPM skeptics.

4 KEY TERMS

beta A statistical measure of the systematic risk of an asset. Betas for traded assets such as stocks and bonds may be estimated using linear regression of asset returns on market returns. *See also* systematic risk.

capital asset pricing model (CAPM) An economic model that relates the expected return on a risky asset to the expected return on the overall market. *See also* security market line.

capital market line (CML) The line in μ-σ space passing through the risk-free asset and the tangency point on the efficient frontier. The slope of the CML represents the maximum attainable expected return per unit of risk; that is, it maximizes the Sharpe Ratio.

diversification In investing, the reduction of portfolio risk that is due to combining two or more assets that are less than perfectly positively correlated.

efficient frontier The subset of possible combinations of two or more risky assets that maximizes expected return for a given level of risk.

equity market risk premium (EMRP) The amount by which the return on the portfolio of all risky assets (defined as "the market") is expected to exceed the risk-free return over a given holding period: $(E(r_{market}) - r_f)$.

excess return The portion of an investment's expected return that compensates investors for risk, that is, the expected return in excess of the risk-free rate.

portfolio improvement rule A rule stating that investors should add a new security to their portfolios only if doing so increases the risk-adjusted return of their portfolios (i.e., the portfolios' Sharpe Ratio).

required expected return The lowest expected return at which investors are willing to hold a risky asset.

risk premium The return of an asset (or portfolio) over and above the risk-free rate. *See also* excess return.

security market line (SML) In the CAPM, the SML relates an asset's expected return to its systematic risk, or beta, in equilibrium. According to the SML, an asset's expected return equals the risk-free rate plus a premium; the premium equals the asset's beta times the EMRP. The equation for the SML is $\mathrm{E}(r_X) = r_f + \beta_X(\mathrm{EMRP})$.

Sharpe Ratio The expected excess return of an investment or portfolio per unit of risk. See the formula in the text.

systematic (nondiversifiable) risk Risk that cannot be eliminated through diversification; that is, market risk. *See also* beta.

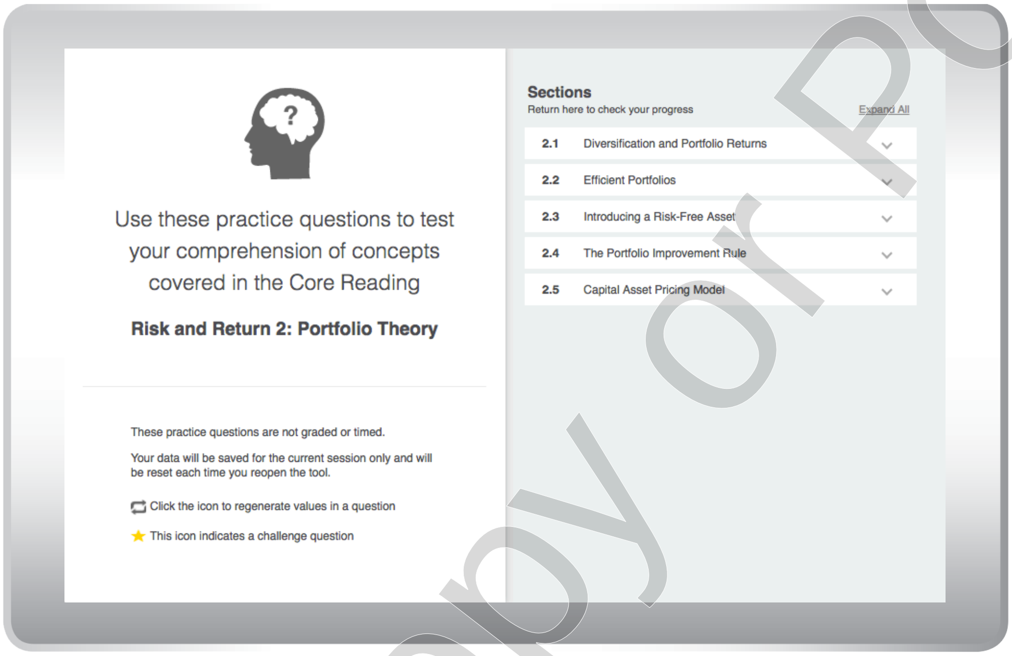
tangency portfolio The portfolio on the efficient frontier that maximizes the Sharpe Ratio. The tangency portfolio is the point at which the line connecting the risk-free asset to the efficient frontier is tangent to the efficient frontier.

5 NOTATION

β_X	the beta of risky asset X with respect to a specific portfolio (usually the market portfolio); a measure of the incremental risk added to a specific portfolio by the addition of asset X
$E(Cf_{X,t})$	the expected future cash flow of asset X at time t
$E(r_p)$	the expected return on a portfolio
$E(r_x)$	expected return on asset X
EMRP	equity market risk premium, $(E(r_{market}) - r_f)$
PV_X	present value of risky asset X
ρ	correlation coefficient
r_f	risk-free rate
r_m, r_{market}	return on the market portfolio
σ	standard deviation
σ^2	variance
μ	expected return or mean value
w	the weight of a security in a given portfolio

6 PRACTICE QUESTIONS

Scan this QR code, click the image, or use this link to access the interactive illustration: bit.ly/hbsp2Ge2Ye1



7 ENDNOTES

1 See Andre F. Perold, "The Capital Asset Pricing Model," *Journal of Economic Perspectives*, 18, no. 4 (2004): 3-24. Perold includes a comprehensive bibliography.

2 See US Department of the Treasury, "Daily Treasury Yield Curve Rates," https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2015 (accessed January 1, 2016).

3 This beta was reported on Yahoo! Finance's website at http://finance.yahoo.com/q?s=MCD&ql=0 (accessed January 1, 2016) but was estimated from data supplied by other vendors, in this case Standard & Poor's Capital IQ.

4 Richard Roll, "A Critique of the Asset Pricing Theory's Tests, Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics* 4, no. 2 (1977): 129–176.

8 INDEX