A fundamental question in finance is how the risk of an investment should affect its expected return. The Capital Asset Pricing Model (CAPM) provided the first coherent framework for answering this question. The CAPM was developed in the early 1960s by William Sharpe (1964), Jack Treynor (1962), John Lintner (1965a, b) and Jan Mossin (1966).

The CAPM is based on the idea that not all risks should affect asset prices. In particular, a risk that can be diversified away when held along with other investments in a portfolio is, in a very real way, not a risk at all. The CAPM gives us insights about what kind of risk is related to return. This paper lays out the key ideas of the Capital Asset Pricing Model, places its development in a historical context, and discusses its applications and enduring importance to the field of finance.

Historical Background

In retrospect, it is striking how little we understood about risk as late as the 1960s—whether in terms of theory or empirical evidence. After all, stock and option markets had been in existence at least since 1602 when shares of the East India Company began trading in Amsterdam (de la Vega, 1688); and organized insurance markets had become well developed by the 1700s (Bernstein, 1996). By 1960, insurance businesses had for centuries been relying on diversification to spread risk. But despite the long history of actual risk-bearing and risk-sharing in organized financial markets, the Capital Asset Pricing Model was developed at a time when the theoretical foundations of decision making under uncertainty were relatively new and when basic empirical facts about risk and return in the capital markets were not yet known.

André F. Perold is the George Gund Professor of Finance and Banking, Harvard Business School, Boston, Massachusetts. His e-mail address is (aperold@hbs.edu).
Rigorous theories of investor risk preferences and decision-making under uncertainty emerged only in the 1940s and 1950s, especially in the work of von Neumann and Morgenstern (1944) and Savage (1954). Portfolio theory, showing how investors can create portfolios of individual investments to optimally trade off risk versus return, was not developed until the early 1950s by Harry Markowitz (1952, 1959) and Roy (1952).

Equally noteworthy, the empirical measurement of risk and return was in its infancy until the 1960s, when sufficient computing power became available so that researchers were able to collect, store and process market data for the purposes of scientific investigation. The first careful study of returns on stocks listed on the New York Stock Exchange was that of Fisher and Lorie (1964) in which they note: “It is surprising to realize that there have been no measurements of the rates of return on investments in common stocks that could be considered accurate and definitive.” In that paper, Fisher and Lorie report average stock market returns over different holding periods since 1926, but not the standard deviation of those returns. They also do not report any particular estimate of the equity risk premium—that is, the average amount by which the stock market outperformed risk-free investments—although they do remark that rates of return on common stocks were “substantially higher than safer alternatives for which data are available.” Measured standard deviations of broad stock market returns did not appear in the academic literature until Fisher and Lorie (1968). Carefully constructed estimates of the equity risk premium did not appear until Ibbotson and Sinquefield (1976) published their findings on long-term rates of return. They found that over the period 1926 to 1974, the (arithmetic) average return on the Standard and Poor’s 500 index was 10.9 percent per annum, and the excess return over U.S. Treasury bills was 8.8 percent per annum.¹ The first careful study of the historical equity risk premium for UK stocks appeared in Dimson and Brealey (1978) with an estimate of 9.2 percent per annum over the period 1919–1977.

In the 1940s and 1950s, prior to the development of the Capital Asset Pricing Model, the reigning paradigm for estimating expected returns presupposed that the return that investors would require (or the “cost of capital”) of an asset depended primarily on the manner in which that asset was financed (for example, Bierman and Smidt, 1966). There was a “cost of equity capital” and a “cost of debt capital,” and the weighted average of these—based on the relative amounts of debt and equity financing—represented the cost of capital of the asset.

The costs of debt and equity capital were inferred from the long-term yields of those instruments. The cost of debt capital was typically assumed to be the rate of interest owed on the debt, and the cost of equity capital was backed out from the cash flows that investors could expect to receive on their shares in relation to the current price of the shares. A popular method of estimating the cost of equity this way was the Gordon and Shapiro (1956) model, in which a company’s dividends are

¹ These are arithmetic average returns. Ibbotson and Sinquefield (1976) were also the first to report the term premium on long-term bonds: 1.1 percent per annum average return in excess of Treasury bills over the period 1926–1974.
assumed to grow in perpetuity at a constant rate \( g \). In this model, if a firm’s current dividend per share is \( D \), and the stock price of the firm is \( P \), then the cost of equity capital \( r \) is the dividend yield plus the dividend growth rate: \( r = \frac{D}{P} + g \).\(^2\)

From the perspective of modern finance, this approach to determining the cost of capital was anchored in the wrong place. At least in a frictionless world, the value of a firm or an asset more broadly does not depend on how it is financed, as shown by Modigliani and Miller (1958). This means that the cost of equity capital likely is determined by the cost of capital of the asset, rather than the other way around. Moreover, this process of inferring the cost of equity capital from future dividend growth rates is highly subjective. There is no simple way to determine the market’s forecast of the growth rate of future cash flows, and companies with high dividend growth rates will be judged by this method to have high costs of equity capital. Indeed, the Capital Asset Pricing Model will show that there need not be any connection between the cost of capital and future growth rates of cash flows.

In the pre-CAPM paradigm, risk did not enter directly into the computation of the cost of capital. The working assumption was often that a firm that can be financed mostly with debt is probably safe and is thus assumed to have a low cost of capital; while a firm that cannot support much debt is probably risky and is thus assumed to command a high cost of capital. These rules-of-thumb for incorporating risk into discount rates were ad hoc at best. As Modigliani and Miller (1958) noted: “No satisfactory explanation has yet been provided . . . as to what determines the size of the risk [adjustment] and how it varies in response to changes in other variables.”

In short, before the arrival of the Capital Asset Pricing Model, the question of how expected returns and risk were related had been posed, but was still awaiting an answer.

### Why Investors Might Differ in Their Pricing of Risk

Intuitively, it would seem that investors should demand high returns for holding high-risk investments. That is, the price of a high-risk asset should be bid sufficiently low so that the future payoffs on the asset are high (relative to the price). A difficulty with this reasoning arises, however, when the risk of an investment depends on the manner in which it is held.

To illustrate, consider an entrepreneur who needs to raise $1 million for a risky new venture. There is a 90 percent chance that the venture will fail and end up worthless; and there is a 10 percent chance that the venture will succeed within a year and be worth $40 million. The expected value of the venture in one year is therefore $4 million, or $4 per share assuming that the venture will have a million shares outstanding.

Case I: If a single risk-averse individual were to fund the full $1 million—where

\(^2\) The cost of equity capital in this model is the “internal rate of return,” the discount rate that equates the present value of future cash flows to the current stock price. In the Gordon-Shapiro model, the projected dividend stream is \( D, D(1 + g), D(1 + g)^2 \ldots \). The present value of these cash flows when discounted at rate \( r \) is \( \frac{D}{r - g} \), which when set equal to the current stock price, \( P \), establishes \( r = \frac{D}{P} + g \).
the investment would represent a significant portion of the wealth of that individual—the venture would have to deliver a very high expected return, say 100 percent. To achieve an expected return of 100 percent on an investment of $1 million, the entrepreneur would have to sell the investor a 50 percent stake: 500,000 shares at a price per share of $2.

Case II: If the funds could be raised from someone who can diversify across many such investments, the required return might be much lower. Consider an investor who has $100 million to invest in 100 ventures with the same payoffs and probabilities as above, except that the outcomes of the ventures are all independent of one another. In this case, the probability of the investor sustaining a large percentage loss is small—for example, the probability that all 100 ventures fail is a miniscule 0.003 percent (= 0.9100)—and the diversified investor might consequently be satisfied to receive an expected return of only, say, 10 percent. If so, the entrepreneur would need to sell a much smaller stake to raise the same amount of money, here 27.5 percent (= $1.1 million/$4 million); and the investor would pay a higher price per share of $3.64 (= $1 million/275,000 shares).

Cases I and II differ only in the degree to which the investor is diversified; the stand-alone risk and the expected future value of any one venture is the same in both cases. Diversified investors face less risk per investment than undiversified investors, and they are therefore willing to receive lower expected returns (and to pay higher prices). For the purpose of determining required returns, the risks of investments therefore must be viewed in the context of the other risks to which investors are exposed. The CAPM is a direct outgrowth of this key idea.

Diversification, Correlation and Risk

The notion that diversification reduces risk is centuries old. In eighteenth-century English language translations of Don Quixote, Sancho Panza advises his master, “It is the part of a wise man to . . . not venture all his eggs in one basket.” According to Herbison (2003), the proverb “Do not keep all your eggs in one basket” actually appeared as far back as Torriano’s (1666) Common Place of Italian Proverbs.

However, diversification was typically thought of in terms of spreading your wealth across many independent risks that would cancel each other if held in sufficient number (as was assumed in the new ventures example). Harry Markowitz (1952) had the insight that, because of broad economic influences, risks across assets were correlated to a degree. As a result, investors could eliminate some but not all risk by holding a diversified portfolio. Markowitz wrote: “This presumption, that the law of large numbers applies to a portfolio of securities, cannot be accepted. The returns from securities are too intercorrelated. Diversification cannot eliminate all variance.”

Markowitz (1952) went on to show analytically how the benefits of diversification depend on correlation. The correlation between the returns of two assets measures the degree to which they fluctuate together. Correlation coefficients range between −1.0 and 1.0. When the correlation is 1.0, the two assets are perfectly positively correlated. They move in the same direction and in fixed
proportions (plus a constant). In this case, the two assets are substitutes for one another. When the correlation is \(-1.0\), the returns are perfectly negatively correlated meaning that when one asset goes up, the other goes down and in a fixed proportion (plus a constant). In this case, the two assets act to insure one another. When the correlation is zero, knowing the return on one asset does not help you predict the return on the other.

To show how the correlation among individual security returns affects portfolio risk, consider investing in two risky assets, \(A\) and \(B\). Assume that the risk of an asset is measured by its standard deviation of return, which for assets \(A\) and \(B\) is denoted by \(\sigma_A\) and \(\sigma_B\), respectively. Let \(\rho\) denote the correlation between the returns on assets \(A\) and \(B\); let \(x\) be the fraction invested in Asset \(A\) and \(y (= 1 - x)\) be the fraction invested in Asset \(B\).

When the returns on assets within a portfolio are perfectly positively correlated (\(\rho = 1\)), the portfolio risk is the weighted average of the risks of the assets in the portfolio. The risk of the portfolio then can be expressed as

\[
\sigma_p = x\sigma_A + y\sigma_B.
\]

The more interesting case is when the assets are not perfectly correlated (\(\rho < 1\)). Then there is a nonlinear relationship between portfolio risk and the risks of the underlying assets. In this case, at least some of the risk from one asset will be offset by the other asset, so the standard deviation of the portfolio \(\sigma_p\) is always less than the weighted average of \(\sigma_A\) and \(\sigma_B\). Thus, the risk of a portfolio is less than the average risk of the underlying assets. Moreover, the benefit of diversification will increase the farther away that the correlation \(\rho\) is from 1.

These are Harry Markowitz’s important insights: 1) that diversification does not rely on individual risks being uncorrelated, just that they be imperfectly correlated; and 2) that the risk reduction from diversification is limited by the extent to which individual asset returns are correlated. If Markowitz were restating Sancho Panza’s advice, he might say: It is safer to spread your eggs among imperfectly correlated baskets than to spread them among perfectly correlated baskets.

Table 1 illustrates the benefits of diversifying across international equity markets. The table lists the world’s largest stock markets by market capitalization as of December 31, 2003, the combination of which we will call the world equity market.

---

\(3\) The portfolio standard deviation, \(\sigma_p\), can be expressed in terms of the standard deviations of assets \(A\) and \(B\) and their correlation using the variance formula:

\[
\sigma_p^2 = x^2\sigma_A^2 + y^2\sigma_B^2 + 2xy\rho\sigma_A\sigma_B.
\]

This expression can be algebraically manipulated to obtain

\[
\sigma_p^2 = (x\sigma_A + y\sigma_B)^2 - 2xy(1 - \rho)\sigma_A\sigma_B.
\]

When \(\rho = 1\), the final term disappears, giving the formula in the text. When \(\rho < 1\), then the size of the second term will increase as \(\rho\) declines, and so the standard deviation of the portfolio will fall as \(\rho\) declines.
portfolio, labeled in the table as WEMP. The capitalization of the world equity market portfolio was about $30 trillion—comprising over 95 percent of all publicly traded equities—with the United Statese representing by far the largest fraction. Table 1 includes the standard deviation of monthly total returns for each country over the ten-year period ending December 31, 2003, expressed on an annualized basis.

Assuming that the historical standard deviations and correlations of return are good estimates of future standard deviations and correlations, we can use this data to calculate that the standard deviation of return of the WEMP—given the capitalization weights as of December 2003—is 15.3 percent per annum. If the country returns were all perfectly correlated with each other, then the standard deviation of the WEMP would be the capitalization-weighted average of the standard deviations.

Table 1

<table>
<thead>
<tr>
<th>Market Capitalization ($) Billions, 12/31/03</th>
<th>Capitalization Weight</th>
<th>S.D. of Return</th>
<th>Beta vs. WEMP</th>
<th>Correlation vs. WEMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>$14,266</td>
<td>47.8%</td>
<td>16.1%</td>
<td>1.00</td>
</tr>
<tr>
<td>Japan</td>
<td>2,953</td>
<td>9.9%</td>
<td>22.3%</td>
<td>0.83</td>
</tr>
<tr>
<td>UK</td>
<td>2,426</td>
<td>8.1%</td>
<td>14.3%</td>
<td>0.78</td>
</tr>
<tr>
<td>France</td>
<td>1,403</td>
<td>4.7%</td>
<td>19.3%</td>
<td>1.00</td>
</tr>
<tr>
<td>Germany</td>
<td>1,079</td>
<td>3.6%</td>
<td>21.7%</td>
<td>1.10</td>
</tr>
<tr>
<td>Canada</td>
<td>910</td>
<td>3.0%</td>
<td>19.9%</td>
<td>1.15</td>
</tr>
<tr>
<td>Switzerland</td>
<td>727</td>
<td>2.4%</td>
<td>17.1%</td>
<td>0.75</td>
</tr>
<tr>
<td>Spain</td>
<td>726</td>
<td>2.4%</td>
<td>21.5%</td>
<td>0.92</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>713</td>
<td>2.4%</td>
<td>29.2%</td>
<td>1.33</td>
</tr>
<tr>
<td>Italy</td>
<td>615</td>
<td>2.1%</td>
<td>23.9%</td>
<td>0.90</td>
</tr>
<tr>
<td>Australia</td>
<td>586</td>
<td>2.0%</td>
<td>18.4%</td>
<td>0.93</td>
</tr>
<tr>
<td>China</td>
<td>513</td>
<td>1.7%</td>
<td>43.3%</td>
<td>1.26</td>
</tr>
<tr>
<td>Taiwan</td>
<td>379</td>
<td>1.3%</td>
<td>33.0%</td>
<td>1.15</td>
</tr>
<tr>
<td>Netherlands</td>
<td>368</td>
<td>1.2%</td>
<td>19.5%</td>
<td>1.02</td>
</tr>
<tr>
<td>Sweden</td>
<td>320</td>
<td>1.1%</td>
<td>24.3%</td>
<td>1.25</td>
</tr>
<tr>
<td>South Korea</td>
<td>298</td>
<td>1.0%</td>
<td>47.7%</td>
<td>1.55</td>
</tr>
<tr>
<td>India</td>
<td>279</td>
<td>0.9%</td>
<td>26.7%</td>
<td>0.63</td>
</tr>
<tr>
<td>South Africa</td>
<td>261</td>
<td>0.9%</td>
<td>26.9%</td>
<td>1.09</td>
</tr>
<tr>
<td>Brazil</td>
<td>235</td>
<td>0.8%</td>
<td>43.6%</td>
<td>1.81</td>
</tr>
<tr>
<td>Russia</td>
<td>198</td>
<td>0.7%</td>
<td>76.9%</td>
<td>2.34</td>
</tr>
<tr>
<td>Belgium</td>
<td>174</td>
<td>0.6%</td>
<td>17.2%</td>
<td>0.65</td>
</tr>
<tr>
<td>Malaysia</td>
<td>168</td>
<td>0.6%</td>
<td>38.6%</td>
<td>0.81</td>
</tr>
<tr>
<td>Singapore</td>
<td>149</td>
<td>0.5%</td>
<td>28.6%</td>
<td>1.04</td>
</tr>
<tr>
<td>Mexico</td>
<td>123</td>
<td>0.4%</td>
<td>35.1%</td>
<td>1.40</td>
</tr>
<tr>
<td>WEMP</td>
<td>$29,870</td>
<td>100%</td>
<td>15.3%</td>
<td>1.00</td>
</tr>
<tr>
<td>S.D. of WEMP assuming perfect correlation</td>
<td></td>
<td></td>
<td>19.9%</td>
<td></td>
</tr>
<tr>
<td>S.D. of WEMP assuming zero correlation</td>
<td></td>
<td></td>
<td>8.4%</td>
<td></td>
</tr>
</tbody>
</table>

Notes: WEMP stands for World Equity Market Portfolio. S.D. is standard deviation expressed on an annualized basis. Calculations are based on historical monthly returns obtained from Global Financial Data Inc.
which is 19.9 percent per annum. The difference of 4.6 percent per annum represents the diversification benefit—the risk reduction stemming from the fact that the world’s equity markets are imperfectly correlated. Also shown in Table 1 is that the standard deviation of the WEMP would be only 8.4 percent per annum if the country returns were uncorrelated with one another. The amount by which this figure is lower than the actual standard deviation of 15.3 percent per annum is a measure of the extent to which the world’s equity markets share common influences.

**Portfolio Theory, Riskless Lending and Borrowing and Fund Separation**

To arrive at the CAPM, we need to examine how imperfect correlation among asset returns affects the investor’s tradeoff between risk and return. While risks combine nonlinearly (because of the diversification effect), expected returns combine linearly. That is, the expected return on a portfolio of investments is just the weighted average of the expected returns of the underlying assets. Imagine two assets with the same expected return and the same standard deviation of return. By holding both assets in a portfolio, one obtains an expected return on the portfolio that is the same as either one of them, but a portfolio standard deviation that is lower than any one of them individually. Diversification thus leads to a reduction in risk without any sacrifice in expected return.

Generally, there will be many combinations of assets with the same portfolio expected return but different portfolio risk; and there will be many combinations of assets with the same portfolio risk but different portfolio expected return. Using optimization techniques, we can compute what Markowitz coined the “efficient frontier.” For each level of expected return, we can solve for the portfolio combination of assets that has the lowest risk. Or for each level of risk, we can solve for the combination of assets that has the highest expected return. The efficient frontier consists of the collection of these optimal portfolios, and each investor can choose which of these best matches their risk tolerance.

The initial development of portfolio theory assumed that all assets were risky. James Tobin (1958) showed that when investors can borrow as well as lend at the risk-free rate, the efficient frontier simplifies in an important way. (A “risk-free” instrument pays a fixed real return and is default free. U.S. Treasury bonds that adjust automatically with inflation—called Treasury inflation-protected instruments, or TIPS—and short-term U.S. Treasury bills are considered close approximations of risk-free instruments.)

To see how riskless borrowing and lending affects investors’ decision choices, consider investing in the following three instruments: risky assets $M$ and $H$, and the riskless asset, where the expected returns and risks of the assets are shown in Table 2. Suppose first that you had the choice of investing all of your wealth in just one of these assets. Which would you choose? The answer depends on your risk tolerance. Asset $H$ has the highest risk and also the highest expected return. You would choose
Asset $H$ if you had a high tolerance for risk. The riskless asset has no risk but also the lowest expected return. You would choose to lend at the risk-free rate if you had a very low tolerance for risk. Asset $M$ has an intermediate risk and expected return, and you would choose this asset if you had a moderate tolerance for risk.

Suppose next that you can borrow and lend at the risk-free rate, that you wish to invest some of your wealth in Asset $H$ and the balance in riskless lending or borrowing. If $x$ is the fraction of wealth invested in Asset $H$, then $1 - x$ is the fraction invested in the risk-free asset. When $x < 1$, you are lending at the risk-free rate; when $x > 1$, you are borrowing at the risk-free rate. The expected return of this portfolio is $(1 - x)r_f + xE_H$, which equals $r_f + x(E_H - r_f)$, and the risk of the portfolio is $x\sigma_H$. The risk of the portfolio is proportional to the risk of Asset $H$, since Asset $H$ is the only source of risk in the portfolio.

Risk and expected return thus both combine linearly, as shown graphically in Figure 1. Each point on the line connecting the risk-free asset to Asset $H$ represents a particular allocation ($x$) to Asset $H$ with the balance in either risk-free lending or risk-free borrowing. The slope of this line is called the Sharpe Ratio—the risk premium of Asset $H$ divided by the risk of Asset $H$:

$$\text{Sharpe Ratio} = \frac{(E_H - r_f)}{\sigma_H}.$$ 

The Sharpe Ratio of Asset $H$ evaluates to 0.175 ($= (12 \text{ percent} - 5 \text{ percent})/40 \text{ percent}$) and all combinations of Asset $H$ with risk-free borrowing or lending have this same Sharpe Ratio.

Also shown in Figure 1 are the risks and expected returns that can be achieved by combining Asset $M$ with riskless lending and borrowing. The Sharpe Ratio of Asset $M$ is 0.25, which is higher than that of Asset $H$, and any level of risk and return that can be obtained by investing in Asset $H$ along with riskless lending or borrowing is dominated by some combination of Asset $M$ and riskless lending or borrowing. For example, for the same risk as Asset $H$, you can obtain a higher expected return by investing in Asset $M$ with 2:1 leverage. As shown in Figure 1, the expected return of a 2:1 leveraged position in Asset $M$ is 15 percent (that is, $(2 \times 10 \text{ percent}) - (1 \times 5 \text{ percent})$), which is higher than the 12 percent expected return of Asset $H$. If you could hold only one risky asset along with riskless lending or borrowing, it unambiguously would be Asset $M$.

Being able to lend and borrow at the risk-free rate thus dramatically changes

<table>
<thead>
<tr>
<th>Expected return</th>
<th>Risk (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riskless asset</td>
<td>5% ($r_f$)</td>
</tr>
<tr>
<td>Asset $M$</td>
<td>10% ($E_M$)</td>
</tr>
<tr>
<td>Asset $H$</td>
<td>12% ($E_H$)</td>
</tr>
</tbody>
</table>
our investment choices. The asset of choice—if you could choose only one risky asset—is the one with the highest Sharpe Ratio. Given this choice of risky asset, you need to make a second decision, which is how much of it to hold in your portfolio. The answer to the latter question depends on your risk tolerance.

Figure 2 illustrates the approach in the case where we can invest in combinations of two risky assets, $M$ and $H$, plus riskless lending and borrowing. The correlation between the returns of assets $M$ and $H$ is assumed to be zero. In the figure, the curve connecting assets $M$ and $H$ represents all expected return/standard deviation pairs that can be attained through combinations of assets $M$ and $H$. The combination of assets $M$ and $H$ that has the highest Sharpe Ratio is 74 percent in Asset $M$ and 26 percent in Asset $H$ (the tangency point). The expected return of this combination is 10.52 percent, and the standard deviation is 18.09 percent. The Sharpe Ratio evaluates to 0.305, which is considerably higher than the Sharpe Ratios of assets $M$ and $H$ (0.25 and 0.175, respectively). Investors who share the same estimates of expected return and risk all will locate their portfolios on the tangency line connecting the risk-free asset to the frontier. In particular, they all will hold assets $M$ and $H$ in the proportions 74/26.

The optimal portfolio of many risky assets can be found similarly. Figure 3 offers a general illustration. Use Markowitz’s algorithm to obtain the efficient frontier of portfolios of risky assets. Find the portfolio on the efficient frontier that has the highest Sharpe Ratio, which will be the point where a ray stretching up from the risk-free point is just tangent to the efficient frontier. Then, in accordance with your risk tolerance, allocate your wealth between this highest Sharpe Ratio portfolio and risk-free lending or borrowing.

This characterization of the efficient frontier is referred to as “fund separation.” Investors with the same beliefs about expected returns, risks and correlations all will invest in the portfolio or “fund” of risky assets that has the highest Sharpe
Ratio, but they will differ in their allocations between this fund and risk-free lending or borrowing based on their risk tolerance. Notice in particular that the composition of the optimal portfolio of risky assets does not depend on the investor’s tolerance for risk.

**Market-Determined Expected Returns and Stand-Alone Risk**

Portfolio theory prescribes that investors choose their portfolios on the efficient frontier, given their beliefs about expected returns and risks. The Capital
Asset Pricing Model, on the other hand, is concerned with the pricing of assets in equilibrium. CAPM asks: What are the implications for asset prices if everyone heeds this advice? In equilibrium, all assets must be held by someone. For the market to be in equilibrium, the expected return of each asset must be such that investors collectively decide to hold exactly the supply of shares of the asset. The Capital Asset Pricing Model will tell us how investors determine those expected returns—and thereby asset prices—as a function of risk.

In thinking about how expected return and risk might be related, let us ask whether, as a rule, the expected return on an investment could possibly be a function of its stand-alone risk (measured by standard deviation of return). The answer turns out to be “no.” Consider the shares of two firms with the same stand-alone risk. If the expected return on an investment was determined solely by its stand-alone risk, the shares of these firms would have the same expected return, say 10 percent. Any portfolio combination of the two firms would also have an expected return of 10 percent (since the expected return of a portfolio of assets is the weighted average of the expected returns of the assets that comprise the portfolio). However, if the shares of the firms are not perfectly correlated, then a portfolio invested in the shares of the two firms will be less risky than either one stand-alone. Therefore, if expected return is a function solely of stand-alone risk, then the expected return of this portfolio must be less than 10 percent, contradicting the fact that the expected return of the portfolio is 10 percent. Expected returns, therefore, cannot be determined solely by stand-alone risk.

Accordingly, any relationship between expected return and risk must be based on a measure of risk that is not stand-alone risk. As we will soon see, that measure of risk is given by the incremental risk that an asset provides when added to a portfolio, as discussed in the next section.

**Improving the Sharpe Ratio of a Portfolio**

Suppose you were trying to decide whether to add a particular stock to your investment portfolio of risky assets. If you could borrow and lend at the risk-free rate, you would add the stock if it improved the portfolio’s Sharpe Ratio. It turns out there is a simple rule to guide the decision—a rule that can be derived by understanding the two special cases: 1) when the additional stock is uncorrelated with the existing portfolio, and 2) when the additional stock is perfectly correlated with the existing portfolio. The rule will lead us directly to the equilibrium risk-return relationship specified by the Capital Asset Pricing Model.

In what follows, it will be helpful to think in terms of “excess return,” the return of an instrument in excess of the risk-free rate. The expected excess return is called the risk premium.

**Adding a Stock that is Uncorrelated with the Existing Portfolio**

When should a portfolio be diversified into an uncorrelated stock? If the excess returns on the stock and existing portfolio are uncorrelated, adding a small
amount of the stock has almost no effect on the risk of the portfolio. At the margin, therefore, the stock is a substitute for investing in the risk-free asset. Including the stock will increase the portfolio’s Sharpe Ratio if the stock’s expected return $E_S$ exceeds the risk-free rate $r_f$. Said another way, the additional stock should be included in the portfolio if its risk premium $E_S - r_f$ is positive.

**Adding a Stock that is Perfectly Correlated with the Existing Portfolio**

If the stock and portfolio excess returns are perfectly correlated, investing in the stock becomes a substitute for investing in the portfolio itself. To see this, recall that a perfect correlation means that the stock and the portfolio excess returns move together in a fixed ratio plus a constant. The fixed ratio is called beta, denoted by $\beta$, and the constant is called alpha, denoted by $\alpha$. In other words, the excess return of the stock is equal to alpha plus beta times the excess return of the portfolio. It also follows that the expected excess return of the stock is alpha plus beta times the expected excess return on the portfolio—that is, $E_S - r_f = \alpha + \beta(E_p - r_f)$. The constant alpha is therefore given by the difference between the risk premium of the stock and beta times the risk premium of the portfolio. Since the stock and the portfolio move together in a fixed proportion, beta is given by the ratio of stock to portfolio standard deviations of excess return: $\beta = \sigma_S/\sigma_P$.

Compare now an investment of $1 in the stock with the following “mimicking” strategy: invest $\beta$ in the portfolio and the balance $(1 - \beta)$ in the risk-free asset, assuming that $\beta < 1$. For example, if beta is 0.5, then investing $0.50$ in the portfolio and $0.50$ in the riskless asset is a strategy that will gain or lose 0.5 percent of excess return for every 1 percent gain or loss in the portfolio excess return. The excess return of the mimicking strategy is beta times the excess return of the portfolio. The mimicking strategy will behave just like the stock up to the constant difference alpha. The mimicking strategy can be thought of as a “stock” with the given beta but an alpha of zero.

Similarly, if $\beta > 1$, the mimicking strategy involves investing $\beta$ in the portfolio of which $(\beta - 1)$ is borrowed at the riskless rate. For example, if beta is 3, the mimicking portfolio involves investing $3$ in the portfolio of which $2$ is borrowed at the risk-free rate. This strategy will gain or lose 3 percent of excess return for every 1 percent gain or loss in the portfolio excess return. Again, the mimicking strategy will behave just like the stock up to the constant difference alpha.

When should a stock be added to the portfolio if its return is perfectly correlated with that of the portfolio? Since, up to the constant alpha, the stock is just a substitute for the portfolio, adding $\beta$ of the stock to the portfolio amounts to owning $\beta$ more of the portfolio. But owning more of the portfolio by itself does not change its Sharpe Ratio. Therefore, adding the stock will increase the portfolio’s Sharpe Ratio if the

---

4 Assume that you have $1 of wealth invested in the portfolio. Then, adding an investment of $x$ in shares of the stock increases the portfolio variance to $\sigma_P^2 + x^2\sigma_S^2$, where $\sigma_P^2$ is the variance of the portfolio and $x^2\sigma_S^2$ is the variance of the additional stock, weighted by the number of dollars invested in the stock. Remember, the variance of a combination of uncorrelated risks equals the sum of the variances of the individual risks. The increase in portfolio risk (standard deviation as well as variance) is proportional to $x^2$, which implies that the change in portfolio risk is negligible for small $x$. The $x$ needed to purchase the shares can come from holding less of the risk-free asset or by borrowing at the risk-free rate.
stock’s expected excess return exceeds that of the mimicking portfolio. This occurs if
\[ \alpha > 0 \] or equivalently if \( E_S - r_f > \beta (E_P - r_f) \), meaning that the stock’s risk premium
must exceed beta times the portfolio risk premium.

**The General Case: Adding a Stock that is Imperfectly Correlated with the Existing Portfolio**

Suppose next that the returns on the stock and the portfolio are correlated to
some degree \( (0 < \rho < 1) \). In this case, the stock’s return can be separated into a return
component that is *perfectly correlated* with the portfolio and a return component that is *uncorrelated*
with the portfolio. Since the standard deviation of the stock is \( \sigma_S \), the
standard deviation of the perfectly correlated component of the stock’s return is \( \rho \sigma_S \).\(^5\)
Thus, the beta of the perfectly correlated component of the stock’s excess return to the
portfolio’s excess return is given by the ratio of standard deviations:
\[ \beta = \frac{\rho \sigma_S}{\sigma_P} \]

As just discussed, the component of the stock’s return that is perfectly correlated
with the portfolio is a substitute for the portfolio itself and can be mimicked
through an investment of \( \beta \) in the portfolio and \((1 - \beta)\) in the riskless asset. The
component of the stock’s excess return that is uncorrelated with the portfolio can,
at the margin, be diversified away and will thus have no effect on the risk of the
portfolio. This component of return can be mimicked through an investment in
the risk-free asset. We can therefore conclude that adding the stock to the portfolio
will improve the Sharpe Ratio if the stock’s risk premium exceeds the sum of the
risk premia of the two mimicking portfolios: \( \beta (E_P - r_f) \) for the perfectly correlated
return component and zero for the uncorrelated return component.

This insight establishes a rule for improving the portfolio. Adding a marginal
share of stock to a portfolio will increase the portfolio’s Sharpe Ratio if the stock’s
alpha is positive, that is, if its risk premium satisfies
\[ E_S - r_f > \beta (E_P - r_f). \]

Conversely, selling short a marginal share of the stock will increase the portfolio’s
Sharpe Ratio if the alpha is negative, \( E_S - r_f < \beta (E_P - r_f) \). The portfolio has the
highest attainable Sharpe Ratio if \( E_S - r_f = \beta (E_P - r_f) \) for every stock—that is, if
the risk premium for each stock is equal to beta times the risk premium for the
portfolio as a whole.

**The Capital Asset Pricing Model**

The rule for improving the Sharpe Ratio of a portfolio allows us to derive the
Capital Asset Pricing Model in a straightforward and intuitive way. We begin with four
assumptions. First, investors are risk averse and evaluate their investment portfolios

\(^5\) The correlation coefficient \( \rho \) is the “R” in “R-squared”—the fraction of the stock’s variance that is
attributable to movements in the portfolio. If \( \rho < 0 \), the standard deviation of the perfectly correlated
component is \( |\rho|\sigma_S \).
solely in terms of expected return and standard deviation of return measured over the same single holding period. Second, capital markets are perfect in several senses: all assets are infinitely divisible; there are no transactions costs, short selling restrictions or taxes; information is costless and available to everyone; and all investors can borrow and lend at the risk-free rate. Third, investors all have access to the same investment opportunities. Fourth, investors all make the same estimates of individual asset expected returns, standard deviations of return and the correlations among asset returns.

These assumptions represent a highly simplified and idealized world, but are needed to obtain the CAPM in its basic form. The model has been extended in many ways to accommodate some of the complexities manifest in the real world. But under these assumptions, given prevailing prices, investors all will determine the same highest Sharpe Ratio portfolio of risky assets. Depending on their risk tolerance, each investor will allocate a portion of wealth to this optimal portfolio and the remainder to risk-free lending or borrowing. Investors all will hold risky assets in the same relative proportions.

For the market to be in equilibrium, the price (that is, the expected return) of each asset must be such that investors collectively decide to hold exactly the supply of the asset. If investors all hold risky assets in the same proportions, those proportions must be the proportions in which risky assets are held in the market portfolio—the portfolio comprised of all available shares of each risky asset. In equilibrium, therefore, the portfolio of risky assets with the highest Sharpe Ratio must be the market portfolio.

If the market portfolio has the highest attainable Sharpe Ratio, there is no way to obtain a higher Sharpe Ratio by holding more or less of any one asset. Applying the portfolio improvement rule, it follows that the risk premium of each asset must satisfy $E_s - r_f = \beta (E_M - r_f)$, where $E_s$ and $E_M$ are the expected return on the asset and the market portfolio, respectively, and $\beta$ is the sensitivity of the asset’s return to the return on the market portfolio.

We have just established the Capital Asset Pricing Model: In equilibrium, the expected return of an asset is given by

$$E_s = r_f + \beta (E_M - r_f).$$

This formula is the one that Sharpe, Treynor, Lintner and Mossin successfully set out to find. It is the relationship between expected return and risk that is consistent with investors behaving according to the prescriptions of portfolio theory. If this rule does not hold, then investors will be able to outperform the market (in the sense of obtaining a higher Sharpe Ratio) by applying the portfolio improvement rule, and if sufficiently many investors do this, stock prices will adjust to the point where the CAPM becomes true.

Another way of expressing the CAPM equation is

$$\text{Sharpe Ratio of Asset } S = \rho \times \text{Sharpe Ratio of the Market Portfolio.}^{6}$$

---

6 Using the fact that $\beta = \rho \sigma_s / \sigma_M$, the equation $E_s = r_f + \beta (E_M - r_f)$ can be rearranged to give $(E_s - r_f) / \sigma_s = \rho (E_M - r_f) / \sigma_M$, which is the expression in the text.
In other words, in equilibrium, the Sharpe Ratio of any asset is no higher than the Sharpe Ratio of the market portfolio (since $\rho \leq 1$). Moreover, assets having the same correlation with the market portfolio will have the same Sharpe Ratio.

The Capital Asset Pricing Model tells us that to calculate the expected return of a stock, investors need know two things: the risk premium of the overall equity market $E_M - r_f$ (assuming that equities are the only risky assets) and the stock’s beta versus the market. The stock’s risk premium is determined by the component of its return that is perfectly correlated with the market—that is, the extent to which the stock is a substitute for investing in the market. The component of the stock’s return that is uncorrelated with the market can be diversified away and does not command a risk premium.

The Capital Asset Pricing Model has a number of important implications. First, perhaps the most striking aspect of the CAPM is what the expected return of an asset does not depend on. In particular, the expected return of a stock does not depend on its stand-alone risk. It is true that a high beta stock will tend to have a high stand-alone risk because a portion of a stock’s stand-alone risk is determined by its beta, but a stock need not have a high beta to have a high stand-alone risk. A stock with high stand-alone risk therefore will only have a high expected return to the extent that its stand-alone risk is derived from its sensitivity to the broad stock market.

Second, beta offers a method of measuring the risk of an asset that cannot be diversified away. We saw earlier that any risk measure for determining expected returns would have to satisfy the requirement that the risk of a portfolio is the weighted average of the risks of the holdings in the portfolio. Beta satisfies this requirement. For example, if two stocks have market betas of 0.8 and 1.4, respectively, then the market beta of a 50/50 portfolio of these stocks is 1.1, the average of the two stock betas. Moreover, the capitalization weighted average of the market betas of all stocks is the beta of the market versus itself. The average stock therefore has a market beta of 1.0.

On a graph where the risk of an asset as measured by beta is on the horizontal axis and return is on the vertical axis, all securities lie on a single line—the so-called Securities Market Line shown in Figure 4. If the market is in equilibrium, all assets must lie on this line. If not, investors will be able to improve upon the market portfolio and obtain a higher Sharpe Ratio. In contrast, Figure 3 presented earlier measured risk on the horizontal axis as stand-alone risk, the standard deviation of each stock, and so stocks were scattered over the diagram. But remember that not all of the stand-alone risk of an asset is priced into its expected return, just that portion of its risk, $\rho \sigma_S$, that is correlated with the market portfolio.

Third, in the Capital Asset Pricing Model, a stock’s expected return does not depend on the growth rate of its expected future cash flows. To find the expected return of a company’s shares, it is thus not necessary to carry out an extensive financial analysis of the company and to forecast its future cash flows. According to the CAPM, all we need to know about the specific company is the beta of its shares, a parameter that is usually much easier to estimate than the expected future cash flows of the firm.
Is the CAPM Useful?

The Capital Asset Pricing Model is an elegant theory with profound implications for asset pricing and investor behavior. But how useful is the model given the idealized world that underlies its derivation? There are several ways to answer this question. First, we can examine whether real world asset prices and investor portfolios conform to the predictions of the model, if not always in a strict quantitative sense, and least in a strong qualitative sense. Second, even if the model does not describe our current world particularly well, it might predict future investor behavior—for example, as a consequence of capital market frictions being lessened through financial innovation, improved regulation and increasing capital market integration. Third, the CAPM can serve as a benchmark for understanding the capital market phenomena that cause asset prices and investor behavior to deviate from the prescriptions of the model.

Suboptimal Diversification

Consider the CAPM prediction that investors all will hold the same (market) portfolio of risky assets. One does not have to look far to realize that investors do not hold identical portfolios, which is not a surprise since taxes alone will cause idiosyncratic investor behavior. For example, optimal management of capital gains taxes involves early realization of losses and deferral of capital gains, and so taxable investors might react very differently to changes in asset values depending on when they purchased the asset (Constantinides, 1983). Nevertheless, it will still be a positive sign for the model if most investors hold broadly diversified portfolios. But even here the evidence is mixed. On one hand, popular index funds make it possible for investors to obtain diversification at low cost. On the other hand, many workers hold concentrated ownership of company stock in employee retirement savings plans and many executives hold concentrated ownership of company stock options.

One of the most puzzling examples of suboptimal diversification is the so-
called home bias in international investing. In almost all countries, foreign ownership of assets is low, meaning that investors tend to hold predominantly home country assets. For example, in 2003, foreign ownership accounted for only 10 percent of publicly traded U.S. equities and 21 percent of publicly traded Japanese equities. Japanese investor portfolios therefore deviate significantly from the world equity market portfolio: they own the vast majority of their home country equities, but only a tiny share of U.S. equities. By contrast, and as shown in Table 1, an investor holding the world equity market portfolio would be invested 48 percent in U.S. equities and only 10 percent in Japanese equities.

Why is suboptimal diversification so pervasive? Common explanations are that obtaining broad diversification can be costly, in terms of direct expenses and taxes, and that investors are subject to behavioral biases and lack of sophistication. None of these reasons, if valid, would mean that the CAPM is not useful. The CAPM tells us that investors pay a price for being undiversified in that they are taking risks for which they are not being compensated. Thus, there exists the potential for portfolio improvement, which in turn creates opportunities for investor education and financial innovation. Indeed, foreign ownership of equities in many countries has more than doubled over the last 20 years, most likely due to the increased availability of low-cost vehicles to invest globally and greater investor appreciation of the need for diversification. Investors today seem to be much better diversified than in decades past, a trend that appears likely to continue.

Performance Measurement

One of the earliest applications of the Capital Asset Pricing Model was to performance measurement of fund managers (Treynor, 1965; Sharpe, 1966; Jensen, 1968). Consider two funds, A and B, that are actively managed in the hope of outperforming the market. Suppose that the funds obtained returns of 12 percent and 18 percent, respectively, during a period when the risk-free rate was 5 percent and the overall market returned 15 percent. Assume further that the standard deviation of funds A and B were 40 percent per annum and 30 percent per annum, respectively. Which fund had the better performance?

At first glance, fund A had greater risk and a lower return than fund B, so fund B would appear to have been the better performing fund. However, we know from the CAPM that focusing on stand-alone risk is misleading if investors can hold diversified portfolios. To draw a firmer conclusion, we need to know how these funds are managed: Suppose that fund A consists of a high-risk but “market-neutral” portfolio that has long positions in some shares and short positions in others, with a portfolio beta of zero. Fund B, on the other hand, invests in selected high beta stocks, with a portfolio beta of 1.5.

Instead of investing in funds A and/or B, investors could have held corresponding mimicking or “benchmark” portfolios. For fund A, since its beta is zero, the benchmark portfolio is an investment in the risk-free asset; for fund B, the benchmark is a position in the market portfolio leveraged 1.5:1 with borrowing at the risk-free rate. The benchmark portfolios respectively would have returned 5 percent and 20 percent (= 5 percent + 1.5 × (15 percent − 5 percent)). Fund
A thus outperformed its benchmark by 7 percent, while fund B underperformed its benchmark by 2 percent, as shown in Table 3. In terms of the CAPM framework, funds A and B had alphas of 7 percent and −2 percent, respectively, where alpha is the difference between a fund’s performance and that predicted given the beta of the fund. Appropriately risk adjusted, fund A’s performance (alpha = 7 percent) exceeded that of fund B (alpha = −2 percent). An investor who held the market portfolio would, at the margin, have obtained a higher return for the same risk by allocating money to fund A rather than to fund B.7

The key idea here is that obtaining high returns by owning high beta stocks does not take skill, since investors can passively create a high beta portfolio simply through a leveraged position in the market portfolio. Obtaining high returns with low beta stocks is much harder, however, since such performance cannot be replicated with a passive strategy. Investors therefore need to assess performance based on returns that have been appropriately risk adjusted. The CAPM provides a clear framework for thinking about this issue.

The CAPM and Discounted Cash Flow Analysis

According to the CAPM, the appropriate discount rate for valuing the expected future cash flows of a company or of a new investment project is determined by the risk-free rate, the market risk premium and the beta versus the market of the company or project. Accuracy in estimating these parameters matters greatly for real world decisionmaking since, for long-dated cash flows, an error in the discount rate is magnified manyfold when calculating the net present value. Beta is usually estimated with use of linear regression analysis applied to historical stock market returns data. Beta can in many circumstances be accurately measured this way even over a relatively short period of time, provided that there is sufficient high-frequency data. When the company or project being valued is not publicly traded or there is no relevant return history, it is customary to infer beta from comparable entities whose betas can be estimated. But measurement issues can arise even if the availability of market returns data is not an issue, for example when the covariance with

---

7 This assumes that the beta of the overall portfolio is held constant—by holding more of the market portfolio if money is allocated to fund A and less of the market portfolio if money is allocated to fund B.
the market is time varying and when local stock market indexes are used as proxies for
the broad market portfolio because the latter is not well specified.

The hardest of all parameters to estimate is usually the market risk premium. The
historical risk premium is estimated from the average of past returns and, unlike
variance-related measures like beta, average returns are very sensitive to the
beginning and ending level of stock prices. The risk premium must therefore be
measured over long periods of time, and even this may not be sufficient if the risk
premium varies over time.

None of these measurement questions poses a problem for the CAPM per se,
however. The market risk premium is common to all models of cash flow valuation,
and its estimation needs to be performed regardless of the difficulty of the task.
Provided that the CAPM is the “right” model, beta too needs to be estimated, irrespective of difficulty.

Extensions of the CAPM

The Capital Asset Pricing Model has been extended in a variety of ways. Some of
the best-known extensions include allowing heterogenous beliefs (Lintner, 1969; Mert-
ton, 1987); eliminating the possibility of risk-free lending and borrowing (Black, 1972);
having some assets be nonmarketable (Mayers, 1973); allowing for multiple time
periods and investment opportunities that change from one period to the next
(Merton, 1973; Breeden, 1979); extensions to international investing (Solnik, 1974;
Stulz, 1981; Adler and Dumas, 1983); and employing weaker assumptions by relying on
arbitrage pricing (Ross, 1976). In most extensions of the CAPM, no single portfolio of
risky assets is optimal for everyone. Rather, investors allocate their wealth differentially
among several risky portfolios, which across all investors aggregate to the market
portfolio.

To illustrate, consider the International Capital Asset Pricing Model. This
model takes into account that investors have consumption needs particular to the
country in which they are resident. Thus, British investors will worry about the
purchasing power of pounds while American investors worry about the purchasing
power of dollars, which means that British and American investors will differently
assess the incremental contribution that any particular asset makes to portfolio risk.
As a result, they will hold somewhat different portfolios.8 In the basic CAPM,
investors care about only one risk factor—the overall market. In this international
version of the model, they are also concerned about real currency fluctuations. This
insight leads to a model of expected returns involving not only the beta of an asset
versus the overall market, but also the betas of the asset versus currency movements
and any other risk that is viewed differently by different investor segments.

Almost all variants of the CAPM have a multi-beta expression for expected

---

8 British investors who own American assets will hedge a portion of their real pound/dollar exchange
rate exposure by borrowing in dollars and lending in pounds, and American investors who own British
assets will hedge a portion of their real dollar/pound exchange rate exposure by borrowing in pounds
and lending in dollars. British and American investors thus will lend to and borrow from each other, and
they will have opposite exposures to the dollar/pound exchange rate.
return. They are derived from the same basic notions: 1) investors will hold portfolios that are optimized given their specific needs, constraints and risk preferences; 2) in equilibrium, asset prices reflect these demands; and 3) assets with high expected returns are those that are correlated with any risk that a significant group of investors has been unable to eliminate from their portfolios.

Whether the basic CAPM or one of its multifactor extensions is the “correct” model of asset prices is ultimately an empirical question, one that is discussed in detail by Fama and French in their companion paper in this journal. Initial tests of the CAPM by Black, Jensen and Scholes (1972) and Fama and MacBeth (1973) supported the theory in that high beta stocks were found to have had higher returns than low beta stocks. However, the relationship between beta and average returns was not as steep as indicated by the theoretical Securities Market Line.

Since this early work, a vast body of research has looked for additional risk factors that affect expected returns. Most notably, Fama and French (1992) find that adding a “value” factor and a “size” factor (in addition to the overall market) greatly improves upon the explanatory power of the CAPM. The pervasiveness of these findings in follow-up research across time and other countries provides strong evidence that more than one systematic risk factor is at work in determining asset prices. However, the value and size factors are not explicitly about risk; at best, they are proxies for risk. For example, size per se cannot be a risk factor that affects expected returns, since small firms would then simply combine to form large firms. Another criticism of the Fama-French findings is that their value effect is based on giving equal weight to small and large companies and is much stronger than observed in capitalization-weighted value indexes. Until the risks that underlie the Fama-French factors are identified, the forecast power of their model will be in doubt and the applications will be limited.

**Conclusion**

The Capital Asset Pricing Model is a fundamental contribution to our understanding of the determinants of asset prices. The CAPM tells us that ownership of assets by diversified investors lowers their expected returns and raises their prices. Moreover, investors who hold undiversified portfolios are likely to be taking risks for which they are not being rewarded. As a result of the model, and despite its mixed empirical performance, we now think differently about the relationship between expected returns and risk; we think differently about how investors should allocate their investment portfolios; and we think differently about questions such as performance measurement and capital budgeting.

- I thank Josh Coval, Mihir Desai, Craig French, Ken Froot, Jim Hines, Elon Kohlberg, Adam Perold, Melissa Perold, Andrei Shleifer, Bill Sharpe, René Stulz, Timothy Taylor, Luis Viceira and Michael Waldman for helpful discussions and comments.
References


