The capital asset pricing model and the market model

“The concept of reward to equity market risk (or beta) is a theoretical insight that, in my view, is likely to endure.”

Barr Rosenberg

(Reprinted with permission from The Journal of Portfolio Management)

Is Beta Dead?” (Wallace [1980]) and other recent articles have asked whether broad consequences, disastrous to modern investment technology, would result from misspecification of the Capital Asset Pricing Model (CAPM), or worse yet, from falsehood of the model. The criticisms have cited imprecise specification of the market portfolio as a misapplication of the CAPM, and have emphasized the difference between the “efficient portfolio” and the market portfolio when the CAPM is false. The purpose of this article is to evaluate these criticisms.

Many of the constructs of the “market model” are widely used in investment: “market portfolio,” “systematic risk and return,” “residual” or “diversifiable risk and return,” “alpha,” “beta.” These ideas play an important role in the methods of “modern portfolio theory.”

The Capital Asset Pricing Model of Sharpe, Lintner, and Mossin is the origin of these formal constructs. The constructs of the CAPM are important building blocks that retain validity in numerous applications, even where the CAPM fails. Sharpe’s [1963, 1964] clear demonstration of the CAPM stimulated diverse quantitative methods in investment. Most of them, however, turn out to be justified by other arguments and not by the CAPM at all.

The CAPM is theory, but, paradoxically, the role of the CAPM as “theory” leading to application has been less important than its role in mobilizing attention and in defining constructs. We should keep in mind that the CAPM is not “true,” since many of its assumptions are not exactly satisfied...
in the real world. Indeed, the CAPM rules out active management and investment research, and thus abolishes most applications at the stroke of a pen, by virtue of the unrealistic assumptions that it makes.

Some common applications do depend upon the correctness of the simple CAPM, or its extensions, in describing equilibrium returns. For these cases, one reaction to recent criticism can be paraphrased as follows: Is the CAPM true? No. Is imprecise knowledge of the market portfolio an important factor in this? No. Does the approximated market portfolio retain an important role in the reconstructed applications that emerge from recognition of the falsehood of the CAPM? In my judgment, yes, but there is controversy on this point (Ross [1977, 1978]). Hence, I would view the particular criticism centered upon imperfect knowledge of the market portfolio as a “red herring,” distracting us from more useful criticisms that open out interesting paths of inquiry.

Applications of theory to active management, in contrast, can have meaning only if the CAPM is false! With regard to each application, the central questions are: Since the application does not depend on the CAPM, what is the justification? Does the market portfolio play an important role in the application? If so, how sensitive is the application to the kinds of error we are likely to make in specifying the market portfolio? Should we seek an “efficient portfolio,” so as to do the application more wisely?

The conclusion seems to be that the market portfolio does play a natural role, and that likely specification errors are relatively unimportant. By contrast, the “true efficient portfolio” is not a useful construct: If we knew the efficient portfolio, the need for the application would disappear! Consequently, it is a logical contradiction to rely on a hypothetical efficient portfolio to improve application that is rendered void by that hypothesis.

In summary, I shall argue that criticisms of applications of theory that are based on imperfect knowledge of the market and the efficient portfolio are not very productive. Nevertheless, these critiques may have been fruitful since they have led to widespread discussion of subtleties that might otherwise have been glossed over.
The plan of this article as follows. The first section reviews the simple CAPM, its unrealistic assumptions, and its provocative implications. The second section progressively relaxes the assumptions in the direction of greater realism, and sketches some consequences.

The third section examines applications of the “market model” that depend upon the importance of the “market factor”: market timing, performance attribution to market timing, research that breaks out market forecasts from forecasts of other components of return, studies of reward to market risk exposure, and representation of the market return in an asset allocation decision. The fourth section considers applications in which the market portfolio is important because it is the average of investors’ portfolios: Index funds, the “universal performance benchmark,” and the idea of nonconsensus forecasts. The applications discussed through this point are not dependent on the CAPM, and are little affected by criticisms of the model.

The fifth section considers the extension of the CAPM into a multiple-factor context in which several factors may be rewarded, and the prediction of equilibrium rewards in this context. The sixth touches briefly upon the investment decision of the client who has, or considers having, multiple managers, and the requisites for performance analysis on behalf of such a client. The final section addresses applications where the falsehood of the CAPM is an important factor: the setting of risk-adjusted rates of return for security valuation, capital budgeting, and rate regulation.

THE ORIGINAL CAPITAL ASSET PRICING MODEL

Well before the CAPM, Markowitz [1959] pioneered the application of decision theory to investment. His was the crucial insight that portfolio optimization is characterized by a trade-off of the reward (expected return) of the individual security against the contribution of that security to portfolio risk. The key aspect of the security’s risk is the contribution to portfolio risk, rather than its own risk. The optimal trade-off of expected reward against the contribution to portfolio risk is “the Markowitz condition;” this condition (and its extensions when investor goals are more complex) remains the central
core of portfolio optimization. Indeed, portfolio optimization systems exist independently of the CAPM.

The CAPM studies a capital market in which all investors independently optimize and achieve the Markowitz condition for their portfolios. The CAPM characterizes the equilibrium condition of the market, when all individuals optimize their circumstances. The CAPM considers supply and demand in the capital market. It exploits the market-clearing condition that, at equilibrium, demand equals supply.

To obtain a neat equilibrium solution, the basic CAPM uses simplified assumptions: (1) All investors have identical expectations about security rewards; (2) all investors have identical expectations about security risks; (3) investors experience identical net returns (taxes and investment expenses are identical); (4) there are no investment constraints (no limits on borrowing or lending, no short-selling restrictions, no upper bounds on holdings); (5) there is a risk-free asset, which is borrowed or lent at identical rates; (6) all investors maximize mean/variance utility functions over a common investment horizon and are risk-averse; (7) investors experience risk only from the investment portfolio (there are no risky assets or liabilities excluded from the problem); (8) markets are perfect (each investor is a price-taker who does not believe he can influence price, there are no transaction costs and no costs of acquiring information).

Evidently, these statements rule out many aspects of diversity and assume away the process of information search and forecasting. The conclusions following from the assumptions are consequently clear-cut: (1) Each individual investor’s portfolio satisfies the Markowitz condition; (2) each investor’s portfolio of risky assets has the same composition as all other investors’; (3) the market portfolio, which is the aggregate of all portfolios, therefore has this same composition; (4) hence, the market portfolio is efficient for all investors, the unique “mutual fund” of all risky assets that exactly suits the needs of all investors; (5) since the market portfolio is efficient, any other portfolio of risky assets is inferior; (6) investors price each security in the market so that its expected reward compensates for its contribution to risk in the market portfolio (i.e. the equilibrium equation is the familiar “security market line,” with expected excess security
return being proportional to beta — in other words, alpha is zero for all securities); (7) hence, every portfolio also has an alpha of zero and every portfolio other than the market portfolio is inferior to the market portfolio because it has incremental diversifiable risk — not because it has a negative alpha.

THE TRANSITION FROM THE CAPM TO APPLICATION

If the CAPM were strictly true, there would be no active management. All investor expectations would be identical, and all investors would hold a single “consensus portfolio.” The correct prices for asset sets would emerge magically as the consequence of costlessly materialized expectations. There would be no investment research.

We will find it interesting to trace a sequence of relaxations of the assumptions, which adds realism and consequently leads to more relevant and fruitful predictions. In what follows, I have tried to list successive relaxations in approximate order of importance, beginning with the most significant.

1. Investors actually experience different earned, after-tax returns due to differential tax law. Because different classes of investors face different tax laws, they hold grossly different portfolios. For example, municipal bonds are held by taxable investors, not by tax-exempt investors, and preferred stocks are held more by corporate investors than by individuals. Each investor favors those securities for which he has a comparative advantage (a lower tax, relative to other investors). Features of assets that are important to tax law, including the distribution of return between yield and capital gain, become important to each investor’s portfolio decision. Therefore, these same features are important to aggregate demand and consequently figure in market equilibrium (Brennan [1970]).

2. Investors have diverse expectations, obtained by a research process and influenced by the actions of other investors in the market. Since expectations are diverse, there is no set of “true expectations” revealed to any market participant that he may use to define the true ex ante efficient portfolio. In practice, there is no such thing. Instead, the research process of each investor builds a set of expectations that constitute best judgment. The efficient portfolio of the investor, defined with
respect to his or her expectations, is not efficient in any absolute sense. The market clearing process now reflects, not “true expectations” as in the CAPM, but a “consensus expectation, “ which is an average of investors’ expectations (Lintner [1969]).

The opportunity to do valuable research creates a competitive research contest among market participants. The need to evaluate diverse skill generates performance analysis. Since profit from research is gained through portfolio revision, transactions tend to disclose beliefs of market participants, which results in a competitive trading process.

Costs of information and research must be debited against investment returns, except in the case of the passive investor who accepts market prices as “fair” at all times. Securities with higher information costs per dollar invested hence require a higher return to compensate. For smaller companies, the magnitude of potential investment is so small that the investor may require a significant premium return to offset the minimum cost of effectively monitoring the security. This argument suggests that small companies may offer higher gross returns in equilibrium, before deduction of the information costs.¹

3. A group of like securities, such as equities, is often styled as “the market” when it is actually only “a market” among multiple markets of risk assets. Analysis within one such market is actually only one component of a larger optimization problem: For example, when analyzing equities separately from bonds, or when separately analyzing individual countries in a multinational portfolio. Moreover, investors are exposed to risks arising from nonfinancial assets (claims on labor income and personal property). Such risks are tied to personal skills and preferences, the risks are often uninsurable, and the assets and liabilities may not even be marketable. These omissions can be classed as “excluded assets and liabilities” (Rudd and Rosenberg [1980]). Investors are concerned not only with the variance of their risky-asset portfolio within a “market,” but also with the covariance of the risky-asset portfolio with other risky events in the economy.

¹. Footnotes appear at the end of the article.
Investors’ attitudes toward these factors of covariance are diverse and express a need to hedge within the capital market those individual risk exposures that are outside the capital market. Investors’ holdings within the market differ for this reason. “Optimization” in any one market, ignoring excluded opportunities, is suboptimal.

4. There are constraints upon investment. Costs of borrowing and short positions are typically higher than the returns from lending or long positions. Moreover, widespread institutional barriers against short positions persist. Often, barriers against high concentration persist as well. The result is that any typical investor is a true “marginal investor” on only a fraction of issues; on other issues, the position is already at a constrained bound and is not altered by some changes in expected security return. The “consensus” appraisal of a security responds differently, and less sensitively, to changes in constrained investor expectations.

5. There are transaction costs arising from commissions and spreads as well as frequent and substantial transaction costs or benefits associated with tax effects and book-value accounting. As a result, many positions are “grandfathered,” so that the investor requires a substantial change in expected reward to induce a trade. Transaction costs also influence equilibrium returns, because the investor must allow for the expected cost of transaction. There are also investment costs associated with custody of securities and accounting of returns. Securities having relatively high expected investment and transaction expenses must have higher equilibrium rewards to compensate. For example, there are higher transaction costs in the purchase and eventual sale of securities having illiquid markets; such securities will be traded less often, but since they must be traded occasionally, a higher gross return may be required in equilibrium to amortize this expense.

There are unusually large surveillance and accounting costs when a security’s trading is suspended due to extreme uncertainty or reorganization; hence, companies with higher probability of bankruptcy may exhibit higher equilibrium gross returns, to provide a cushion to cover possible expenses of this kind.
6. Unpredictable inflation causes assets with fixed nominal returns to have risky “real” returns (returns expressed in purchasing power). Unless inflation-indexed, default-free bonds exist, there is no true risk-free asset. This fact makes minimum-risk portfolios of risky assets a more relevant investment vehicle.

7. Investors have diverse goals. We can approximate these moderately well by mean/variance utility functions defined over the return (or the logarithm of return) over a short holding period, but this “induced myopic utility function” cannot capture all of the subtleties of the multiperiod decision problem, particularly when returns are themselves serially dependent, as is the case for nominal returns on bonds, or when there are investment vehicles with highly asymmetric distributions of returns, as is the case with options.

Also, some investors are obligatory holders of certain securities: The most prevalent causes are retention of voting control of the corporation and incentive compensation for management and employees. In some cases, large fractions of outstanding shares may be closely held and “disappear from the market” for extended periods.

8. When the theory calls for us to compute the portfolio of all outstanding assets, there are problems in specifying this portfolio. Since these problems are central to the ambiguity of the “market portfolio,” it is important to go into them carefully. In the best of circumstances, a security is unambiguous in definition and publicly recorded: For example, a common stock registered with the SEC. Here, there may be problems in finding the number of securities outstanding, but such information is being steadily collected worldwide for most categories of financial assets. The difficulty here is to identify cross ownership (one corporation owning another’s stock); outstanding securities that are held as an asset of some other security should not be double-counted; fortunately, holdings of public companies in excess of 5% are registered in the U.S.

Other assets are unambiguous in definition, but hard to find. These include unregistered common stocks, privately placed debt, and nonfinancial assets (homogenous assets such
as commodities are less of a problem than heterogeneous assets such as real estate and antiques).

Finally, there are assets that are ambiguous in their very definition, such as the present value of labor income (an unmarketable asset) and government debt. Government debt is an investors’ asset that is offset by taxpayers’ liability in the form of the obligation to pay future taxes. As we pursue the reasoning implied by the aggregate social balance sheet, more and more assets tend to be offset by liabilities, so that the risk associated with them is a risk of redistribution, rather than a risk to the society’s aggregate portfolio. Redistributive risks arising from the political process are very difficult to model.

The response to these problems is to separate assets into broad homogeneous classes or “markets” and to use an index of approximate outstandings of more prominent securities in each market as a surrogate for the total market. Weaknesses of this solution are unrepresentativeness within the market, which we will argue is of small import, and exclusion of assets and liabilities, the problem mentioned above.

THE MARKET FACTOR

In view of the deficiencies of the simple CAPM, its great effect upon application must be explained through the vitality of its constructs and the manner of thinking that it has engendered.

One major contribution has been to call attention to the distinction between market-related and residual return. The “market model” expresses the return on every security as the sum of a systematic (market-related) component and a residual component that is uncorrelated with the market. The security’s response to the market is its systematic risk coefficient, or beta. The mean and variance of the market return determine, through the beta coefficient, the systematic mean and variance of a security or portfolio. The expected value of the residual component and the residual variance are important properties of the security’s residual return, and covariance among residual returns is important for portfolio residual variance.

The CAPM thus emphasizes the return on the market portfolio (or a surrogate for it). Empirical studies have shown that we can use the return on any stock market index to explain
a large fraction of the variance of individual security returns and a still larger fraction of the variance of portfolio returns. This confirms long-standing recognition of “market movements.”

Within each investment “market,” such as equities, bonds, or real estate, securities tend to move up and down together. The statement, “The market is up,” or “The market is down,” would be meaningless otherwise. Numerous studies have since confirmed that in each market one “prominent factor” accounts for a far greater proportion of the variability of security returns than any other single factor, and that all — or almost all — security returns respond to the factor in the same direction. Following common parlance, this can be called the “market factor.”

Studies in equity and bond markets confirm that broad-based indexes of returns within each market are highly correlated, even though the included securities and index weights are different. The correlation is so high because any widely based and correctly computed index tends to show up the prominent factor and becomes a surrogate for it.

Since this factor is so prominent, we should naturally take it into account in the investment process. Active investors almost universally attempt to forecast the movements of the market, although many do not make “market timing” an important element of their investment policy. In many organizations, “top-down” guides to security analysts are provided in the form of market forecasts; individual analysts forecast individual security returns conditional upon the market forecast.

We can use any widely based index to define the market. There are reasons, however, for using a “market portfolio” that is a capitalization-weighted average of all outstanding securities in the market. For one thing, investment return on this portfolio is the aggregate of return for all investors and therefore a natural variable in a macroeconomic model. Furthermore, since the market portfolio is the weighted average of all investors’ holdings, weighted by their wealth, its value reflects a weighted average of all investors’ valuations, or, in some sense, a “consensus valuation.”
An important element of performance analysis is the attribution of return among various aspects of investment strategy. Because of the prominence of the market factor, strategy with respect to that factor is usually the first aspect to be emphasized, and Fama [1972] suggested that market timing be distinguished from selectivity of individual stocks as an element of performance, as did Rosenberg [1978]. This performance decomposition requires specification of a market return. The natural surrogate to use is the market index that is being forecast and with respect to which strategy is defined. Since the concern is to identify investment strategy, the index that is called for should be the index the investor is using.

The broad-based index, as market factor surrogate, is also important in historical studies of the reward to market factor risk exposure. Using any definition of the market factor, it is a legitimate question to ask how security returns have related to the security’s exposure to that factor. Of course, this is only a pure test of the CAPM if the market-factor surrogate return is identical to the market portfolio return.

Putting aside such niceties, we can still consider the important question of the actual historical pattern of compensation. Several studies have shown that there has been higher historical average reward for higher beta stocks. Moreover, the comparative studies to date have found little change in the estimated amount of reward when we vary the definition of the market index among broad-based indexes. In other words, the exact definition of the market portfolio has not had an important effect upon the estimated reward for exposure to risk of the equity market factor.

Another important application of the “market factor” is as a surrogate for investment opportunities in that market, for consideration in a decision problem by an investor who is “allocating investment” across various markets. In this application, what is required is a representative index for each of several markets. Again, the “market portfolio” for each market is a natural index to use, since it is typical of all investors in the market.

As soon as the CAPM was publicized, research showed that securities had different degrees of responsiveness to the market factor (different betas). Later, we found that significant
differences in beta could be consistently predicted. Of course, the exact definition of beta follows from the chosen surrogate for the market factor.

Precisely because of the great prominence of the market factor in all broadly based indexes, however, substitution of one such index for another changes the definition of beta very little. What occurs is largely a change in the scale of betas (as from Fahrenheit to Centigrade), with little relative change in individual betas. Of course, real changes do emerge when one index is significantly biased relative to another, so as to have importantly different exposure to some secondary common factor in the market. On the other hand, these are second-order changes, because the market factor is so much the most prominent. Therefore, beta becomes a meaningful and predictable characteristic of a security.

In these applications, expected residual return is, by definition, that element of expected return that is not due to the market. Expected residual return is therefore a key description of the individual security in any investment process where we single out the market return. The distinction between residual and market related return is a consequence of the structure of the investment process (by virtue of the market return being a distinct construct), and not a result of the CAPM. Since we express expected residual return conventionally in terms of the intercept of the security market line (zero-beta return) and the individual security’s alpha, alpha is the natural way of describing the security’s desirability for investment purposes, net of attractiveness arising from exposure through its beta to zero beta and market returns. There is nothing mysterious about this alpha; it is simply an expression of judgment on the security’s expected return. For the same reason as with beta, alpha changes little when one broad-based market index is substituted for another.

Note that none of the applications discussed in this section have been dependent upon the CAPM, nor are they importantly influenced by the exact definition of the market portfolio.

THE AGGREGATE OF INVESTORS’ PORTFOLIOS

The CAPM rests upon the market-clearing condition that aggregate demand must equal aggregate supply. Aggregate demand is the sum of all investors’ portfolios. Aggregate
supply is the ensemble of securities, which, when viewed as the portfolio of all outstanding assets, is the market portfolio. Therefore the average of all investors’ portfolios, weighted by the values of their investments, equals the market portfolio. The investment-weighted average of the returns on investors’ portfolios similarly equals the market portfolio return. This simple relation, an accounting identity, has profound consequences.

One consequence is an argument for a passive investment strategy equal to the market portfolio. Such strategies have come to be called “index funds.” If carefully constructed, the fund earns a gross return equal to the market portfolio return. The net return is less, due to small costs of transactions and a small passive management fee. The market portfolio return is also the average of all investors’ gross returns. Average net returns of all investors are lower due to transaction and management costs, and these expenses are significantly greater for active investors than for the passive fund.

Hence, the net return of the index fund is higher than the average net return of all investors. Moreover, since active investors’ positions diverge from one another in the attempt to profit from diverse expectations, investors, on average, take more risk than is present in the market portfolio. Consequently, the passive “market portfolio” strategy earns an above-average net return at a below-average risk.

Imprecise specifications of the market portfolio can damage this argument only if the error causes a failure to capture the average return of investors. Since data on institutional investors’ holdings are in the public domain, we can compute and approximate the average holdings of this population of investors. There seems little chance that the ambiguity of stock and bond market portfolios is a significant obstacle to attaining above-average net performance through a broadly based passive “market fund.”

The CAPM asserts that the “market portfolio” is not just average in gross return, but also “efficient”: The market-portfolio strategy is the perfect strategy for all investors. But the valid claim of above-average net return is more important than the problematic claim of perfection. My impression is that the CAPM based argument of efficiency has been peripheral in
the marketing of index funds. If it were crucial, surely there would have been efforts to adhere to the exact requirements of the CAPM by making equity index funds closely representative of the capitalization weighted equity sector. Instead, the first passive fund followed an equal-weighted index, and most strategies since that time have matched conspicuous indexes such as the S&P 500 or the Dow Jones Industrials, rather than a broader index such as the NYSE.

Universal performance comparison with the market portfolio is another application that follows upon recognition that the market portfolio return is an average gross investor return. The market portfolio return defines the average gross payoff of the investment “game” in any market. In other words, average residual performance, relative to the market return, is zero. Comparison with the market portfolio defines a zero-sum game.

One widespread use of performance measurement is to array the accomplishments of many managers in competition with one another. For this purpose, there might seem to be no reason for the inclusion of the market portfolio return as a universal benchmark: After all, the managers’ returns can simply be ranked in deciles. Nevertheless, an index that is a good surrogate for the market portfolio return is a desirable benchmark for several reasons.

First, the index does define the average gross return in the competition to the extent that average holdings of the investors approximate the weights of the index. Second, the index provides an unambiguous and unbiased benchmark. By contrast, percentiles of comparison populations tend to be biased, due to selective survivorship and due to retrospective inclusion of favorable past history of new entrants. Third, if passive funds track the index, the index represents a conspicuous investment alternative and is interesting for this reason.

Misunderstanding has been widespread to the effect that performance comparisons versus the market portfolio are undermined by the falsehood of the CAPM. According to this argument, it is the efficiency of the market portfolio that makes such comparison interesting, rather than its average character.
We must dispose of this misconception. In order to know the efficient portfolio (or the efficient frontier), one must have absolute foreknowledge of the true properties of all portfolios — both expected reward and risk. Computation of an estimate of those properties, based upon some performance numbers, is then a meaningless exercise: If the computations produce a different answer, this must be due to statistical noise. In short, knowledge of the efficient portfolio renders performance analysis meaningless. Nor can any paradigm for performance analysis be based upon hypothetical knowledge of the efficient portfolio, unless the purpose of the exercise is to reject that hypothesis, and by so doing, to deny meaning to the claim of efficiency.

Since the true efficient portfolio is unavailable and irrelevant to performance measurement, the next question is whether a benchmark that is believed to be more efficient that the average (market) portfolio is a meaningful possibility. For performance analysis by a single investor, reflecting his special circumstances, this is a valuable step, discussed below. As a device for universal comparison, however, any such benchmark destroys itself. Any candidate portfolio other than the average holding, which is believed to be more efficient for the average client, is intrinsically self-disfulfilling. As soon as the candidate portfolio is accepted as being more efficient, managers naturally attempt to move the average of their holdings away from the present average portfolio toward the candidate. As market prices adjust to changing demand, the efficiency advantage of the candidate portfolio must erode; the process of adjustment cannot cease until the average portfolio becomes efficient and thereby supplants the candidate.

The self-destruction argument relies on disclosure of the candidate portfolio to all investors. If the portfolio is known in advance to some, but kept secret from most, it may not be self-disfulfilling, but then its usefulness in current universal comparison disappears. And if the portfolio is only arrived at ex post, its retrospective use is subject to all of the legitimate criticisms directed against hindsight, as well as to inevitable controversy over the fairness of a retrospective standard. In short, the return on the average portfolio is uniquely singled out as a benchmark for comparison in a universal population,
and the idea of greater efficiency seems to have little relevance.

A third important application that follows from the “averageness” of the market portfolio is the concept of “nonconsensus” forecasts. For aggregate demand to equal aggregate supply, each security must be priced so that the “average marginal investor” will hold it. The security’s price settles where it is “fairly priced,” in the view of marginal investors. It follows that any one investor should favor a security to the extent that he finds reasons to believe that it is more desirable to him than to the marginal investors.

This line of reasoning demonstrates the usefulness of the information disclosed by market price. To the extent that manipulation is absent and all investors are just about as well informed as any one investor, market price discloses to that investor a meaningful consensus appraisal. Moreover, if the investor is typical of the marginal investors who hold the asset and if he finds no reason to differ from their judgments, then the correct position is probably the average of their holdings, which may be close to the proportion in the market portfolio. He can then justify deviation from the market proportion only by nonconsensus beliefs.

The first two applications discussed in this section, passive management and performance comparison, relied on the “averageness” of the market portfolio. The third application begins to rely on a more subtle property: The market portfolio, as an average of portfolios, is optimal with respect to a similar average of investor’s expectations, or “consensus expectations.” The nature of the averaging process is made precise below.

THE MARKET PORTFOLIO IN EXTENSIONS OF THE CAPITAL ASSET PRICING MODEL

As the CAPM assumptions are abandoned in favor of more realistic ones, multiple features of securities may influence expected returns at equilibrium. Features that enter into the tax law are one category. Features that determine correlations with risks outside the market are a second. Features influencing information and investment costs are a third.
Thus, equilibrium expected reward may depend upon several features. The next point is that the equilibrium reward for any particular feature may change, due to changes in the market environment; when a change occurs, there is a one-time windfall return on securities in proportion to the amount of that feature that they possess. Such windfalls constitute an uncertain factor in market return related to the feature. The expected reward and uncertain windfall are combined into a factor of return: Expected factor return rewards the feature, and risk of the factor introduces the uncertainty from possible changes in reward.

Moreover, asset features associated with likely rewards generally correlate with fundamental circumstances of the issuer. For example, a common stock’s (1) yield, (2) size, (3) probability of bankruptcy, and (4) covariability with the bond market are not only technical features of the stock, which may be rewarded at equilibrium, but also relate strongly to the fundamental operating circumstances of the company and its industry. Outside economic events that produce returns in proportion to these circumstances therefore cause investment returns that align with the features, and which — when viewed as returns — become further variability of the factors.

The cumulative effect is an environment in which multiple features have associated common factors, with possibly nonzero rewards and definitely nonzero risk. There is a widespread misunderstanding that the market model implies a single-factor model that therefore rules out multiple factors. In fact, the distinction between market and residual return is quite separate from, and coexists with, the distinction between multiple common factors and specific return. A multiple-factor and specific-return model implies, for any given market portfolio, a market- and residual-return model that is superimposed upon the multiple-factor model (Rosenberg [1974]). For some applications of the basic multiple-factor model, there is no need to distinguish between market and residual return. In many other applications — in particular, all those where the market return is a distinct element in forecasting and strategy — we must distinguish the market factor, and we can then express other factors as residual factors.
When we admit heterogeneous expectations such a thing as “true expected return” no longer exists. Nevertheless, we can express the equilibrium for the capital market in terms of investment-weighted averages of expectations that thereby define “consensus expectations.” Equilibrium continues to require that, in addition to the other multiple features that may be rewarded, covariance with the market portfolio, or systematic risk, is rewarded. Reward and covariance are here defined in terms of consensus expectations.

When we take into account restricted borrowing and lending and constraints on holdings, the averaging process that underlies equilibrium becomes more complex. In particular, some investors’ attitudes toward some securities have no direct impact on those securities’ prices, because the investors, constrained against adjusting their holdings, cease to be marginal investors. When the average is defined across those investors who are truly “marginal” for every security, however, it is again true in equilibrium that covariance with an average portfolio is rewarded in the consensus view.

It is difficult to imagine a market equilibrium in which covariance with the equity market portfolio, or some risk measure that is closely akin to this, would not be rewarded. A single, highly risky, prominent factor exists in the equity market and appears to constitute a societal risk rather than a redistributive risk. Covariance with the market portfolio is a surrogate for risk exposure to this factor and hence a conspicuous candidate for reward. Moreover, at equilibrium, the market must clear: Any element of risk must be compensated for, in aggregate, in proportion to that element’s covariance with portfolios of investors who hold that security.

It is possible for the subset of investors holding a particular asset to have portfolios such that the covariance of that asset with their portfolios is different from its covariance with the market, but I have rarely seen plausible examples, and then only in cases of excluded risks. For almost all elements of risk, it does seem probable that the aggregate of portfolios exposed to that element will covary with it similarly to the market portfolio, so that the security’s covariance with the market portfolio is a reasonable guide to investor’s risk exposure that must be compensated. In sum, equilibrium considerations do
suggest that market factor and market portfolio covariances are natural candidates for reward.

The historical studies of which I am aware tend to confirm the existence of reward for market index covariance, although smaller in magnitude than the CAPM would imply. Future studies will no doubt give more precise information on historical rewards to multiple factors, although factor variability inescapably obscures the historical expected reward. The fact remains that a historical study is exactly that — a historical study.

To expect predictive content for the future, one must presume (1) that pricing relationships in the market are stable and (2) that pricing will not change as a result of the study. Stability suggests that the relation must be an equilibrium one. For disclosure of the study not to be self-disfulfilling, the factor compensations found by the study must be consistent with equilibrium. Thus, the task of predicting factor rewards cannot be separated from characterization of equilibrium factor rewards through economic analysis of market circumstances. The single most important tool of microeconomics has always been the insight that the market must clear; the market portfolio is the construct that implements this condition.

**PERFORMANCE ANALYSIS**

Performance analysis includes all aspects of the study of historical performance for the purpose of predicting managers’ skill and usefulness in the future. Universal performance comparison is the simplest framework, with limited usefulness for the client’s decision problem. The purpose of this section is to touch briefly on the roles of the market portfolio, market model, and efficient portfolio in the performance-analysis process. The approach taken here is quite different from recent literature (Roll [1978, 1979b], Mayers and Rice [1979], Cornell [1979]).

The money management client attempts to construct a best investment strategy, built upon the services of one or more money managers. Most large pools of funds now apportion their assets among multiple managers. The client’s problem is to choose a portfolio of managers, just as an investor would construct a portfolio of securities. The client’s decision has also additional dimensions because of the potential flexibility
that the manager has to restructure the portfolio and management fee.

It is extremely useful, before tackling the question of performance measurement, to consider what would be the optimal portfolio for the client if management were completely passive. Passive management would make no use of special information but, instead, would consider only the relatively permanent aspects of the capital market — those aspects that characterize equilibrium. The client’s equilibrium portfolio is shaded toward those assets that are relatively favorable for him, in comparison with the average investor. The client’s goal, when constructing an optimal portfolio in a given market, may be atypical, due to special tax circumstances, excluded assets and liabilities not in that market, and multiperiod investment goals. The equilibrium portfolio reflects the unique circumstances of the investor but not the special information that is generated in the competitive research progress of active managers.

The outcome of this process is, at least, a specification of this equilibrium portfolio, a judgment as to equilibrium rewards of security features, and a specification of risk aversion for the client, not only with respect to the risk of the market factor, but also with respect to various other elements of risk that may have had special disutility, or even utility in hedging. In view of the best description of equilibrium in the capital markets available to the investor, this equilibrium portfolio is the efficient portfolio: It may coincide with the market portfolio, as the CAPM would suggest, but it may not.

There are several advantages to constructing the equilibrium portfolio. First, it is itself an investment alternative: The client can commit some funds to a passively managed equilibrium strategy, with low management fee, low transaction costs, and without active risk. The equilibrium portfolio exploits the disclosure of consensus expectations through the capital markets and takes into account the atypical needs of the client as well. Second, the equilibrium portfolio, in the client’s judgment, gives the highest possible utility that can be achieved without superior expectations. Therefore, it is a benchmark for evaluation of managers’ skill.
For example, for the typical tax-exempt investor, the equilibrium portfolio of equities may be shaded toward high-yielding stocks, according to the still controversial argument that this reflects the equilibrium distribution of equities between tax exempt and taxable investors. This portfolio may be typical of the equilibrium portfolios of all tax-exempt investors, and, interestingly, one way of determining the equilibrium portfolio is to take the average of the portfolios of all investors with similar status; this presumes that the average of the peers have knowledgeably determined their investment strategy.

An equilibrium portfolio shaded toward high-yielding stocks is historically a slightly stiffer performance benchmark than the market portfolio. By advocating this equilibrium adjustment, the client asserts that, from the point of view of tax-exempt returns, market equilibrium permits one to outperform the market portfolio without superior information. (Conversely, from the taxable investor’s point of view, one can outperform the market portfolio in after-tax return by shading toward low-yielding, growth oriented stocks. The taxable investor would evaluate the after-tax, net performance of the manager, using a growth-shaded benchmark.) The equilibrium benchmark reflects the opportunities, built into market equilibrium, to maximally serve the client with consensus expectations.

Next, suppose, further, that the client is a highly levered financial company. The pension portfolio of the company is only one of the assets of the pension fund; the fund’s main asset is the claim on contributions from the ongoing earnings of the company itself. This is an “excluded asset” in the problem of the beneficiaries. Because of the company’s great exposure to the financial markets, the needs of the pension fund portfolio (and of the PBGC, as insurer) include the goal of hedging as much of the company’s risk as possible within the pension portfolio. The resulting equilibrium portfolio may be quite atypical in its equity holdings of financial firms and in asset allocation.

Taken in isolation, this equilibrium portfolio is not efficient and constitutes a benchmark that is easier to beat than the market portfolio. When performance is correctly measured, however, allowing for the special disutility of covariance with
the financial markets, the equilibrium portfolio is again the best portfolio employing only consensus judgments. The special circumstances of the client are implemented, not only by a special equilibrium portfolio, but also by computation of the disutility of risk appropriately for the client.

In studying the performance of any single manager, we can compute the utility contribution from the relative performance of that manager, compared to the equilibrium portfolio. This procedure is not identical to any of the classic procedures of performance analysis, even when the client’s disutility for risk is homogeneous, but it is analogous. In a sense, the client’s equilibrium portfolio takes the market portfolio’s role: We substitute a more efficient portfolio (the client’s equilibrium) for a less efficient one (the market portfolio — which is less efficient for this client).

The treatment of performance analysis is not yet fully developed, however. Each money manager’s portfolio may differ from this equilibrium portfolio for two reasons. First, the normal investment emphasis of the manager may differ from the client’s equilibrium portfolio. The manager’s style, investment specialization, or accustomed habitat may cause the normal or neutral holdings of the manager to be atypical, and the equilibrium or neutral position of the client may be atypical also. This “normal difference,” when the manager’s norm and the client’s equilibrium are compared, does not reflect the manager’s judgment. Second, the actual portfolio of the manager differs from his normal by an “active portfolio,” resulting from the current set of active judgments of the manager. The active portfolio (Treynor and Black [1973]) is a “hedge portfolio” (a portfolio with zero dollar value) that manifests the manager’s skillful judgment.

The equilibrium portfolio is the ideal circumstance for the client if consensus expectations are correct. Yet, the client hopes to find managers whose expectations are superior to the consensus. The ability to develop superior expectations is the active advantage of the manager, which he reflects in the active portfolio; it thereby redounds to the benefit of the client. Clients should apply performance analysis to the active portfolios of managers — that is, to the performance difference between actual and normal portfolios.
When the active portfolio is studied, the manager’s normal portfolio is functioning as a benchmark for his performance. The manager chooses his own benchmark, since the normal portfolio is his now *ex ante* description of his neutral point. The manager’s normal portfolio thus plays another traditional role of the market portfolio, in parallel to the equilibrium portfolio.

This might seem to undermine performance analysis, since the manager can choose any benchmark. Since the client is informed of the normal portfolio in advance, however, the client uses this information to construct a stable of managers whose aggregate normal positions sum to the client’s equilibrium. If necessary, a portion of the client’s funds can be managed passively to bring the aggregate normal position in line, by compensating or hedging normal biases of active managers. In so doing, the client immunizes himself against the normal biases of the managers, which become irrelevant and disappear from the performance analysis. The greater the diversity among managers’ normal positions and among clients’ equilibrium portfolios, the greater is the need for treatment of normal bias.

The advantage of the normal benchmark is that it allows the manager’s skill to be isolated and most accurately estimated. The active portfolio, in the manager’s view, is the exact reflection of his judgments. The normal bias of the manager, which does not reflect an active decision, introduces incidental noise that is best eliminated.

The normal portfolio also plays a key role because it represents the “rest point” for the manager. The range of “variably aggressive” portfolios, beginning with a passive portfolio identical with the normal and moving out along the active frontier with increasing emphasis on active judgments, are all potential and valid outputs of the manager’s production process. In principle, it is open to manager and client to determine the location along this frontier at which the client’s portfolio is to be managed and the fee that will be charged. This open dimension is increasingly acknowledged in portfolio management. The active portfolio defines the character of the active frontier, something that is hidden in any analysis of the actual portfolio that omits a normal benchmark.
The performance of the active portfolio is evaluated so as to reflect the client’s special disutilities for aspects of risk. The key problem is not just to compute historical performance; rather, it is to predict future performance. For this purpose, performance attribution to various aspects of investment strategy is valuable. As mentioned before, one important distinction is between market timing and selectivity, for managers whose investment process singles out the market forecast. When either the client’s risk attitude toward the market factor is distinctive (presumably due to excluded assets that are correlated with this), or when the manager singles out this factor, the market factor (and hence the market portfolio as the natural surrogate for the factor) plays an important role in performance analysis of the active portfolio.

The client predicts future management skill based upon (1) past performance, (2) external evidence concerning the quality of the manager’s investment process, and (3) prior skepticism derived from the competitive nature of the investment process. The last perspective arises because, if the manager’s portfolio were the result of random selection within his normal universe (stratified so as to produce his normal portfolio on average), then skill (the advantage relative to consensus) would be zero. This is also the case when the manager, with the best of intentions, is not capable of improving upon consensus expectations. It is clear that the average manager does not outperform the consensus (which is the average). Hence, prior skepticism takes the form of expecting the performance of the manager to match the passive performance of his normal benchmark, implying an active expected return (akin to alpha) of zero. Use of the passive normal portfolio as benchmark causes the adjustment for prior skepticism to take this simple form. This is a third advantage of the normal benchmark.

Based upon much information, including past performance, the client constructs a portfolio — which may be an admixture of multiple active and passive management processes — that is believed to be efficient for the future. In doing this, the predicted active skills of managers are incorporated, so that the expected return of the portfolio is incremented to reflect superiority relative to the consensus.

The question of using a “more efficient portfolio” in performance analysis can not be reopened. The equilibrium
The portfolio is efficient for this client with respect to consensus expectations. Why not go a step further and use the existing aggregate portfolio — efficient with respect to active beliefs — as the comparison benchmark?

The answer to this question explains why it is an equilibrium portfolio, and not an active portfolio, that should be used in performance comparison. The equilibrium portfolio, being a reflection of the consensus, has stable properties that are not influenced by the active management of existing managers. In this sense, it is a stable prior perspective. Also, it implies a clear benchmark for prior skepticism. By contrast, when another manager is compared to the active portfolio, the prior expectation of alpha is negative, there is less reason to expect stability in the relative skill of a new candidate, and performance attribution is clouded by the presence of the investment strategies of existing active managers.

Finally, a new manager’s pattern of active research and investment may well be related to and correlated with the active research and investment of one or more of the existing managers. If so, the existing portfolio should be reconstituted when the new manager is added (Rosenberg [1977], Sharpe [1980]), and the information needed to do this can only be established by analyzing all managers (existing and new) with respect to the equilibrium portfolio.

**DISCOUNT RATES ADJUSTED FOR FEATURES**

Valuation methods for financial securities based upon the discounting of expected future cash flows are common. For equities, there is the dividend-discount paradigm; for bonds, the computation of present value through the term structure of interest rates. These discount rates may vary with time and with the risk of the security.

The discount rate is also the required rate of return, since it translates future flows into present value. As such, it can serve the corporation for project valuation and as the “hurdle” in capital budgeting. The same discount rate is the “fair rate of return,” a construct with an increasingly important role in rate regulation.

The CAPM asserts that the discount rate for a risky project equals the risk-free rate, plus beta times the excess return on
the market portfolio. In practice, the “market portfolio” has been implemented as the S&P 500 or other broad equity market index. As we relax CAPM assumptions, the discount rate becomes a more complex function of security features.

The first natural extension is to free the “intercept” of the relationship, so that it is estimated from the data. This is done in fitting the *ex ante* capital market line and in historical studies of capital asset returns. Usually, the intercept is found to be higher than the risk-free rate, implying that the compensation for market risk is less than the full excess return on the market portfolio.

The next extension is to insert other features as candidates for reward. The security market line then becomes a “security market plane,” with the slope along each axis being the reward for the corresponding feature. *Ex ante* security market planes have been estimated with rewards for several features, through the dividend-discount model. Historical studies of returns have estimated rewards to such features as specific risk, total risk, yield, co-skewness, and size.

The central problem is to forecast rewards for these features. We can draw these forecasts from current valuation of securities, solving for rates that equate discounted forecast cash flows to current price. *Ex ante* relationships fitted to forecasted dividends for a large population of stocks can come up with quite precise estimates of the co-efficients of the fitted plane, but this is only a true characterization of the *ex ante* security market plane to the extent that the input dividend forecasts are effective proxies for consensus forecasts. Predicted rewards can also be extrapolated from historical studies of equilibrium returns. Finally, by modeling investors’ demand functions, rewards can be computed as the market-clearing conditions of a general equilibrium model. Whichever method(s) are used, there will inevitably be a large element of judgment in the predictions.

The question of rewards for factors other than equity market risk has been the subject of active study and controversy for a decade — and no doubt will continue to be so in the decades to come. Nevertheless, no one has refuted the existence of equilibrium reward for equity market risk; indeed, it has rarely been questioned, although the magnitude has been
in doubt. The concept of reward to equity market risk (or beta) is a theoretical insight, that, in my view, is likely to endure.

1 Smallness of the total company not effectively limit the holding of an individual investor, so the small investor might have a comparative advantage investing in small companies. Investors having personal dealings with the company may have an information advantage as well. Such investors might then become the dominant investors in small companies, due to their comparative advantage. The equilibrium returns of small companies would then reflect primarily the portfolio optimization of such investors, with respect to their expectations and wealth.

BIBLIOGRAPHY


