BARRA

Single Country Equity

RISK MODEL HANDBOOK

0112 O 03/98

Contents

About BARRA

In recent years the investment management industry has adjusted to continuing changes—theoretical advances, technological developments, and market growth. To address these challenges, investment managers and financial institutions require the most advanced and powerful analytical tools available.

A pioneer in risk management

As the leading provider of global investment decision tools, BARRA has responded to these industry changes by providing quantitative products and services that are both flexible and efficient. Since our founding in 1975, BARRA has been a leader in modern financial research and techniques.

Initially, our services focused on risk analysis in equity markets. Our U.S. Equity Model set a standard of accuracy that BARRA continues to follow. BARRA uses the best data available to develop econometric financial models. In turn, these models are the basis of software products designed to enhance portfolio performance through returns forecasting, risk analysis, portfolio construction, transaction cost analysis, and historical performance attribution.

In 1979, BARRA expanded into the fixed income area with the release of U.S. bond valuation and risk models. In the mid-1980s we developed a global tactical asset allocation system: The BARRA World Markets Model™. More recently, the Total Plan Risk™ approach was developed to provide multi-asset-class value-at-risk (VAR) analyses.

BARRA now has offices around the world and products that cover most of the world's traded securities. By 1997, our clients comprised approximately 1,200 financial institutions worldwide managing over $7 trillion in assets. They rely on BARRA's investment technology and consulting services to strengthen their financial analysis and investment decision-making.

Introduction

In this handbook

This handbook contains a general discussion of equity risk and return, and the methods BARRA uses to model portfolio risk.

Chapter 1. Why Risk is Important gives an overview of why financial professionals should care about risk.

Chapter 2. Defining Risk outlines the basic statistical concepts underlying risk analysis, and traces the history of equity risk theory.

Chapter 3. Modeling and Forecasting Risk discusses the application of multiple-factor modeling (MFM) to the equity risk analysis problem.

Chapter 4. Modern Portfolio Management and Risk relates the various types of active and passive equity management to the use of a risk model.

Chapter 5. BARRA Multiple-Factor Modeling details the process of creating and maintaining a BARRA equity MFM.

The Glossary and Index are useful resources for clarifying terminology and enhancing the handbook's usefulness.

Further references

BARRA has a comprehensive collection of articles and other materials describing the models and their applications. To learn more about the topics contained in this handbook, consult the following references or our extensive Publications Bibliography, which is available from BARRA offices and from our Web site at *http://www.barra.com*.

Books

Andrew Rudd and Henry K. Clasing, *Modern Portfolio Theory: The Principles of Investment Management*, Orinda, CA, Andrew Rudd, 1988.

Richard C. Grinold and Ronald N. Kahn: *Active Portfolio Management: Quantitative Theory and Applications*, Probus Publishing, Chicago, IL, 1995.

1. Why Risk is Important

Superior investment performance is the product of careful attention to four elements:

- forming reasonable return expectations
- controlling risk so that the pursuit of opportunities remains tempered by prudence
- controlling costs so that investment profits are not dissipated in excessive or inefficient trading
- controlling and monitoring the total investment process to maintain a consistent investment program

These four elements are present in any investment management problem, be it a strategic asset allocation decision, an actively managed portfolio, or an index fund—managed bottom-up or top-down, via traditional or quantitative methods.

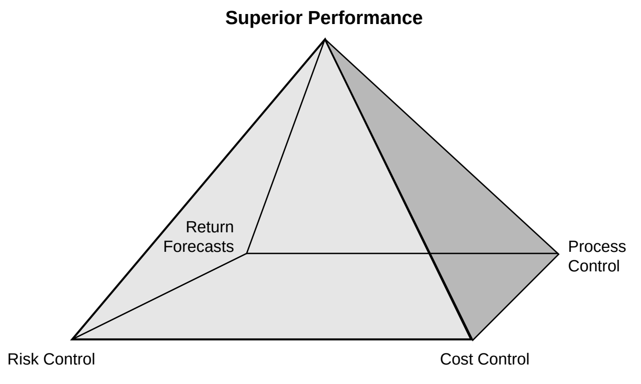


Figure 1-1
The Performance Pyramid

In a simpler view, return and risk are the protagonist and antagonist of investing. According to an old adage, the tradeoff between return and risk is the tradeoff between eating well and sleeping well. Clearly, risk doesn't just matter to quants!

Ignoring risk is hazardous to your portfolio. The optimal strategy ignoring risk places the entire portfolio in one stock. But no institutional investor follows this strategy. Hence risk considerations must impact every institutional portfolio. Unfortunately, they sometimes do not impact them enough.

We need not look far to find examples of financial disasters that arose through lack of sufficient risk control. The debacles of Orange County, Barings Bank, and the Piper Jaffray Institutional Government Income Fund all testify to the dangers of ignoring or poorly understanding risk.

But risk analysis is more than avoiding disasters—it can in fact enhance opportunities. Peter Bernstein has argued that a lack of understanding of risk holds back economic development.[1] Modern economic growth requires understanding risk.

What are the expected returns to a new venture? What are the risks? Do the returns outweigh the risks? Can I hedge the risks? In modern economies, the future is not beyond management, not simply subject to the whims of many gods. In fact, the period which marked the development of probability and statistics (during and after the Renaissance) also marked a time of profound growth in trade, exploration, and wealth. The ideas of risk management enabled the modern economic world, according to Bernstein. Risk analysis enhanced opportunities.

While Bernstein's argument may seem inspiring—though not of day-to-day relevance—in fact the goal of risk analysis is not to minimize risk but to properly weigh risk against return. Sometimes risk analysis leads to taking more risk.

The goal of risk analysis

Risk is important. It is a critical element of superior investment performance. Good risk analysis should provide not only a number—a quantification of risk—but insight, especially insight into the "Performance Pyramid."

We have illustrated superior performance as a three-dimensional object. A single risk number is only one-dimensional. So what do we mean by insight?

---

1. See Peter L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, John Wiley & Sons, New York, 1996.

Risk analysis should uncover not just overall risk, but the largest and smallest bets in the portfolio. Do the largest bets correspond to the highest expected returns? They should. If they do not, the portfolio isn't properly balancing return and risk. Are the bets too large or too small? What is the "worst case" scenario? How will the portfolio compare to its benchmark?

Robust risk analysis can provide answers to all these questions as well as insight to all investors. In this volume we will discuss the history and current practice of equity risk modeling for single country markets. Other methods are used for different securities, such as bonds or currencies, and for different market structures, such as the global stock market. The underlying message is clear: The investor armed with superior methods of assessing and controlling risk possesses a significant competitive edge in modern capital markets.

2. Defining Risk

Some basic definitions

In an uncertain investment environment, investors bear risk. Risk is defined as the total dispersion or volatility of returns for a security or portfolio. Further, risk reflects uncertainty about the future.

We will define risk as the standard deviation of the return. Risk is an abstract concept. An economist considers risk to be expressed in a person's preferences. What is perceived as risky for one individual may not be risky for another.[1]

We need an operational and therefore universal and impersonal definition of risk. Institutional money managers are agents of pension fund trustees and other asset sponsors, who are themselves agents of the sponsoring organization and, ultimately, the beneficiaries of the fund. In that setting we cannot hope to have a personal view of risk.

We need a symmetric view of risk. Institutional money managers are judged relative to a benchmark or relative to their peers. The money manager suffers as much if he does not hold a stock and it goes up as if he held a larger than average amount of the stock and it goes down.

We need a flexible definition of risk. Our definition of risk should apply to individual stocks and to portfolios. We should be able to talk about realized risk in the past, and forecast risk over any future horizon.

The definition of risk that meets these criteria of being universal, symmetric, and flexible is the standard deviation of return.[2,3] If R_P is a portfolio's total return, then the portfolio's standard deviation of

1. There is a vast literature on this subject. The books of Arrow, Raiffa, and Borch are a good introduction.
2. An economist would call the standard deviation a measure of uncertainty rather than risk.
3. There is something of a debate currently over using measures of "downside" risk instead of volatility. Given the symmetric nature of active returns, U.S. institutional investors' avoidance of strategies like portfolio insurance which skew portfolio returns, and the practical limitations of analyzing large portfolios (>100 names), downside risk is inappropriate or irrelevant for active management. For a discussion, *see* Ronald Kahn and Dan Stefek, "Heat, Light, and Downside Risk," 1997.

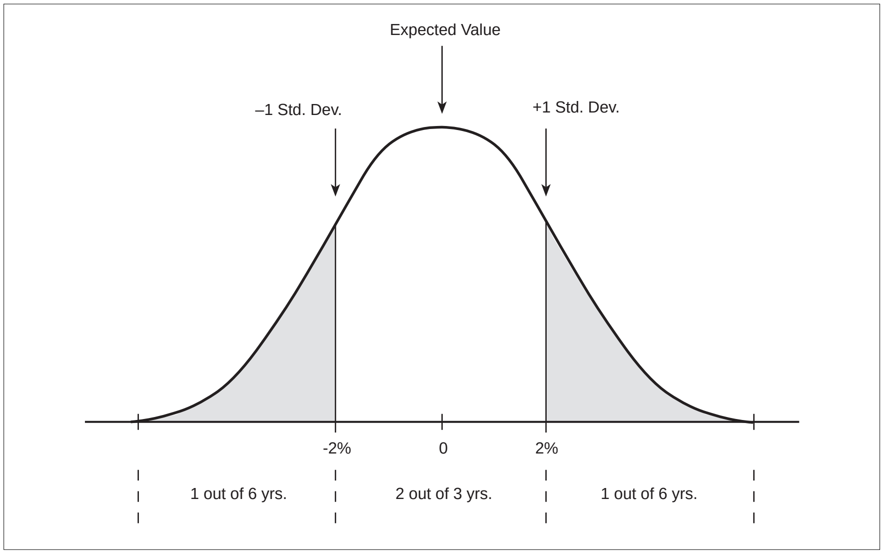
return is denoted by $\sigma_P \equiv Std[R_P]$. A portfolio's excess return r_P differs from the total return R_P by a constant R_F, so the risk of the excess return is equal to the risk of the total return. We will typically quote this risk, or standard deviation of return, on a percent per year basis. We will also occasionally refer to this quantity as volatility.[1]

The rough interpretation of standard deviation is that the return will be within one standard deviation of its expected value two-thirds of the time and within two standard deviations nineteen times out of twenty. Figure 2-1 graphically illustrates this fact.

Figure 2-1

Risk: The Dispersion of Returns

The standard deviation is a statistical measure of dispersion around an expected value—in this case, zero.



1. For a more detailed discussion of these concepts, please see Richard C. Grinold and Ronald N. Kahn, *Active Portfolio Management: Quantitative Theory and Applications*, Probus Publishing, Chicago, IL, 1995.

Risk measurement

A related risk measure is variance, the standard deviation squared. The formulae are:

$$Std[\tilde{r}] = \sqrt{Var[\tilde{r}]} \qquad \text{(EQ 2-1)}$$

$$Var[\tilde{r}] = \mathrm{E}\left[(\tilde{r} - \bar{r})^2\right] \qquad \text{(EQ 2-2)}$$

where:

$\tilde{r}$ = return,

$\bar{r}$ = expected or mean return,

$Std[x]$ = standard deviation of x,

$Var[x]$ = variance of x, and

$\mathrm{E}[x]$ = expected value of x.

The standard deviation is the more common risk indicator since it is measured in the same units as return. Of course, if the standard deviation is known, the variance can be easily computed and vice versa. Other measures, including value-at-risk and shortfall risk, can be easily computed from the standard deviation.

An example

The standard deviation has some interesting characteristics. In particular it does *not* have the portfolio property. The standard deviation of a stock portfolio is not the weighted average of the standard deviations of the component stocks.

For example, suppose the correlation between the returns of Stocks 1 and 2 is ρ_{12}. If we have a portfolio of 50% Stock 1 and 50% Stock 2, then:

$$\sigma_P = \sqrt{(0.5 \cdot \sigma_1)^2 + (0.5 \cdot \sigma_2)^2 + 2 \cdot (0.5 \cdot \sigma_1)(0.5 \cdot \sigma_2) \cdot \rho_{12}} \qquad \text{(EQ 2-3)}$$

and $\sigma_P \le 0.5 \cdot \sigma_1 + 0.5 \cdot \sigma_2$ (EQ 2-4) with such equality being maintained only if the two stocks are perfectly correlated ($\rho_{12} = 1$). For risk, the whole is less than the sum of its parts. This is the key to portfolio diversification.

Figure 2-2
Risk Reduction Benefits of Diversification:
A Two-Stock Example

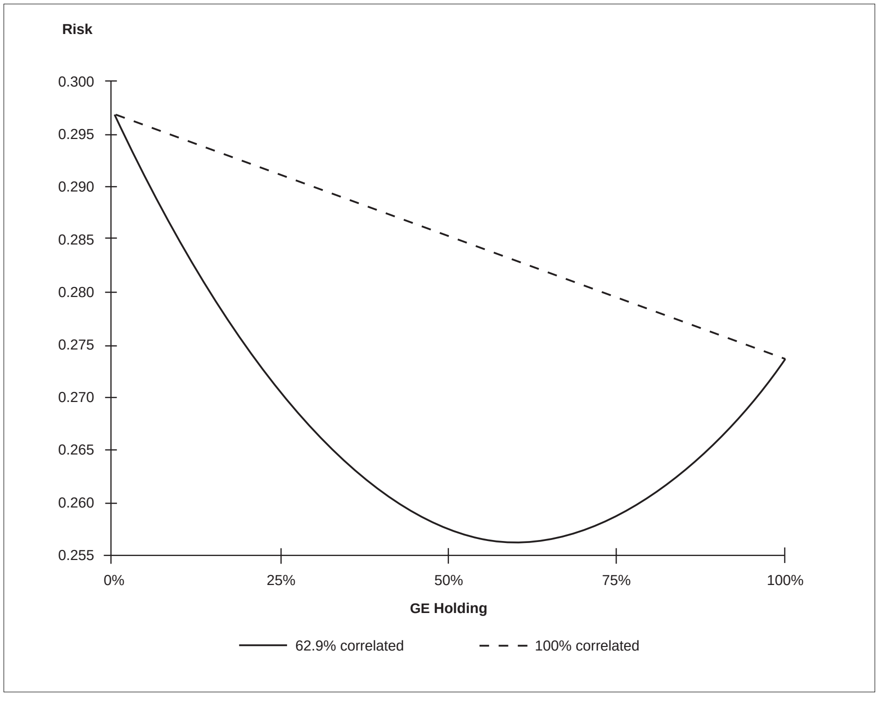


Figure 2-2 shows a simple example. The risk of a portfolio made up from IBM and General Electric is plotted against the fraction of GE stock in the portfolio. The curved line represents the risk of the portfolio; the straight line represents the risk that we would obtain if the returns on IBM and GE were perfectly correlated. The risk of GE is 27.4% per year; the risk of IBM is 29.7% per year; and the two returns are 62.9% correlated. The difference between the two lines is an indication of the benefit of diversification in reducing risk.

Risk reduction through diversification

We can see the power of diversification in another example. Given a portfolio of N stocks, each with risk σ and uncorrelated returns, the risk of an equal-weighted portfolio of these stocks will be:

$$\sigma_P = \frac{\sigma}{\sqrt{N}} \quad \text{(EQ 2-5)}$$

Note that the average risk is σ, while the portfolio risk is $\sigma/\sqrt{N}$.

For a more useful insight into diversification, assume now that the correlation between the returns of all pairs of stocks is equal to ρ. Then the risk of an equally weighted portfolio is:

$$\sigma_P = \sigma \cdot \sqrt{\frac{1+\rho \cdot (N-1)}{N}} \tag{EQ 2-6}$$

In the limit that the portfolio contains a very large number of correlated stocks, this becomes:

$$\sigma_P \Rightarrow \sigma \cdot \sqrt{\rho} \tag{EQ 2-7}$$

To get a feel for this, consider the example of an equal-weighted portfolio of the 20 Major Market Index constituent stocks. In December 1992, these stocks had an average risk of 27.8%, while the equal-weighted portfolio has a risk of 20.4%.[1] Equation 2-6 then implies an average correlation between these stocks of 0.52.

Risks don't add across stocks and risks don't add across time. However, variance will add across time if the returns in one interval of time are uncorrelated with the returns in other intervals of time. The assumption is that returns are uncorrelated from period to period. The correlation of returns across time (called *autocorrelation*) is close to zero for most asset classes. This means that variances will grow with the length of the forecast horizon and the risk will grow with the square root of the forecast horizon. Thus, a 5% annual active risk is equivalent to a 2.5% active risk over the first quarter or a 10% active risk over four years. Notice that the variance over the quarter, year, and four-year horizon (6.25, 25, and 100) remains proportional to the length of the horizon.

We use this relationship every time we "annualize" risk—i.e., standardize our risk numbers to an annual period. If we examine monthly returns to a stock and observe a monthly return standard deviation of $\sigma_{monthly}$, we convert this to annual risk according to:

$$\sigma_{annual} = \sqrt{12} \cdot \sigma_{monthly} \tag{EQ 2-8}$$

1. These are predicted volatilities from BARRA's U.S. Equity Model.

Drawbacks of simple risk calculations

The mathematical calculation of risk using standard deviation of returns is therefore straightforward and can be extended to any number of securities. However, this approach suffers from several drawbacks:

- Estimating a robust covariance matrix of returns requires data for as many periods as we have securities to analyze. For large markets, such as the U.S. stock market, these long returns histories may simply not be available.

- Estimation error may occur in any one period due to spurious asset correlations that are unlikely to repeat in a systematic fashion.

- A simple covariance matrix of returns offers little in the way of economic analysis. In other words, it is largely a "black box" approach with little intuitive basis or forecastability.

For all these reasons, financial economists have sought for many years to model investment risk in more nuanced ways. We will now turn to a brief history of these efforts.

Evolution of concepts

The development of equity risk concepts has evolved from the modest and unscientific guesswork of early investment theory to the quantitative analysis and technical sophistication of modern financial tools.

"Buy a stock. If it goes up, sell it. If it goes down, don't buy it."
Will Rogers, 1931

Before the 1950s, there was no concept of systematic, or market-related, return. Return was a rise in the value of a stock and risk was a drop in the value of a stock. The investor's primary investment tools were intuition and insightful financial analysis. Portfolio selection was simply an act of assembling a group of "good" stocks.

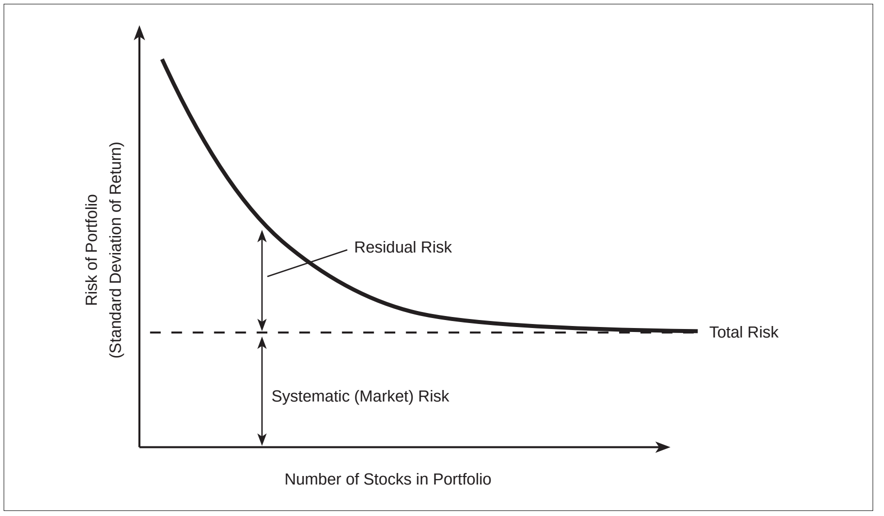


Figure 2-3

Diversification and Risk

As a portfolio manager increases the number of stocks in a portfolio, residual—or non-market-related—risk is diversified. Market risk is undiversifiable.

Financial theorists became more scientific and statistical in the early 1950s. Harry Markowitz was the first to quantify risk (as standard deviation) and diversification. He showed precisely how the risk of the portfolio was less than the risk of its components. In the late 1950s, Breiman and Kelly derived mathematically the peril of ignoring risk. They showed that a strategy that explicitly accounted for risk outperformed all other strategies in the long run.[1]

"Diversification is good."
Harry Markowitz, 1952

We now know how diversification affects risk exposures. It averages factor-related risk, such as industry exposures, and significantly reduces security-specific risk. However, diversification does not eliminate all risk because stocks tend to move up and down together with the market. Therefore, systematic, or market, risk cannot be eliminated by diversification.

Figure 2-3 shows the balance between residual risk and market risk changing as the number of different stocks in a portfolio rises. At a certain portfolio size, all residual risk is effectively removed, leaving only market risk.

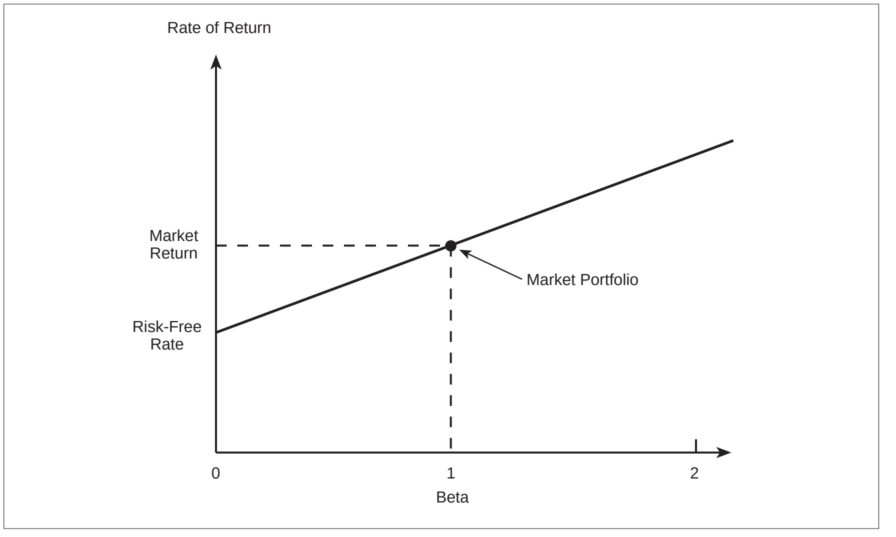
As investment managers became more knowledgeable, there was a push to identify the conceptual basis underlying the concepts of risk, diversification, and returns. The Capital Asset Pricing Model (CAPM) was one approach that described the equilibrium relationship between return and systematic risk. William Sharpe earned the Nobel Prize in Economics for his development of the CAPM.

1. See, for example, Leo Breiman, "Investment Policies for Expanding Businesses Optimal in a Long-Run Sense," *Naval Research Logistics Quarterly*, Vol. 7, No. 4, December 1960, pp. 647–651.

Figure 2-4

The Capital Asset Pricing Model

The Capital Asset Pricing Model asserts that the expected excess return on securities is proportional to their systematic risk coefficient, or beta. The market portfolio is characterized by a beta of unity.



The central premise of CAPM is that, on average, investors are not compensated for taking on residual risk. CAPM asserts that the expected residual return is zero while the expected systematic return is greater than zero and linear (*see* Figure 2-4).

The measure of portfolio exposure to systematic risk is called *beta* (β). Beta is the relative volatility or sensitivity of a security or portfolio to market moves. More simply, beta is the numerical value of an asset's systematic risk. Returns, and hence risk premiums, for any stock or portfolio will be related to beta, the exposure to undiversifiable systematic risk. Equation 2-9 states this linear relationship.

$$\mathrm{E}[\tilde{r}_i] - r_F = \beta_i \mathrm{E}[\tilde{r}_M - r_F] \qquad \text{(EQ 2-9)}$$

where:

"Only undiversifiable risk should earn a premium."

William F. Sharpe, 1964

Capital Asset Pricing Model

$\tilde{r}_i$ = return on asset i,

r_F = risk-free rate of return,

$\tilde{r}_M$ = return on market portfolio, and

$$\beta_i = \frac{Cov[\tilde{r}_i, \tilde{r}_M]}{Var[\tilde{r}_M]}$$

The CAPM is a model of return. Underlying it are equilibrium arguments and the view that the market is efficient because it is the portfolio that every investor on average owns. The CAPM does not require that residual returns be uncorrelated. But it did inspire Sharpe to suggest a one-factor risk model that does assume uncorrelated residual returns. This model has the advantage of simplicity. It is quite useful for back of the envelope calculations. But it ignores the risk that arises from common factor sources, such as industries, capitalization, and yield.

By the 1970s, the investment community recognized that assets with similar characteristics tend to behave in similar ways. This notion is captured in the Arbitrage Pricing Theory (APT). APT asserts that security and portfolio expected returns are linearly related to the expected returns of an unknown number of underlying systematic factors.

"The arbitrage model was proposed as an alternative to the mean variance capital asset pricing model."

Stephen A. Ross, 1976

Arbitrage Pricing Theory

The focus of APT is on forecasting returns. Instead of equilibrium arguments, Ross and others used arbitrage arguments to assert that expected specific returns are zero, but expected common factor returns (including the market and other factors) need not be zero. Just like the CAPM, APT inspired a class of risk models: the multiple-factor model (MFM). The basic premise of an MFM is that many influences act on the volatility of a stock, and these influences are often common across many stocks. A properly constructed MFM is able to produce risk analyses with more accuracy and intuition than a simple covariance matrix of security returns or the CAPM.

"Since the factors can represent the components of return as seen by the financial analyst, the multiple-factor model is a natural representation of the real environment."

Barr Rosenberg, 1974

Multiple Factor Models

Multifactor models of security market returns can be divided into three types: macroeconomic, fundamental, and statistical factor models. Macroeconomic factor models use observable economic time series, such as inflation and interest rates, as measures of the pervasive shocks to security returns. Fundamental factor models use the returns to portfolios associated with observed security attributes such as dividend yield, the book-to-market ratio, and industry membership. Statistical factor models derive their factors from factor analysis of the covariance matrix of security returns. BARRA research has confirmed that of these three, fundamental factor models outperform the other two types in terms of explanatory power.[1]

We now turn to a discussion of fundamental MFMs in more detail.

1. Gregory Connor, "The Three Types of Factor Models: A Comparison of Their Explanatory Power," *Financial Analysts Journal*, May/June 1995.

3. Modeling and Forecasting Risk

Through the years, theoretical approaches to investment analysis have become increasingly sophisticated. With more advanced concepts of risk and return, investment portfolio models have changed to reflect this growing complexity. The multiple-factor model (MFM) has evolved as a helpful tool for analyzing portfolio risk.

What are MFMs?

Multiple-factor models (MFMs) are formal statements about the relationships among security returns in a portfolio. The basic premise of MFMs is that similar stocks should display similar returns. This "similarity" is defined in terms of ratios, descriptors, and asset attributes which are based on market information, such as price and volume, or fundamental data derived from a company's balance sheet and income statement.

MFMs identify common factors and determine return sensitivity to these factors. The resulting risk model incorporates the weighted sum of common factor return and specific return. The risk profile will respond immediately to changes in fundamental information.

How do MFMs work?

We derive MFMs from securities patterns observed over time. The difficult steps are pinpointing these patterns and then identifying them with asset factors that investors can understand. Asset factors are characteristics related to securities price movements, such as industry membership, capitalization, and volatility.

Once model factors are chosen and assigned to individual assets in the proper proportions, cross-sectional regressions are performed to determine the returns to each factor over the relevant time period. This allows the model to be responsive to market changes in a timely fashion.

Risk calculation is the final step in constructing a sound and useful model. Variances, covariances, and correlations among factors are estimated and weighted. We then use these calculations to describe the risk exposure of a portfolio.

Investors rely on risk exposure calculations to determine stock selection, portfolio construction, and other investment strategies. They base their decisions on information gleaned from MFM analysis as well as their risk preferences and other information they possess.

Advantages of MFMs

There are several advantages to using MFMs for security and portfolio analysis.

- MFMs offer a more thorough breakdown of risk and, therefore, a more complete analysis of risk exposure than other methods such as simple CAPM approaches.
- Because economic logic is used in their development, MFMs are not limited by purely historical analysis.
- MFMs are robust investment tools that can withstand outliers.
- As the economy and individual firms change, MFMs adapt to reflect changing asset characteristics.
- MFMs isolate the impact of individual factors, providing segmented analysis for better informed investment decisions.
- From an applications viewpoint, MFMs are realistic, tractable, and understandable to investors.
- Lastly, MFMs are flexible models allowing for a wide range of investor preferences and judgment.

Of course, MFMs have their limitations. They predict much, but not all, of portfolio risk. In addition, a model does not offer stock recommendations; investors must make their own strategy choices.

A simple MFM

To illustrate the power of MFMs, let's begin with a simple example.

Accurate characterization of portfolio risk requires an accurate estimate of the covariance matrix of security returns. A relatively simple way to estimate this covariance matrix is to use the history of security returns to compute each variance and covariance. This approach, however, suffers from two major drawbacks:

- Estimating a covariance matrix for, say, 3,000 stocks requires data for at least 3,000 periods. With monthly or weekly estimation horizons, such a long history may simply not exist.

- It is subject to estimation error: in any period, two stocks such as Weyerhaeuser and Ford may show very high correlation—higher than, say, GM and Ford. Our intuition suggests that the correlation between GM and Ford should be higher because they are in the same line of business. The simple method of estimating the covariance matrix does not capture our intuition.

This intuition, however, points to an alternative method for estimating the covariance matrix. Our feeling that GM and Ford should be more highly correlated than Weyerhaeuser and Ford comes from Ford and GM being in the same industry. Taking this further, we can argue that firms with similar characteristics, such as their line of business, should have returns that behave similarly. For example, Weyerhaeuser, Ford, and GM will all have a common component in their returns because they would all be affected by news that affects the stock market as a whole. The effects of such news may be captured by a stock market component in each stock's return. This common component may be the (weighted) average return to all U.S. stocks. The degree to which each of the three stocks responds to this stock market component depends on the sensitivity of each stock to the stock market component.

Additionally, we would expect GM and Ford to respond to news affecting the automobile industry, whereas we would expect Weyerhaeuser to respond to news affecting the forest and paper products industry. The effects of such news may be captured by the average returns of stocks in the auto industry and the forest and paper products industry. There are, however, events that affect one stock without affecting the others. For example, a defect in the brake system of GM cars, that forces a recall and replacement of the system, will likely have a negative impact on GM's stock price. This event, however, will most likely leave Weyerhaeuser and Ford stock prices unaltered.

These arguments lead us to the following representation for returns:

$$\tilde{r}_{GM} = \mathrm{E}[\tilde{r}_{GM}] + \beta_{GM} \cdot [\tilde{r}_M - \mathrm{E}[\tilde{r}_M]] \\ + 1 \cdot [\tilde{r}_{AUTO} - \mathrm{E}[\tilde{r}_{AUTO}]] + 0 \cdot [\tilde{r}_{FP} - \mathrm{E}[\tilde{r}_{FP}]] + u_{GM} \tag{EQ 3-1}$$

where:

$\tilde{r}_{GM}$ = GM's realized return,

$\tilde{r}_M$ = the realized average stock market return,

$\tilde{r}_{AUTO}$ = realized average return to automobile stocks,

$\tilde{r}_{FP}$ = the realized average return to forest and paper products stocks,

E[.] = expectations,

β_{GM} = GM's sensitivity to stock market returns, and

u_{GM} = the effect of GM specific news on GM returns.

This equation simply states that GM's realized return consists of an expected component and an unexpected component. The unexpected component depends on any unexpected events that affect stock returns in general $[\tilde{r}_M - \mathrm{E}[\tilde{r}_M]]$, any unexpected events that affect the auto industry $[\tilde{r}_{AUTO} - \mathrm{E}[\tilde{r}_{AUTO}]]$, and any unexpected events that affect GM alone (u_{GM}). Similar equations may be written for Ford and Weyerhaeuser.

The sources of variation in GM's stock returns, thus, are variations in stock returns in general, variations in auto industry returns, and any variations that are specific to GM. Moreover, GM and Ford returns are likely to move together because both are exposed to stock market risk and auto industry risk. Weyerhaeuser and GM, and Weyerhaeuser and Ford, on the other hand, are likely to move together to a lesser degree because the only common component in their returns is the market return. Some additional correlation would arise, however, because auto and forest and paper products industry returns may exhibit some correlation.

By beginning with our intuition about the sources of co-movement in security returns, we have made substantial progress in estimating the covariance matrix of security returns. What we need now is the covariance matrix of common sources in security returns, the variances of security specific returns, and estimates of the sensitivity of security returns to the common sources of variation in their returns. Because the common sources of risk are likely to be much fewer than the number of securities, we need to estimate a much smaller covariance matrix and hence a smaller history of returns is required. Moreover, because similar stocks are going to have larger sensitivities to similar common sources of risk, similar stocks will be more highly correlated than dissimilar stocks: our estimated correlation for GM and Ford will be larger than that for Ford and Weyerhaeuser.

The decomposition of security returns into common and specific sources of return is, in fact, a multiple-factor model of security returns. We now turn to a generalized discussion of this process for many factors.

Model mathematics

MFMs build on single-factor models by including and describing the interrelationships among factors. For single-factor models, the equation that describes the excess rate of return is:

$$\tilde{r}_j = X_j \tilde{f} + \tilde{u}_j \qquad \text{(EQ 3-2)}$$

where:

$\tilde{r}_j$ = total excess return over the risk-free rate,

X_j = sensitivity of security j to the factor,

$\tilde{f}$ = rate of return on the factor, and

$\tilde{u}_j$ = nonfactor (specific) return on security j.

We can expand this model to include K factors. The total excess return equation for a multiple-factor model becomes:

$$\tilde{r}_j = \sum_{k=1}^{K} X_{jk} \tilde{f}_k + \tilde{u}_j \qquad \text{(EQ 3-3)}$$

where:

X_{jk} = risk exposure of security j to factor k, and

$\tilde{f}_k$ = rate of return on factor k.

Note that when K=1, the MFM equation reduces to the earlier single-factor version. For example, the CAPM is a single-factor model in which the "market" return is the only relevant factor.

When a portfolio consists of only one security, Equation 3-3 describes its excess return. But most portfolios comprise many securities, each representing a proportion, or weight, of the total portfolio. When weights h_{P1}, h_{P2},...,h_{PN} reflect the proportions of N securities in portfolio P, we express the excess return in the following equation:

$$\tilde{r}_P = \sum_{k=1}^{K} X_{Pk} \tilde{f}_k + \sum_{j=1}^{N} h_{Pj} \tilde{u}_j \qquad \text{(EQ 3-4)}$$

where:

$$X_{Pk} = \sum_{j=1}^{N} h_{Pj} X_{jk}$$

This equation includes the risk from all sources and lays the groundwork for further MFM analysis.

Risk prediction with MFMs

Investors look at the variance of their total portfolios to provide a comprehensive assessment of risk. To calculate the variance of a portfolio, you need to calculate the covariances of all the constituent components.

Without the framework of a multiple-factor model, estimating the covariance of each asset with every other asset is computationally burdensome and subject to significant estimation errors. For example, using an estimation universe of 1,400 assets, there are 980,700 covariances and variances to calculate (*see* Figure 3-1).

$$V(i, j) = \text{Covariance}\,[r(\tilde{i}),\ r(\tilde{j})]$$

where $V(i, j)$ = asset covariance matrix, and
i, j = individual stocks.

$$V = \begin{bmatrix} V(1,1) & V(1,2) & \cdots & V(1,N) \\ V(2,1) & V(2,2) & \cdots & V(2,N) \\ V(3,1) & V(3,2) & \cdots & V(3,N) \\ \vdots & \vdots & & \vdots \\ V(N,1) & V(N,2) & \cdots & V(N,N) \end{bmatrix}$$

Figure 3-1
Asset Covariance Matrix
For N=1,400 assets, there are 980,700 covariances and variances to estimate.

An MFM simplifies these calculations dramatically. This results from replacing individual company profiles with categories defined by common characteristics (factors). Since the specific risk is assumed to be uncorrelated among the assets, only the factor variances and covariances need to be calculated during model estimation (*see* Figure 3-2).

$$\tilde{r} = X\tilde{f} + \tilde{u}$$

where $\tilde{r}$ = vector of excess returns,
X = exposure matrix,
$\tilde{f}$ = vector of factor returns, and
$\tilde{u}$ = vector of specific returns.

$$\begin{bmatrix} \tilde{r}(1) \\ \tilde{r}(2) \\ \vdots \\ \tilde{r}(N) \end{bmatrix} = \begin{bmatrix} X(1,1) & X(1,2) & \cdots & X(1,K) \\ X(2,1) & X(2,2) & \cdots & X(2,K) \\ \vdots & \vdots & & \vdots \\ X(N,1) & X(N,2) & \cdots & X(N,K) \end{bmatrix} \begin{bmatrix} \tilde{f}(1) \\ \tilde{f}(2) \\ \vdots \\ \tilde{f}(K) \end{bmatrix} + \begin{bmatrix} \tilde{u}(1) \\ \tilde{u}(2) \\ \vdots \\ \tilde{u}(N) \end{bmatrix}$$

Figure 3-2
Factor Return Calculation
Using an MFM greatly simplifies the estimation process. Figure 3-2 depicts the multiple-factor model in matrix terms.

By using a multiple-factor model, we significantly reduce the number of calculations. For example, in the U.S. Equity Model (US-E3), 65 factors capture the risk characteristics of equities. This reduces the number of covariance and variance calculations to 2,145 (*see* Figure 3-3). Moreover, since there are fewer parameters to determine, they can be estimated with greater precision.

Figure 3-3

Factor Covariance Matrix

For K=65 factors, there are 2,145 covariances and variances to estimate. Quadrant I includes the covariances of risk indices with each other; quadrants II and III are mirror images of each other, showing the covariances of risk indices with industries; and quadrant IV includes covariances of industries with each other.

$$F(k, m) = \text{Covariance}\left[\tilde{f}(k), \tilde{f}(m)\right]$$

where $F(k, m)$ = factor covariance matrix, and
k, m = common factors.

$$F = \left[\begin{array}{ccc:ccc}
F(1,1) & \cdots & F(1,13) & F(1,14) & \cdots & F(1,65) \\
\vdots & \text{I} & \vdots & \vdots & \text{II} & \vdots \\
F(13,1) & \cdots & F(13,13) & F(13,14) & \cdots & F(13,65) \\
\hdashline
F(14,1) & \cdots & F(14,13) & F(14,14) & \cdots & F(14,65) \\
\vdots & \text{III} & \vdots & \vdots & \text{IV} & \vdots \\
F(65,1) & \cdots & F(65,13) & F(65,14) & \cdots & F(65,65)
\end{array}\right]$$

We can easily derive the matrix algebra calculations that support and link the above diagrams by using an MFM. From Figure 3-2, we start with the MFM equation:

$$\tilde{r}_i = X\tilde{f} + \tilde{u} \qquad \text{(EQ 3-5)}$$

where:

$\tilde{r}_i$ =excess return on asset i,

X =exposure coefficient on the factor,

$\tilde{f}$ =factor return, and

$\tilde{u}$ =specific return.

Substituting this relation in the basic equation, we find that:

$$Risk = Var\left(\tilde{r}_j\right) \qquad \text{(EQ 3-6)}$$

$$= Var\left(X\tilde{f} + \tilde{u}\right) \qquad \text{(EQ 3-7)}$$

Using the matrix algebra formula for variance, the risk equation becomes:

$$Risk = XFX^T + \Delta \qquad \text{(EQ 3-8)}$$

where:

X = exposure matrix of companies upon factors,

F = covariance matrix of factors,

X^T = transpose of X matrix, and

Δ = diagonal matrix of specific risk variances.

This is the basic equation that defines the matrix calculations used in risk analysis in the BARRA equity models.

4. Modern Portfolio Management and Risk

In the previous chapters we observed that risk modeling is essential to successful portfolio management and that the standard deviation of returns is the best numerical risk measure. We also traced the evolution of risk concepts from portfolio standard deviation of security returns, through the CAPM and APT, to the current application of multiple-factor models (MFMs) in the portfolio management problem. In this chapter we will briefly explore the components of portfolio management and how BARRA MFMs can assist the manager at various points.

Portfolio management—two types

An equity portfolio manager must choose between two management methods: passive or active. BARRA MFMs are used to facilitate both methods.

Passive management

Passive management is an outgrowth of CAPM logic. In its broadest sense, passive management refers to any management strategy that does not rely on the possession of superior information. More specifically, disclosure of a passive investment strategy offers no competitive information that would undermine the strategy's validity.

One type of passive management is indexing, tracking the performance of a particular index. An example is the "buy-and-hold" philosophy which exposes the portfolio only to systematic risk. The second form of passive management is constructing a portfolio to match prespecified attributes or constraints. The portfolio may be yield-biased with a selected beta or match an index within certain parameters. This is often called *enhanced indexing*.

Benchmark

A *benchmark* is the standard of comparison for investment performance and risk analysis. It is widely used to evaluate and track performance of investment managers. The benchmark is also known as the *normal portfolio*—that is, the asset basket a manager would hold in the absence of any judgmental information. It reflects the manager's particular style and biases.

Tracking Error

Tracking error is a measure of risk exposure. It is the annualized standard deviation of the difference between portfolio return and benchmark return.

Because it provides a relative measure of risk exposure, tracking error is a useful evaluation tool, particularly for passive portfolios. Moreover, it offers relevant performance comparisons because the benchmark is selected based on portfolio characteristics and investor objectives.

Passive management procedures are distinguished by the following attributes:

- They exclude any transactions in response to judgments about security valuations and the market as a whole.
- They contain relatively minimal residual risk with respect to the benchmark or index.
- They often involve industry or sector weighting.

BARRA MFMs facilitate passive management by providing robust portfolio risk estimates versus passive benchmarks. The indexed portfolio can be readily compared with the benchmark to determine the magnitude of active risk (or tracking error) and its composition. Corrective action can be taken based on individual holding analysis which reveals those securities contributing the most active risk. Optimizing software can also be used to automate the rebalancing process.

Active management

Alpha

Alpha (α) generally refers to the expected exceptional return of a portfolio, factor, or individual asset. The use of alphas is a distinction of active management. They indicate that a manager believes a portion of expected return is attributable to particular factors.

Historical alpha is the difference between actual performance and the performance of a diversified market portfolio with identical systematic risk over the same period. *Judgmental*, or *predicted*, *alpha* is the expected value of subsequent extraordinary return based on a return forecast.

Active management refers to investment strategies designed to increase return by using superior information. In other words, the active manager seeks to profit from information that would lose its value if all market participants interpreted it in the same way. If, for example, an investment manager observed that securities with certain characteristics performed better (or worse) than expected, the manager could hold a larger (or smaller) proportion of that security to increase the subsequent value of the portfolio.

By following active management strategies, investors can add value to their portfolio if they predict returns better than the consensus expectations of the market. Information is obtained through ongoing research to forecast such things as yield curve changes, factor and industry returns, and transitory mispricing. At any given time, portfolio construction should reflect the tradeoff between risk and return—that is, any contribution to risk should be offset by the contribution to reward.

There are several basic types of active investment strategies. They include *market timing*, *sectoral emphasis*, and *stock selection*.

Market timing is the process of altering market risk exposure based on short-term forecasts in order to earn superior returns. The manager seeks to sell before the market goes down and buy before the market goes up. However, this strategy increases the variability in the portfolio beta, inducing increased systematic risk through time. BARRA MFMs assist market timing by giving the investor a robust beta estimate for any portfolio and indicating the most efficient way to take on or reduce market risk, including the use of futures.

The second type of active management is *sectoral emphasis*. Sectoral emphasis can be thought of as a combination of the other active strategies. It is both factor timing and a broad version of stock selection. The manager attempts to increase residual return through manipulating common factor exposures. For example, the manager can bet on an industry of high-yield stocks. Because several sectors can be emphasized at any given time, diversification is possible. BARRA MFMs possess detailed industry and risk index exposure information that can be utilized for any combination of sectoral tilts.

Lastly, *stock selection* is a portfolio allocation strategy based on picking mispriced stocks. It uses security alphas to identify over- and undervalued stocks. The manager can then adjust the asset proportions in the portfolio to maximize specific return. These active holdings, in both positive and negative directions, increase residual risk and portfolio alpha. The primary objective of this strategy is to manage asset holdings so that any change in incremental risk is compensated by a comparable change in return. BARRA MFMs facilitate stock selection by extending the risk model down to the individual equity level.

Figure 4-1 illustrates the typical prevalence of these various types of risk in a single stock, a small portfolio, and a multiple-portfolio situation. In each case, the manager's goal is to earn a superior return with minimum risk. The use of a multiple-factor model permits the manager to pursue these active management strategies with maximum precision.

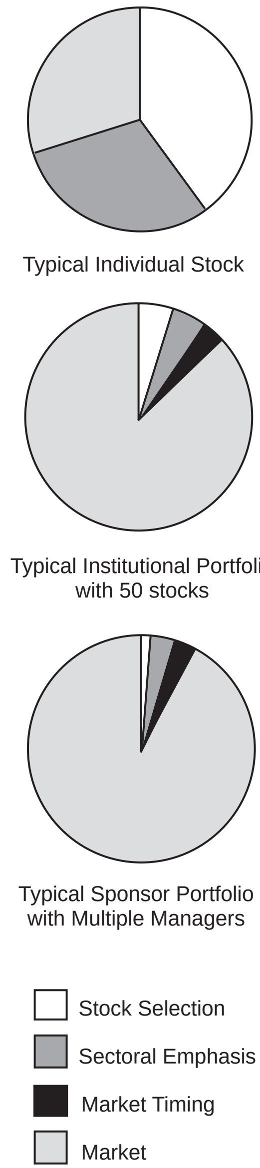


Figure 4-1

Diversification and Risk

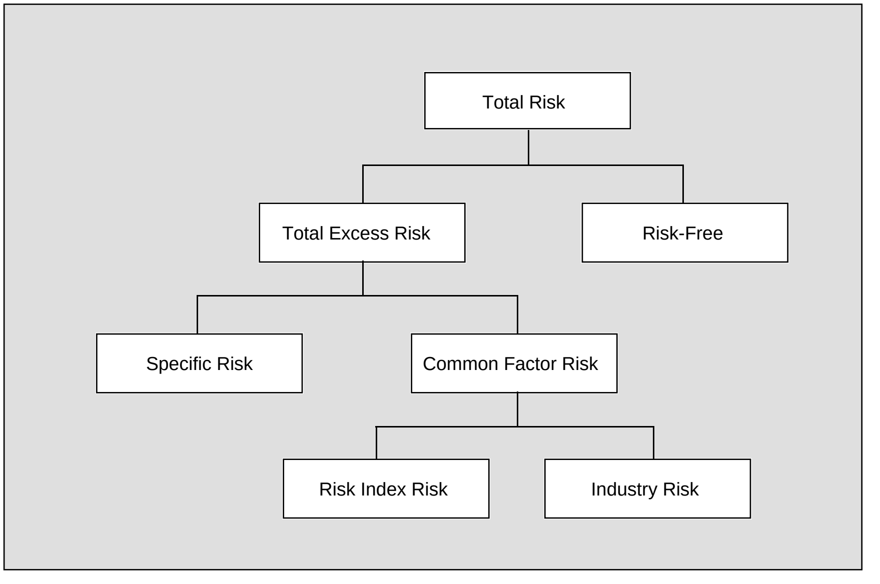
Market risk grows as a proportion of total risk as portfolio size increases.

Decomposing risk

BARRA's equity models isolate components of risk based on correlations across securities sharing similar characteristics. There are several ways to break down a portfolio's risk. BARRA uses any of four decompositions of risk, each reflecting a different perspective on portfolio management. These four decompositions are used in different ways by active and passive managers to provide insight and enhance performance.

Total Risk Decomposition

Figure 4-2
Total Risk Decomposition



The simplest risk decomposition, **Total Risk** (*see* Figure 4-2), is a basic breakdown into specific and common factor risk. There is no concept of a market, or systematic, portfolio. The risk is attributed purely to common factor and security-specific influences.

❖ *Common factor risk* is portfolio risk that arises from assets' exposures to common factors, such as capitalization and industries.

❖ *Specific risk* is unique to a particular company and thus is uncorrelated with the specific risk of other assets. For a portfolio, specific risk is the weighted sum of all the holdings' specific risk values.

This risk decomposition is useful for managers who wish to minimize total risk, or for managers such as hedge funds which may be pursuing market-neutral or other long/short strategies.

Systematic-Residual Risk Decomposition

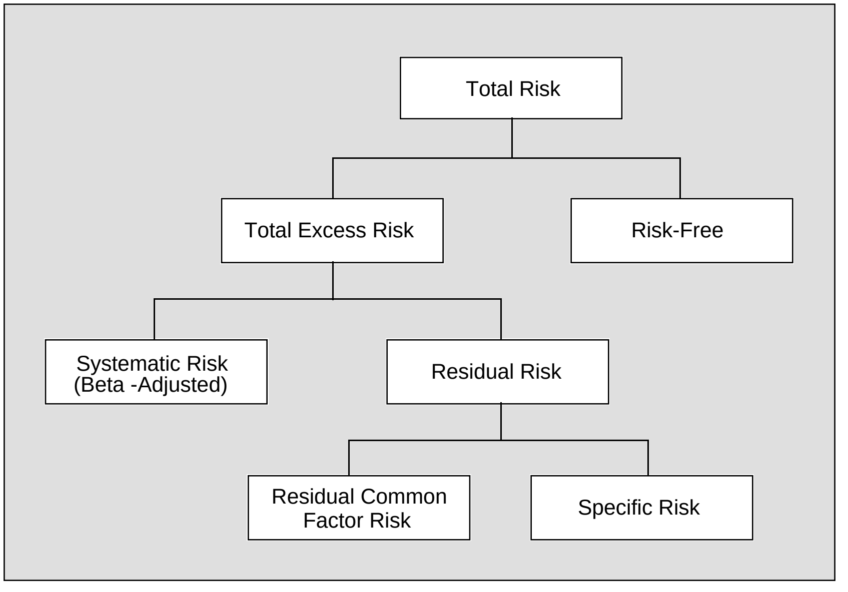


Figure 4-3
Systematic-Residual Risk Decomposition

This decomposition introduces the market into risk analysis (*see* Figure 4-3). This perspective partitions risk into the familiar categories of systematic (market) and residual risk.

❖ *Systematic risk* is the component of risk associated with the market portfolio. It is linked to the portfolio beta, which is the weighted average of the portfolio's asset betas.

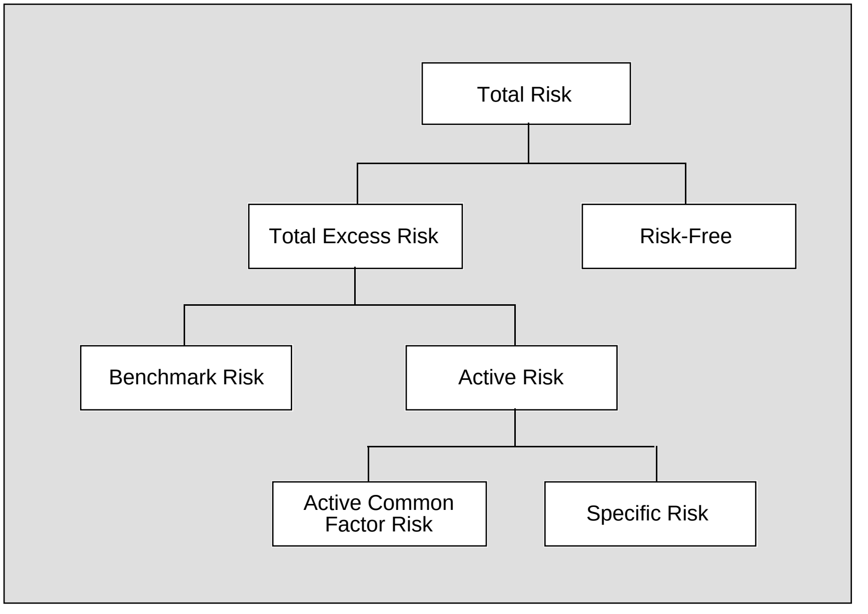
❖ *Residual risk* is the component of risk uncorrelated with the market portfolio. It is further divided into specific and common factor sources. Residual risk can be diversified to a negligible level.

This type of risk decomposition is the Capital Asset Pricing Model (CAPM) approach. With this approach, you can compare your managed portfolio against a broad-based market portfolio. A benchmark portfolio never comes into play. Risk is partitioned into *residual* and *systematic* components, and residual risk is further divided into specific and common factor sources.

This risk decomposition would be most useful to market timers or other managers who "tilt" away from the market portfolio on an opportunistic basis.

Active Risk Decomposition

Figure 4-4
Active Risk Decomposition



In this type of decomposition, the concepts of benchmark and active risk are superimposed on the common factor and specific risks itemized in Total Risk Decomposition (*see* Figure 4-4).

- *Benchmark risk* is the risk associated with the benchmark portfolio.

- *Active risk* is the risk that arises from the manager's effort to outperform the benchmark. It is further divided into active specific and active common factor sources. Active risk is also known as tracking error.

This perspective is most commonly used in analyzing index funds as well as traditionally managed active portfolios.

In this type of decomposition, there is no concept of a market portfolio. The analysis concentrates solely on the managed portfolio against the benchmark that you select. However, for most managers, market risk is a component of active risk; these managers might prefer Active Systematic-Active Residual Risk decomposition, the fourth and last type.

Active Systematic-Active Residual Risk Decomposition

Finally, Active Systematic-Active Residual Risk decomposition (*see* Figure 4-5), the most complete perspective, expands the previous decomposition by including systematic sources of risk. Both this method and Active Risk are helpful in performance evaluation and analysis because they consider the *benchmark portfolio*, which reveals management style.

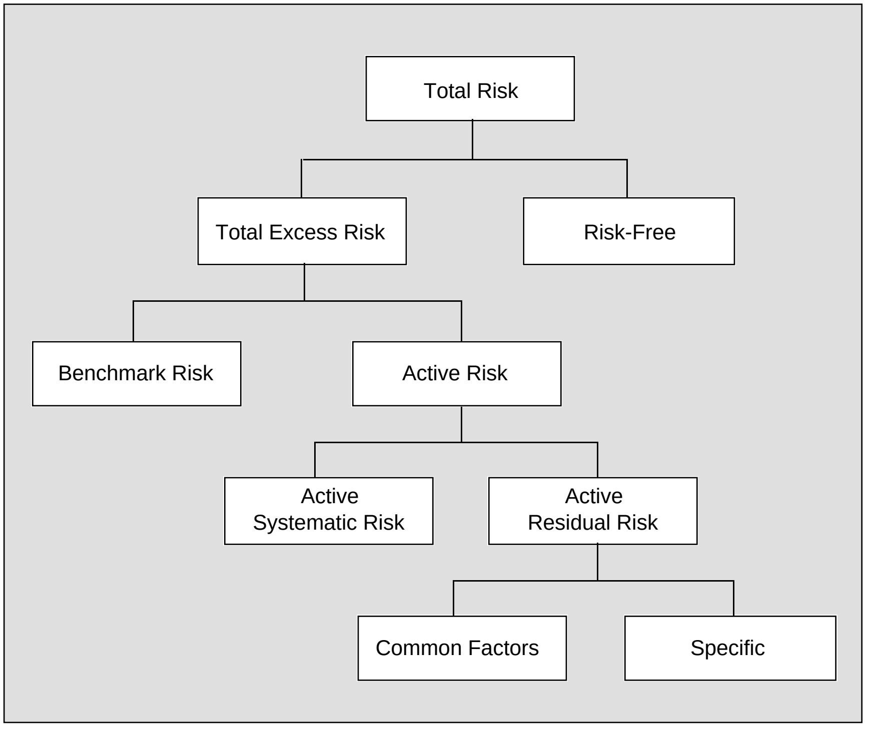


Figure 4-5
Active Systematic-Active Residual Risk Decomposition

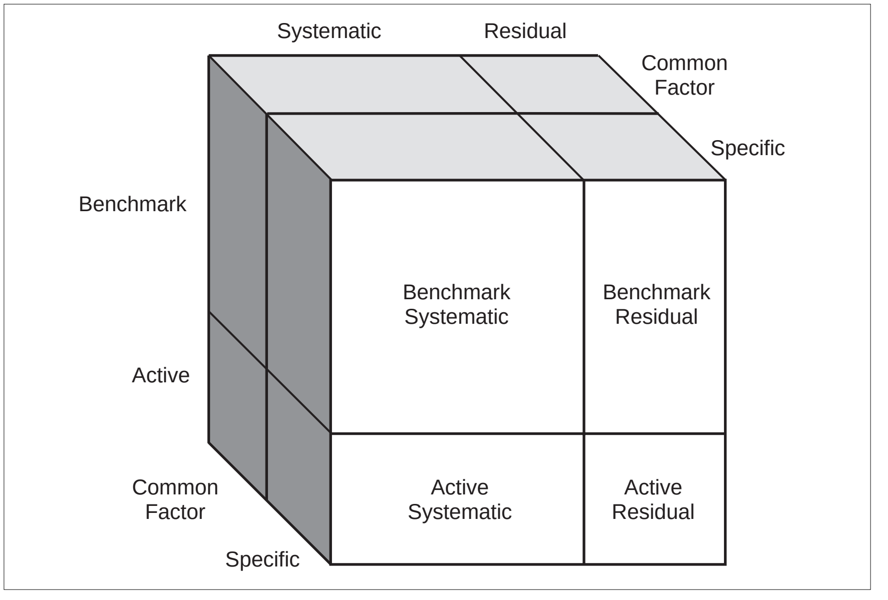
This risk decomposition is useful for managers who overlay a market-timing strategy on their stock selection process and don't want market risk considerations to affect their analysis.

Summary of risk decomposition

These methods of risk decomposition represent compatible perspectives. Figure 4-6 shows how the four methods are different ways to slice the same pie—specific/common factor, systematic/residual, and benchmark/active.

Figure 4-6

Risk Decomposition Overview



Performance attribution

Performance attribution completes the portfolio management process by applying the MFM to past portfolio activity. Performance attribution is the process of matching return with its sources. Return is attributed to common factors, market timing, and asset selection. Using benchmark comparisons to judge performance, the value of each investment decision can be determined.

BARRA's performance attribution programs decompose return into its major components using any of the four methods of risk decomposition listed above. Performance can then be evaluated with respect to customized or industry-standard benchmark portfolios designed to compare managers with their own standards.

Summary

In this chapter we have outlined the general management approaches available to equity managers and discussed how BARRA MFMs can be utilized at various stages of the management process. In the next chapter we will describe in detail the process of building an MFM.

5. BARRA Multiple-Factor Modeling

A BARRA equity risk model is the product of a thorough and exacting model estimation process. This section provides a brief overview of model estimation.

Figure 5-1
Model Estimation Process

1. *Data acquisition and cleaning*
2. *Descriptor selection and testing*
3. *Descriptor standardization*
4. *Risk index formulation*
5. *Industry allocation*
6. *Factor return estimation*
7. *Covariance matrix calculation:*
    a. Exponential weighting
    b. Market volatility: GARCH
8. *Specific risk forecasting*
9. *Model updating*

Overview

The creation of a comprehensive equity risk model is an extensive, detailed process of determining the factors that describe asset returns. Model estimation involves a series of intricate steps that is summarized in Figure 5-1.

The first step in model estimation is acquiring and cleaning data. Both market information (such as price, dividend yield, or capitalization) and fundamental data (such as earnings, sales, or assets) are used. Special attention is paid to capital restructurings and other atypical events to provide for consistent cross-period comparisons.

Descriptor selection follows. This involves choosing and standardizing variables which best capture the risk characteristics of the assets. To determine which descriptors partition risk in the most effective and efficient way, the descriptors are tested for statistical significance. Useful descriptors often significantly explain cross-sectional returns.

Risk index formulation and assignment to securities is the fourth step. This process involves collecting descriptors into their most meaningful combinations. Though judgment plays a major role, a variety of techniques are used to evaluate different possibilities. For example, cluster analysis is one statistical tool used to assign descriptors to risk indices.

Along with risk index exposures, industry allocations are determined for each security. In most BARRA models a single industry exposure is assigned, but multiple exposures for conglomerates are calculated in a few models, including the U.S. and Japan models.

Next, through cross-sectional regressions, factor returns are calculated and used to estimate covariances between factors, generating the covariance matrix used to forecast risk. Exponential weighting of data observations may be used if testing indicates it improves accuracy. This matrix may be further modified to utilize GARCH techniques.

Specific returns are separated out at this stage and *specific risk* is forecast. This is the portion of total risk that is related solely to a particular stock and cannot be accounted for by common factors. The greater an asset's specific risk, the larger the proportion of return attributable to idiosyncratic, rather than common, factors.

Lastly, the model undergoes final testing and updating. Risk forecasts are tested against alternative models. Tests compare *ex ante* forecasts with *ex post* realizations of beta, specific risk, and active risk. New information from company fundamental reports and market data is incorporated, and the covariance matrix is recalculated.

Figure 5-2 summarizes these steps.

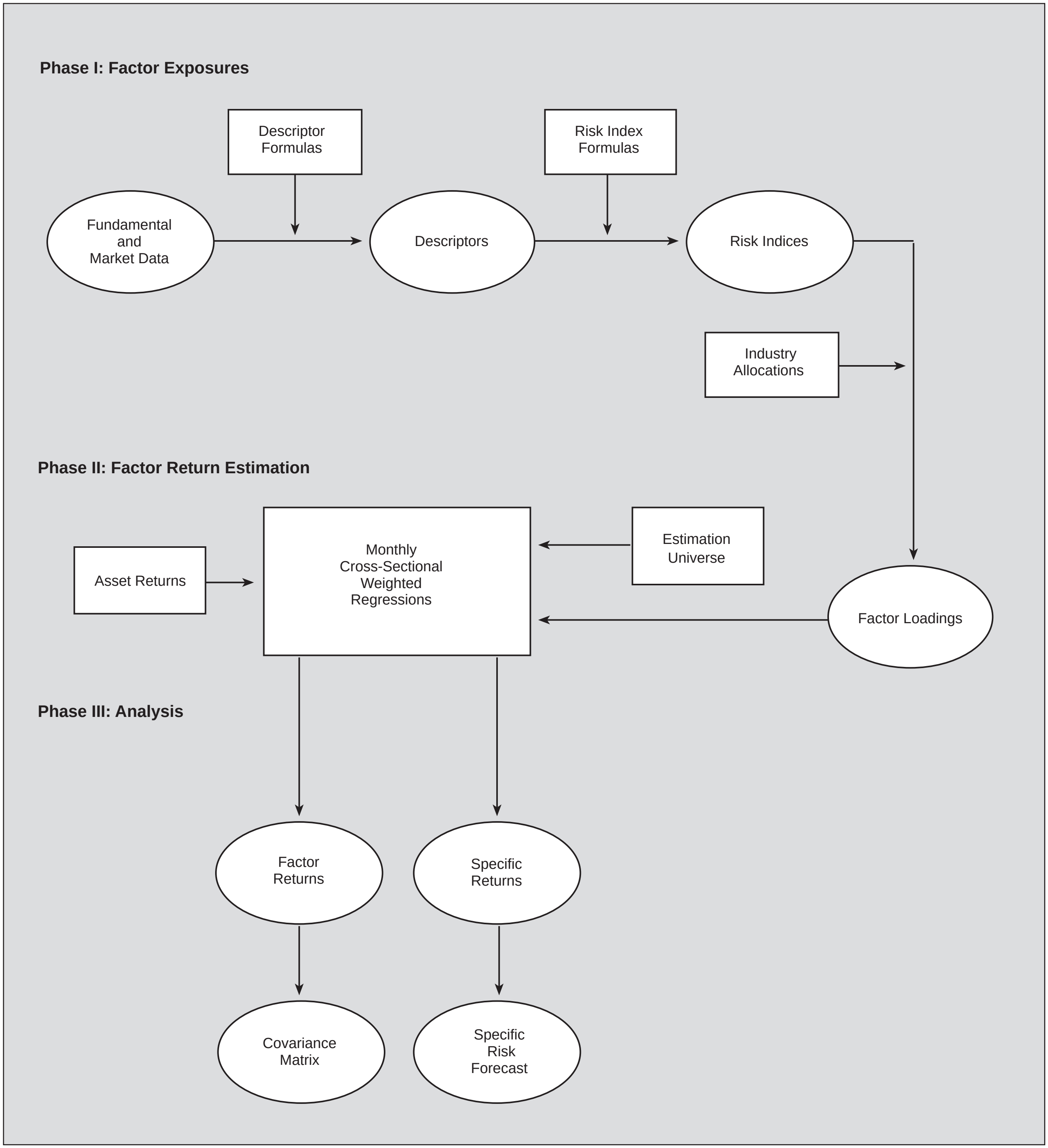


Figure 5-2

Data Flow for Model Estimation

This figure depicts the model estimation process. The oval shapes mark the data flow throughout model estimation, while the rectangular shapes show manipulations of and additions to the data.

Descriptor selection and testing

Descriptor candidates are drawn from several sources. Market information, such as trading volume, stock prices, and dividends, is available daily. Fundamental company data—such as earnings, assets, and industry information—are derived from quarterly and annual financial statements. For some descriptors, market and fundamental information is combined. An example is the earnings to price ratio, which measures the relationship between the market's valuation of a firm and that firm's earnings.

Descriptor selection is a largely qualitative process that is subjected to rigorous quantitative testing. First, preliminary descriptors are identified. Good descriptor candidates are individually meaningful; that is, they are based on generally accepted and well-understood asset attributes. Furthermore, they should divide the market into well-defined categories, providing full characterization of the portfolio's important risk features. BARRA has more than two decades of experience identifying important descriptors in equity markets worldwide. This experience informs every new model we build.

Selected descriptors must have a sound theoretical justification for inclusion in the model. They must be useful in predicting risk and based on timely, accurate, and available data. In other words, each descriptor must add value to the model. If the testing process shows that they do not add predictive power, they are rejected.

Normalization

Normalization is the process of setting random variables to a uniform scale. Also called *standardization*, it is the process by which a constant (usually the mean) is subtracted from each number to shift all numbers uniformly. Then each number is divided by another constant (usually the standard deviation) to shift the variance.

Descriptor standardization

The risk indices are composed of descriptors designed to capture all the relevant risk characteristics of a company. The descriptors are first normalized; that is, they are standardized with respect to the estimation universe. This is done to allow combination of descriptors into meaningful risk factors, known as *risk indices*. The normalization process is summarized by the following relation:

$$[\textit{normalized descriptor}] = \frac{[\textit{raw descriptor}] - [\textit{mean}]}{[\textit{standard deviation}]}$$

Risk index formulation

We regress asset returns against industries and descriptors, one descriptor at a time, after the normalization step. Each descriptor is tested for statistical significance. Based on the results of these calculations and tests, descriptors for the model are selected and assigned to risk indices.

Risk index formulation is an iterative process. After the most significant descriptors are added to the model, remaining descriptors are subjected to stricter testing. At each stage of model estimation, a new descriptor is added only if it adds explanatory power to the model beyond that of industry factors and already-assigned descriptors.

Industry allocation

For most BARRA equity models, companies are allocated to single industries. For the U.S. and Japan, however, sufficient data exist to allocate to multiple industries.

For the U.S. and Japan, industry exposures are allocated using industry segment data (i.e., operating earnings, assets, and sales). The model incorporates the relative importance of each variable in different industries. For example, the most important variable for oil companies would be assets; for retail store chains, sales; and for stable manufacturing companies, earnings. For any given multi-industry allocation, the weights will add up to 100%.

Multiple industry allocation provides more accurate risk prediction and better describes market conditions and company activity. BARRA's multiple-industry model captures changes in a company's risk profile as soon as new business activity is reported to shareholders. Alternative approaches can require 60 months or more of data to recognize changes that result from market prices.

Factor return estimation

The previous steps have defined the exposures of each asset to the factors at the beginning of every period in the estimation window. The factor excess returns over the period are then obtained via a cross-sectional regression of asset excess returns on their associated factor exposures:

$$\tilde{r}_t = X_t \tilde{f}_t + u_t \qquad \text{(EQ 5-1)}$$

where:

$\tilde{r}_t$ = excess returns to each asset

X_t = exposure matrix of assets to factors

$\tilde{f}_t$ = factor returns to be estimated

u_t = specific returns

The resulting factor returns are robust estimates which can be used to calculate a factor covariance matrix to be used in the remaining model estimation steps.

Covariance matrix calculation

The simplest way to estimate the factor covariance matrix is to compute the sample covariances among the entire set of estimated factor returns. Implicit in this process is the assumption that we are modeling a stable process and, therefore, each point in time contains equally relevant information.

There is evidence, however, that correlations among factor returns change. Moreover, a stable process implies a stable variance for a well-diversified portfolio with relatively stable exposures to the factors. There is considerable evidence that, in some markets, the volatility of market index portfolios changes. For example, periods of high volatility are often followed by periods of high volatility. The changing correlations among factor returns, and the changing volatility of market portfolios, belie the stability assumption underlying a simple covariance matrix.

For certain models we relax the assumption of stability in two ways (*see* Table 5-1 at the end of this chapter for details). First, in computing the covariance among the factor returns, we assign more weight to recent observations relative to observations in the distant past. Second, we estimate a model for the volatility of a market index portfolio—for example, the S&P 500 in the U.S. and the TSE1 in Japan—and scale the factor covariance matrix so that it produces the same volatility forecast for the market portfolio as the model of market volatility.

Exponential weighting

Suppose that we think that observations that occurred 60 months ago should receive half the weight of the current observation. Denote by T the current period, and by t any period in the past, $t = 1,2,3,\ldots,T-1,T$, and let $\delta = .5^{1/60}$. If we assign a weight of δ^{T-t} to observation t, then an observation that occurred 60 months ago would get half the weight of the current observation, and one that occurred 120 months ago would get one-quarter the weight of the current observation. Thus, our weighting scheme would give *exponentially declining weights* to observations as they recede in the past.

Our choice of sixty months was arbitrary in the above example. More generally, we give an observation that is *HALF-LIFE* months ago one-half the weight of the current observation. Then we let:

$$\delta = (.5)^{\frac{1}{HALFLIFE}} \qquad \text{(EQ 5-2)}$$

and assign a weight of:

$$w(t) = \delta^{T-t}. \qquad \text{(EQ 5-3)}$$

The length of the *HALF-LIFE* controls how quickly the factor covariance matrix responds to recent changes in the market relationships between factors. Equal weighting of all observations corresponds to *HALF-LIFE* = ∞. Too short a *HALF-LIFE* effectively "throws away" data at the beginning of the series. If the process is perfectly stable, this decreases the precision of the estimates. Our tests show that the best choice of *HALF-LIFE* varies from country to country. Hence, we use different values of *HALF-LIFE* for different single country models.

Computing market volatility: Extended GARCH models

There is considerable evidence that, in some markets, market volatility changes in a predictable manner. We find that returns that are large in absolute value cluster in time, or that volatility persists. Moreover, periods of above-normal returns are, on average, followed by lower volatility, relative to periods of below-normal returns. Finally, we find that actual asset return distributions exhibit a higher likelihood of extreme outcomes than is predicted by a normal distribution with a constant volatility.

Variants of GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models capture these empirical regularities by allowing volatility to increase following periods of high realized volatility, or below-normal returns, and allowing volatility to decrease following periods of low realized volatility, or above-normal returns.

The following discussion lays out the general theory of extended GARCH modeling. Variants of this approach will be applied as appropriate to BARRA single country models over time. *See* Table 5-1 of this chapter for current coverage.

Formally, denote by $\tilde{r}_t$ the market return at time t, and decompose it into its expected component, $\mathrm{E}(\tilde{r}_t)$, and a surprise, ε_t:

$$\tilde{r}_t = \mathrm{E}(\tilde{r}_t) + \varepsilon_t \tag{EQ 5-4}$$

The observed persistence in realized volatility indicates that the variance of the market return at t, $Var(\tilde{r}_m)_t$, can be modeled as:

$$Var(\tilde{r}_m)_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta\, Var(\tilde{r}_m)_{t-1} \tag{EQ 5-5}$$

This equation, which is referred to as a GARCH(1,1) model, says that current market volatility depends on recent realized volatility via ε_{t-1}^2, and on recent forecasts of volatility via $Var(\tilde{r}_m)_{t-1}$. If α and β are positive, then this period's volatility increases with recent realized and forecast volatility.

GARCH(1,1) models have been found to fit many financial time series. Nevertheless, they fail to capture relatively higher volatility following periods of below-normal returns. We can readily extend the GARCH(1,1) model to remedy this shortcoming by modeling market volatility as:

$$Var(\tilde{r}_m)_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta\, Var(\tilde{r}_m)_{t-1} + \theta\varepsilon_{t-1} \tag{EQ 5-6}$$

If θ is negative, then returns that are larger than expected are followed by periods of lower volatility, whereas returns that are smaller than expected are followed by higher volatility.

Having satisfactorily fit a GARCH model to the volatility of a market proxy portfolio, it is used to scale the factor covariance matrix so that the matrix gives the same risk forecast for the market portfolio as the GARCH model. In implementing the scaling, however, we scale only the systematic part of the factor covariance matrix.

Specific risk modeling

Overview

Referring to the basic factor model:

$$\tilde{r}_i = \sum_k X_{ik} \tilde{f}_k + \tilde{u}_i \tag{EQ 5-7}$$

The specific risk of asset *i* is the standard deviation of its specific return, $Std(\tilde{u}_i)$. The simplest way to estimate the specific risk matrix is to compute the historical variances of the specific returns. This, however, assumes that the specific return distribution is stable over time. Rather than use historical estimates, we *model* specific risk to capture fluctuations in the general level of specific risk and the relationship between specific risk and asset fundamental characteristics.

An asset's specific risk is the product of the *average* level of specific risk that month across assets, and each asset's specific risk *relative* to the average level of specific risk. Moreover, our research has shown that the relative specific risk of an asset is related to the asset's fundamentals. Thus, developing an accurate specific risk model involves a model of the average level of specific risk across assets, and a model that relates each asset's relative specific risk to the asset's fundamental characteristics.

Methodology

Denote by S_t the average level of specific risk across assets at time t, and by V_{it} asset i's specific risk relative to the average.

In equation form:

$$Std(\tilde{u}_{it}) = S_t(1 + V_{it}) \qquad \text{(EQ 5-8)}$$

where:

$Std(\tilde{u}_{it})$ = asset specific risk,

S_t = average level of specific risk at time t, and

V_{it} = asset i's specific risk relative to the average.

We estimate a model for S_t via time-series analysis, in which the average level of realized specific risk is related to its lagged values and to lagged market returns. Similarly, we estimate a model of relative specific risk by regressing realized relative specific risks of all firms, across all periods, on the firm fundamentals, which include the BARRA risk factors.

Modeling the average level of specific risk

We model the average level of specific risk via a time series model, where the average specific risk is related to its own lagged values, as well as to lagged market excess returns:

$$S_t = \alpha + \sum_{i=1,k} \beta_i S_{t-i} + \beta_{k+1} r_{m_{t-1}} + \varepsilon_t \qquad \text{(EQ 5-9)}$$

where:

S_t = average specific risk at time t,

α = expected level of average specific risk across assets,

β_i = sensitivity of average specific risk to lagged values,

S_{t-i} = lagged average specific risk,

β_{k+1} = sensitivity of average specific risk to market returns,

$r_{m_{t-1}}$ = market return at time t-1, and

ε_t = residual level of specific risk.

This simple model captures mean reversion or trends and persistence in volatility, as well as lower average specific risk following market rises, and higher specific risk following market declines (provided β_{k+1} is negative). We evaluate the performance of alternative weighting schemes by regressing realized average specific risk against predicted specific risk out of sample:

$$S_t = a + b\hat{S}_t + e_t \qquad \text{(EQ 5-10)}$$

The results of these tests help us determine the appropriate coefficients in Equation 5-9.

Modeling the relative level of specific risk

To model the relative level of specific risk, we first identify factors that may account for the variation in specific risk among assets. These factors will vary depending on the country model and may include:

- Industry membership
- Risk index exposures

Having identified these factors, we then estimate the relationship by performing the "pooled" *cross-sectional* regression:

$$V_{it} = \sum_{k=1,K} X_{ikt}\gamma_k + e_{it} \qquad \text{(EQ 5-11)}$$

where:

V_{it} = relative specific risk for asset i at time t,

t = 1 to T months,

i = 1 to N assets,

X_{ikt} = the exposure of asset i to factor k at time t, and

γ_k = factor k's contribution to relative specific risk.

In estimating this relationship, we mitigate the effects of outliers on the estimators by Windsorizing descriptors and risk indices.

Estimating the scaling coefficients

Average specific risk can vary widely over the full capitalization range of an equity market. To correct for this effect, scaling coefficients are used to adjust for any specific risk bias. To estimate the scaling coefficients, we divide the set of securities into capitalization deciles and, for each decile, we compute the bias in predicted specific risk from the previous steps. The scaling coefficients are then computed as a piece-wise linear function of an asset's relative capitalization in a manner that makes the average bias in predicted specific risk zero for each capitalization decile.

Final specific risk forecast

The above three components—average level of specific risk, relative level of specific risk, and decile scaling coefficients—are combined to produce the final asset specific risk forecast:

$$Std(\varepsilon_{it}) = scale_{it} \cdot \hat{S}_t \cdot (1 + \hat{V}_{it}) \quad \text{(EQ 5-12)}$$

where:

ε_{it} = specific risk of asset ***i*** at time ***t***,

$scale_{it}$ = scale coefficient for the decile that asset ***i*** falls into at time ***t***,

$\hat{S}_t$ = average level of specific risk across all assets, and

$\hat{V}_{it}$ = relative level of specific risk for asset ***i*** at time ***t***.

Updating the model

Model updating is a process whereby the most recent fundamental and market data are used to calculate individual stock exposures to the factors, to estimate the latest month's factor returns, and to recompute the covariance matrix.

The latest data are collected and cleaned. Descriptor values for each company in the database are computed, along with risk index exposures and industry allocations. Next, a cross-sectional regression is run on the asset returns for the previous month. This generates factor returns which are used to update the covariance matrix. Finally, this updated information is distributed to users of BARRA's applications software.

Comparison of risk model features

Table 5-1 summarizes BARRA's single country equity models and the features of each as of January 1998.

Country Model	Number of Industries	Number of Risk Indices	Industry Allocation Method: Multiple/Single	Exponential Smoothing Half-Life: Number of Months	GARCH: Yes/No
Australia (AUE2)	24	9	Single	90	No
Canada (CNE3)	40	11	Single	60	No
France (FRE3)	12	9	Single	48	No
Germany (GRE2)	17	10	Single	90	No
Germany—Trading Model (GRTM)	17	10	Single	90	Yes
Hong Kong (HKE1)	13	10	Single	48	No
Japan (JPE2)	40	12	Multiple	60	Yes
Korea (KRE1)	28	12	Single	24	No
Malaysia (MLE1)	14	10	Single	36	No
Netherlands (NLE1)	8	7	Single	60	No
New Zealand (NZE1)	6	5	Single	48	Yes
South Africa (SAE1)	23	11	Single	36	No
Sweden (SNE3)	20	10	Single	90	Yes
Switzerland (SWE2)	12	8	Single	90	No
Taiwan (TWE1)	25	10	Single	36	No
Thailand (THE1)	32	9	Single	36	No
U.K. (UKE5)	38	12	Single	90	No
U.K.—Trading Model (UKTM)	38	12	Single	90	Yes
U.S. (USE2)	55	13	Multiple	90	Yes
U.S. (USE3)	52	13	Multiple	90	Yes
U.S.—Small Cap (USSC)	42	11	Single	90	No
U.S.—Short-Term Risk (STRM)	55	3	Multiple	40 trading days	Yes

Table 5-1 BARRA Single Country Equity Risk Models* †

* As of 1/98.
† See text for definitions of model features.

Glossary

active management	The pursuit of investment returns in excess of a specified benchmark.
active return	Return relative to a benchmark. If a portfolio's return is 5%, and the benchmark's return is 3%, then the portfolio's active return is 2%.
active risk	The risk (annualized standard deviation) of the active return. Also called tracking error.
alpha	The expected residual return. Beyond the pages of this book, alpha is sometimes defined as the expected exceptional return and sometimes as the realized residual or exceptional return.
arbitrage	To profit because a set of cash flows has different prices in different markets.
Arbitrage Pricing Theory (APT)	Developed in the late 1970s, the theory which asserts that securities and portfolio returns are based on the expected returns attributable to an unknown number of underlying factors. APT provides a complementary alternative to its precursor, the Capital Asset Pricing Model.
benchmark	A reference portfolio for active management. The goal of the active manager is to exceed the benchmark return.
beta	The sensitivity of a portfolio (or asset) to a benchmark. For every 1% return to the benchmark, we expect a $\beta \cdot 1\%$ return to the portfolio.
beta, historical	Historical measure of the response of a company's return to the market return, ordinarily computed as the slope coefficient in a 60-month historical regression.
beta, predicted	Predicted systematic risk coefficients (predictive of subsequent response to market return) that are derived, in whole or in part, from the fundamental operating characteristics of a company. Also called fundamental beta.

breadth

The number of independent forecasts available per year. A stock picker forecasting returns to 100 stocks every quarter exhibits a breadth of 400, assuming each forecast is independent (based on separate information).

Capital Asset Pricing Model (CAPM)

The simplest version states that the expected excess return on securities will be exactly in proportion to their systematic risk coefficient, or beta. The CAPM implies that total return on any security is equal to the risk-free return, plus the security's beta, multiplied by the expected market excess return.

certainty equivalent return

The certain (zero risk) return an investor would trade for a given (larger) return with an associated risk. For example, a particular investor might trade an expected 3% active return with 4% risk for a certain active return of 1.4%.

characteristic portfolio

A portfolio which efficiently represents a particular asset characteristic. For a given characteristic, it is the minimum risk portfolio with portfolio characteristic equal to 1. For example, the characteristic portfolio of asset betas is the benchmark. It is the minimum risk beta = 1 portfolio.

coefficient of determination (R^2)

See R-squared.

common factor

An element of return that influences many assets. According to multiple-factor risk models, the common factors determine correlations between asset returns. Common factors include industries and risk indices.

constraint

In portfolio optimization, a limitation imposed upon the portfolio so that it will have desired characteristics.

correlation

A statistical term giving the strength of linear relationship between two random variables. It is a pure number, ranging from -1 to +1: +1 indicates a perfect positive linear relationship; -1 indicates a perfect negative linear relationship; 0 indicates no linear relationship. For jointly distributed random variables, correlation is often used as a measure of strength of relationship, but it fails when a nonlinear relationship is present.

covariance	The tendency of different random investment returns to have similar outcomes, or to "covary." When two uncertain outcomes are positively related, covariance is positive, and conversely, negatively related outcomes have negative covariances. The magnitude of covariance measures the strength of the common movement. For the special case of a return's covariance with itself, the simplified name of variance is used. Covariance can be scaled to obtain the pure number, correlation, that measures the closeness of the relationship without its magnitude.
descriptor	A variable describing assets, used as an element of a risk index. For example, a volatility risk index, distinguishing high volatility assets from low volatility assets, could consist of several descriptors based on short-term volatility, long-term volatility, systematic and residual volatility, etc.
Dividend Discount Model (DDM)	A model of asset pricing based on discounting the future expected dividends.
dividend yield	The dividend per share divided by the price per share. Also known as the **yield**.
earnings yield	The earnings per share divided by the price per share.
efficient frontier	A set of portfolios, one for each level of expected return, with minimum risk. We sometimes distinguish different efficient frontiers based on additional constraints, e.g., the fully invested efficient frontier.
exceptional return	Residual return plus benchmark timing return. For a given asset with beta equal to 1, if its residual return is 2%, and the benchmark portfolio exceeds its consensus expected returns by 1%, then the asset's exceptional return is 3%.
excess return	Return relative to the risk-free return. If an asset's return is 3% and the risk-free return is 0.5%, then the asset's excess return is 2.5%.
factor portfolio	The minimum risk portfolio with unit exposure to the factor and zero exposures to all other factors. The excess return to the factor portfolio is the **factor return**.

factor return — The return attributable to a particular common factor. We decompose asset returns into a common factor component, based on the asset's exposures to common factors times the factor returns, and a specific return.

information coefficient — The correlation of forecast return with their subsequent realizations. A measure of skill.

information ratio — The ratio of annualized expected residual return to residual risk. A central measurement for active management, value added is proportional to the square of the information ratio.

market — The portfolio of all assets. We typically replace this abstract construct with a more concrete benchmark portfolio.

modern portfolio theory (MPT) — The theory of portfolio optimization which accepts the risk/reward tradeoff for total portfolio return as the crucial criterion. Derived from Markowitz's pioneering application of statistical decision theory to portfolio problems, optimization techniques and related analysis are increasingly applied to investments.

multiple-factor model (MFM) — A specification for the return process for securities. This model states that the rate of return on any security is equal to the weighted sum of the rates of return on a set of common factors, plus the specific return on the security, where the weights measure the exposures (or sensitivity) of the security to the factor. These exposures are identified with microeconomic characteristics, or descriptors of the firms (*see* descriptor).

Several simplifications of this model have been used historically. If there is only one factor, it becomes a *single-factor model*; if this one factor is identified with an index, it is called a *single-index model*; if the single-factor is identified with the market factor, it becomes the *market model*. Depending on the statistical specification, some of these could become a *diagonal model*, which simply indicates that the covariance matrix between security returns is (or can easily be transformed into) a diagonal matrix.

normal — A benchmark portfolio.

normalization — The process of transforming a random variable into another form with more desirable properties. One example is standardization in which a constant (usually the mean) is subtracted from each number to shift all numbers uniformly. Then each number is divided by another constant (usually the standard deviation) to shift the variance.

optimization — The best solution among all the solutions available for consideration. Constraints on the investment problem limit the region of solutions that are considered, and the objective function for the problem, by capturing the investor's goals correctly, providing a criterion for comparing solutions to find the better ones. The optimal solution is that solution among those admissible for consideration which has the highest value of the objective function. The first-order conditions for optimality express the tradeoffs between alternative portfolio characteristics to provide the optimum solution.

outlier — A data observation that is very different from other observations. It is often the result of an extremely rare event or a data error.

passive management — Managing a portfolio to match (not exceed) the return of a benchmark.

payout ratio — The ratio of dividends to earnings. The fraction of earnings paid out as dividends.

performance analysis — Evaluation of performance in relation to a standard or benchmark with the purpose of assessing manager skill.

performance attribution — The process of attributing portfolio returns to causes. Among the causes are the normal position for the portfolio, as established by the owner of funds or the manager, as well as various active strategies, including market timing, common factor exposure, and asset selection. Performance attribution serves an ancillary function to the prediction of future performance, in as much as it decomposes past performance into separate components that can be analyzed and compared with the claims of the manager.

R-squared — A statistic usually associated with regression analysis, where it describes the fraction of observed variation in data captured by the model. It varies between 0 and 1.

regression	A data analysis technique that optimally fits a model based on the squared differences between data points and model fitted points. Typically, regression chooses model coefficients to minimize the (possibly weighted) sum of these squared differences.
residual return	Return independent of the benchmark. The residual return is the return relative to beta times the benchmark return. To be exact, an asset's residual return equals its excess return minus beta times the benchmark excess return.
residual risk	The risk (annualized standard deviation) of the residual return.
risk	The uncertainty of investment outcomes. Technically, risk defines all uncertainty about the mean outcome, including both upside and downside possibilities. The more intuitive concept for risk measurement is the standard deviation of the distribution, a natural measure of spread. Variance, the square of the standard deviation, is used to compare independent elements of risk.
risk-free return	The return achievable with absolute certainty. In the U.S. market, short maturity Treasury bills exhibit effectively risk-free returns. The risk-free return is sometimes called the time premium, as distinct from the risk premium.
risk index	A common factor typically defined by some continuous measure, as opposed to a common industry membership factor defined as 0 or 1. Risk index factors include Volatility, Momentum, Size, and Value.
risk premium	The expected excess return to the benchmark.
score	A normalized asset return forecast. An average score is 0, with roughly two-thirds of the scores between -1 and 1. Only one-sixth of the scores lie above 1.
security market line	The linear relationship between asset returns and betas posited by the Capital Asset Pricing Model.
Sharpe ratio	The ratio of annualized excess returns to total risk.

significance (statistical significance)	A statistical term which measures the spread or variability of a probability distribution. The standard deviation is the square root of variance. Its intuitive meaning is best seen in a simple, symmetrical distribution, such as the normal distribution, where approximately two-thirds of all outcomes fall within ± 1 standard deviation of the mean, approximately 95 percent of all outcomes fall within ± 2 standard deviations, and approximately 99 percent of all outcomes fall within ± 2.5 standard deviations. The standard deviation of return—or, more properly, of the logarithm of return, which is approximately symmetrically distributed—is very widely used as a measure of risk for portfolio investments.
skill	The ability to accurately forecast returns. We measure skill using the information coefficient.
specific return	The part of the excess return not explained by common factors. The specific return is independent of (uncorrelated with) the common factors and the specific returns to other assets. It is also called the idiosyncratic return.
specific risk	The risk (annualized standard deviation) of the specific return.
standard error	The standard deviation of the error in an estimate. A measure of the statistical confidence in the estimate.
standardization	Standardization involves setting the zero point and scale of measurement for a variable. An example might be taken from temperature, where the centigrade scale is standardized by setting zero at the freezing point of water and establishing the scale (the centigrade degree) so that there are 100 units between the freezing point of water and the boiling point of water. Standardization for risk indices and descriptors in BARRA equity models sets the zero value at the capitalization-weighted mean of the companies in the universe and sets the unit scale equal to one cross-sectional standard deviation of that variable among the estimation universe.
systematic return	The part of the return dependent on the benchmark return. We can break excess returns into two components: systematic and residual. The systematic return is the beta times the benchmark excess return.
systematic risk	The risk (annualized standard deviation) of the systematic return.

t-statistic — The ratio of an estimate to its standard error. The t-statistic can help test the hypothesis that the estimate differs from zero. With some standard statistical assumptions, the probability that a variable with a true value of zero would exhibit a t-statistic greater than 2 in magnitude is less than 5%.

tracking error — *See* active risk.

transaction costs — The costs incurred for a portfolio when securities are changed for other securities. Transaction costs are deducted from the value of the portfolio directly, rather than paid as fees to the money manager. These costs arise from three sources: (1) commissions and taxes paid directly in cash; (2) the typical "dealer's spread" (or one-half of this amount) earned by a dealer, if any, who acts as an intermediary between buyer and seller; and (3) the net advantage or disadvantage earned by giving or receiving accommodation to the person on the other side of the trade. The third component averages out to zero across all trades, but it may be positive or negative, depending on the extent to which a trader, acting urgently, moves the market against the selected strategy.

universe — The list of all assets eligible for consideration for inclusion in a portfolio. At any time, some assets in the universe may be temporarily ruled out because they are currently viewed as overvalued. However, the universe should contain all securities that might be considered for inclusion in the near term if their prices move to such an extent that they become undervalued. Universe also defines the normal position of a money manager, equating the normal holding with the capitalization-weighted average of the securities in the universe or followed list.

utility — A measure of the overall desirability or goodness of a person's situation. In the theory of finance, utility is the desirability of a risky series of outcomes. The utility (or expected utility) of a set of risky outcomes is assumed to measure its goodness, so that a package with higher utility is always preferred to one with lower utility. In portfolio theory, utility is almost always defined by a function of the mean and variance of portfolio outcomes, which is then called a mean/variance utility function. The further assumption that the utility function is linear in its two arguments (mean and variance) results in a linear mean/variance utility function (LMVU).

value added — The utility, or risk-adjusted return, generated by an investment strategy: the return minus a risk aversion constant times the variance. The value added depends on the performance of the manager and the preferences of the owner of the funds.

variance — A statistical term for the variability of a random variable about its mean. The variance is defined as the expected squared deviation of the random variable from its mean—that is, the average squared distance between the mean value and the actually observed value of the random variable. When a portfolio includes several independent elements of risk, the variance of the total arises as a summation of the variances of the separate components.

volatility — A loosely-defined term for risk. Here we define volatility as the annualized standard deviation of return.

yield — *See* dividend yield.

Index

T

U

V

W

Y

Contributors

Editors

Ronald N. Kahn
Pamela Brougham
Paul Green

Consultants

Ken Hui
Tom Leimkuhler
Scott Scheffler
Aamir Sheikh

Production Services

Susan McIntosh