CHAPTER 8

# Multivariate Models

> Model: A simplified description of a system or process . . . that assists calculations and predictions.
>
> —*Oxford English Dictionary*

Perhaps the defining characteristic of value-at-risk (VAR) systems is large-scale aggregation. VAR models attempt to measure the total financial risk of an institution. The scale of the problem requires the application of multivariate models to simplify the system. In many cases, it would be too difficult, and unnecessary, to model all positions individually as risk factors. Many positions are driven by the same set of risk factors and can be aggregated into a smaller set of exposures without loss of risk information.

Chapter 7 discussed the simple case where the number of positions is the same as the number of risk factors. Thus, if we had $N$ assets, we would use $N$ risk factors whose joint movement is described by an $N$ by $N$ covariance matrix. In general, however, we will choose fewer risk factors than the number of assets. This chapter provides tools for this simplification.

The fact that VAR is a large-scale portfolio aggregation has important consequences that too often are ignored. With large portfolios, the total risk depends heavily on correlations, even more so than on volatilities. Thus it is important to devote resources to model comovements between risk factors. The key challenge for the risk manager is to build a risk measurement system based on a parsimonious specification that provides a good approximation of the portfolio risk.

Multivariate models are most useful in situations where the risk manager requires internally consistent risk estimates for a portfolio of assets.

This is required, for instance, when the history of the current portfolio does not provide sufficient information to build a distribution of values. This is the case, for example, for distributions involving credit losses, such as those for collateralized debt obligations. Even when such a distribution exists, the multivariate approach is useful because it does not require reestimating the model for portfolios that differ from the current positions. Finally, multivariate models provide much better understanding of the structural drivers of losses by explicitly modeling joint movements in the risk factors.

Section 8.1 explains why the covariance matrix needs simplification. Factor models provide guidance for deciding how many risk factors are appropriate and are presented in Section 8.2. As we will see, an important role for the risk manager is to decide on the risk-factor structure. Using too many risk factors is unwieldy. Using too few, however, may create risk holes. This choice should be guided by the type of portfolio and trading strategy. Section 8.3 then discusses how to build joint distributions of the risk factors using a recently developed methodology called *copulas*. This allows more realistic modeling of the risk factors, in particular situations where markets experience extreme losses, as unfortunately is sometimes the case.

## 8.1. WHY SIMPLIFY THE COVARIANCE MATRIX

Chapter 7 examined the simple case where the number of assets $N$ is the same as the number of considered risk factors. Their joint movement then is described by the covariance matrix $\Sigma$. This assumes that all the risk factors provide useful information. In practice, this may not be the case.

Examination of the covariance matrix can help us to simplify the risk structure. Correlations, or covariances, are essential driving forces behind portfolio risk. When the number of assets $N$ is large, however, measurement of the covariance matrix becomes increasingly difficult. The covariance matrix has two dimensions, and the number of entries increases with the square of $N$. With 10 assets, for instance, we need to estimate $N \times (N + 1)/2 = 10 \times 11/2 = 55$ different variance and covariance terms. With 100 assets, this number climbs to 5050.

For large portfolios, this causes real problems. Correlations may not be estimated imprecisely. As a result, we could even have situations where the calculated portfolio variance is not positive, which makes no economic sense.

Define the portfolio weights as $w$. In practice, the covariance matrix is estimated from historical data. In the simplest method, VAR is derived from the portfolio variance, computed as

$$\sigma_p^2 = w'\Sigma w \tag{8.1}$$

The question is, Is the number resulting from this product guaranteed to be always positive? Unfortunately, not always. For this to be the case, we need the matrix $\Sigma$ to be *positive definite* (abstracting from the obvious case where all elements of $w$ are zero).

Negative values can happen, for instance, when the number of historical observations $T$ is less than the number of assets $N$. In other words, if a portfolio consists of 100 assets, there must be at least 100 historical observations to ensure that the portfolio variance will be positive. This is also an issue when the covariance matrix is estimated with decaying weights, as in the GARCH method explained in Chapter 9. If the weights decay too quickly, the number of effective observations can be less than the number of assets, rendering the covariance matrix nonpositive definite.

Problems also occur when the series are linearly correlated. This happens, for example, when two assets are identical ($\rho = 1$). In this situation, a portfolio consisting of \$1 on the first asset and −\$1 on the second will have exactly zero risk. In practice, this problem is more likely to occur with a large number of assets that are highly correlated, such as zero-coupon bonds or currencies fixed to each other.

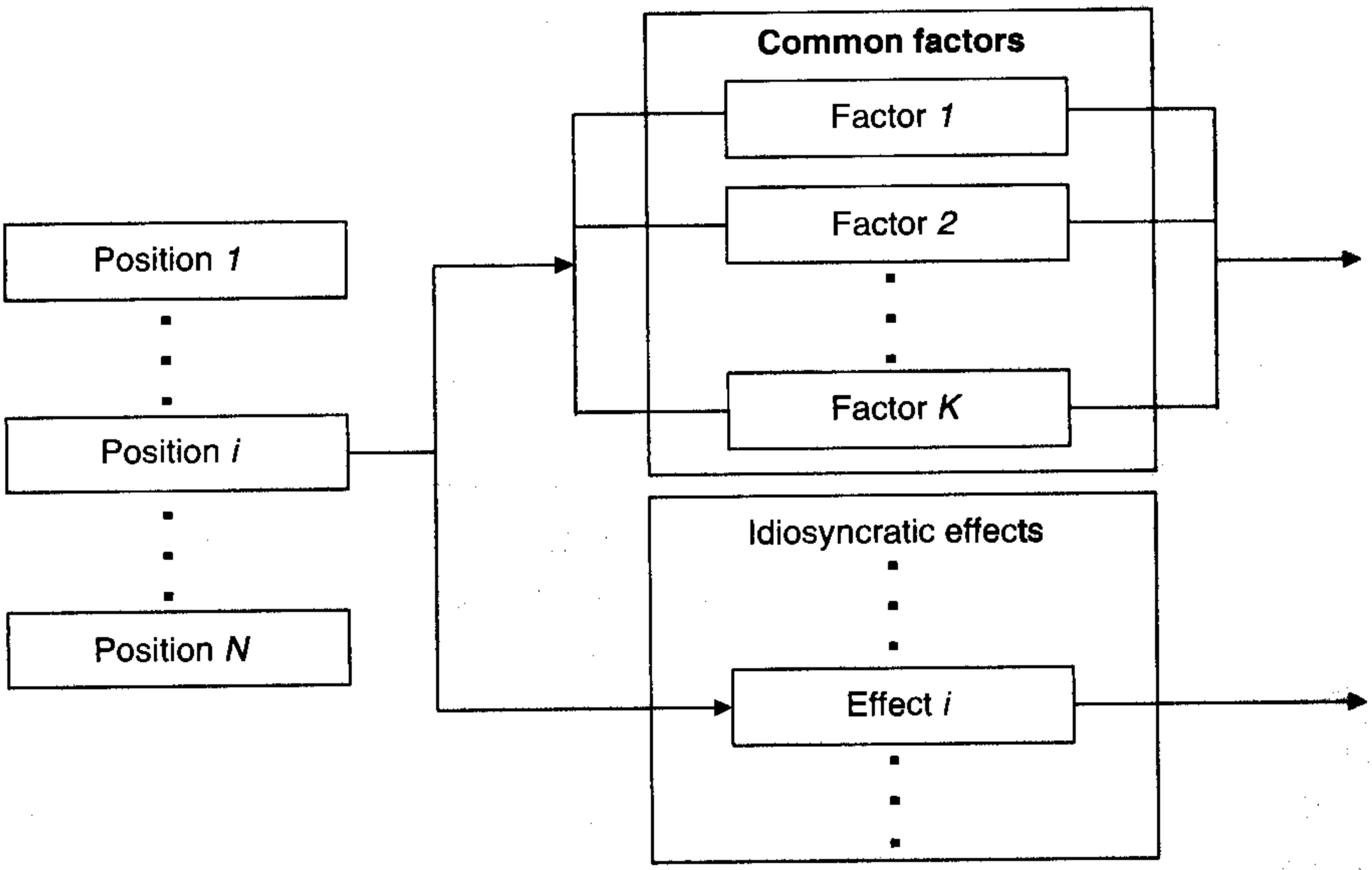
Most of the time this problem will not show up, and the portfolio variance will be positive. This may not be the case, however, if the portfolio has been *optimized* on the basis of the covariance matrix itself. Such optimization is particularly dangerous because it can create positions that are very large yet apparently offset each other with little total risk.

Such situations do arise in practice, however. As we shall see in Chapter 21, it largely explains the failure of the hedge fund Long-Term Capital Management. In practice, simple rules of thumb can help. If users notice that VAR measures appear abnormally low in relation to positions, they should check whether small changes in correlations lead to large changes in their VARs.

Alternatively, the risk structure can be simplified. Figure 8-1 shows how movements in $N$ asset values can be decomposed into a small number of common risk factors $K$ and asset-specific or idiosyncratic effects that are uncorrelated with each other. As we shall see, this structure reduces the number of required parameters substantially and is more robust than is using

**FIGURE 8-1**

Simplifying the risk structure.




a full covariance matrix. In addition, it lends itself better to an economic intuition, which helps to understand the results.

The framework described in Figure 8-1 can be extended to idiosyncratic effects that are correlated or have a more complex joint distribution, which can be modeled using the copula approach. It is also very flexible because it allows time variation in the comovements of the common factors.

## 8.2 FACTOR STRUCTURES

### 8.2.1 Simplifications

These issues become more troublesome as the number of assets increases. Assume that we want to select stocks from the entire universe of listed equities. These number more than 38,000. It is impossible to construct a covariance matrix for these assets that is positive definite.

This problem can be alleviated by the use of simpler structures for the covariance matrix. One example would be to have the same correlation coefficient across all pairs of assets. In this case, the sample is said

to be *homogeneous*. The Basel II rules are based on such a model, with a correlation coefficient of 0.20. This may be too simplistic, however, because it does not allow much differentiation between risk factors.

### 8.2.2 Diagonal Model

Another simple model is the *diagonal model,* originally proposed by Sharpe in the context of stock portfolios. The assumption is that the common movement in all assets is due to one common factor only, the stock market index, for example. The return on a stock $R_i$ is regressed on the return on the stock market index $R_m$, giving an unexplained residual $\epsilon_i$. Formally, the model is

$$R_i = \alpha_i + \beta_i R_m + \epsilon_i \tag{8.2}$$

with assumptions

$$E(\epsilon_i) = 0 \qquad E(\epsilon_i R_m) = 0 \qquad E(\epsilon_i \epsilon_j) = 0 \tag{8.3}$$

where $\beta_i$ is the exposure, or *loading*, on the market factor. For stocks, *beta* is also called *systematic risk* when the factor is the stock market index. The fixed intercept $\alpha_i$ can be ignored in what follows because it is not random and hence does not contribute to risk. Finally, define the variances as $\sigma_i^2 = V(R_i^2)$, $\sigma_m^2 = V(R_m^2)$, and $V(\epsilon_i^2) = \sigma_{\epsilon,i}^2$. In Equation (8.2), the $\beta_i R_m$ term is called *general market risk*, and the second term $\epsilon_i$, *specific risk*.

There are two key assumptions in Equation (8.3). First, the errors are uncorrelated with the common factor by construction. We have $\text{cov}(\epsilon_i, R_m) = E(\epsilon_i R_m) - E(\epsilon_i) E(R_m) = 0$. Second, the errors are uncorrelated across each other because $E(\epsilon_i \epsilon_j) = 0$.

The return on asset $i$ is driven by the market return $R_m$ and an idiosyncratic term $\epsilon_i$, which is not correlated with the market or across assets. As a result, the variance of stock $i$'s return can be decomposed as

$$\begin{aligned}\sigma_i^2 = V(\beta_i R_m + \epsilon_i) &= \beta_i^2 \sigma_m^2 + 2\,\text{cov}(\beta_i R_m, \epsilon_i) + V(\epsilon_i) \\ &= \beta_i^2 \sigma_m^2 + \sigma^2{}_{\epsilon,i}\end{aligned} \tag{8.4}$$

because $R_m$ and $\epsilon_i$ are uncorrelated. The covariance between two assets is

$$\sigma_{i,j} = \text{cov}(\beta_i R_m + \epsilon_i, \beta_j R_m + \epsilon_j) = \beta_i \beta_j \sigma_m^2 \tag{8.5}$$

which is solely due to the common factor because all the other terms are zero owing to Equation (8.3).

As a result, we can construct the full covariance matrix as

$$\Sigma = \begin{bmatrix} \beta_1\beta_1\sigma_m^2 + \sigma_{\varepsilon,1}^2 & \cdots & \beta_1\beta_N\sigma_m^2 \\ \vdots & & \vdots \\ \beta_N\beta_1\sigma_m^2 & \cdots & \beta_N\beta_N\sigma_m^2 + \sigma_{\varepsilon,N}^2 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} \begin{bmatrix} \beta_1 \cdots \beta_N \end{bmatrix} \sigma_m^2 + \begin{bmatrix} \sigma_{\varepsilon,1}^2 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \sigma_{\varepsilon,N}^2 \end{bmatrix}$$

Written in matrix notation, the covariance matrix has two components, a matrix composed of the outer product of the vector $\beta$ and a diagonal matrix with entries on the diagonal and zeroes elsewhere, that is,

$$\Sigma = \beta\beta'\sigma_m^2 + D_\epsilon \tag{8.6}$$

Since the matrix $D_\epsilon$ is diagonal, the number of parameters is reduced from $N \times (N+1)/2$ to $2N + 1$ ($N$ for the betas, $N$ in $D$, and one for $\sigma_m$). With 100 assets, for instance, the number of parameters in the covariance matrix is reduced from 5050 to 201, which is a considerable improvement.

Furthermore, we can use this simplification to compute the risk of a portfolio $p$, represented by weights $w$ on the assets, that is,

$$R_p = \sum_{i=1}^{N} w_i R_i = w'R \tag{8.7}$$

Using Equation (8.6), the variance of the portfolio reduces to

$$V(R_p) = V(w'R) = w'\Sigma w = w'(\beta\beta'\sigma_m^2 + D_\epsilon)w = (w'\beta)(\beta'w)\sigma_m^2 + w'D_\epsilon w \tag{8.8}$$

Both terms on the right-hand side of this equation must be scalars, that is, single numbers. The first term contains $\beta_p = w'\beta = \Sigma_{i=1}^N w_i\beta_i$, which is the beta of the overall portfolio. The second term is $\Sigma_{i=1}^N w_i^2\sigma^2{}_{\epsilon i}$.

Now, consider what happens when the number of assets $N$ increases and the portfolio is well diversified, which means that the weights decrease at the rate of $1/N$. The second term then becomes negligible as $N$ increases. For instance, if all the residual variances are identical and have equal weights, this second term is $[\Sigma_{i=1}^N(1/N)^2]\sigma_\epsilon^2 = [N(1/N)^2]\sigma_\epsilon^2 = [(1/N)]\sigma_\epsilon^2$, which converges to 0 as $N$ increases. Therefore, the variance of the portfolio converges to

$$V(R_p) \rightarrow (w'\beta\beta'w)\,\sigma_m^2 = (\beta_p\sigma_m)^2 \tag{8.9}$$

where $\beta_p = w'\beta$ is the portfolio beta. This simplified model is called the *beta model*. Thus, in large portfolios, specific risk becomes unimportant for the purpose of measuring VAR. This is a very important result. It means that the risk of large and well-diversified portfolios is dominated by the common factor.

## Example

As an example, consider three stocks, General Motors (GM), Ford, and Hewlett Packard (HPQ). The top panel in Table 8-2 displays the full covariance matrix. Numbers are reported in percentage monthly returns. This matrix can be simplified by estimating a regression of each stock on the U.S. stock market. These regressions are displayed in the second panel of the table, which shows betas of 0.806, 1.183, and 1.864, respectively. GM has the lowest beta; HPQ has the highest systematic risk. The market variance is $V(R_m) = 11.90$, which implies a monthly volatility of 3.45 percent, or 12 percent annually.

The bottom panel in the table reconstructs the covariance matrix using the diagonal approximation. For instance, the variance for GM is taken as $\beta_1^2 \times V(R_m) + V(\epsilon_1)$, which is $0.806^2 \times 11.90 + 64.44 = 7.73 + 64.44 = 72.17$. Ignoring specific risk, the beta model would forecast a variance of 7.73. The covariance between GM and Ford is $\beta_1\beta_2 V(R_m)$, which is $0.806 \times 1.183 \times 11.90 = 11.35$.

**TABLE 8-2**

The Diagonal Model

| | Covariances | | | Correlations | | |
|---|---|---|---|---|---|---|
| | **GM** | **Ford** | **HPQ** | **GM** | **Ford** | **HPQ** |
| Full matrix | | | | | | |
| GM | 72.17 | | | 1 | | |
| Ford | 43.92 | 66.12 | | 0.636 | 1 | |
| HPQ | 26.32 | 44.31 | 90.41 | 0.326 | 0.573 | 1 |
| Regression | | | | | | |
| $\beta_i$ | 0.806 | 1.183 | 1.864 | | | |
| $V(R_i)$ | 72.17 | 66.12 | 90.41 | | | |
| $V(\epsilon_i)$ | 64.44 | 49.46 | 49.10 | | | |
| $\beta_i^2 V(R_m)$ | 7.73 | 16.65 | 41.32 | | | |
| Diagonal model | | | | | | |
| GM | 72.17 | | | 1 | | |
| Ford | 11.35 | 66.12 | | 0.164 | 1 | |
| HPQ | 17.87 | 26.23 | 90.41 | 0.221 | 0.339 | 1 |

The last three columns in the table report the correlations between pairwise stocks. Actual correlations are all positive, as are those under the diagonal model. Although the diagonal-model matrix resembles the original covariance matrix, the approximation is not perfect. For instance, the actual correlation between GM and Ford is 0.636. Using the diagonal model, the correlation is driven by exposure to the market only. The estimated correlation is 0.164, which is lower than the true correlation. This is so because both stocks have relatively low betas, which is the only source of common variation.

Now let us compute VAR for a portfolio of $100 million invested equally in the three stocks. VAR is computed over a monthly horizon at the 95 percent confidence level. The first line in Table 8-3 shows the VAR of each individual stock, which ranges from 13.41 to 15.68 percent.

**TABLE 8-3**

Computing the VAR of a $100 Million Stock Portfolio
(Monthly VAR at 95 Percent Level)

| | | Covariance Matrix | | | |
|---|---|---|---|---|---|
| | **Position** | **GM** | **Ford** | **HPQ** | **VAR** |
| VAR (%) | | 14.01% | 13.41% | 15.68% | |
| Cov. matrix | | | | | |
| Full model | | | | | |
| GM | $33.33 | 72.17 | 43.92 | 26.32 | $11.76 |
| Ford | $33.33 | 43.92 | 66.12 | 44.31 | |
| HPQ | $33.33 | 26.32 | 44.31 | 90.41 | |
| Diagonal model | | | | | |
| GM | $33.33 | 72.17 | 11.35 | 17.87 | $10.13 |
| Ford | $33.33 | 11.35 | 66.12 | 26.23 | |
| HPQ | $33.33 | 17.87 | 26.23 | 90.41 | |
| Beta model | | | | | |
| GM | $33.33 | 7.73 | 11.35 | 17.88 | $7.30 |
| Ford | $33.33 | 11.35 | 16.65 | 26.24 | |
| HPQ | $33.33 | 17.88 | 26.24 | 41.32 | |
| Undiversified model | | | | | |
| GM | $33.33 | 72.17 | 69.08 | 80.78 | $14.37 |
| Ford | $33.33 | 69.08 | 66.12 | 77.32 | |
| HPQ | $33.33 | 80.78 | 77.32 | 90.41 | |

Next, the table displays four covariance matrices: the full matrix, the diagonal model, the beta model, and the undiversified model. The last is obtained by assuming unit correlation coefficients. The full matrix gives the true risk measure; the others are approximations.

Their respective VARs are \$11.76, \$10.13, \$7.30, and \$14.37. These numbers indicate that the diagonal model provides a good approximation of the actual portfolio VAR, although slightly on the low side. The beta model, in contrast, substantially underestimates the true VAR because it ignores residual risk. Finally, the undiversified VAR, obtained from adding the individual VARs, is much too high.

### 8.2.3 Multifactor Models

If a one-factor model is not sufficient, better precision can be obtained with multiple factors. Equation (8.2) can be generalized to $K$ factors, that is,

$$R_i = \alpha_i + \beta_{i1} f_1 + \cdots + \beta_{iK} f_K + \epsilon_i \tag{8.10}$$

where $R_1, \ldots, R_N$ are the $N$ asset returns, and $f_1, \ldots, f_K$ are *factors* independent of each other. In the previous three-stock example, the covariance matrix model can be improved with a second factor, such as transportation industry, that would pick up the higher correlation between GM and Ford. The key assumption again is that the residuals $\epsilon_i$ are uncorrelated across each other. All the common movements between the asset returns $R_i$ have been picked up by the multiple factors.

Extending Equation (8.6), the covariance matrix acquires a richer structure, that is,

$$\Sigma = \beta_1 \beta_1' \sigma_1^2 + \cdots + \beta_K \beta_K' \sigma_K^2 + D_\epsilon \tag{8.11}$$

The total number of parameters is $(N \times K + K + N)$, which is considerably fewer than for the full model. With 100 assets and 5 factors, for instance, the number is reduced from 5050 to 605, which is not a minor decrease.

Extending Equation (8.9) to multiple factors, we have

$$V(R_p) \to (\beta_{1p}\sigma_1)^2 + \ldots + (\beta_{Kp}\sigma_K)^2 \tag{8.12}$$

**Factor model result** Assuming that asset returns are driven by a small number of common factors and that residual movements are uncorrelated, the risk of portfolios that are well diversified with a large number of assets will be dominated by the common factors.

The next question is, How do we choose these common factors? Two methods. The first prespecifies factors that we think are important. This requires a good knowledge of markets and economic factors that drive them. The second derives factors from asset returns themselves through statistical techniques applied to the covariance matrix. The factors then can be given an economic interpretation after proper transformation.

One technique is *principal components analysis* (PCA). PCA attempts to find a series of independent linear combinations of the original variables that provide the best explanation of diagonal terms in the matrix. The methodology is summarized in Appendix 8.A. Another statistical method is *factor analysis* (FA). FA differs from PCA in that it focuses on the off-diagonal elements of the correlation matrix. This is important for applications where correlations are critical, such as differentials swaps, because the volatility of a difference involves a correlation.

### 8.2.4 Application to Bonds

Multifactor models are important because they can help the risk manager to decide on the number of VAR building blocks for each market. Consider, for instance, a government bond market that displays a continuum of maturities ranging from 1 day to 30 years. The question is, How many VAR building blocks do we need to represent this market adequately?

Before we start, note that the price of a bond $P$ is a nonlinear function of its yield to maturity $y$.[1] Taking the first derivative of this price function with respect to the yield gives

$$(dP/P) = -D^* \times (dy) \tag{8.13}$$

where $D^*$ is defined as the bond's *modified duration*. Equation (8.13) relates the relative change in the bond price, or return, to the change in yield using a linear approximation. Thus we can use the bond return or the yield change interchangeably as the risk factor. Generally, risk managers prefer to use yields as risk factors because their interpretation is more intuitive and also because yields have better statistical properties. In terms of volatility, we can write

$$\sigma(dP/P) = |D^*| \times \sigma(dy) \tag{8.14}$$

[1] For an excellent introduction to fixed-income tools, see Tuckman (2002).

Closed-form expressions can be derived for the modified duration of most bonds. For zero-coupon bonds with maturity $T$, for example, this is simply

$$D^* = \frac{T}{1+y}$$

So far, Equation (8.14) describes the relationship between bond prices and yields for each maturity, taken individually. It imposes no restrictions on movements in yields across maturities. In practice, the risk structure is often simplified to a one-factor model. The *duration model* assumes that the yield curve experiences parallel moves, either up or down. This implies that the volatility of yield changes is the same across all maturities and that correlations between yield changes are all equal to one. The issue is whether these simplifications fit the data.

Table 8-4 presents monthly VARs for 11 zero-coupon bonds for maturities going from 1 to 30 years in the U.S. Treasury bond market. For simplicity, assume normal distributions, so that VAR is proportional to the volatility, VAR = $\alpha\sigma$. The first column reports the VAR for returns, or VAR($dP/P$). Based on Equation (8.14), this can be related to VAR for yields

**TABLE 8-4**

Risk of U.S. Bonds (Monthly VAR at 95 Percent Level)

| Term (year) | Returns VAR (%) | Yield (%) | Modified Duration | Yield VAR (%) |
|---|---|---|---|---|
| 1 | 0.470 | 5.83 | 0.945 | 0.497 |
| 2 | 0.987 | 5.71 | 1.892 | 0.522 |
| 3 | 1.484 | 5.81 | 2.835 | 0.523 |
| 4 | 1.971 | 5.89 | 3.777 | 0.522 |
| 5 | 2.426 | 5.96 | 4.719 | 0.514 |
| 7 | 3.192 | 6.07 | 6.599 | 0.484 |
| 9 | 3.913 | 6.20 | 8.475 | 0.462 |
| 10 | 4.250 | 6.26 | 9.411 | 0.452 |
| 15 | 6.234 | 6.59 | 14.072 | 0.443 |
| 20 | 8.146 | 6.74 | 18.737 | 0.435 |
| 30 | 11.119 | 6.72 | 28.111 | 0.396 |

$$\text{VAR}(dP/P) = |D^*| \times \text{VAR}(dy) \tag{8.15}$$

With strictly parallel moves in the term structure, VAR($dy$) should be constant across maturities. Indeed, the last column in the table shows that yield VARs are similar across maturities. Longer maturities, however, display slightly less yield volatility than short maturities. The 30-year zero, for instance, has a yield VAR of 0.396 percent. This is lower than the yield VAR for the 1-year zero of 0.497 percent.[2] Thus the volatility of yield changes is fairly constant across maturities, except for a slight decrease toward the long end.

Next, Table 8-5 displays the correlation matrix. The correlations are high, suggesting the presence of common factors behind bond returns. Correlations are very high for close maturities but tend to decrease with the spread between maturities. The lowest value, 0.644, is obtained between the 1- and 30-year zeroes. Could this pattern of correlation be simplified to just a few common factors?

**TABLE 8-5**

Correlation Matrix of U.S. Bonds

| Term (year) | 1Y | 2Y | 3Y | 4Y | 5Y | 7Y | 9Y | 10Y | 15Y | 20Y | 30Y |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | | | | | | | | | | |
| 2 | 0.897 | 1 | | | | | | | | | |
| 3 | 0.886 | 0.991 | 1 | | | | | | | | |
| 4 | 0.866 | 0.976 | 0.994 | 1 | | | | | | | |
| 5 | 0.855 | 0.966 | 0.988 | 0.998 | 1 | | | | | | |
| 7 | 0.825 | 0.936 | 0.965 | 0.982 | 0.990 | 1 | | | | | |
| 9 | 0.796 | 0.909 | 0.942 | 0.964 | 0.975 | 0.996 | 1 | | | | |
| 10 | 0.788 | 0.903 | 0.937 | 0.959 | 0.971 | 0.994 | 0.999 | 1 | | | |
| 15 | 0.740 | 0.853 | 0.891 | 0.915 | 0.930 | 0.961 | 0.976 | 0.981 | 1 | | |
| 20 | 0.679 | 0.791 | 0.832 | 0.860 | 0.878 | 0.919 | 0.942 | 0.951 | 0.991 | 1 | |
| 30 | 0.644 | 0.761 | 0.801 | 0.831 | 0.853 | 0.902 | 0.931 | 0.943 | 0.975 | 0.986 | 1 |

[2] This number can be translated into an annualized yield volatility, which is (0.497/1.65) $\sqrt{12}$ = 1.04 percent per year.

**TABLE 8-6**

Principal Components of Correlation Matrix: U.S. Bonds

| Maturity (year) | Eigenvectors: Factor 1 $\beta_1$ | Factor 2 $\beta_2$ | Factor 3 $\beta_3$ | Percentage of Variance Explained by: Factor 1 | Factor 2 | Factor 3 | Total Variance Explained |
|---|---|---|---|---|---|---|---|
| 1 | 0.27 | 0.52 | 0.79 | 72.2 | 17.9 | 9.8 | 99.8 |
| 2 | 0.30 | 0.34 | −0.17 | 89.7 | 7.8 | 0.5 | 98.0 |
| 3 | 0.31 | 0.26 | −0.22 | 94.3 | 4.5 | 0.7 | 99.5 |
| 4 | 0.31 | 0.18 | −0.26 | 96.5 | 2.2 | 1.0 | 99.7 |
| 5 | 0.31 | 0.13 | −0.24 | 97.7 | 1.1 | 0.9 | 99.7 |
| 7 | 0.31 | −0.01 | −0.17 | 98.9 | 0.0 | 0.4 | 99.3 |
| 9 | 0.31 | −0.10 | −0.11 | 98.2 | 0.7 | 0.2 | 99.1 |
| 10 | 0.31 | −0.13 | −0.08 | 98.1 | 1.2 | 0.1 | 99.4 |
| 15 | 0.30 | −0.28 | 0.11 | 94.1 | 5.3 | 0.2 | 99.6 |
| 20 | 0.29 | −0.41 | 0.24 | 87.2 | 11.0 | 0.9 | 99.1 |
| 30 | 0.29 | −0.47 | 0.24 | 83.6 | 14.5 | 0.9 | 99.0 |
| Average | 0.30 | 0.00 | 0.01 | 91.9 | 6.0 | 1.4 | 99.3 |
| Eigenvalue | 10.104 | 0.662 | 0.156 | | | | |

Table 8-6 displays the results of the PCA applied to the correlation matrix in Table 8-5.[3] Appendix 8.A gives more detail on the method.[4] With PCA, the factors are linear combinations of the data. Redefining the yield change $dy$ as $R$, to shorten notations, the first principal component is defined as

$$z_1 = \beta_{11}R_1 + \cdots + \beta_{N1}R_N = \beta_1' R \tag{8.16}$$

Here, $\beta_1$ is called the first *eigenvector*, which represents the coefficients in the linear combination of the original variables that make up the first principal component. It is scaled so that the sum of its squared elements is 1. We observe from Table 8-6 that the first factor has similar coefficients across maturities. Thus it can be defined as a yield *level* factor.

[3] Note that PCA is sensitive to the values of the variances, or diagonal coefficients in the covariance matrix. Here we apply PCA to the correlation matrix. Using the correlation matrix gives equal weight to all risk factors.

[4] For the first application of PCA to the bond market, see Garbade (1986). Golub and Tilman (2000) provide a good overview of risk management with PCA.

**TABLE 8-7**

Correlation Matrix Fitted by First Component

| Term (year) | 1Y | 2Y | 3Y | 4Y | 5Y | 7Y | 9Y | 10Y | 15Y | 20Y | 30Y |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.722 | | | | | | | | | | |
| 2 | 0.805 | 0.897 | | | | | | | | | |
| 3 | 0.825 | 0.920 | 0.943 | | | | | | | | |
| 4 | 0.835 | 0.931 | 0.954 | 0.965 | | | | | | | |
| 5 | 0.840 | 0.936 | 0.959 | 0.971 | 0.977 | | | | | | |
| 7 | 0.845 | 0.942 | 0.965 | 0.977 | 0.983 | 0.989 | | | | | |
| 9 | 0.842 | 0.939 | 0.962 | 0.974 | 0.979 | 0.985 | 0.982 | | | | |
| 10 | 0.842 | 0.938 | 0.962 | 0.973 | 0.979 | 0.985 | 0.981 | 0.981 | | | |
| 15 | 0.824 | 0.919 | 0.942 | 0.953 | 0.959 | 0.965 | 0.961 | 0.961 | 0.941 | | |
| 20 | 0.793 | 0.884 | 0.906 | 0.917 | 0.923 | 0.928 | 0.925 | 0.925 | 0.906 | 0.872 | |
| 30 | 0.777 | 0.866 | 0.888 | 0.898 | 0.904 | 0.909 | 0.906 | 0.906 | 0.887 | 0.854 | 0.836 |

The bottom of the table shows the associated *eigenvalues*, defined as the variance of $z_1$. For the first factor, this is $\sigma^2(z_1) = 10.104$.

Table 8-7 displays the correlation matrix fitted by the first principal component. This is constructed as $\beta_1\beta_1'\sigma^2(z_1)$. This matrix reproduces fairly well the large off-diagonal entries. Note that this matrix is very much simplified. In particular, it cannot be a true correlation matrix because the diagonal elements are not unity.

Going back to Table 8-6, the percentage of variance explained represents the fraction of the diagonal element explained by each principal component. For instance, this is 72.2 percent for the first risk factor and first maturity. This is also the first diagonal element in the fitted correlation matrix. Across all maturities, the average is 91.9 percent. Thus the first factor has high average explanatory power.

In economic terms, the *level* factor provides an excellent fit to movements of the term structure. This also explains why the *duration model* provides a good measure of interest-rate risk. The PCA approach, however, is slightly more general than duration because duration assumes first that all elements of the first eigenvector are identical and second that all yield volatilities are equal.

The second factor explains an additional 6.0 percent of movements. Because it has the highest explanatory power and highest loadings for short and long maturities, it describes the *slope* of the term structure. Finally, the last factor is much less important. It seems to be most related to 1-year rates, perhaps because of different characteristics of money-market instruments. Together, these three factors explain an impressive 99.3 percent of all return variation.

We now illustrate how PCA can be used to compute risk for sample portfolios. Because we use the correlation matrix of changes in yields, we need to convert these positions into dollar exposures on these normalized risk factors. Define these dollar exposures as $x$. Using Equation (8.14), each entry is

$$x = D^* \times \sigma(dy) \times P = D^* \times [\text{VAR}(dy)/1.65] \times \text{P} \qquad (8.17)$$

defining $P$ as the market value of the position on each risk factor. The portfolio variance is then given by $x'\Sigma_\rho x$, where $\Sigma_\rho$ is the correlation matrix of changes in yields.

With two factors, the portfolio variance is, from Equation (8.12),

$$\sigma^2(R_p) = \beta_{1p}^2\, \sigma^2\,(z_1) + \beta_{2p}^2 \sigma^2(z_2) \qquad (8.18)$$

where $\beta_{1p} = x'\beta_1$ is the portfolio exposure to the first factor, $\beta_{2p} = x'\beta_2$, to the second factor.

Consider first a portfolio investing $P$=\$100 million each in 1-year and 30-year bonds. As shown in Table 8-8, the first-factor exposure is $\beta_{1p} = 0.285 \times 0.2673 + 6.747 \times 0.2877 = 2.017$. The second is $\beta_{2p} = -3.005$. The portfolio variance is

$$\sigma^2(R_p) = (2.017^2 \times 10.104) + (-3.005^2 \times 0.662) = 41.099 + 5.977 = 47.076$$

Taking the square root and multiplying by $\alpha = 1.65$, this gives a two-factor portfolio VAR of \$11.32 million. Using the full 11 factors gives a VAR of \$11.44 million. The first factor alone would have given a VAR of $\sqrt{111.892}$ = \$10.58 million, which is close. Thus, for this simple portfolio, using one principal component only would provide a good approximation to the true risk.

Table 8-8 analyzes another portfolio with \$100 million invested in the 10-year bond, \$40 million short the 30-year bond, and \$60 million short the 1-year bond. Because of the long and short positions, this is

**TABLE 8-8**

Risk Analysis by Principal Components

| Maturity (year) | Modified Duration $D^*$ | Yield VAR (%) VAR($dy$) | Portfolio 1 Position ($ million) $P$ | Portfolio 1 Position ($ million) $x$ | Portfolio 2 Position ($ million) $P$ | Portfolio 2 Position ($ million) $x$ |
|---|---|---|---|---|---|---|
| 1 | 0.945 | 0.497 | +100 | 0.285 | −60 | −0.171 |
| 2 | 1.892 | 0.522 | 0 | 0 | 0 | 0 |
| 3 | 2.835 | 0.523 | 0 | 0 | 0 | 0 |
| 4 | 3.777 | 0.522 | 0 | 0 | 0 | 0 |
| 5 | 4.719 | 0.514 | 0 | 0 | 0 | 0 |
| 7 | 6.599 | 0.484 | 0 | 0 | 0 | 0 |
| 9 | 8.475 | 0.462 | 0 | 0 | 0 | 0 |
| 10 | 9.411 | 0.452 | 0 | 0 | +100 | 2.578 |
| 15 | 14.072 | 0.443 | 0 | 0 | 0 | 0 |
| 20 | 18.737 | 0.435 | 0 | 0 | 0 | 0 |
| 30 | 28.111 | 0.396 | +100 | 6.747 | −40 | −2.699 |
| Exposure | | | $\beta_{1p} = +2.017$<br>$\beta_{2p} = -3.005$ | | $\beta_{1p} = -0.019$<br>$\beta_{2p} = +0.829$ | |

largely hedged against the first factor, with $\beta_{1p} = -0.019$ only. The two-factor risk analysis gives

$$\sigma^2(R_p) = (-0.019^2 \times 10.104) + (0.829^2 \times 0.662) = 0.004 + 0.455 = 0.459$$

Using the one-factor model generates a VAR of \$0.10 million, which is too low. The two-factor model provides a better approximation, a VAR of \$1.12 million that is close to the true VAR of \$1.42 million. Here we need at least a two-factor model.

This decomposition shows that for some purposes, the risk of a bond portfolio can be usefully summarized by its exposure to a very small number of factors. Whether this is sufficient depends on the structure of the portfolio being modeled.[5]

[5] As an example, Jamshidian and Zhu (1997) use information from PCA to design simulations that are very efficient. The basic idea is to simulate more data points for the first factor and fewer for remaining factors. Gibson and Prisker (2001), however, show that this approximation may fail, especially for hedged portfolios, and propose refinements to correct these shortcomings.

**BOX 8-1**

## RISK MODELS AT PIMCO

The *Total Return Fund*, run by asset manager Pacific Investment Management Company (PIMCO), is the largest bond mutual fund in the world, with close to $100 billion in assets. The portfolio has more than 10,000 different positions in fixed-income instruments, including derivatives. It would be impossible for the portfolio manager, Bill Gross, to keep track mentally of all these positions. This is where risk models can help.

PIMCO reduces the dimensionality of the problem by focusing on a small number of risk factors. These include (1) the level of the yield curve, (2) the slope between the 2- and 10-year maturities on the yield curve, (3) the slope between the 10- and 30-year maturities, (4) the spread between mortgages and Treasuries, and (5) the spread between corporates and Treasuries. Each position is expressed in terms of its exposure to these risk factors. These exposures are totted up across the entire portfolio, giving summary measures of exposures to these principal risk factors. The portfolio manager then can translate bets on risk factors into exposures and positions.

In 2002, PIMCO received the *asset management risk manager of the year award*. In describing this prestigious award, *Risk* noted that the "firm's risk-centric decision-making has allowed it to consistently beat its benchmarks." Indeed, over the previous 10 years, the Total Return Fund has rewarded investors with an average value added of 1.5 percent annually.

### 8.2.5 Comparison of Methods

To illustrate this important point, Table 8-9 presents VAR calculations for three portfolios.[6] The first is a diversified portfolio with $1 million equally invested in 10 stocks. The second consists of a $1 million portfolio with 10 stocks all in the same industry (high technology). The third expands on the diversified portfolio but is market-neutral, with long positions in the first five stocks and short the others. In other words, this is a *hedge fund* with zero net position in stocks.

[6] The diversified portfolio consists of positions in Ford, Hewlett-Packard, General Electric, Procter & Gamble, AT&T, Boeing, General Motors, Disney, Microsoft, and American Express. These are spread among 6 of the 10 industrial sectors in the market. The long-short portfolio is long the first five and short the others. The market index is taken as the Standard & Poor's (S&P) 500. VAR is measured with a 1-month horizon at the 95 percent level of confidence using historical data from 1990 to 1999.

**TABLE 8-9**

Comparison of VAR Methods

| | Portfolio | | |
|---|---|---|---|
| | **Diversified** | **High Tech** | **Long-Short** |
| **Net Position** | **$1,000,000** | **$1,000,000** | **$0** |
| **VAR** | | | |
| Index mapping | $63,634 | $63,634 | $0 |
| Beta mapping | $70,086 | $84,008 | $298 |
| Industry mapping | $69,504 | $90,374 | $7,388 |
| Diagonal model | $81,238 | $105,283 | $41,081 |
| Individual mapping (exact) | $78,994 | $118,955 | $32,598 |

Five methods are examined:

- *Index mapping* replaces each stock by a like position in the index $m$, that is,

$$\text{VAR}_1 = \alpha W \sigma_m$$

- *Beta mapping* only considers the net beta of the portfolio, that is,

$$\text{VAR}_2 = \alpha W (\beta_p \sigma_m)$$

- *Diagonal model* considers both the beta and specific risk, that is,

$$\text{VAR}_3 = \alpha W \sqrt{(\beta_p \sigma_m)^2 + w' D_\epsilon w}$$

- *Industry mapping* replaces each stock by a like position in an industry index $I$, that is,

$$\text{VAR}_4 = \alpha W \sqrt{w_I' \sum_I w_I}$$

- *Individual mapping* uses the full covariance matrix of individual stocks and provides an exact VAR measure over this sample period, that is,

$$\text{VAR}_5 = \alpha W \sqrt{w' \sum w}$$

The table shows that the quality of the approximation depends on the structure of the portfolio. This is an important conclusion. For the first portfolio, all measures are in a similar range, $60,000–$80,000. The diagonal model provides the best approximation, followed by the beta and industry-mapping models.

The second portfolio is concentrated in one industry and, as a result, has higher VAR. The index-mapping model now seriously underestimates the true risk of the portfolio. In addition, the beta and industry-mapping models also fall short because they ignore the portfolio concentration. The diagonal model is closest to the exact value, as before.

Finally, the third portfolio shows the dangers of simple mapping methods. The index-mapping model, given a zero net investment in stocks, predicts zero risk. With beta mapping, the risk measure, driven by the net beta, is close to zero, which is highly misleading. The best approximation is again provided by the diagonal model, which considers specific risks. In conclusion, the best risk model depends on the portfolio. This requires risk managers to have a thorough understanding of the investment process.

## 8.3 COPULAS

The traditional approach to multivariate analysis is based on the joint multivariate normal distribution for the risk factors. This implies that expected returns are linearly related to each other, as described by correlation coefficients and that, in addition, the probability of seeing extreme observations for many risk factors is low. A growing body of empirical research, however, indicates that these assumptions may be suspect. And this matters: The joint tail behavior of risk factors drives the shape of the tails of the portfolio distribution. Thus, using a normal assumption could lead to a serious underestimation of value at risk.

### 8.3.1 What Is a Copula?

This is where the concept of copulas comes to the rescue. To simplify, consider two risk factors only, 1 and 2. Their joint distribution can be split up into two statistical constructs. First is the marginal distribution for the two variables, $f_1(x_1)$ and $f_2(x_2)$. Second is the way in which the two marginals are *attached* to each other. This is done with a *copula*, which is a function that links marginal distributions into a joint distribution. Formally, the copula is a function of the marginal (cumulative) distributions $F(x)$,

which range from 0 to 1. In the bivariate case, it has two arguments plus parameters $\theta$, that is,

$$c_{12}\,[F_1(x_1),\, F_2(x_2);\theta] \tag{8.19}$$

The link between the joint and marginal distributions is made explicit by *Sklar's theorem*, which states that for any joint density there exists a copula that links the marginal densities, that is,

$$f_{12}(x_1, x_2) = f_1(x_1) \times f_2(x_2) \times c_{12}[F_1(x_1), F_2(x_2);\theta] \tag{8.20}$$

Consider, for example, a multivariate *normal* distribution. This can be split into two normal marginals and a normal copula. Assume that all variables are standardized, that is, have zero mean and unit standard deviation. Define $\Phi$ as the normal probability density function, $N$ as the cumulative normal function, $c^N$ as the normal copula, and $\rho$ as its correlation coefficient. This gives

$$f_1(x_1) = \Phi(x_1) \qquad f_2(x_2) = \Phi(x_2) \tag{8.21}$$

and

$$f_{12}(x_1,x_2) = \Phi(x_1,x_2;\, \rho) = \Phi(x_1) \times \Phi(x_2) \times c_{12}^{N}\,[N(x_1), N(x_2);\, \rho] \tag{8.22}$$

This shows that a bivariate normal density is constructed from two normal marginal densities and a normal copula. The bivariate density has one parameter, the correlation coefficient, which only appears in the copula.

Thus the copula contains all the information on the nature of the dependence between the random variables but gives no information on the marginal distributions. This allows a neat separation between the marginals and dependence. More complex dependencies can be modeled with different copulas.

### 8.3.2 Marginals and Copulas

In general, the copula can be any function that satisfies the appropriate restrictions behind Equation (8.20). It can be derived from the joint density function, for example, the normal or the student $t$. The student distribution is interesting because it displays fatter tails than the normal and greater dependences in the tails. We could mix and match the normal and student marginals with the normal and student copulas to represent the data better.

**FIGURE 8-2**

Combination of marginals and copulas.

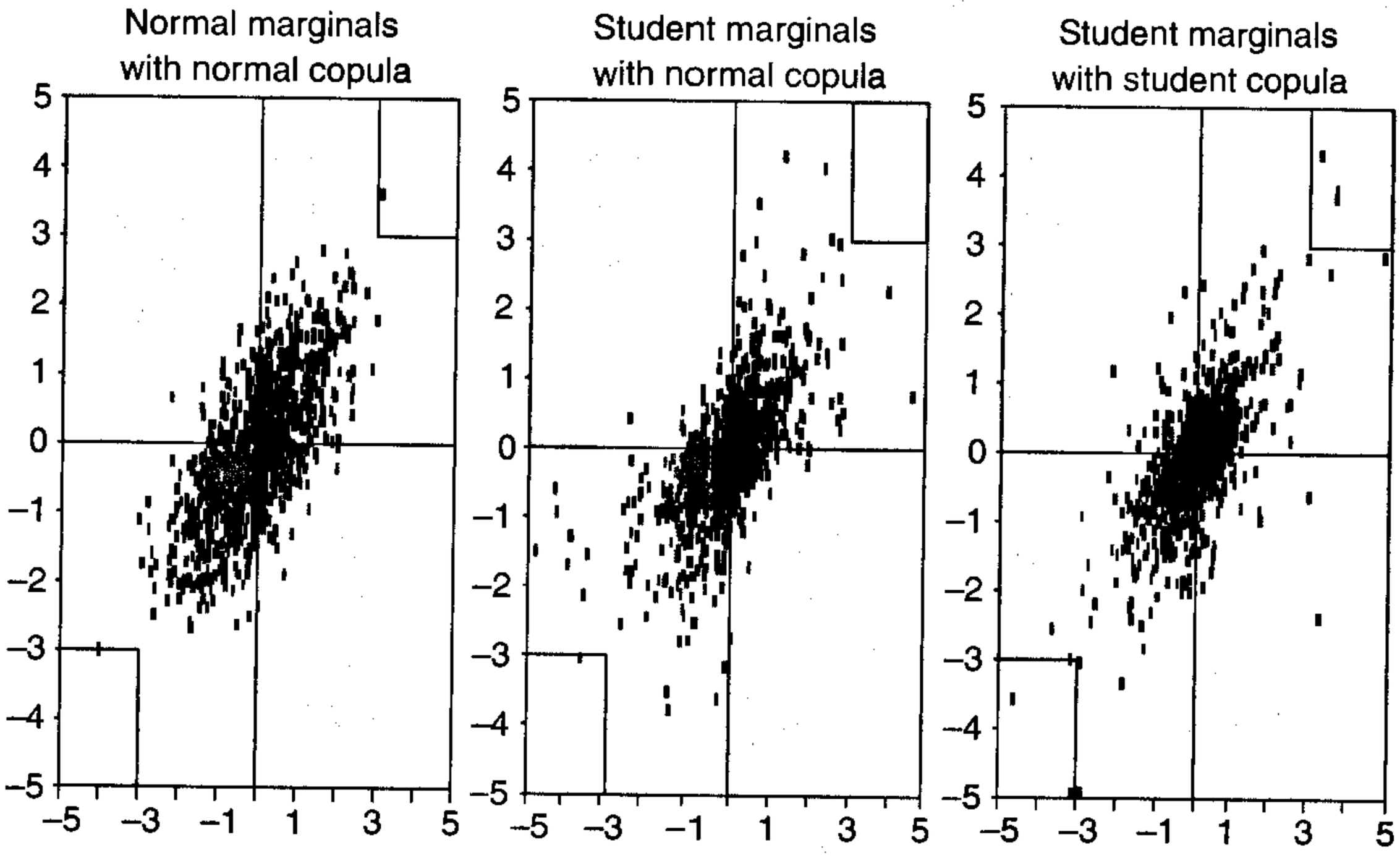



Figure 8-2 describes a plot of two variables generated with (1) normal marginals with a normal copula, (2) student marginals with a normal copula, and (3) student marginals with a student copula parametrized with 3 degrees of freedom. Now observe the boxes including extreme observations, either both above +3 or both below −3. If the two series are returns on two stock markets, for instance, we should be worried about situations where the two markets fall at the same time because this increases the risk of portfolios that have long positions in the two markets.

With the normal marginals, the dispersion of each variable is limited. Indeed, the probability of a move beyond +3 or −3 is only 0.003 percent or, on average, 3 observations from the 1000 in this sample. With the student marginals, there is greater dispersion of each variable, reflecting the fatter tails of the distribution. The two left panels in Figure 8-2 are based on the normal copula. This does not generate many dependencies in the tails. In the panel on the left we only have two cases with joint extreme observations; in the middle panel, there is only one.

The panel on the right combines the student marginals with the student copula. This has many more observations in the tails, three in the top

box and three in the lower box. These comovements increase the portfolio risk sharply. As an example, consider a portfolio equally weighted in the two risk factors. The VAR at the 99 percent confidence level increases from 2.1 to 2.3 to 2.4 when going from the left to the right panel. In this case, assuming a normal distribution understates risk by more than 10 percent, which is substantial. Furthermore, this bias will worsen with a greater number of risk factors. If the student copula is a better reflection of reality, it should be used instead of the normal copula.

To summarize, the risk-modeling process works in several steps. First, the risk manager has to choose the best functional form for the marginal distributions and the copula function. Second, the parameters of these functions must be estimated. The third step then consists of running simulations that generate random variables that mimic the risk factors. The current portfolio can be modeled as a series of positions on the risk factors. In the final step, the risk manager constructs the distribution of returns for the current portfolio. This can be summarized by VAR using a quantile of the distribution.

### 8.3.3 Applications

The preceding section has illustrated the use of *elliptical* copulas, which are symmetric around the mean. These imply the same probability of joint positive or negative movements, assuming positive correlations.

More generally, the copula can be asymmetric, with greater probability of joint moves in one direction or another. Geman and Kharoubi (2003), for example, wanted to examine the association between stocks and hedge-fund strategies. They fit several copulas to the joint movements between historical series. They found that for most categories of hedge funds, the *Cook-Johnson copula* provides the best fit. This is an asymmetric copula with greater probability of joint down moves for the two risk factors. This means that when stock markets drop precipitously, it is likely that some hedge-fund strategies will lose money as well. As a result, some categories of hedge funds provide much less diversification with stocks than hoped for.

Copulas are bound to be used increasingly in financial risk management because they can be used to build joint distributions of risk factors. They are finding a wide range of applications, as illustrated in Box 8-2. Another application, detailed in Chapter 21, will be the integration of market, credit, and operational risk at the highest level of the financial institution.

**BOX 8-2**

## COPULAS IN FINANCE

*Collateralized debt obligations* (CDOs) are pooled investments in debt instruments that offer ready-made diversification. The total cash flows are directed to different classes of claims, or *tranches*, according to predefined priority rules. Losses owing to default hit first the lowest-rated tranches, then the middle-rated tranches (called *mezzanine*), and then the senior tranches. To ascertain expected losses to each tranche, we need to construct the entire distribution of portfolio values.

Payoffs on CDO tranches depend heavily on correlations among defaults in the underlying credit portfolio. Low correlations make the senior tranches safer. On the other hand, if all underlying bonds default at the same time, the senior tranches could face serious losses. David Li (2000) is widely credited with having developed the first commercial model for CDO pricing, using the concept of copula functions.

Since then, the standard industry model has been the normal copula because of its simplicity. CreditMetrics, for instance, generates a joint distributions in asset values using a multivariate normal distribution, which implies a normal copula.

Like all models, these are just approximations of reality. Sometimes these approximations work poorly. On May 5, 2005, the credit-rating agencies downgraded the debt issued by General Motors (GM) and Ford to below investment grade. This event, however, was specific to these two firms and did not affect others. Many investors had tried to hedge GM and Ford's debt by shorting other bonds, based on the relationships predicted by the normal copula. They lost millions of dollars during this episode.

## 8.4 CONCLUSIONS

Risk management systems typically involve large-scale aggregation. Because of the number of risk factors, simplifications are often required. This chapter has provided tools to model the multivariate distribution of risk factors.

This involves choosing the shape of the joint density and its parameters. Generally, normal joint densities are used merely because of convenience. Such densities, however, do not generate the joint movements in the tails that we seem to observe in empirical data. This is important because the possibility of large simultaneous drops in prices means

that the portfolio risk can be very high. Such tail dependences can be modeled, for instance, using the student copula.

The covariance matrix, or correlation matrix, also needs special attention. In large samples, portfolio risk is driven primarily by correlations. With a large number of assets, however, there are too many parameters to estimate. The covariance matrix needs simplifications. Factor models help to reduce the dimensionality of the problem.

A particularly interesting application is that of principal component analysis. This approach simplifies the risk measurement process considerably and gives a better understanding of the underlying economics. The choice of number of risk factors, however, is driven by a tradeoff between parsimony and accurate risk measurement. Ultimately, the choice of the joint distribution should be made by the risk manager based on market experience and a solid understanding of these multivariate models.