# Principals of Risk:
# Finding Value-at-Risk Through Factor-Based Interest Rate Scenarios

Jon Frye
NationsBanc-CRT
233 South Wacker Drive, 27th Floor
Chicago, IL 60606
312-234-2644
Fax: 234-4009
Email: jonf@crt.com

April, 1997

---

The author wishes to thank Mark Bennett, Bill DeRonne, John Fuqua, Doug Huggins, Carol Lobbes, George McClintick, Ivan Marcotte, Mike O'Neill and Tim Weithers for their extensive comments on earlier versions, and many others at NationsBanc-CRT and the Chicago Risk Management Workshop for their helpful observations and suggestions. The views expressed reflect those of the author and the errors are his.

**Principals of Risk:**
**Finding Value-at-Risk Through Factor-Based Interest Rate Scenarios**


Abstract

The Factor-Based approach to calculating VAR begins with a principal components analysis of the yield curve. This decomposes yield curve movements into a small number of underlying factors including a “Shift” factor that allows rates to rise or fall and a “Twist” factor that allows the curve to steepen or flatten. Combining these factors produces specific yield curve scenarios used to estimate hypothetical portfolio profit or loss. The greatest loss among these scenarios provides an intuitive and rapid VAR estimate that tends to provide a conservative estimate of the nominal percentile of the loss distribution.

## Principals of Risk:
## Finding Value-at-Risk Through Factor-Based Interest Rate Scenarios

This article describes a method to assess the value-at-risk (VAR) of portfolios containing interest rate sensitive instruments. The Factor-Based Scenario Method calculates profit or loss for the entire portfolio under several specially constructed hypothetical interest rate scenarios that in turn derive from a principal components analysis of the yield curve. The VAR estimate equals the greatest loss that results in any of the scenarios.

The Factor-Based Scenario Method provides the user with several advantages relative to alternative calculation strategies. First, it includes options in its analysis without modifications. By contrast, options can provoke extremely misleading results in VAR calculations that linearize profit as a function of market variables. Second, the Scenario Method produces estimates of risk quickly, unlike methods that depend on extensive simulation. A VAR estimate has little value unless it arrives in time to allow adjustments to the portfolio. Third, the method provides a useful and easily understood summary of the qualitative nature of the risk facing a portfolio. One can immediately see that the portfolio is sensitive, say, to rising interest rates or to a steepening yield curve. Fourth, the method identifies whether an additional trade will increase or decrease risk, which provides insight into hedge strategies. Finally, the method allows the straightforward aggregation of risks across portfolios maintained and valued on different computer systems. The response to a given scenario of the combined portfolio is simply the sum of the responses of the individual portfolios.

The Factor-Based Scenario Method is not foolproof, however, and the user must judge the appropriateness of the technique for the portfolio at hand. Nonetheless the method works well for a class of portfolios that is quite important in practice—portfolios that display a concave response to changing market prices, such as portfolios dominated by short positions in standard options. The negative gamma of such positions creates particular concern for risk control purposes.

The following section considers four alternative measures of VAR in the context of a single underlying market. The next two sections provide an orientation to the Factor-Based Scenario Method and to the statistical technique of principal components analysis (PCA). The subsequent section places PCA in the more familiar context of regression analysis. “From Components to Scenarios” presents the steps that convert the generic results of PCA to specific scenarios useful for risk control. The final sections of the article examine the performance of the Factor-Based Scenario Method applied both to randomly generated sample portfolios and to an actual trading portfolio.

## Measuring Risk From a Single Market

VAR is usually defined as a specific percentile of a distribution of loss over a specific length of time. We will consider a one-day, $99^{th}$ percentile VAR:

$$P[\text{one-day loss} > \text{VAR}] = 1\%$$

Suppose we seek VAR for a portfolio tied to only a single market price. Figure 1 depicts a delta-hedged short position in a short-term, out-of-the-money call option on a Treasury Bond futures contract. Possible changes to the futures price have been stated in standard deviations. Consider four alternative VAR calculations: Monte Carlo simulation, historical simulation, variance-covariance, and the Factor-Based Scenario Method.

The Monte Carlo approach randomly selects a hypothetical movement for the Bond futures price and then values the portfolio. Many such hypothetical valuations reveal a distribution of profit and loss. The $99^{th}$ percentile of the loss distribution provides the VAR estimate.

This approach has some significant drawbacks in practice. A portfolio of fixed income instruments might require minutes of computer time to value. Even if several machines are devoted to risk calculations, they can complete only a few hundred calculations in the allotted production cycle. This relatively low number of simulations does not allow for the accurate estimation of the higher percentiles of the loss distribution. The loss that might occur one time in a thousand—one day in four years, on average—has considerable practical importance but would be estimated quite inaccurately. Second, the result of a Monte Carlo simulation is a random number that depends on random inputs. In a trading situation where a position may need adjustment, decision making is not enhanced by the idea that a second run might have a different result. Finally, Monte Carlo results may be difficult to interpret and communicate when dissimilar random scenarios result in nearly the same loss.

One suggestion to speed the calculation and make it non-random is to use actual historical price movements—perhaps the last 100 days of price changes—rather than simulated movements. This obviously puts a cap on the number of simulation runs, but again it cannot provide good estimates of high percentiles of the loss distribution.

Perhaps the most frequently discussed method for the VAR calculation is the variance-covariance method. The variance-covariance method approximates the response of the portfolio by a process called ‘mapping,’ by which the risk of a position is represented as quantities of standardized instruments. The standardized instruments respond linearly to underlying markets and have no optionality. Essentially, the portfolio response is approximated by a linear function.

If the portfolio has a nonlinear response to market variables such as seen in Figure 1, this approach can produce very misleading results. In particular, it assigns zero risk to a delta-hedged short option position. Far from having zero risk, this position loses considerable money when the market moves either up or down a large amount. It is exactly this sort of position that risk controllers need to limit in order to avoid large losses when the market moves strongly.

The Factor-Based Scenario Method first devises the scenarios, generally with reference to the quantiles of the normal distribution. It may try scenarios at various levels of intensity such as ±2.33 standard deviations, ±4 standard deviations, and ±6 standard deviations. (A different limit would apply to each set of scenarios.) For the portfolio of Figure 1 there is only one "factor," the price of the T-Bond contract. Second, it revalues the portfolio under each scenario. Figure 1 shows valuations under the "Up 2.33 SDs" and "Down 2.33 SDs" scenarios. For this portfolio, both the Up scenario and the Down scenario produce losses. The Up scenario produces the greater loss, $1.45 million, taken as the VAR estimate.

The Factor-Based Scenario Method performs a complete revaluation of the portfolio under each of the hypothetical scenarios it selects. In this way it resembles the Monte Carlo method more strongly than the variance-covariance approach. Yet the Scenario Method produces its result much more quickly than Monte Carlo because the number of revaluations it employs is small—two in this case.

Something is lost in this otherwise favorable bargain: the resulting estimate does not exactly equal the 99$^{th}$ percentile of the loss distribution. This is partly because the T-Bond futures price is not normally distributed; for example, it tends to rise by 2.33 standard deviations more than 1% of the time. Even treating the market as normal, the factor-based estimate understates the 99$^{th}$ percentile loss. That is because there is 1% chance that the price rises by more than 2.33 standard deviations (causing a loss greater than $1.45 million), and in addition there is a 0.000000006% chance that price falls by more than 6.4 standard deviations (this occurs to the left beyond Figure 1 and also causes a loss greater than $1.45 million). Where the 0.01 quantile was sought, the 0.01000000006 quantile was found. Should the portfolio show the same loss under both the Up scenario and the Down scenario, the resulting estimate represents the 98$^{th}$ percentile of loss rather than the 99$^{th}$ percentile. The actual 99$^{th}$ percentile loss would correspond to 2.58 standard deviations (two-tailed 1%) rather than 2.33 standard deviations (one-tailed 1%), a 9.7% error. The potential for this degree of error is the cost of the several advantages of the Factor-Based Scenario Method for a portfolio of hedged short options in the case of one market.

One can easily construct portfolios where the estimation error is large. The simplest is when the portfolio is long a put and long a call as in Figure 2. Valued at either the Up scenario or the Down scenario, this portfolio makes money, so the Scenario Method sets VAR equal to $0 despite what can be seen in the diagram. If the profit function tends to have important interior minima such as this, the Scenario Method is not appropriate and the remainder of this article assumes the portfolio at hand lacks interior minima.

## Interest Rate Scenarios and Principal Components Analysis

The most common yield curve scenarios posit a shift of a given number of basis points at all maturities. However, portfolio value usually depends on the steepness of the yield curve as well as its overall level. This recognition might lead to additional scenarios: a

"curve steepening" scenario, a "curve flattening" scenario, and a "curve rising and steepening" scenario. More scenarios might be devised to reveal the risk of more complicated positions, and so forth. Unfortunately, this ad-hoc process quickly loses contact with the probabilities revealed in interest rate data.

The Factor-Based Scenario Method takes a more rigorous route to derive interest rate scenarios. The method views movements in the yield curve as stemming from the movements of several underlying factors. As in Litterman and Scheinkman (Journal of Fixed Income), the factors emerge from the use of the multivariate statistical technique of Principal Components Analysis (PCA). Of the several types of factor analysis described in the statistical literature, PCA results in an explicit statement of the factors themselves. It also supplies the associated standard deviations to measure the relative importance of each component in describing the data. The first principal component has the greatest effect on the data and the greatest standard deviation; the other principal components follow in turn.

The first few principal components of the yield curve data represent the most typical movements of the yield curve. The first principal component appears similar to a "level shift" factor. This corresponds to the intuition that the most important event affecting the yield curve on a given day is whether yields generally tend to rise or to fall. The second, less important, component appears similar to a "curve steepening" factor. This factor allows the yield curve to pivot steeper or flatter. The third principal component allows intermediate-term yields to rise (fall) while both short-term and long-term rates fall (rise).

Nothing about PCA is intrinsically difficult. Still, several facts conspire to make it unfamiliar even to market participants familiar with statistical regression. The technique of PCA is usually taught in multivariate statistics courses where most of the applications come from social sciences such as psychology, rather than finance. Different texts on PCA do not agree on the meanings of some key terms. The geometric intuition imparted by some texts may not seem relevant. Nonetheless, PCA applies directly and simply to the problem at hand because it reduces the apparent complexity of the interest rate world while retaining its essential richness. Anyone acquainted with linear regression can understand and interpret the results of PCA.

## Principal Components Results

This section provides an orientation to the PCA results produced by standard software. The data consists of 1543 daily observations at ten points along the yield curve: six and one quarter years (1/89 through 3/95) of daily changes in constant-maturity U.S. Treasury yields (in basis points) observed at tenors of 3, 6 and 12 months and 2, 3, 4, 5, 7, 10, and 30 years.

The first step in the analysis removes the downward drift of interest rates—about 200 basis points (less than 0.2 basis points per day) at each point on the curve over the sample period. The mean of each adjusted series is zero. (This has no effect on the results

produced by standard statistical packages, but facilitates the analogy to regression analysis presented below.) Table 1 shows the mean-adjusted data for the first ten dates.

Several statistical software packages offer PCA. Running any of them produces a wealth of results that include (1) the loadings[1] of each principal component, (2) the factor scores, and (3) the standard deviation of each component. Table 2 presents selected PCA results for the interest rate data.

Table 2 shows the loadings of all ten principal components that arise from the analysis of the ten points on the yield curve. The first three have the character of typical movements of the yield curve as understood by market participants. PC1 causes every interest rate to rise, PC2 causes the yield curve to steepen, and PC3 causes one-year to five-year rates to fall while both long-term and short-term rates rise. Thus the first few principal components correspond to the intuitions that traders have about how the market moves on a daily basis.

The factor scores restate each day's yield curve movement as a combination of the movements of principal components. For example, Table 1 shows that on 1/5/89 all interest rates rose and that short term rates rose more than long term rates. Table 2 restates this movement as a combination of factors: PC1 enters with a large positive score (the curve shifts up) and PC2 enters with a large negative score (the curve flattens). The scores of the first two principal components shown in Table 2 have nearly as much information as all ten data points shown in Table 1.

The standard deviations of the ten factor score series determine their ordering. In Table 2, the volatility of the scores of PC1 exceeds that of any other component, the volatility of the scores of PC2 exceeds that of any component except PC1, and so forth. The squares of the standard deviations—the variances of the components—have an interesting property. The sum of the variances of the components equals the sum of the variances of the original data. The total variance in the original data is thus

$$17.49^2 + 6.05^2 + 3.10^2 + 2.17^2 + .... = 367.9$$

and the proportion of this data explained by the first factor is $17.49^2/367.9 = 83.1\%$. The proportion explained by the first two principal components is

$$(17.49^2 + 6.05^2)/367.9 = 93.1\%$$

Since all ten of the components explain 100% of the variation in the original data, the principal component scores provide an exact restatement of each day's activity that is equivalent to the original data. The principal components description is simpler, though, because most of the variance in the original data concentrates in the first few principal components. A movement in the first principal component explains movement at all points of the yield curve and explains as much of that movement as possible.

## Analogy to Linear Regression Analysis

A comparison with regression analysis will elucidate the previous statement. In regression, one has a dependent variable that one wishes to explain with the independent variable. Most interest centers on the regression coefficient and on $R^2$. The least-squares procedure maximizes $R^2$ by minimizing the sum of the squares of the errors produced by the regression equation.

Intuitively, PCA "makes up" the data on the right hand side of the regression equation. Instead of beginning with known data, PCA finds the data that provides the best fit. If one had only one day of yield curve data to explain, the best explanation would be the data itself. The statistical puzzle that PCA solves is to find the best explanatory variable for all days in the sample period.

Continuing the analogy to regression analysis reveals some insights. It is helpful to think of PCA as operating by iteration. The iteration begins with a guess: the uniform-shift variable [1,1,1,1,1,1,1,1,1,1]. This variable serves on the right-hand side of each of N=1543 regression equations. Each regression separately estimates its coefficient to minimize the sum of squared errors within that regression. The total of the N sums of squared errors from these regressions measures the overall explanatory power of the independent variable. A good explanatory variable will produce a lower cumulated sum than a poor one. One can imagine the iterative procedure improving on the uniform shift variable until no further improvement is possible. *The cumulated sum of squared errors reaches its minimum when the explanatory variable equals the first principal component.*[2] No other variable explains the data as well as the first principal component. Starting with nothing, PCA finds the optimal independent variable.[3]

Continuing the regression analogy, *the coefficients that arise from the N regressions on the first principal component equal the factor scores as produced by PCA.* For example, regressing the data for 1/5/89 on the PC1 column of the loadings matrix results in a regression coefficient of 23.6––the factor score for PC1 that appears in Table 2 for 1/5/89. One could say that the market on 1/5/89 experienced 23.6 "units" of PC1. The restatement of market movement provided by the factor scores has the character of a regression relationship where the components play the role of the explanatory variables and the factor scores play the role of regression coefficients.

Finally, *the standard deviation of the set of coefficients that arise from the N regressions equals the standard deviation of the first principal component.* Any day's coefficient estimates how strongly the first component affects that day's interest rate movements. The standard deviation of the 1543 coefficients tells the volatility of the component as revealed by the data. In the case of PC1, the standard deviation equals 17.49. Thus we can say that on 1/5/89 PC1 took the unusual value of 23.6/17.49 = 1.35 standard deviations.

If asked to write down the explanatory variable that would do the best job explaining daily variation in the yield curve, many practitioners would choose a level shift in which all rates rise or fall by the same amount. The first principal component of the yield curve

data improves on the intuitive idea of "all rates go up or down by the same amount" to provide a calibrated and non-uniform shift in rates.

Though the first principal component gives the best single explanation of daily movements, it leaves some variation unexplained. The remaining variation can act as data in a second set of N regressions. The best explanatory variable for these regressions equals the second principal component. The second principal component of the yield curve data improves on the intuitive idea "the yield curve tends to steepen or flatten" to provide a precise quantification of the relationships along the curve as it steepens or flattens.

The first two principal components taken together explain over 93% of the variation in the original data. Combinations of them provide a rather accurate impression of yield curve movements on most days. Since these two factors provide an approximate description of the historical data, it is natural to use them to develop an idea of the sorts of yield curve movements that may arise in the future. Combinations of the principal components produce yield curve scenarios that reveal the interest rate risk of a portfolio of transactions.

## From Components to Scenarios

This section derives four scenarios from the first two principal components. No formal statistical test can judge that two components adequately capture yield curve movements. Instead, we pursue the practical test that the resulting scenarios are rich enough to reflect both the movement of individual interest rates and the movement of various spreads between interest rates. To capture the variation of certain spreads, sixteen scenarios produced from the first four components appear needed, and this is the model finally adopted.

The upper section of Table 3 shows the calculation of the "Shift" factor and the "Twist" factor, the two building blocks for four scenarios. Shift equals the standard deviation of PC1 times the loadings of PC1, and Twist equals the standard deviation of PC2 times the loadings of PC2. Thus the Shift factor moves the thirty-year yield by 17.49 x 0.25 = 4.4 basis points, and the Twist factor moves the thirty-year yield by 6.05 x 0.33 = 2.0 basis points. The lower section of Table 3 shows the resulting scenarios as combinations of Shift and Twist. In the UpUp scenario, both factors operate in the positive direction. The yield of a thirty-year instrument rises by (4.4 + 2.0) x 2.33 = 14.9 basis points, where ±2.33 standard deviations corresponds to the $1^{st}$ and $99^{th}$ percentiles of a normal distribution. Repeating the calculation of scenario UpUp for each point on the curve, all yields rise, though long-term yields rise more than short-term yields. In the UpDn scenario, all yields rise, but short-term yields rise more than long-term yields; the thirty-year yield rises by (4.4 - 2.0) x 2.33 = 5.6 basis points.

The bold lines in Figure 3 show the 1st and 99th percentiles of the data. For example, on only one day in 100 does the three-month rate rise by more than 15.2 basis points. By contrast, scenario UpDn (interest rates generally move higher while the yield curve flattens) moves the three-month rate higher by 16.4 basis points. Thus one of the scenarios approximates the 99th percentile at the three-month point on the curve. The same is true at every point of the curve: the amount of movement represented by the 1st or 99th percentile is approximated by one of the four scenarios. If matching the movement of individual rates on the yield curve were the only test, four scenarios would appear adequate. However, if the portfolio to be analyzed responds to differences between rates, one would prefer also to check the modeling of spreads along the curve.

To investigate the adequacy of two factors to explain spreads along the yield curve, scatterplots appear to provide the most insight. Figure 4 compares the daily changes in the three-month yield and the six-month yield to four points corresponding to the four factor-based scenarios. For example, the scenario DnUp moves the three-month rate down 16.4 basis points and moves the six-month rate down 17.6 basis points; it appears as the lowest of the four scenario points on the graph. A polygon connects the four scenario points in Figure 4.

The immediate impression is that many data points lie outside the polygon. The outliers tend to clump along the broad sides of the polygon rather than at the "points" of the polygon in the first and third quadrants. That is, the scenario points reflect the cases where the three-month and six-month rates rise together or fall together, but fail to adequately reflect cases where the rates move in different directions. While the first two principal components describe the broad movements of the yield curve, they fail to describe movements in the six-month/three-month interest rate spread. This suggests adding a third or fourth factor.

The third factor, "Bow-1," equals PC3 times its standard deviation. Adding that to Shift and Twist results in eight scenarios. Each of the previous four scenarios gives rise to two offspring, one in which Bow-1 rises and one in which Bow-1 falls. Adding "Bow-2" (PC4 times its standard deviation) results in the sixteen scenarios displayed in Figure 5 along with the market data of Figure 4. Eight of the scenario points define the periphery subject to stress, and eight others lie within the area. (The latter are not wasted; they appear on the periphery of scatterplots involving other points on the curve.) The resulting polygon appears to adequately contain the historical data on three-month and six-month rates, and a similar conclusion follows from inspection of each of the other forty-four scatterplots involving pairs of variables.

Figure 6 displays the loadings of the first four principal components. PC1 appears as nearly a uniform shift and PC2 appears as nearly a linear twist. PC3 allows the short end of the yield curve to steepen (flatten) while the long end flattens (steepens). PC4 introduces contrast between the three-month and the six-month, and between the one-year and the two-year points on the curve. As seen in Figures 4 and 5, these contrasts

provide an improvement in explaining some spreads that may have significance for certain portfolios.

Figure 7 shows eight of the sixteen scenarios that stem from the first four factors. These eight contain the Shift factor operating in the positive direction. The other eight scenarios would appear as mirror opposites of these. The message of Figure 7 is simply that the scenarios resulting from the first four factors display a very wide range of yield curve movements. Even if a portfolio contains long and short positions and depends on spreads and subtle interrelations on the yield curve, its risks should be revealed by one or another of these scenarios.

## Non-Option Portfolios

Most of the issues that affect the accuracy of the Factor-Based Scenario Method make themselves felt within simple portfolios that do not contain options. Still, the Factor-Based Method approximates the percentile of loss observed in historical data. The error of approximation tends to be one of overstatement—the conservative result from the perspective of most users of risk control. For example, using 2.33 standard deviations to construct scenarios tends to result in a VAR estimate greater than the $99^{th}$ percentile of the loss distribution.

The basic conservatism of the Factor-Based Method stems from the interaction of three competing influences on the precision of the resulting estimate. The first two of these influences would lead, in isolation, to an understatement of risk and the third to an overstatement:

(1) The factors may have a non-normal distribution with "fat tails." A factor would exceed 2.33 standard deviations in more than 1% of cases. Other things being equal, this influence would cause the Factor-Based estimate of VAR to understate risk.

(2) The portfolio may depend on factors of higher order than those used in constructing the scenarios. Scenarios based only on the first four factors would underestimate the risk of such a portfolio.

(3) The portfolio may depend on two or more of the first four factors.

The last situation occurs in the majority of cases and warrants further exploration. To allow visualization and to simplify the discussion, assume that only two factors matter, Shift and Twist, and that they have a bivariate normal distribution. Figure 8 displays the scatterplot of the factor scores for the 1543 days of data and the four scenario points. If the portfolio in question responds linearly to the two factors, within the region of a rectangle such as that in Figure 8, the portfolio will achieve its maximum loss at a corner. Suppose the greatest loss occurs at point UU. The linearity of the profit function implies the set of points at which loss is equal to that at UU will appear as a straight line through UU.

Consider three cases. In the first, a portfolio depends only on the Shift factor and not on the Twist factor. Its loss will exceed the loss at point UU only if the factor score for Shift exceeds 2.33, that is, only if the data point appears to the right of the vertical line passing through point UU. The probability that this occurs equals 1% under the normality assumption. Therefore the portfolio loss that arises from scenario UU will represent the 99$^{th}$ percentile loss. Second, consider a portfolio that depends only on the Twist factor. Its loss will exceed that at point UU only if the factor score for Twist exceeds 2.33 and the point appears above the horizontal line through UU. Again, the probability equals 1% and the portfolio value at point UU will again represent the 99$^{th}$ percentile loss. Finally, consider a portfolio that depends on both Shift and Twist. It will have a loss greater than at point UU only if the combination of Shift and Twist results in a point located above a line such as AB in Figure 8. A slanted line such as AB is farther, on average, from the center of the distribution than either the horizontal or vertical line through UU and isolates less of the probability space. The probability that a point will be above AB is less than 1%, so the value at UU represents a percentile of the loss distribution above the 99$^{th}$ percentile. The Factor-Based Method therefore tends to overstate risk for portfolios that depend on more than one factor.

Figure 9 examines the practical importance of these three sources of estimation error. Each point in Figure 9 represents a randomly generated portfolio of fixed income instruments.[4] Each random portfolio was subjected to two experiments. First, the daily P&L was simulated for each of the 1543 days of data. The 99$^{th}$ percentile of the distribution of these losses establishes the horizontal position of the point. Second, the VAR of the portfolio was estimated by the Factor-Based Scenario Method. That is, the portfolio was valued under each of the sixteen scenarios (eight of which appear in Figure 7), and the greatest loss among these establishes the vertical position of the point. When a point lies on the 45° line, the Factor-Based Scenario Method has succeeded in estimating the 99$^{th}$ percentile of the loss distribution for that portfolio.

Figure 9 makes several points. First, understatements of risk are few and small. This serves the risk control function by protecting against unpleasant surprises. Qualitatively, the conservatism introduced by influence (3), above, tends to dominate the influences of (1) and (2). Second, the Factor-Based Scenario Method tends to greater accuracy (as a fraction of the 99$^{th}$ percentile historical risk) for the portfolios that have greatest risk. This partly reflects the fact that the most risky portfolio will tend to respond to the most volatile factor rather than to a combination of several factors. Third, the average overstatement of risk (20.2%) and the variation in its overstatement (standard deviation = 15.7%) are within a range that senior managers (if not traders) may find acceptable. When options enter the picture they tend to absorb some, but generally not all, of this inherent conservatism.

## Portfolios Containing Options

Introducing options adds two elements to the analysis. First, option values depend on an additional parameter, volatility, and changes in volatility can affect portfolio value just as

much as changes in interest rates. For the Factor-Based Method, this adds one or more dimensions to those already identified. The resulting scenarios combine an interest rate shift and a volatility shift.

Second, options on fixed-income instruments have a much more nonlinear response to interest rates than swaps, bonds, or futures. The nonlinear response of options presents two challenges to the accuracy of the Factor-Based Scenario Method. Both challenges were met in the one-market environment. The most important challenge is the potential that large losses might occur between scenario points. Through a variety of means the user can gain confidence that this is not the case for the portfolio at hand. Take as a representative portfolio the options held by an institutional participant in the fixed-income option market.[5] The portfolio reflects various market-making, hedging, and position-taking activities and contains options on thirteen contracts of Treasury Bond futures, Treasury Note futures, and Five-Year Treasury futures. Just over $100 billion of securities underlay the options of this portfolio. To focus on the issue of nonlinearity, a hedging set of T-Bond, T-Note, and Five-Year futures was included to minimize the maximum loss under any scenario.[6]

Figure 10 shows the portfolio profit and loss response to independent movements of each of the four factors. The line labeled "Shift" shows the outcome when the Shift factor score takes on various values (measured in standard deviations), while the other factors remain fixed. The portfolio is evidently "long Shift" in that it tends to make money when Shift has a small positive score. However, a large absolute Shift score, either positive or negative, results in a loss. One can see that the portfolio is more sensitive to Shift than any other factor and least sensitive to Bow-2. Most important, the response to each factor has concave curvature. On each plotted line, just as for a short option straddle, the maximum loss in any region occurs at one of the endpoints of the region.

Figure 11 shows the portfolio's concave profit surface as a function of the first two factors. The four corners of the floor of the diagram correspond to the four corners of the rectangle in Figure 8. The concavity of the surface implies that the greatest loss on a rectangular region occurs at one of the corners, specifically the corner where both Shift and Twist decline. This corresponds to scenarios beginning "DD." The portfolio has the character of a short straddle and can also experience large losses in the diagonally opposite case of scenarios beginning "UU."

The second challenge, implicit in the concavity of the profit surface, is that multiple areas of factor space may contribute to the probability that the loss will exceed the VAR estimate. This is analogous to the situation of Figure 1, in which the portfolio suffered a loss exceeding VAR both when the price rose 2.33 standard deviations and when the price fell 6.4 standard deviations. In the four-factor case, Figure 10 suggests the position could lose more than VAR given a large-enough movement in any of the four factors. This might seem to imply the potential for a large understatement of risk, but such is not the case. Consider the story told by Figure 12 for a hypothetical portfolio in the two-factor case. Here, all four scenarios produce the same loss, as do all the points along the circular level curve. Loss exceeds VAR at any point outside the circle. For normally

distributed factors, that probability is 0.5%. Intuitively, the curvature of the level curve causes it to miss the high-density areas between it and the square connecting the scenario points. It seems that the profit function cannot conform closely enough to the square to result in an understatement of risk, despite the chance for extreme movement of either factor to trigger a loss greater than VAR.

The level curve of a portfolio with a concave profit function might look also like that in Figure 13. Clearly, the probability outside such a curve exceeds the probability above the tangent line AB. That tangent line is the level curve of a linearly responding portfolio. Though the VAR estimate for linear portfolios was found to be conservative in Figures 8 and 9, introducing options might reduce confidence in the estimate.

A historical simulation of the representative portfolio provides evidence that the Factor-Based Scenario Method estimate remains conservative, nonetheless. Figure 14 shows the distribution of profit and loss for 300 out-of-sample days of interest rate history, the 99$^{th}$ percentile loss, and the Factor-Based Scenario VAR estimate. The distribution resembles that for a short options position—many small profits and some large losses. The VAR estimate appears conservative relative to this test.

The Factor-Based Scenario Method can provide many benefits in estimating value-at-risk: it produces a quick, easily interpreted result that is valid for option-intensive portfolios with negative gamma. However, the Factor-Based Method cannot guarantee accuracy for any type of portfolio sensitive to any set of risk factors. Generally the method works best when the portfolio appears to depend on a small number of factors. That would be the case for portfolios of fixed income instruments within a particular country or for spot foreign exchange risk among major countries. At the other extreme, a portfolio of options on individual stocks might be a poor candidate for factor-based VAR estimation if, for example, the spread between two gold stocks has a strong effect on profit; such a spread relationship might require many principal components to model adequately. Note that both the nature of the position and the statistical properties of market variables must be taken into any assessment of appropriateness.

## Summary and Conclusion

For portfolios of fixed income instruments dominated by short options positions, the Factor-Based Scenario Method tends to produce conservative estimates of value-at-risk. The method chooses a set of interest rate scenarios and values the portfolio under each. The greatest loss that comes about under any scenario serves as the VAR estimate. The scenario that produces the greatest loss characterizes the market movement most dangerous for the portfolio.

Each scenario combines factors, referred to here as Shift, Twist, Bow-1, and Bow-2. Shift appears as a nearly uniform increase (or decrease) in yields at all points of the yield curve. Twist appears as a pivoting of the yield curve where long-term rates rise (or fall) relative to short-term rates. Bow-1 and Bow-2 induce more complex movements. In each scenario, each factor operates in either the positive direction or the negative direction.

The statistical technique of principal components analysis determines the factors. Principal components analysis can be viewed as equivalent to a regression study that finds optimal independent variables. The first principal component represents the most typical movement of the yield curve. Multiplying by its standard deviation produces the Shift factor. The other factors have similar definitions.

Basing the scenarios on 2.33 standard deviations (corresponding to 99% confidence for a normal distribution), the risk estimate tends to overstate the $99^{th}$ percentile of the loss distribution. That is because the Factor-Based Scenario Method tends to overstate risk for portfolios that depend on more than one factor—and most portfolios depend on multiple factors.

---

## Appendix 1: Proof of a Duality Between PCA and Regression Analysis

Given: an N*p matrix of mean-adjusted (the column means equal 0) data, Y
Define: an arbitrary, normal p*1 vector x (that is, x'x = 1)
Then: The cumulated sum of squares from N independent regressions of the rows of Y on x reaches its minimum when x equals the first principal component of the data Y.

Proof: Let $Y_i'$ symbolize the $i^{th}$ row of the matrix Y. The $i^{th}$ regression takes the form

$$Y_i = \hat{a}\, x + e_i$$

where â is the regression coefficient and $e_i$ is a p*1 vector of residuals. Minimizing $e_i'e_i$ subject to x'x = 1 implies

$$\hat{a} = Y_i'\, x$$

A row vector of regression residuals is then $Y_i' - Y_i'xx'$ and the Nxp matrix of residuals from all regressions is Y-Yxx'. The cumulated sums of squares from all regressions equals

$$\begin{aligned} \text{CumSSE} &= \text{trace}( (Y-Yxx') (Y-Yxx')' ) \\ &= \text{trace}(YY' - Yxx'Y') \\ &= \text{trace}(YY') - \text{trace}(x'Y'Yx) \\ &= \text{trace}(YY') - N\, x'Sx \end{aligned}$$

where S = Y'Y/N is the variance-covariance matrix of Y. By definition, among all normal vectors x, the first principal component of Y maximizes x'Sx . Therefore CumSSE reaches its minimum when x is chosen as the first principal component.

----------End of Appendix 1----------

---

## Appendix 2: Mathematical Definition of Principal Components

Most multivariate statistics texts provide an introduction to principal components analysis. A background in statistics and matrix algebra is helpful in exploring this topic.

Let Y be an N*p matrix of data.

The first principal component of Y is the p*1 vector $x_1$ that maximizes the variance of $x_1'y$ subject to $x_1'x_1 = 1$.

The second principal component of Y is the p*1 vector $x_2$ that maximizes the variance of $x_2'y$ subject to $x_2'x_2 = 1$ and $x_2'x_1 = 0$.

The other principal components are defined similarly. It can be shown that the principal components equal the eigenvectors of S, the variance-covariance matrix of Y. Therefore,

$$S = PD^2P'$$

where P is a matrix having the eigenvectors of S as its columns and $D^2$ is a diagonal matrix of the eigenvalues of S in decreasing order from left to right. In the terminology of this article, the columns of P are the principal components loadings and the columns of PD are factors Shift, Twist, etc.

----------End of Appendix 2----------

Sources

Beckstrom, Rod, and Alyce Campbell, eds. *An Introduction to VAR.* C*ATS Software, 1995.

Hendricks, Darryll. "Evaluation of Value-at-Risk Models Using Historical Data." *Federal Reserve Bank of New York Economic Policy Review* 2, no. 1 (1996): 39-69.

Jordan, James V. and Robert J. Mackay. "Assessing Value At Risk for Equity Portfolios: Implementing Alternative Techniques." Working Paper, Center for Study of Futures and Options Markets, Virginia Polytechnic Institute, 1995.

Litterman, Robert and José Scheinkman. "Common Factors Affecting Bond Returns." *Journal of Fixed Income* 1, no. 1 (1991): 54-61.

Morgan Guaranty Trust Company Global Research. *RiskMetrics ™— Technical Document.* New York: Morgan Guaranty Trust Company, 1994.

Morrison, Donald F. *Multivariate Statistical Methods.* New York: McGraw Hill, 1976.

Pritsker, Matt. "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time." Working Paper, Board of Governors of the Federal Reserve System, 1996.

**Table 1.**
**Sample of Mean-Adjusted Changes in Interest Rates in Basis Points**

| | 3m | 6m | 1y | 2y | 3y | 4y | 5y | 7y | 10y | 30y |
|---|---|---|---|---|---|---|---|---|---|---|
| 1/4/89 | 6.0 | 3.5 | 4.7 | 2.8 | 0.7 | 0.6 | 0.1 | -0.2 | -0.9 | -1.0 |
| 1/5/89 | 19.8 | 7.4 | 8.1 | 7.3 | 8.6 | 6.8 | 6.4 | 6.6 | 5.6 | 2.3 |
| 1/6/89 | -10.4 | 0.3 | 5.7 | 1.0 | 0.9 | 0.5 | -0.3 | -2.1 | -2.1 | -4.1 |
| 1/9/89 | 0.2 | 2.3 | -0.4 | -5.1 | -1.8 | -1.3 | -1.1 | 0.1 | -0.7 | -0.5 |
| 1/10/89 | -8.2 | -5.9 | -4.4 | -1.6 | -1.1 | -0.9 | 0.2 | -0.2 | -0.4 | -0.3 |
| 1/11/89 | 7.6 | 3.5 | 1.3 | 2.9 | -1.8 | -0.4 | -0.7 | -0.2 | 0.1 | 0.4 |
| 1/12/89 | -1.4 | -7.0 | -13.7 | -7.8 | -7.6 | -8.5 | -8.6 | -10.9 | -9.3 | -8.1 |
| 1/13/89 | -6.7 | -8.6 | -9.0 | -7.0 | -8.1 | -9.6 | -9.7 | -10.3 | -8.2 | -8.6 |
| 1/17/89 | 4.4 | 5.7 | 3.0 | 2.8 | 2.8 | 1.6 | 0.9 | 0.1 | 0.3 | 0.7 |
| 1/18/89 | -13.5 | -5.3 | -7.3 | -4.3 | -3.8 | -4.8 | -5.0 | -5.7 | -6.5 | -5.2 |

**Table 2. Selected Results from Principal Components Analysis**

**Principal Component Loadings**

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 3m | 0.21 | -0.57 | 0.50 | 0.47 | -0.39 | -0.02 | 0.01 | 0.00 | 0.01 | 0.00 |
| 6m | 0.26 | -0.49 | 0.23 | -0.37 | 0.70 | 0.01 | -0.04 | -0.02 | -0.01 | 0.00 |
| 12m | 0.32 | -0.32 | -0.37 | -0.58 | -0.52 | -0.23 | -0.04 | -0.05 | 0.00 | 0.01 |
| 2y | 0.35 | -0.10 | -0.38 | 0.17 | 0.04 | 0.59 | 0.56 | 0.12 | -0.12 | -0.05 |
| 3y | 0.36 | 0.02 | -0.30 | 0.27 | 0.07 | 0.24 | -0.79 | 0.00 | -0.09 | -0.00 |
| 4y | 0.36 | 0.14 | -0.12 | 0.25 | 0.16 | -0.63 | 0.15 | 0.55 | -0.14 | -0.08 |
| 5y | 0.36 | 0.17 | -0.04 | 0.14 | 0.08 | -0.10 | 0.09 | -0.26 | 0.71 | 0.48 |
| 7y | 0.34 | 0.27 | 0.15 | 0.01 | 0.00 | -0.12 | 0.13 | -0.54 | 0.00 | -0.68 |
| 10y | 0.31 | 0.30 | 0.28 | -0.10 | -0.06 | 0.01 | 0.03 | -0.23 | -0.63 | 0.52 |
| 30y | 0.25 | 0.33 | 0.46 | -0.34 | -0.18 | 0.33 | -0.09 | 0.52 | 0.26 | -0.13 |

**Sample of Principal Component Scores**

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1/4/89 | 4.6 | -7.4 | 0.0 | 0.1 | -1.9 | -0.1 | 0.8 | 0.1 | 0.0 | -0.3 |
| 1/5/89 | 23.6 | -11.8 | 5.9 | 6.9 | -5.0 | -0.8 | -0.9 | -1.4 | -0.8 | -0.3 |
| 1/6/89 | -1.9 | 1.5 | -10.7 | -6.2 | 2.4 | -1.8 | -0.5 | -0.3 | -0.3 | 0.7 |
| 1/9/89 | -3.1 | -1.3 | 3.1 | -2.1 | 1.3 | -2.7 | -1.8 | -1.2 | 0.5 | -0.4 |
| 1/10/89 | -6.1 | 8.8 | -3.1 | 0.3 | 1.2 | 0.4 | 0.1 | -0.3 | 0.7 | 0.3 |
| 1/11/89 | 3.0 | -6.8 | 3.8 | 1.2 | -1.4 | 1.3 | 2.7 | 0.4 | -0.5 | -0.2 |
| 1/12/89 | -26.7 | -1.7 | 1.3 | 6.8 | 1.9 | 1.5 | -0.4 | 1.4 | 0.4 | 1.0 |
| 1/13/89 | -27.0 | 0.5 | -3.2 | 1.8 | 0.2 | 1.4 | 0.1 | 0.1 | -1.1 | 0.9 |
| 1/17/89 | 6.6 | -5.7 | 0.7 | -0.3 | 1.2 | 0.7 | -0.8 | 0.9 | -0.1 | 0.0 |
| 1/18/89 | -18.1 | 6.4 | -6.8 | -1.5 | 5.1 | 0.9 | -0.6 | 0.6 | 0.6 | -0.4 |

**Principal Component Standard Deviations**

| PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 |
|---|---|---|---|---|---|---|---|---|---|
| 17.49 | 6.05 | 3.10 | 2.17 | 1.97 | 1.69 | 1.27 | 1.24 | 0.80 | 0.79 |

**Table 3. Construction of Two Factors and Four Scenarios**

**Yield Curve Factors in Basis Points**

| | 3mo | 6mo | 12mo | 2yr | 3yr | 4yr | 5yr | 7yr | 10yr | 30yr |
|---|---|---|---|---|---|---|---|---|---|---|
| Shift | 3.6 | 4.6 | 5.6 | 6.2 | 6.3 | 6.3 | 6.2 | 5.9 | 5.4 | 4.4 |
| Twist | -3.5 | -3.0 | -1.9 | -0.6 | 0.1 | 0.8 | 1.0 | 1.6 | 1.8 | 2.0 |

**Yield Curve Scenarios (2.33 SDs) in Basis Points**

| | 3mo | 6mo | 12mo | 2yr | 3yr | 4yr | 5yr | 7yr | 10yr | 30yr |
|---|---|---|---|---|---|---|---|---|---|---|
| UpUp | 0.3 | 3.7 | 8.6 | 12.9 | 15.0 | 16.7 | 16.8 | 17.5 | 17.0 | 14.9 |
| UpDn | 16.4 | 17.6 | 17.5 | 15.8 | 14.5 | 12.9 | 12.2 | 9.9 | 8.4 | 5.6 |
| DnUp | -16.4 | -17.6 | -17.5 | -15.8 | -14.5 | -12.9 | -12.2 | -9.9 | -8.4 | -5.6 |
| DnDn | -0.3 | -3.7 | -8.6 | -12.9 | -15.0 | -16.7 | -16.8 | -17.5 | -17.0 | -14.9 |

**Figure 1. When only one market affects the portfolio, the Factor-Based Scenario Method simply shocks that market up and down by a fixed number of standard deviations. In this case the portfolio is a delta-hedged short position in a short-term, out-of-the-money call option on Treasury Bond futures. Such a position is unaffected by a very small movement in the underlying market but tends to lose money in large moves up or down. In particular, the portfolio loses money under both the Up scenario (underlying price up 2.33 standard deviations) and the Down scenario (underlying price down 2.33 standard deviations). The greatest loss among the scenarios, $1.45 million, provides the VAR estimate. This estimate corresponds to the 98.999999994 percentile because of the small chance that price might fall enough to produce a $1.45 million loss. (This ignores the existence of daily limits on futures price movements.)**

**Figure 2. A portfolio that displays an interior minimum is a poor candidate for using the Factor-Based Scenario Method. In this case a loss is possible despite the fact that both scenarios result in profit. This portfolio is long options, while the portfolio of Figure 1 is short options.**

**Figure 3. The two bold lines correspond to the 1st and 99th percentiles of the historical sample. For example, the 99th percentile of the three-month rate equals 15.2 basis points. The four lighter-weight lines denote the four scenarios based on the factors Shift and Twist. The factor-based scenarios move the curve by an amount approximately equal to the 99th (or 1st) percentile. If this were the only test of adequacy—that a set of scenarios reflects broad overall movements in interest rates—the two-factor, four-scenario model would appear to pass.**

**Figure 4. Historical changes in three-month and six-month yields appear as "+" signs. The four factor-based scenarios are connected by a polygon. Many data points lie outside the polygon, leading to the likelihood that a position sensitive to the six-month/three-month spread may frequently lose more than the worst case among the four scenarios. Evidently, more factors are required to model this spread relationship.**

**Figure 5. The data of Figure 4 with the sixteen scenarios stemming from four factors. The outliers are few and scatter fairly evenly around the periphery of the polygon connecting the scenario points. The coverage in the other 44 scatterplots is equivalent or better.**

**Figure 6. The loadings of the first four principal components have natural interpretations as features of daily changes on the yield curve. The sum of the squares of the loadings of each principal component equals one. The scale of any one loading therefore lacks direct interpretation.**

**Figure 7. Eight of the sixteen scenarios that stem from the first four factors. In these eight, Shift operates in the positive direction. The other eight scenarios equal the negatives of these. This illustrates the richness of the interest rate scenarios produced by the Factor-Based Method.**

**Figure 8. A scatterplot of factor scores for Shift and Twist in standard deviations. Assume that the portfolio responds linearly to these factors, and that the scenario point with greatest loss is point UU. (1) For a portfolio that depends only on Shift, loss will exceed that at UU only for points to the right of the vertical line through UU. This occurs 1% of the time (assuming normality) so the Factor-Based Scenario Method accurately estimates the 99th percentile of the loss distribution. (2) For a portfolio that depends only on Twist, loss will exceed that at UU only for points above the horizontal line through UU. Again the method is accurate. (3) For a portfolio that depends on both Shift and Twist, loss will exceed that at UU only for points above a slanted line such as AB. Fewer points lie above AB than to the right of UU or above UU. Therefore the loss at UU overstates the 99th percentile loss for a portfolio such as (3).**

**Figure 9. This illustrates the inherent conservatism of the Factor-Based Scenario Method. The scatterplot displays results for 1000 randomly generated portfolios. The horizontal axis has the 99th percentile of the loss distribution from the historical data for a given portfolio. The vertical axis has the Factor-Based VAR estimate. The 45° line represents the exact 99th percentile. Points above (below) the 45° line represent overstatement (understatement) of risk. Overstatement of risk is more common, and tends to be of greater magnitude, than understatement, which occurs in about 9% of the 1000 portfolios. The average overstatement equals 20.2% and declines (as a fraction of the 99th percentile historical loss) for portfolios containing greater risk.**

**Figure 10. Each factor score is allowed to have an independent effect on the representative portfolio. Concavity is most apparent for the Shift factor but is present for all factors. As any factor is adjusted through a range of values, the greatest loss occurs at one of the endpoints of the range.**

**Figure 11. The representative portfolio resembles a short straddle. It loses money when factors move strongly either higher or lower, though the worst case is when both factors rise or both factors fall. Such a portfolio has its greatest loss at a corner, not at an interior point.**

**Figure 12. Data points illustrate the probability density of the factor scores and a level curve represents the profit of a hypothetical portfolio. This portfolio experiences the same loss at every scenario point and at every point along the circular line connecting the scenario points. Loss at any point outside the circle will be greater than at the scenario points. Inspection shows few such points; for normally distributed factors the probability outside the circle is 0.5%. Despite the potential for the movement of either factor to cause loss exceeding VAR, the estimate remains conservative.**

**Figure 13. Two level curves: AB represents a linearly responding portfolio such as that of Figure 8 and the oval represents an option portfolio in which a large score for either factor can produce a loss. Clearly, more of the probability density lies outside the oval than above the line AB. This means an options portfolio tends to absorb some of the inherent conservatism of the Factor-Based Scenario Method.**

**Figure 14. The simulated P&L distribution of the representative portfolio over 300 days of history. The distribution resembles that for a short option position, with many small profits and a few large losses. The Factor-Based Scenario Method produces a VAR estimate slightly greater than the 99th percentile loss.**

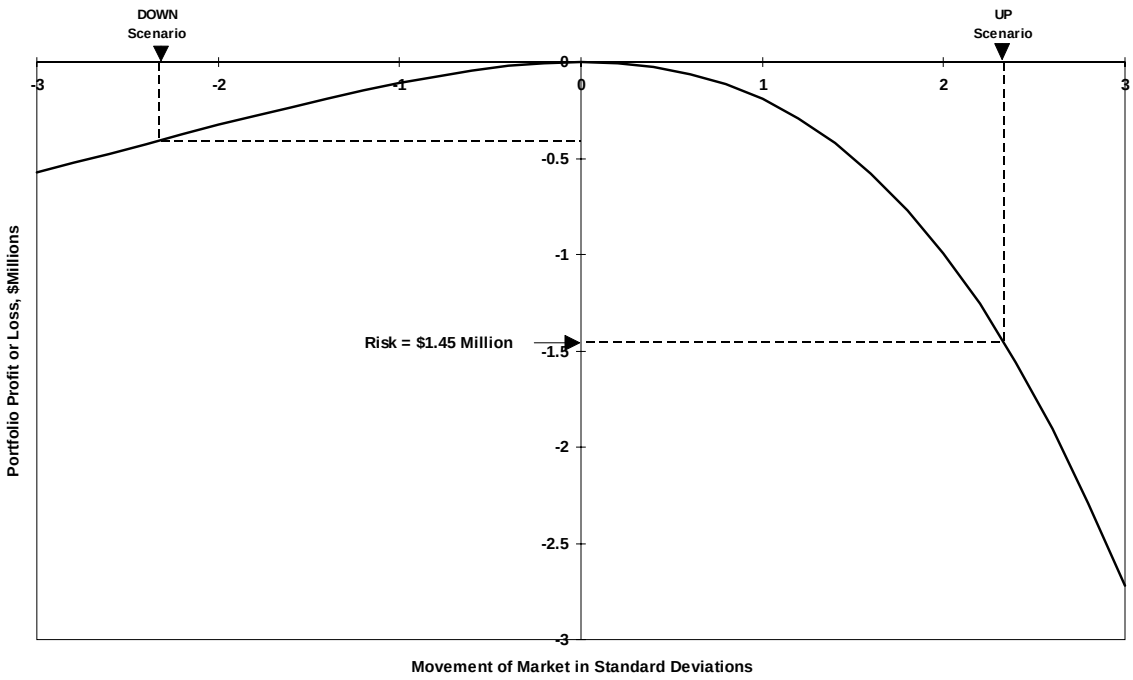


**Figure 15. When only one market affects the portfolio, the scenario method simply shocks that market up and down by fixed numbers of standard deviations. In this case the portfolio is a delta-hedged short position in a short-term, out-of-the-money call option on Treasury Bond futures. Such a position is unaffected by a very small movement in the underlying market but tends to lose money in large moves up or down. In particular, the portfolio loses money under both the Up scenario (underlying price up 2.33 standard deviations) and the Down scenario (underlying price down 2.33 standard deviations). The greatest loss among the scenarios, $1.45 million, provides the VAR estimate. This estimate corresponds to the 98.999999994 percentile because of the small chance that price might fall enough to produce a $1.45 million loss. (This ignores the existence of daily limits on futures price movements.)**

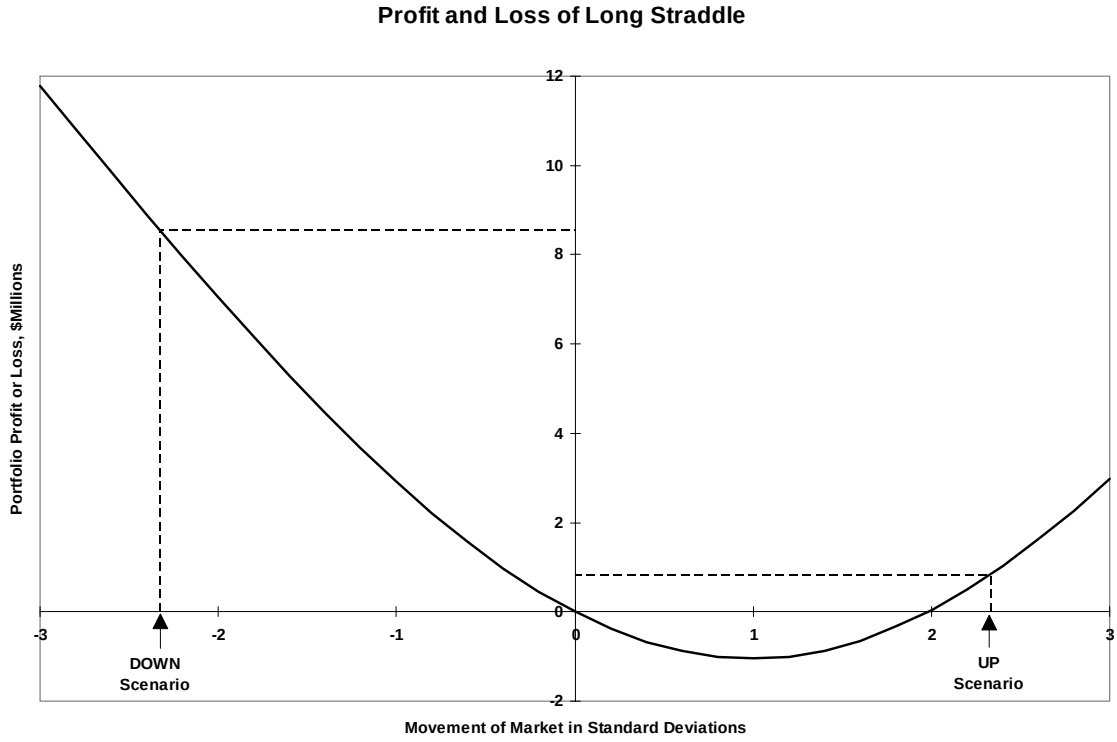


**Figure 16. A portfolio that displays an interior minimum is a poor candidate for using the Factor-Based Scenario Method. In this case a loss is possible despite the fact that both scenarios result in profit. This portfolio is long options, while the portfolio of Figure 1 is short options.**

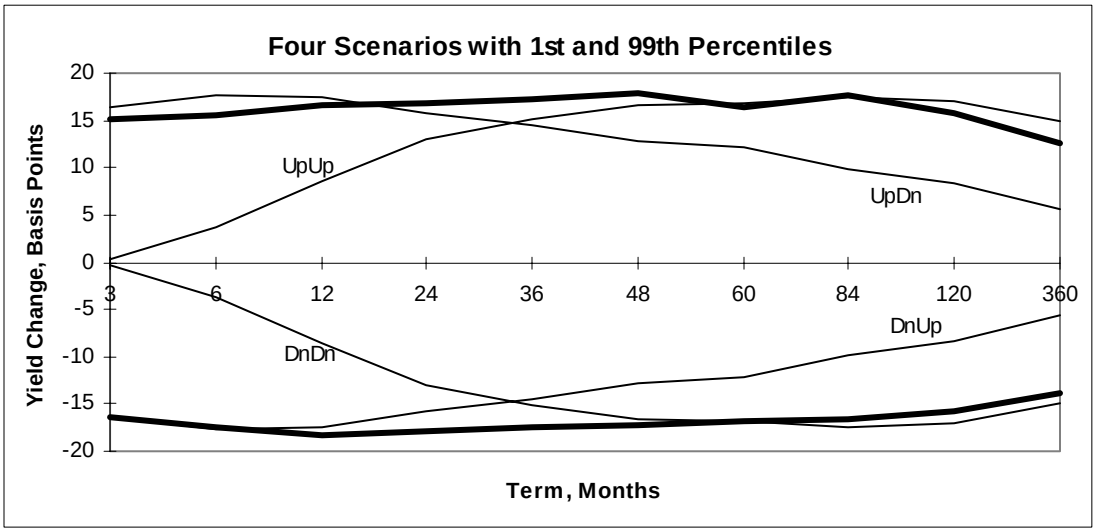


**Figure 17. The two bold lines correspond to the 1st and 99th percentiles of the historical sample. For example, the 99th percentile of the three-month rate equals 15.2 basis points. The four lighter-weight lines denote the four scenarios based on the factors Shift and Twist. The factor-based scenarios move the curve by an amount approximately equal to the 99th (or 1st ) percentile. If this were the only test of adequacy—that a set of scenarios reflect overall movements in interest rates—the two-factor, four-scenario model would pass.**

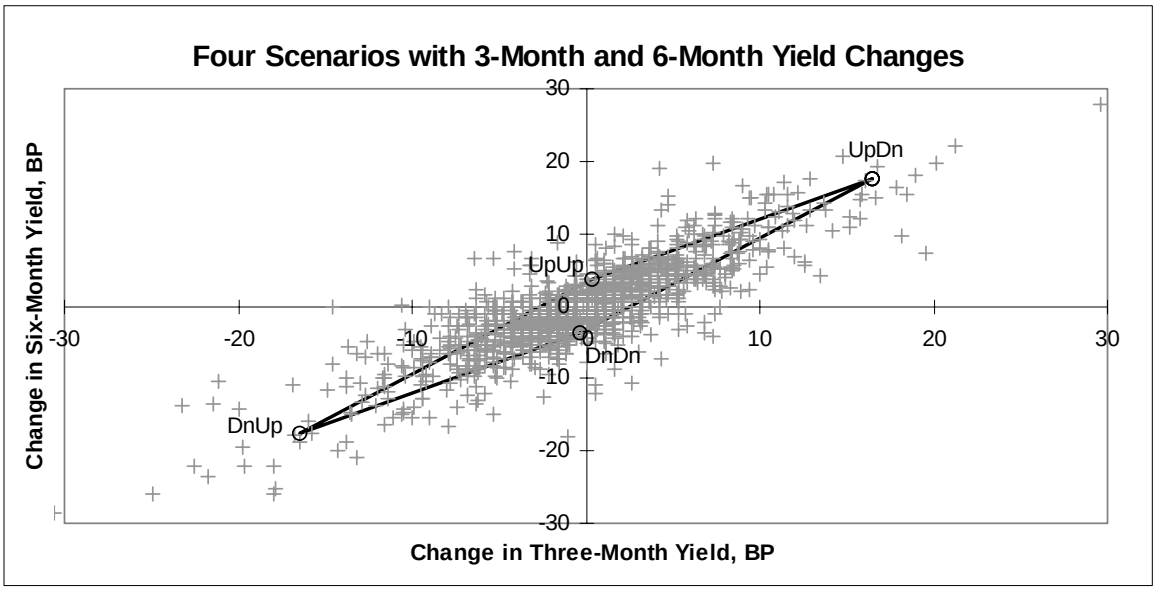


**Figure 18. Historical changes in three-month and six-month yields appear as gray "+" signs. The four factor-based scenarios appear as open circles with a polygon connecting them. Many data points lie outside the polygon, leading to the likelihood that a position sensitive to the six-month/three-month spread may frequently lose more than the worst case among the four scenarios. Evidently, more factors are required to model this spread relationship.**

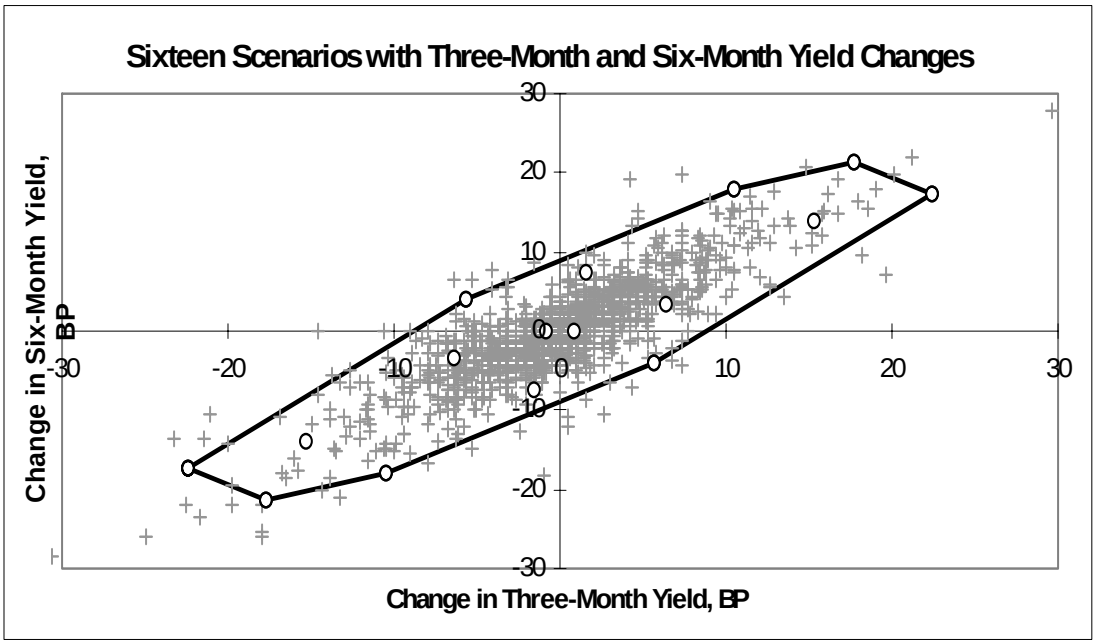


**Figure 19. The data of Figure 4 with the sixteen scenarios stemming from four factors. The outliers are few and scatter fairly evenly around the periphery of the polygon connecting the scenario points. The coverage in the other 44 scatterplots is equivalent or better.**

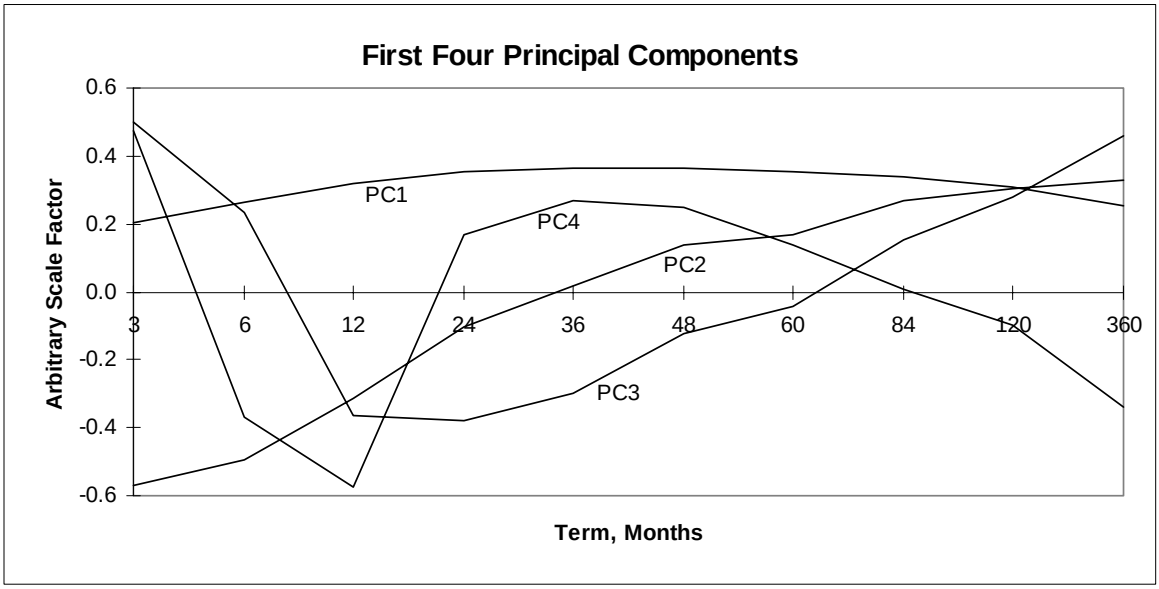


**Figure 20. The loadings of the first four principal components have natural interpretations as features of daily changes on the yield curve. The sum of the squares of the loadings of each principal component equals one. The scale of any one loading therefore lacks direct interpretation.**

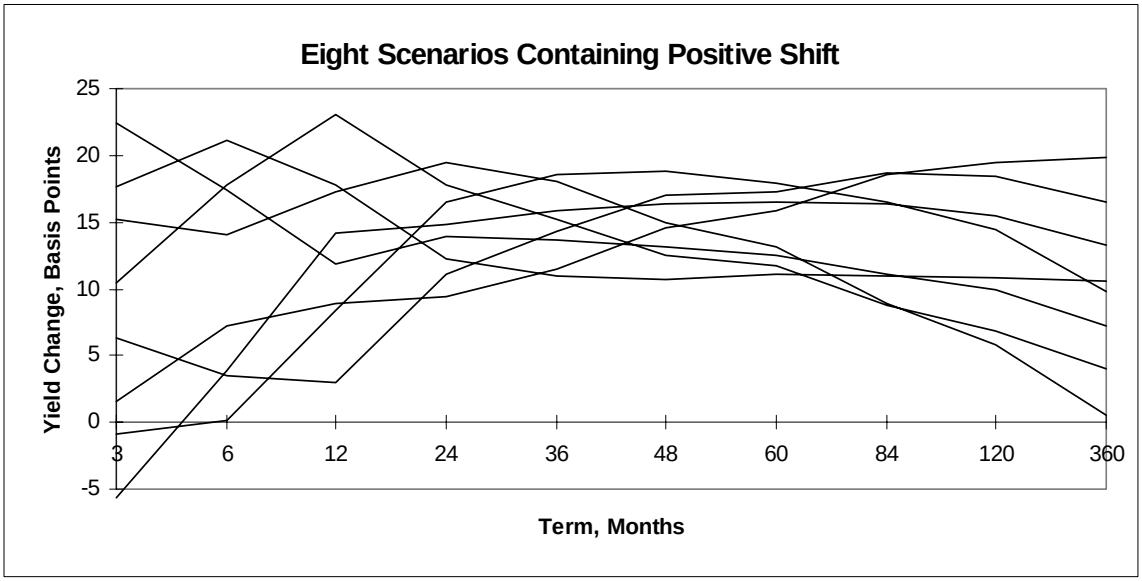


**Figure 21. Eight of the sixteen scenarios that stem from the first four factors. In these eight, Shift operates in the positive direction. The other eight equal the negatives of these. For each maturity, at least one scenario moves the associated rate higher than its 99$^{th}$ percentile, and in scatter plots of any pair of tenors, few data points lie outside the polygon connecting the scenario points.**

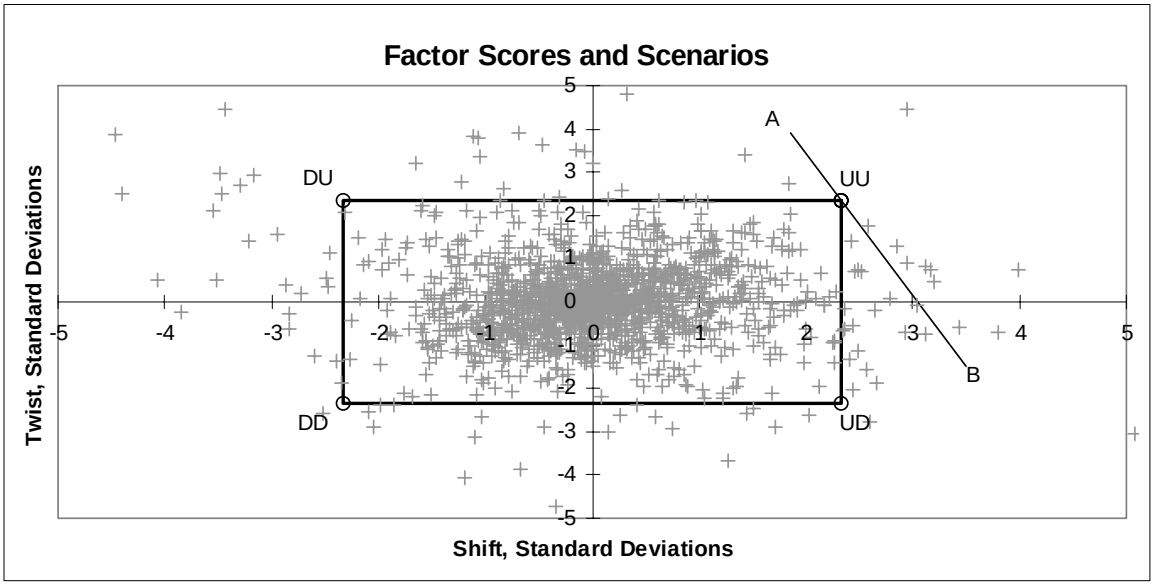


**Figure 22. A scatterplot of factor scores for Shift and Twist in standard deviations. Assume that bivariate normal Shift and Twist fully explain movement in the yield curve, that the portfolio responds linearly to these factors, and that the scenario point with greatest loss is point UU. (1) For a portfolio that depends only on Shift, loss will exceed that at UU only for points to the right of the vertical line through UU. This occurs 1% of the time so the Factor-Based Scenario Method accurately estimates the 99$^{th}$ percentile of the loss distribution. (2) For a portfolio that depends only on Twist, loss will exceed that at UU only for points above the horizontal line through UU. Again the method is accurate. (3) For a portfolio that depends on both Shift and Twist, loss will exceed that at UU only for points above a slanted line such as AB. The nature of the bivariate normal distribution implies, and inspection of Figure 8 suggests, that fewer points lie above AB than to the right of UU or above UU. Therefore the loss at UU overstates the 99$^{th}$ percentile loss for a portfolio such as (3).**

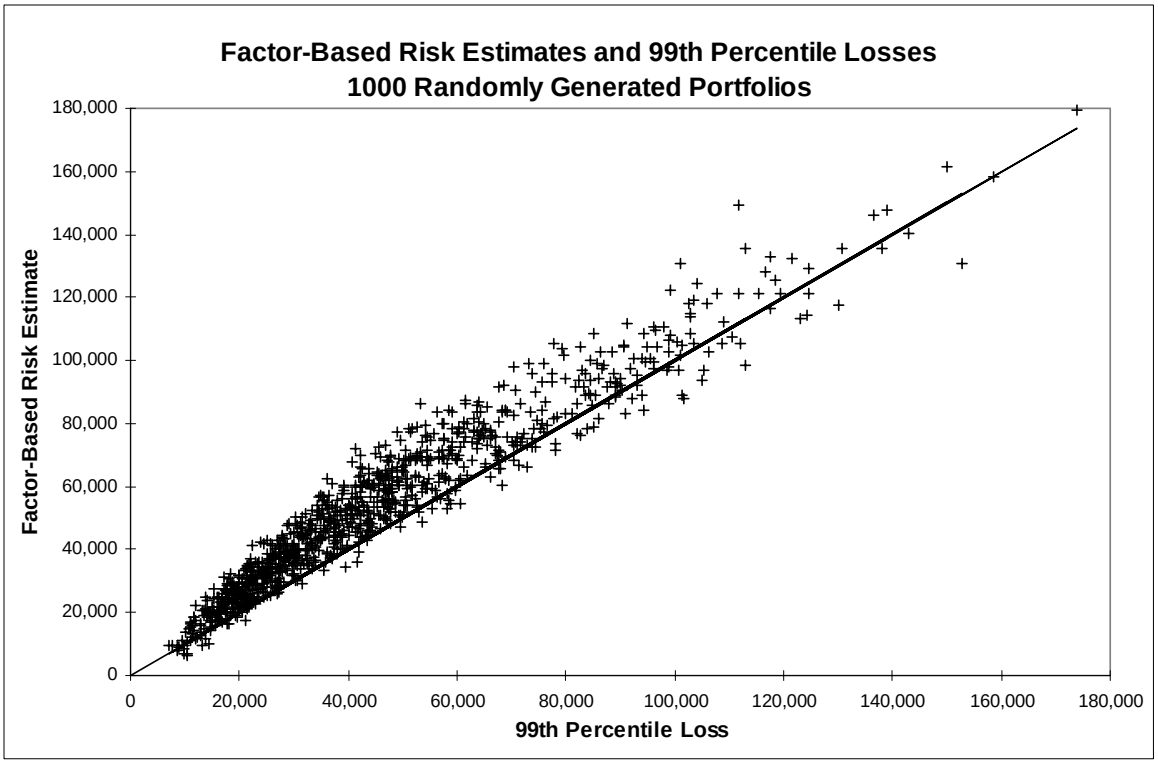


**Figure 23. An illustration of the inherent conservatism of the Factor-Based Scenario Method. The scatterplot displays results for 1000 randomly-generated portfolios. The horizontal axis has the 99$^{th}$ percentile of the loss distribution from the historical data for a given portfolio. The vertical axis has the Factor-Based VAR estimate. The 45° line represents the exact 99$^{th}$ percentile. Points above (below) the 45° line represent overstatement (understatement) of risk. Overstatement of risk is more common, and tends to be of greater magnitude, than understatement, which occurs in about 9% of the 1000 portfolios. The average overstatement equals 20.2% and appears to decline for portfolios containing greater risk.**

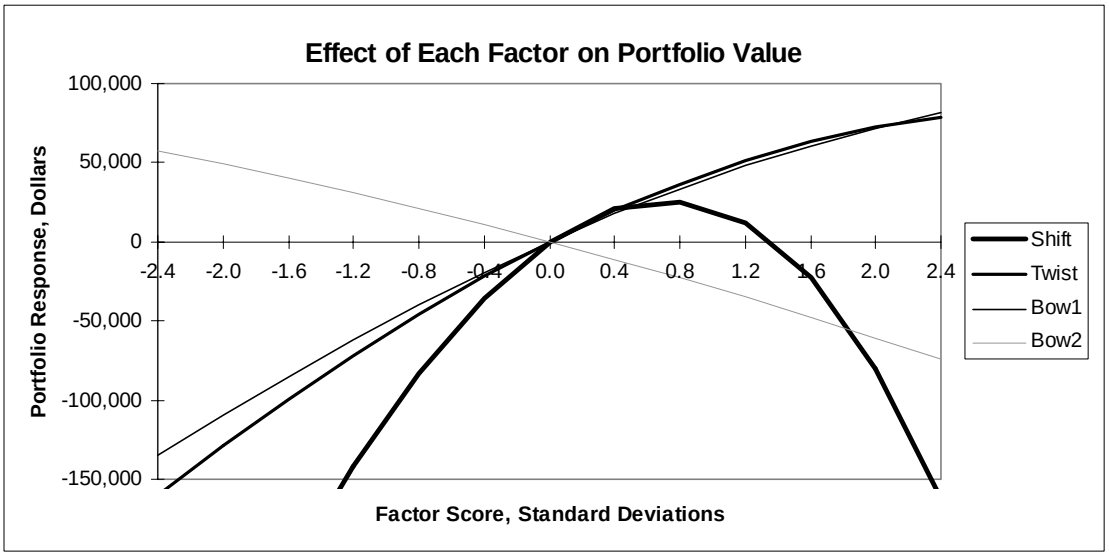


**Figure 24. Each factor score is allowed to independently affect the representative portfolio of an institutional participant in the fixed income options market. Convexity is most apparent for the Shift factor but present for all factors. As any factor is adjusted through a range of values, the greatest loss occurs at one of the endpoints of the range.**

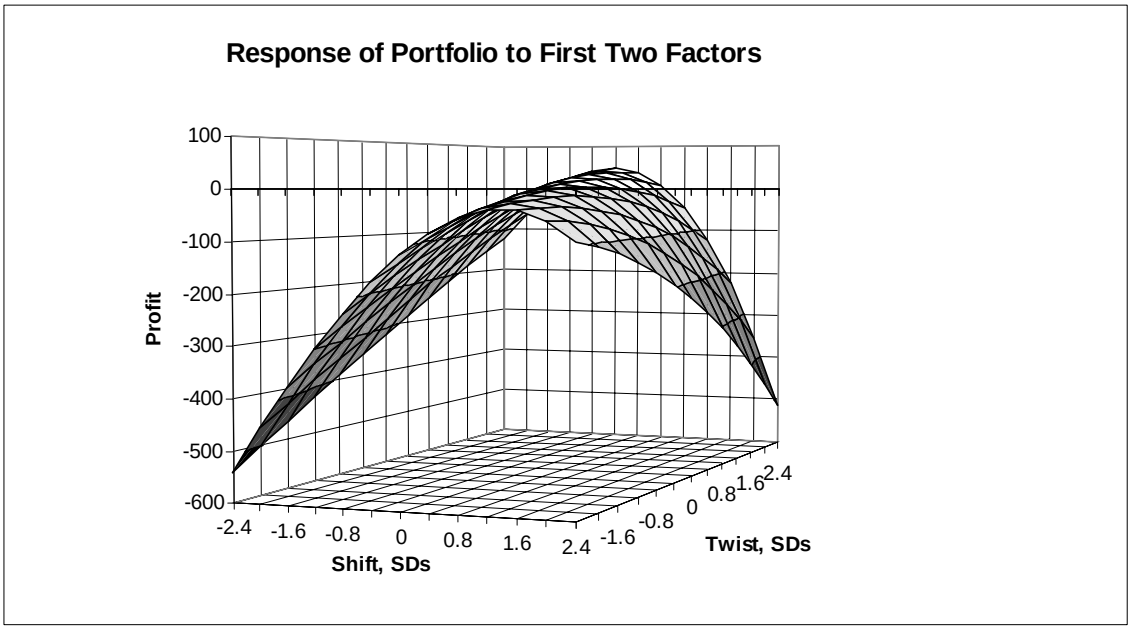


**Figure 25. The representative portfolio resembles a short straddle. It makes money when factor scores are low and loses money especially when the first two factors both rise or both fall. It has its greatest loss at a corner, not an interior point.**

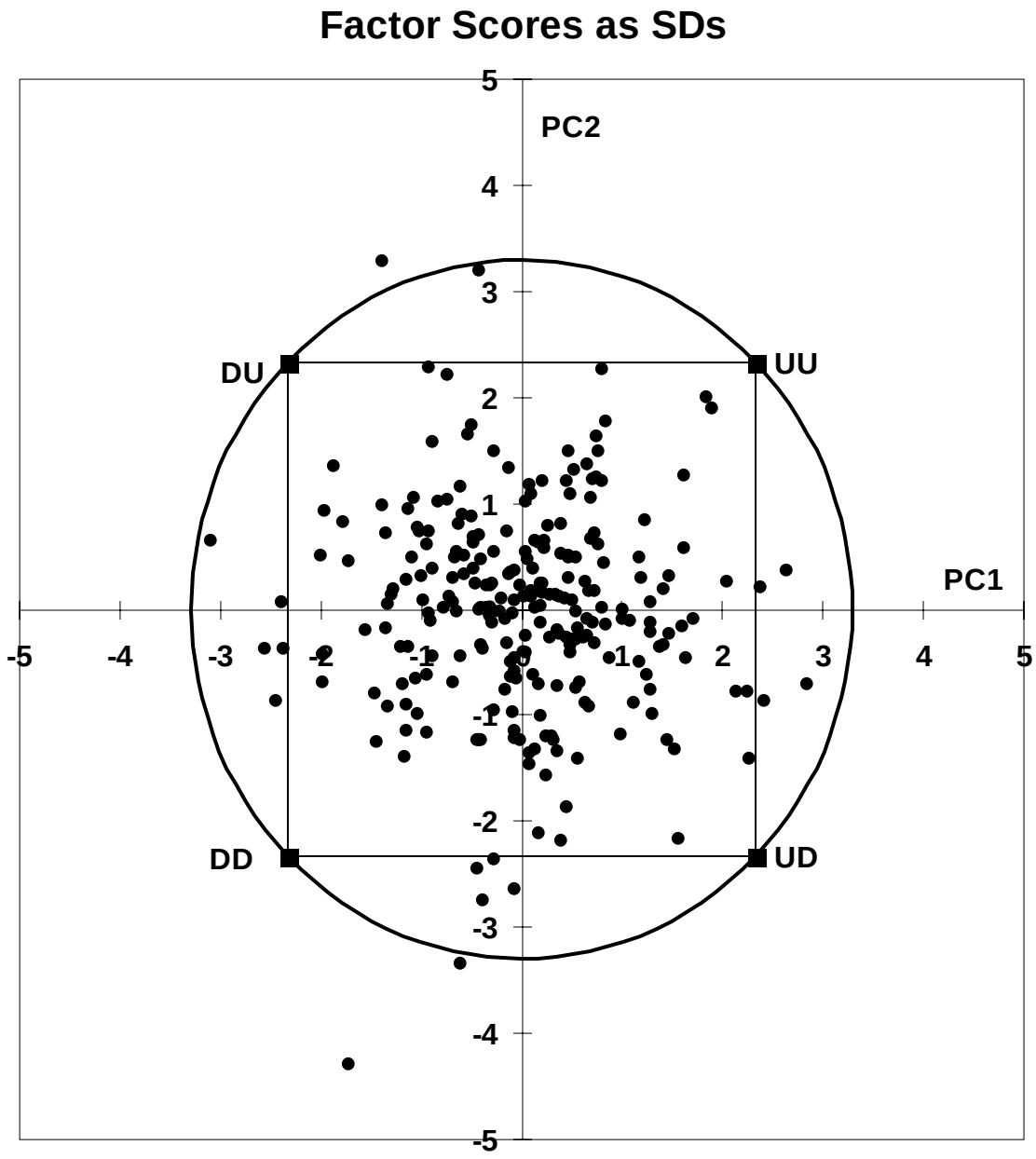


**Figure 26. The level curve of a hypothetical convex portfolio with data points to illustrate the probability density of the factor scores. This portfolio experiences the same loss at every scenario point, and at every point along the circular line connecting them. Loss will be greater than at the scenario points at any point outside the circle. Inspection shows few such points; for normally distributed factors the probability outside the circle is 0.5%. Despite the potential for the movement of either factor to cause loss to exceed VAR, the estimate is conservative.**

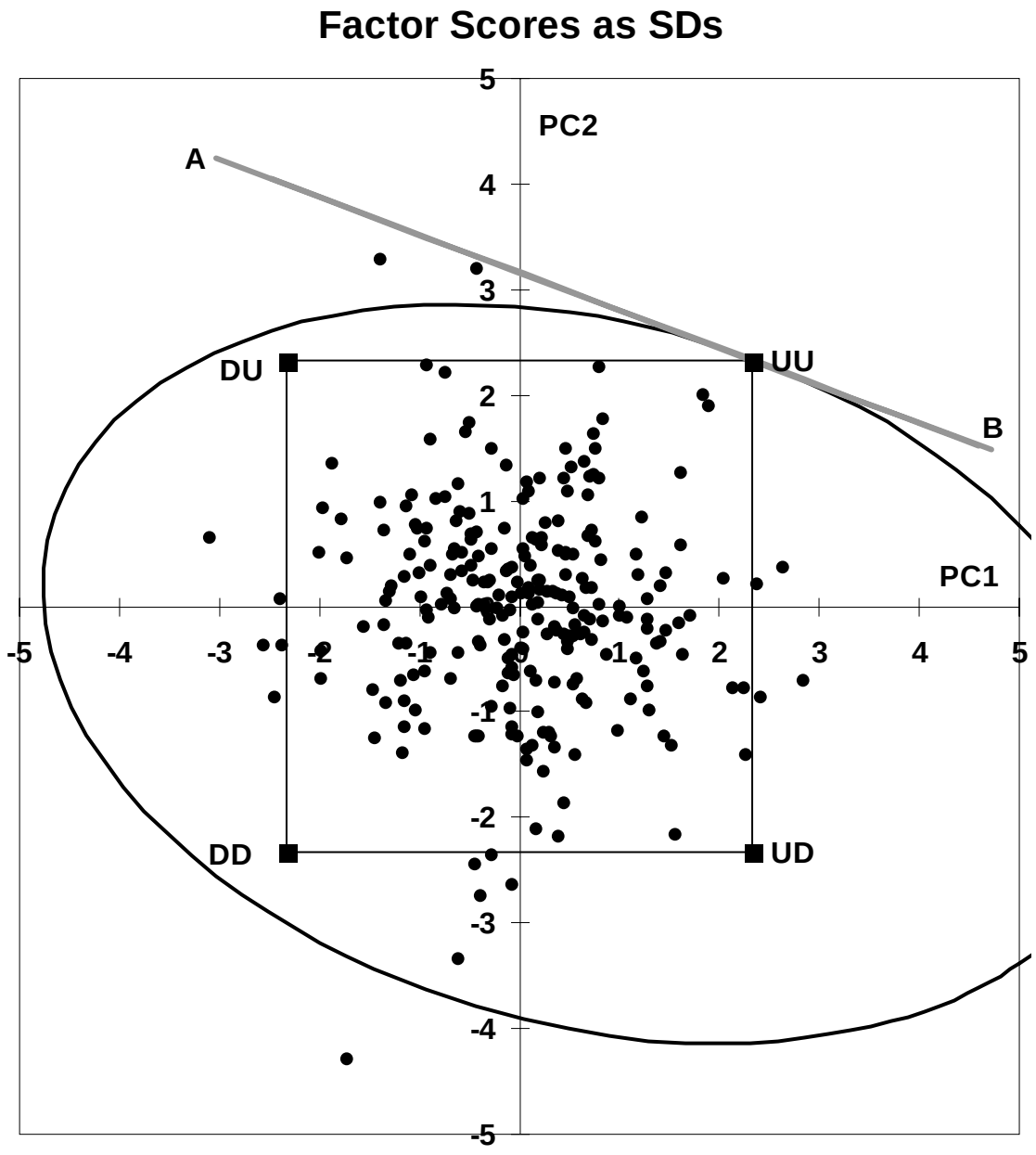


**Figure 27. Two level curves. Level curve AB represents a linearly responding portfolio such as those of Figures 8 and 9. The oval level curve represents a convex options portfolio in which a large score for either factor can produce a large loss. Clearly, more of the probability density lies outside the oval than above the line AB. This means an options portfolio tends to absorb some of the inherent conservatism of the Factor-Based Scenario Method.**

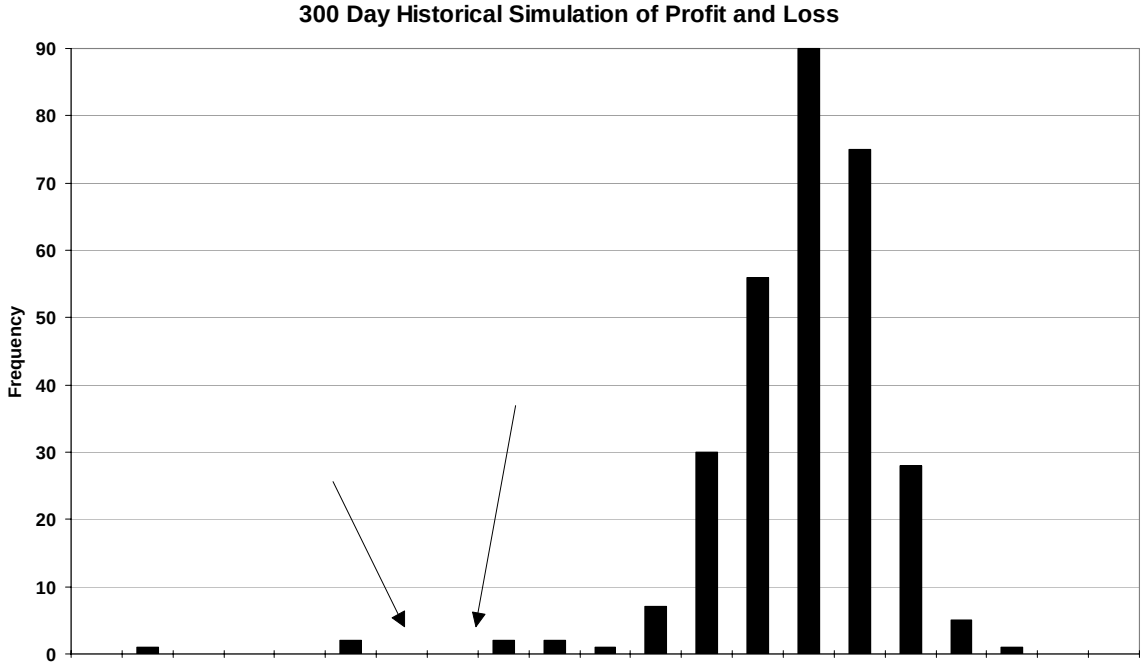


**Figure 28. The simulated P&L distribution of the representative portfolio over 300 days of history. The distribution resembles that for a short option position, with many small profits and a few large losses. The Factor-Based Scenario Method produces a VAR estimate just larger than the 99$^{th}$ percentile loss.**

## Endnotes

[1] This article defines "loading" as an element of an eigenvector of the covariance matrix. Some software packages may use the same word to mean an element of an eigenvector of the *correlation* matrix, a correlation between a data series and a component, or an element of an eigenvector times the square root of its associated eigenvalue. To resolve uncertainty, note that the sum of the squares of the elements of an eigenvector equals one.

[2] For proof, see Appendix I.

[3] Up to a multiplicative factor. Constraining the sum of the squares of the loadings to equal 1.0 results uniquely in the first principal component.

[4] The quantity of a security is a random normal with mean zero and standard deviation equal to $10,000,000 divided by duration. Therefore, all securities have the same expected dollar risk.

[5] Changes have been made to safeguard proprietary information.

[6] The hedge does not insulate the portfolio from movements in futures contracts or factors.