Improving Grid-based Methods for Estimating Value at Risk of Fixed-Income Portfolios

Michael S. Gibson and Matthew Pritsker*

March 23, 2000

*Trading Risk Analysis Section, Division of Research and Statistics, Federal Reserve Board. This paper represents the views of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. We can be reached via email at michael.s.gibson@frb.gov and matthew.pritsker@frb.gov. Postal address: Mail Stop 91, 20th and C Streets NW, Washington, DC 20551. Phone: 1-202-452-2495 (Gibson) and 1-202-452-3534 (Pritsker). Fax: 1-202-452-3819.

Improving Grid-based Methods for Estimating Value at Risk of Fixed-Income Portfolios

Abstract: Jamshidian and Zhu (1997) propose a discrete grid method for simplifying the computation of Value at Risk (VaR) for fixed-income portfolios. Their method relies on two simplifications. First, the value of fixed income instruments is modeled as depending on a small number of risk factors chosen using principal components analysis. Second, they use a discrete approximation to the distribution of the portfolio's value.

We show that their method has two serious shortcomings which imply it cannot accurately estimate VaR for some fixed-income portfolios. First, risk factors chosen using principal components analysis will explain the variation in the yield curve, but they may not explain the variation in the portfolio's value. This will be especially problematic for portfolios that are hedged. Second, their discrete distribution of portfolio value can be a poor approximation to the true continuous distribution.

We propose two refinements to their method to correct these two shortcomings. First, we propose choosing risk factors according to their ability to explain the portfolio's value. To do this, instead of generating risk factors with principal components analysis, we generate them with a statistical technique called partial least squares. Second, we compute VaR with a "Grid Monte Carlo" method that uses continuous risk factor distributions while maintaining the computational simplicity of a grid method for pricing. We illustrate our points with several example portfolios where the Jamshidian-Zhu method fails to accurately estimate VaR, while our refinements succeed.

Keywords: scenario simulation, principal components, partial least squares, Monte Carlo

1 Introduction

Computing value at risk for large portfolios can be computationally intensive. This is especially true for portfolios that contain fixed income derivatives whose values depend nonlinearly on some underlying securities, or for which analytical pricing formulas do not exist. In a recent paper, Jamshidian and Zhu (1997, henceforth JZ) propose a method for significantly simplifying computation of VaR for fixed income portfolios.

The JZ method exploits two insights to simplify VaR computation. First, the value of fixed income instruments is modeled as depending on a small number of risk factors, which are chosen to explain the variation in the default-free yield curve. This step can be termed "dimension reduction." Second, the continuous joint distribution of the risk factors is approximated using a multivariate discrete distribution. The multivariate discrete distribution yields a grid of risk factor realizations with a probability attached to each node. At each node on the grid, the portfolio's value is computed. This produces a discrete approximation to the distribution of the portfolio's value, and VaR is computed by choosing the appropriate quantile of the discrete distribution's cumulative distribution function.[1] In their 1997 paper, JZ give several examples where their method accurately measures VaR with much less computational effort than a full Monte Carlo simulation. The beauty of the JZ method is that VaR is computed while only pricing the portfolio a small number of times.

In this paper we show that the original implementation of the JZ method has two serious shortcomings which imply it cannot accurately estimate VaR for some fixed-income portfolios. The first shortcoming is that JZ pick the risk factors to explain maximal variation in the yield curve and not maximal variation in the portfolio. With poorly chosen risk factors, the JZ method's dimension reduction might drop risk factors that are important in explaining the portfolio's value but not so important in explaining yield curve variation. This can sometimes lead JZ to underestimate the true risk of the portfolio. JZ's method of choosing factors is especially problematic if some of the factors that explain variation in the yield curve have been hedged against in the fixed income portfolio.

The second shortcoming is that a discrete approximation to the distribution of the risk factors can generate a poor approximation to the distribution of portfolio value, even for simple portfolios.

We propose two refinements to the JZ method which correct its two shortcomings. First, we choose risk factors according to their ability to capture changes in the portfolio's value,

[1] JZ also present an extension of their method to multi-currency portfolios. That extension also relies on dimension reduction and discretization.

not their ability to capture variation in the yield curve. Second, we work with a continuous distribution of the derived risk factors while maintaining the computational simplicity of a grid method. We begin, like JZ, by pricing the portfolio exactly for a grid of factor realizations. By interpolating within the grid (and extrapolating outside the grid), we can construct a mapping from factor realizations to portfolio value.[2] We then compute VaR with a Monte Carlo simulation by drawing from the factors' continuous distributions and pricing the portfolio using the grid-based approximation. We call this technique "Grid Monte Carlo."[3]

Because Grid Monte Carlo draws risk factor realizations from a continuous distribution, it avoids the errors from approximating a continuous distribution with a discrete distribution. An added benefit of our approach is that we can optimize the choice of grid points to capture the roughness of the pricing function.[4] By contrast, JZ's choice of grid points is dictated by the discrete approximation of the probability function. As a consequence, JZ's method might miss rough points in the pricing function, which is another potentially important source of error in their method.

Our refinements improve the accuracy of the JZ method. There is a cost to be paid, however, in the form of increased computational burden. The increased burden stems from our need to use portfolio-specific information when choosing the risk factors. The increased burden could be minimal if portfolio-specific risk information (such as deltas and gammas) has already been computed. If not, some costly portfolio revaluations would be needed to produce the information. Because computing prices via interpolation and extrapolation requires no additional costly portfolio revaluations, Grid Monte Carlo is nearly as fast as JZ's discrete approximation.

While we focus our discussion on the JZ method, our two refinements may also be useful improvements to other grid methods used to measure the risk of fixed income portfolios, such as Frye (1996) and Chishti (1999). In fairness, we expect that JZ and other authors of papers advocating grid methods are aware of the issues we raise, though not perhaps of the magnitude of the VaR estimation errors that can occur. Our fear is that not all users of these methods are aware of these issues.

Abken (2000) also discusses problems with the JZ method. He focuses on errors caused by discretization of the distribution of portfolio value, not on the JZ method's potentially

[2] In section 4, we give examples of two specific interpolation schemes, a simple linear interpolation and a more sophisticated second-order Taylor-series expansion.

[3] See Pritsker (1997) for another variant of a Grid Monte Carlo method.

[4] For example, adding grid points on either side of the threshold of a knock-out option.

poor choice of risk factors. For the four sample portfolios of interest rate derivatives he analyzes (over several different time periods), the JZ method's VaR estimates are frequently off by 5–15 percent. He also addresses the presumption that, as the grid becomes finer, the JZ method's VaR estimates should approach the "true" VaR. He shows that, for the density of the grid that is likely to be used in any realistic application, the JZ method's VaR estimates can oscillate erratically around the true VaR. The oscillations come about as the VaR quantile jumps from one point on the discrete probability distribution to another when the grid density is changed.

The next two sections of the paper discuss the problems with the JZ method and our solutions in more detail. The fourth section gives some detailed examples that illustrate our points. In the conclusion we discuss how our method can be implemented as part of the global VaR computation for large financial firms.

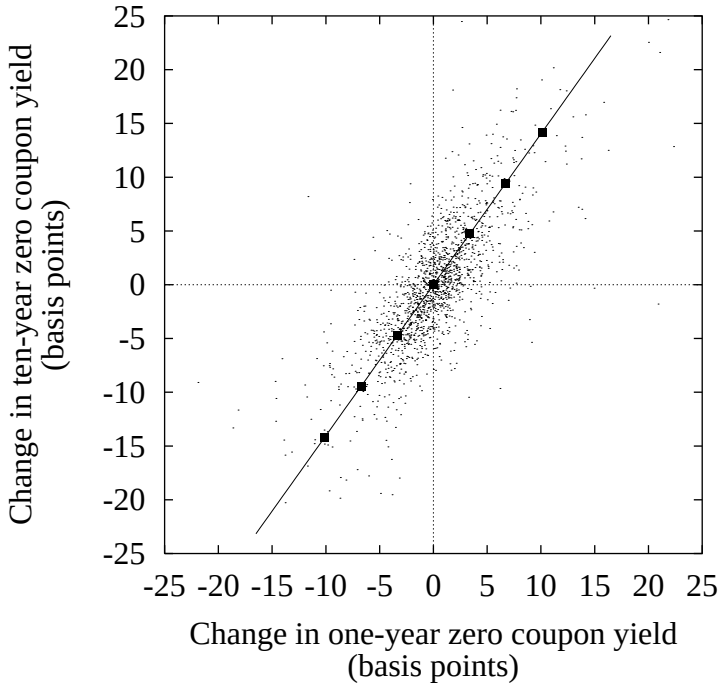
2 Pitfalls with the JZ method

This section explains in more detail why the JZ method can sometimes fail. The two shortcomings of the JZ method are the potentially poor choice of risk factors and errors from discretizing the risk factors' joint distribution.

JZ extract risk factors from the yield curve using principal components analysis (PCA). Using PCA, the first risk factor is the linear combination of yields which explains the most variation in the yield curve. Typically, this factor captures a nearly parallel shift in the yield curve. Each additional PCA factor explains the maximal yield curve variation which was not explained by the previous factor(s). The second and third PCA factors typically look like a "tilt" (a change in the slope) and "twist" (a U-shaped shock) to the yield curve. JZ appeal to the well-known empirical finding that the first three PCA factors capture a large fraction of the historical variation in the USD yield curve to justify their focus on a three-factor model. In a three-factor model, all principal components but the first three are dropped. These three become the risk factors which are used to build a three-dimensional grid for pricing.

We illustrate how PCA factors can be a poor choice for doing risk measurement with a simple example. Our example uses PCA to perform dimension reduction from two dimensions down to one. We limit our example to two dimensions so we can make our point with graphs. Exactly the same logic would apply in the realistic case of dimension reduction from an arbitrarily large number of points on the yield curve down to three factors.

Figure 1: A simple example of dimension reduction and discretization



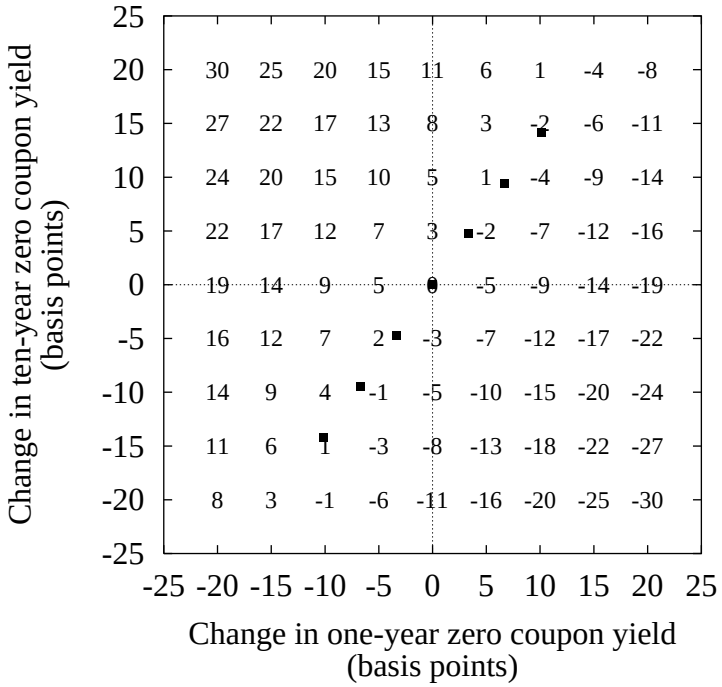
Notes: The scatter plot shows actual one-day changes in one-year and ten-year zero coupon yields from January 1993 to May 1999. The solid line shows the first PCA risk factor. The JZ discretization of the first PCA risk factor is shown as the seven solid squares.

Suppose there are only two risk factors, a one-year zero coupon yield and a ten-year zero coupon yield. We use PCA to extract the first principal component, which we will use as our sole risk factor for risk measurement. Following JZ, we approximate the distribution of the single risk factor with a seven-point discrete distribution. This situation is illustrated in Figure 1.

As the figure shows, the first PCA factor captures most of the variation in the two risk factors. They move together most of the time and the first PCA factor reflects that. The JZ discretization of the first PCA factor encompasses most of the probability mass of the data. However, dimension reduction from two factors down to one in effect assumes that the two yields never move in opposite directions. Any portfolio that takes opposing positions in one-year and ten-year zero coupon bonds will have its VaR underestimated.

One such portfolio consists of going long the one-year zero and short the ten-year. This portfolio will increase in value when the one-year rate falls and the ten-year rate rises. It will lose value when the opposite happens. Neither of these scenarios will be captured by the

Figure 2: A simple example of dimension reduction and discretization, continued



Notes: The numbers show the change in the value of a portfolio that is long the one-year zero coupon bond and short the ten-year. The JZ discretization of the first PCA risk factor is shown as the seven solid squares.

one-factor model. The one factor model only "sees" shocks with the one-year and ten-year rates moving in the same direction (along the line in Figure 1).

The change in value for this portfolio is shown by the numbers on Figure 2, along with the JZ discretization of the first PCA risk factor. It is evident from the figure that the JZ discretization will show a VaR of no more than 2. The true VaR of this portfolio is of course much higher. (The 5 percent VaR is 4.6 and the 1 percent VaR is 7.3.)

This simplistic example has illustrated the first shortcoming of the JZ method: dimension reduction will result in an underestimate of VaR when the dimensions that are dropped are important for the particular portfolio. Although we limited the example to two dimensions so we could explain it graphically, the more realistic situation we have in mind is a risk measurement model based on the first three PCA factors and risk-taking in the direction of the fourth, fifth, or higher PCA factors. The examples in the fourth section of the paper include situations like these.

The second shortcoming of the JZ method is that it uses a discrete distribution to ap-

Table 1: Discrete Distribution Approximations for JZ's Three-Factor Model

Factor 1

f_1	–2.45	–1.63	–0.82	0.00	0.82	1.63	2.45
Prob(f_1)	0.016	0.094	0.234	0.313	0.234	0.094	0.016

Factor 2

f_2	–2.00	–1.00	0.00	1.00	2.00
Prob(f_2)	0.0625	0.25	0.375	0.25	0.0625

Factor 3

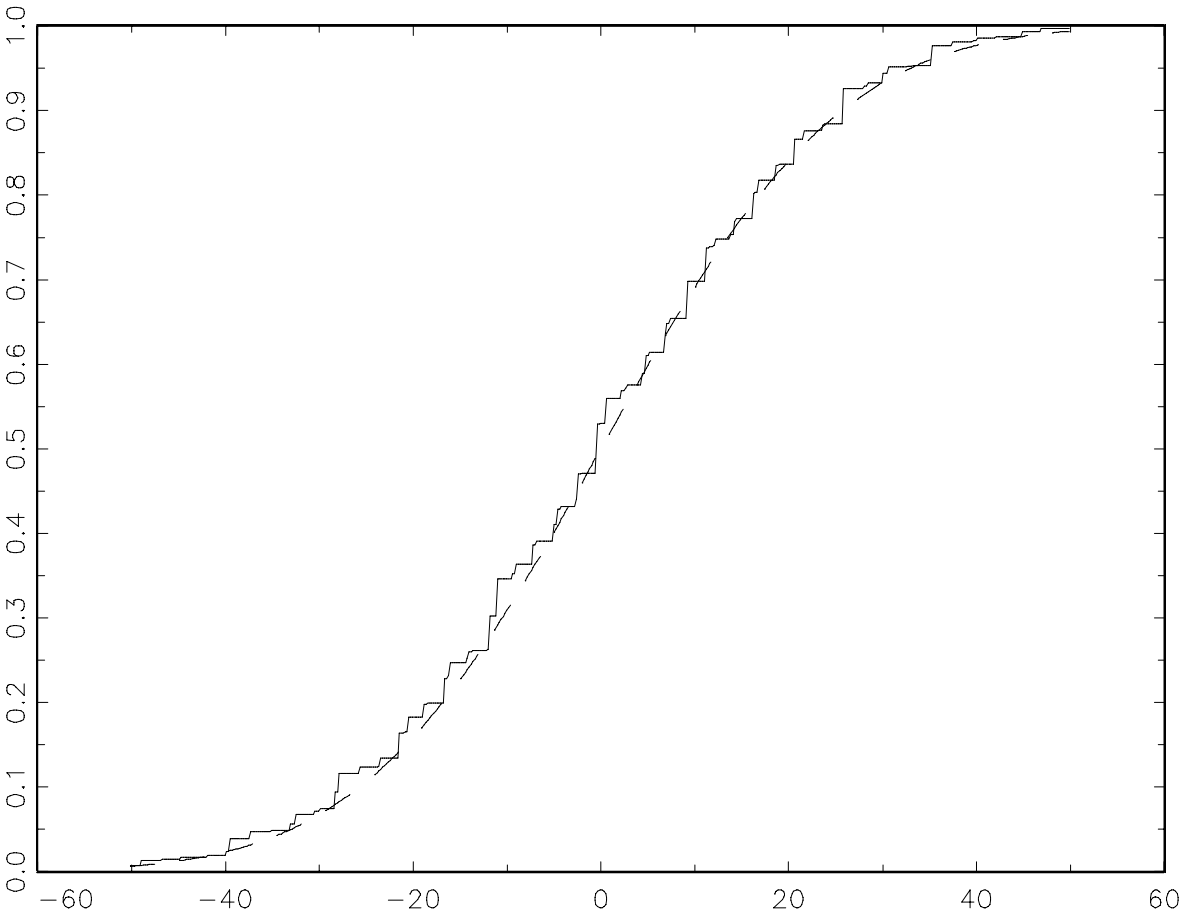
f_3	–1.41	0.00	1.41
Prob(f_3)	0.25	0.50	0.25

proximate the continuous distribution of each factor. The number of points in the discrete approximation for each factor is based on the factor's relative importance in explaining variation in the yield curve. In the three-factor models in JZ's paper, they approximate the distribution of the the first, second, and third factors using binomial distributions with 7, 5, and 3 points respectively. A binomial distribution is chosen because, as the number of points increases, it converges to a Normal distribution. Each of the true factors is assumed to be distributed Normally with mean 0 and variance 1. The discrete distribution approximations for each factor are provided in Table 1.

The joint distribution of the three factors is approximated with a multinomial distribution. There are 105 possible outcomes on the 7-by-5-by-3 grid of risk factor realizations ($105 = 7 \times 5 \times 3$). The PCA factors are orthogonal by construction and are assumed to be independently distributed. Therefore, the probability for each point on the grid is the product of the marginal factor probabilities. The portfolio is priced exactly at each of the 105 points. The end result is a discrete approximation to the cumulative distribution function (CDF) of the portfolio's value, which is then used to compute VaR.

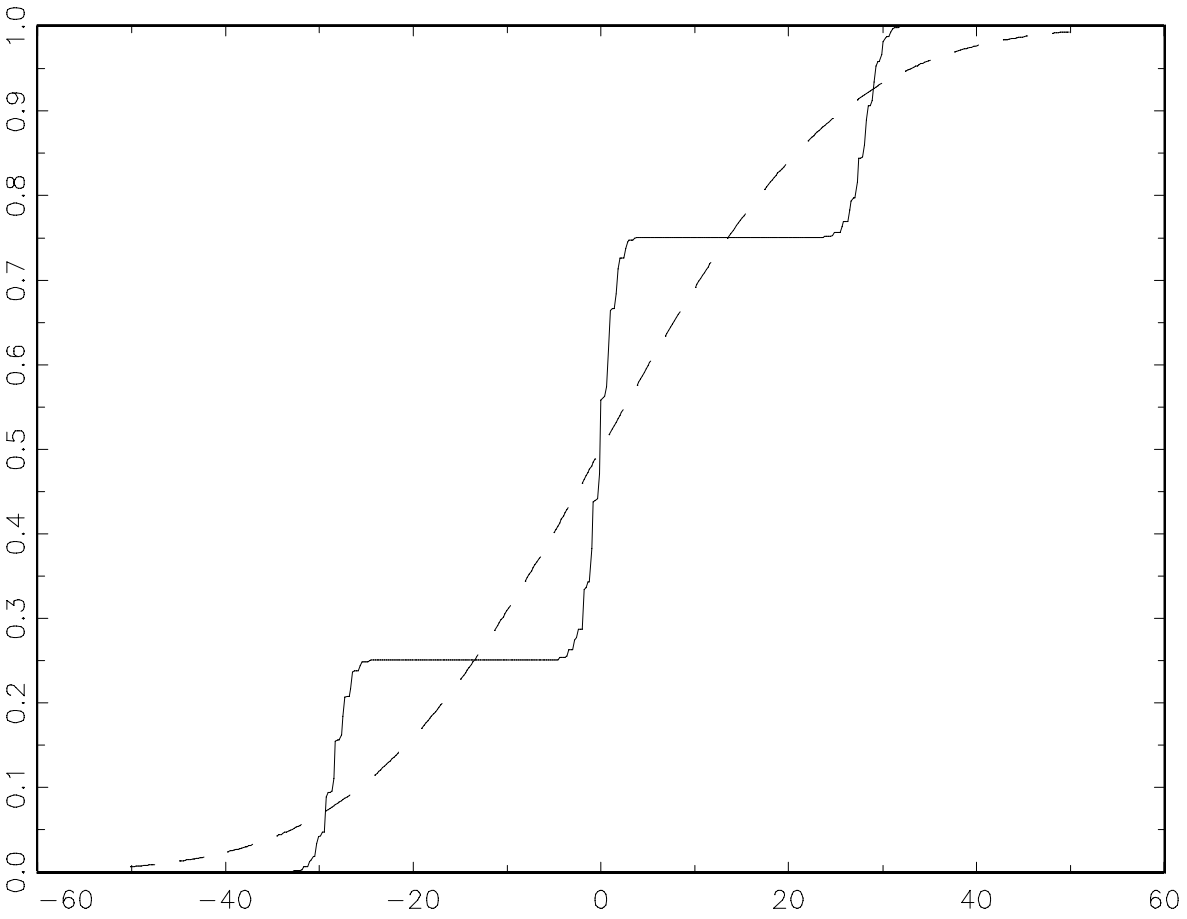
To illustrate how the JZ method can work well for some portfolios, assume that the best factors have been chosen and each factor is distributed Normally with mean 0 and variance 1. For purposes of illustration, we will assume that it is possible to take linear positions in each factor. Consider first a portfolio where \$11.58 is invested in each factor (the "diversified portfolio"). For the diversified portfolio, the true distribution function and the JZ approximation are pretty close (Figure 3, dashed and solid lines respectively). Moreover, the true VaR at the 1 percent confidence level is \$46.64 and JZ's approximation

Figure 3: True CDF (dashed line) and JZ Approximation (solid line): Diversified Portfolio



Note: The diversified portfolio consists of a position of $11.58 invested in each of the three PCA factors. Its 1 percent VaR is $46.64.

Figure 4: True CDF (dashed line) and JZ Approximation (solid line): Concentrated Portfolio



Note: The concentrated portfolio consists of a position of $1 in the first two PCA factors and $20 in the third PCA factor. Its 1 percent VaR is $46.64.

is $49, which is pretty close. For this portfolio, the 105 points that JZ use to approximate the CDF do a pretty good job.

To see what can go wrong with the JZ method, suppose instead that the portfolio is composed of $1 positions in factors 1 and 2, and a $20 position in factor 3 (a "concentrated portfolio"). By construction, the true 1 percent VaR is again $46.64. JZ's approximation is $31.20, understating true VaR by 33 percent. The true distribution function and the JZ approximation are shown in Figure 4. The reason for the understatement of VaR is clear from both Table 1 and Figure 4. The bulk of the risk in this portfolio is concentrated in factor 3. Table 1 shows that risk factor 3 is approximated using a distribution that only has 3 points in its support, and has no probability mass below the 25th percentile. This distributional assumption for factor 3 is simply not adequate for approximating the distribution's first

percentile, which is the relevant point for the 1 percent VaR. Figure 4 illustrates the same point. The figure shows that virtually all of the probability mass clusters around three points. Moreover, the neighborhood near the true 1 percent quantile of the distribution receives no probability mass at all.

3 Solutions to the problems

3.1 Choose the factors in a portfolio-dependent way

The first problem with the JZ method is the way the risk factors are chosen. Factors are chosen based on their ability to closely and parsimoniously mimic the fluctuations in the yield curve. But, the factors that explain yield curve variation may not be the same factors that explain the variation in the value of a fixed income portfolio. The obvious answer is that the factors should be chosen based on their ability to explain the variation in portfolio value.

Partial Least Squares (PLS) is one technique that can be used to choose risk factors by taking into account information about the portfolio. Like PCA, PLS decomposes the data into orthogonal factors which are ranked in order of "importance." Both PCA and PLS can be used to do the dimension reduction that is a key element of grid methods by dropping the "less important" factors. While PCA ranks factors by their ability to explain the *variance* of the data, PLS ranks factors by their ability to explain the *covariance* between the data and some other variable y. In our application, y will be the portfolio's value.

PLS is often used in other scientific disciplines facing data-intensive problems similar to ours. According to Garthwaite (1994),

> The intention of PLS is to form components that capture most of the information in the X variables that is useful for predicting [another variable y], while reducing the dimensionality of the regression problem by using fewer components than the number of X variables. PLS is considered especially useful for constructing prediction equations when there are many explanatory variables and comparatively little sample data.

The situation of many explanatory variables and little data is an accurate description of the problem of measuring VaR for a fixed income portfolio. An arbitrarily high number of yield curve points can be taken as the X data, yet producing the variable y (portfolio value) relies on computationally burdensome pricing algorithms.

Both PCA and PLS can be expressed algebraically as

$$F = X\Lambda$$

where F is a matrix of factor realizations, X is a matrix of yield curve data, and Λ is a matrix of factor loadings. Each row of F and X corresponds to one time series observation (i.e., one day for daily data). The factors can be ordered by their "importance." Without loss of generality the columns of Λ can be ordered so that the first column of F contains the "most important" factor and the rightmost column of F contains the "least important." Going from the high dimensionality of the yield curve data to some smaller dimension requires dropping columns of F and Λ, beginning with the rightmost column and moving left.[5]

The difference between PCA and PLS is that PLS bases its factor decomposition on information on an additional variable y. In our application, y is the portfolio's value for a given set of risk factor shocks X. Using information on portfolio value means the factors can be chosen, and risk measurement done, according to the risks that are important to that particular portfolio.

The additional information that PLS requires may be costly to produce. One way to generate this information would be to choose a set of yield curve shocks X and, for each row of X, revalue the portfolio to generate one element of y. The number of shocks (and portfolio revaluations) should not be too small, or y may not accurately represent the true risks of the portfolio.

Alternatively, we could just use whatever risk information has already been produced for the portfolio, avoiding costly portfolio revaluations. For example, information on the portfolio's deltas and gammas may already be available "for free" to the VaR model. The vector y could then be generated with a delta-gamma approximation. Note that the delta-gamma approximation will not be used to produce VaR numbers directly. It is only used to choose the factors on which VaR estimation will be based. The actual VaR estimation can be done with the JZ method or the Grid Monte Carlo method we discuss in the following section. Neither would use a global delta-gamma approximation to estimate VaR.[6]

Using PLS to choose factors for JZ's scenario simulation method will concentrate the risk measurement effort in the place where it will do the most good. This should improve the accuracy of risk measurement. In addition, using PLS factors will ensure that the grid will

[5]See Stone and Brooks (1990) for a comparison of PCA and PLS. See Garthwaite (1994) for a clear exposition of PLS.

[6]Some variants of Grid Monte Carlo may use local deltas and gammas to interpolate between grid points.

be densest for the risk factors that are most important for the portfolio, reducing but not eliminating the discretization problem. As the next subsection discusses, we believe further action can and should be taken to deal with the discretization problem.

3.2 Base risk measurement on continuous probability distributions and use the grid for pricing

If the full set of risk factors is used in computing VaR, it is well known that pseudo Monte Carlo with full repricing will generate estimates of VaR that converge in probability to true VaR.[7] The problem with using Monte Carlo with full repricing is that the repricing step is too time consuming for this method to be practical for large fixed income portfolios. JZ reduce the number of repricing steps by discretizing the risk factors' joint distribution.

As we noted earlier, discretizing the distributions can generate large errors in the estimates of VaR for some portfolios. Our approach is to avoid the errors from discretizing the distribution of the risk factors by making Monte Carlo draws from the continuous distribution of the risk factors, but simplifying the repricing step. If our simplified pricing method generates accurate prices, but is fast to compute, Monte Carlo using our simplified pricing method will inherit the desirable properties of Monte Carlo with full repricing, but will generate VaR estimates very rapidly.

The repricing algorithm that we use starts with a grid of discrete factor realizations.[8] We exactly reprice the portfolio at each node of the grid. To use the grid to compute VaR, N Monte Carlo draws are made from the risk factors' true distribution. For factor draws that do not fall on the grid, the price of the portfolio is computed by interpolation (or extrapolation). We call this technique "Grid Monte Carlo."[9] In the examples in the next section, we consider two different interpolation methods. If the grid points and interpolation method are chosen well, the interpolation (and extrapolation) errors can be made small and will get progressively smaller as the mesh of the grid is reduced.

[7] This will be true when the risk factors are drawn from their true distributions and the number of Monte Carlo draws is allowed to grow large.

[8] This could be the same grid used by JZ, but it does not have to be.

[9] See Chishti (1999) for a similar Grid Monte Carlo technique.

4 Examples

We have described how the JZ method can fail to measure the risk of some fixed income portfolios. In this section we show examples of portfolios where the JZ method fails and our two refinements improve risk measurement.

We present our examples in two parts. In both parts, we compare the JZ method's estimates of VaR with VaR estimates using one of our refinements. In the first part, we use PLS instead of PCA to choose the factors underlying the JZ method, leaving the discrete grid unchanged. In the second part, we replace the JZ method's discrete pricing grid and probability distribution with a discrete pricing grid used with a continuous probability distribution (Grid Monte Carlo). Note that our refinements will improve any risk measurement technique that is based on the future distribution of portfolio value. We focus on VaR because it is commonly used and familiar.

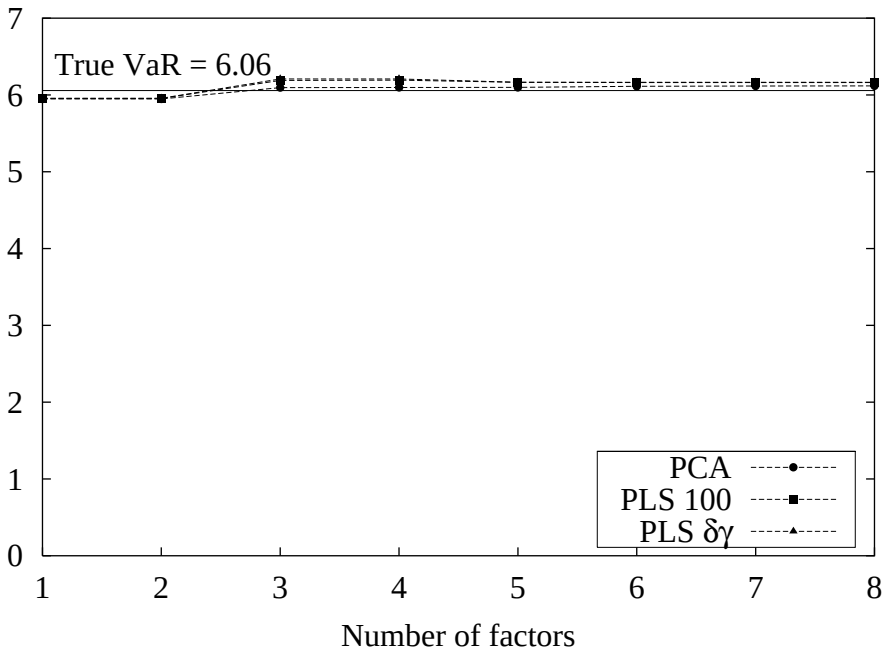
In all our examples, the yield curve data used for risk measurement are 1,000 days of artificial zero-coupon bond yields at nine maturities (1, 3, 5, 7, 9, 10, 15, 20, 30 years). The covariance structure of the artificial data is that of actual U.S. dollar zero-coupon yields over 1993–99, but the artificial data are multivariate normal. We choose to work with multivariate normal data to avoid having to worry about risk measurement errors caused by "fat tails" that are ignored by the JZ method. Our refinements to the JZ method will complement any refinements used to account for "fat tails" in the empirical distribution of zero coupon bond yields. In all examples, the VaR we report is averaged over 100 randomly-drawn 1,000-day datasets to ensure the results do not depend on a particular 1,000-day sample.

4.1 Choosing factors with PLS instead of PCA

In our first set of examples we show the effects of using information on the portfolio to choose the factors for risk measurement. Information on the portfolio is incorporated using PLS. We look at two different ways to produce the information on the portfolio. In some examples, PLS is implemented by randomly choosing some days from the dataset, computing the portfolio's value on those days, and using that vector of portfolio values as the input to the PLS algorithm.[10] In other examples, we assume that the portfolio's deltas and gammas are available "for free" and we use a delta-gamma approximation to generate portfolio values for all 1,000 days in the data. We compare the 5 percent one-day VaR computed with PCA

[10]The PLS algorithm we use is presented in Appendix A.

Figure 5: Value at Risk for Long Bond portfolio



factors and the VaR computed with PLS factors, while leaving all other parts of the JZ method, including the discrete probability distribution, unchanged.

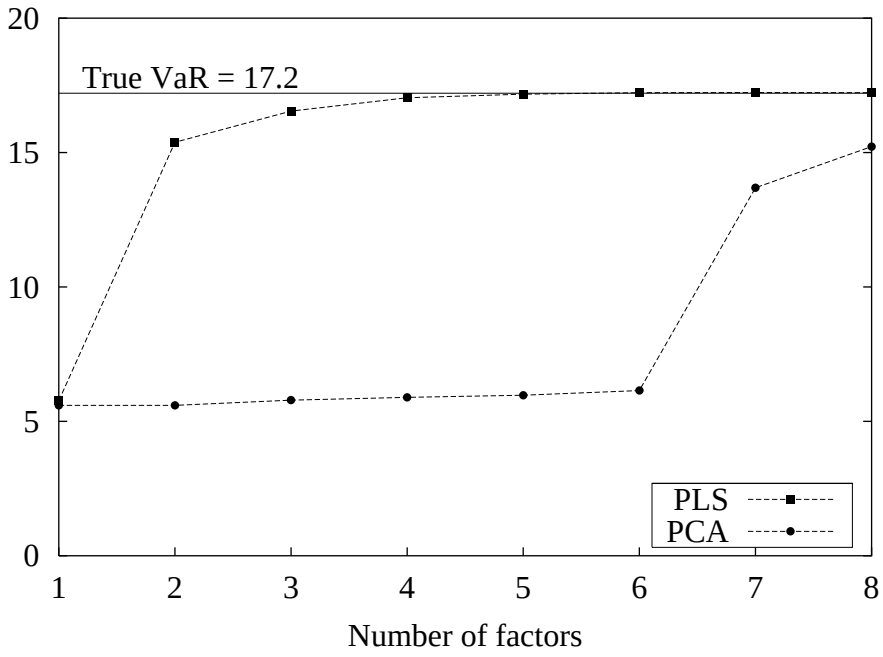
Our first portfolio is a simple portfolio of long bonds spread evenly across the maturity spectrum.[11] We begin with a simple long bond portfolio as a benchmark, to show that the JZ method can work well for some portfolios. As we have discussed, the JZ method should work well if the portfolio's value is sensitive to the PCA yield curve factors.

Estimates of VaR for the Long Bond portfolio are shown in Figure 5. The figure shows the estimated VaR for the original JZ method with PCA-based factors (labeled "PCA") and two versions of our refinement with PLS-based factors (labeled "PLS 100" and "PLS $\delta\gamma$"). The number of factors used in the VaR measurement ranges along the horizontal axis from one to eight. Recall that as the number of factors used in a grid-based method increases, the computational burden grows in a multiplicative fashion.[12] In practice, grid-based methods become unwieldy with more than three or four factors. The number of grid points used in our examples are seven for the first factor, five for the second factor, and three for each additional factor.

[11]The portfolio weights at the nine maturities are (3.6, 0.3, 1.4, 1.4, 0.2, 3.8, 0.6, 2.2, 4.4). They were chosen by drawing nine uniform(0,5) random numbers.

[12]We stopped with eight factors. A nine-factor model would have required excessive computer memory and time to compute.

Figure 6: Value at Risk for Butterfly portfolio



As we expected, the original JZ method with PCA-based factors does a good job of measuring the risk of a simple long bond portfolio, as does our refinement with PLS-based factors. Even with a one-factor model, all three estimated VaRs are close to the true VaR of 6.06.[13]

Our second portfolio provides a more challenging risk measurement test: a "butterfly" portfolio consisting of buying a single 5-year bond and a single 9-year bond and selling two 7-year bonds. This portfolio is intended to capture the spirit of "spread bets" whose risk will likely be missed by PCA-based methods.

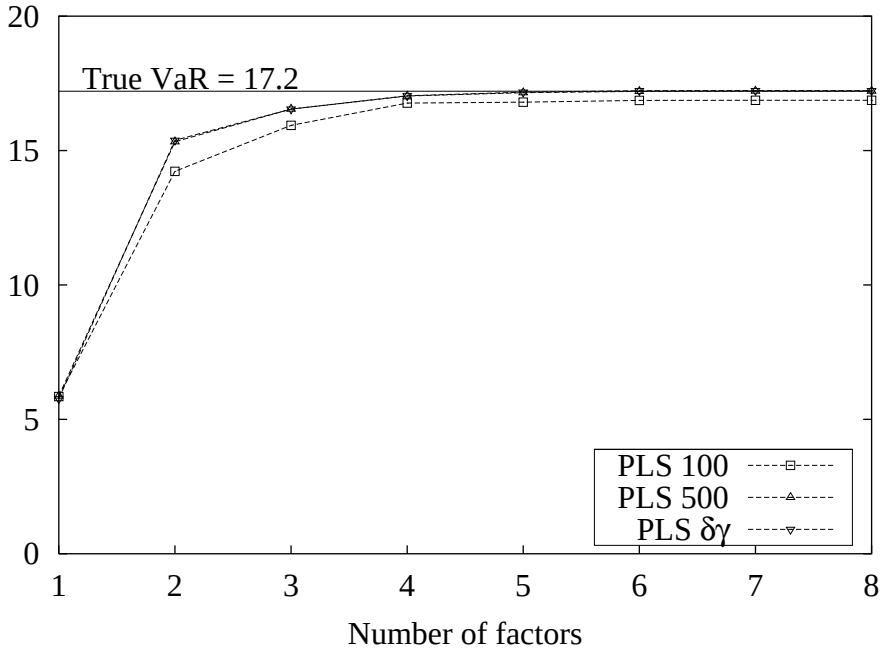
Estimates of VaR for the butterfly portfolio are shown in Figure 6. The PCA-based factors do a poor job of capturing the portfolio's risk, while the PLS-based factors do much better.[14] The first three PLS-based factors come close to the true VaR. With the PCA-based factors, only when the seventh and eighth factors are added does the JZ method begin to approach the true VaR.[15]

[13]In all our examples, "True VaR" is calculated via full Monte Carlo simulation with 1,000,000 draws.

[14]These PLS factors are chosen using a delta-gamma approximation.

[15]The two one-factor models produce about the same VaR estimate because the first PLS factor, like the first PCA factor, has all nine yields moving in the same direction. All maturities are highly correlated, so individually each maturity has about the same ability to explain the variation in the portfolio's value. Thus, each maturity gets about the same weight in making up the first PLS factor.

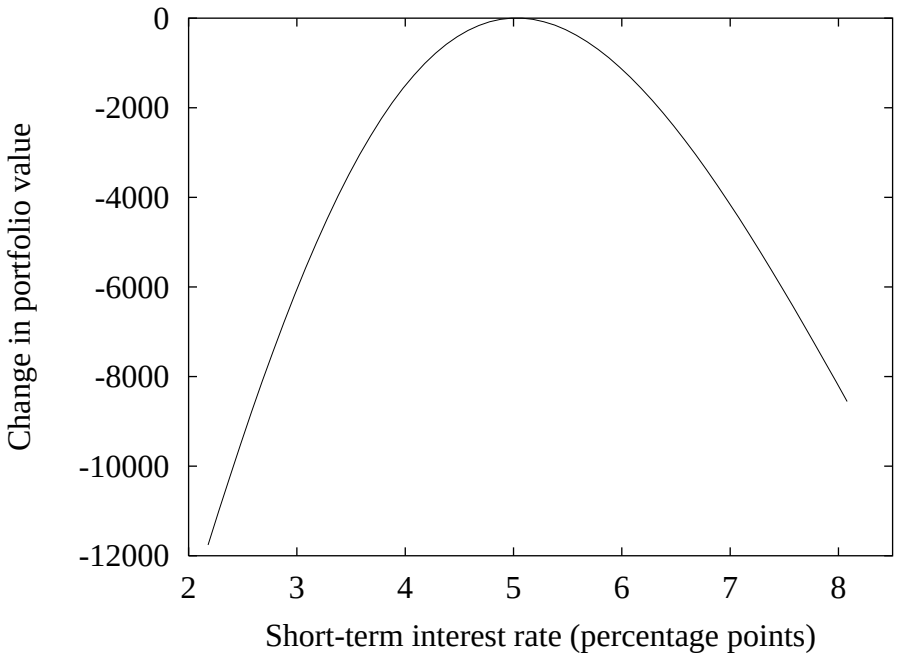
Figure 7: Value at Risk for Butterfly portfolio, with different amounts of preprocessing



One decision that has to be made in order to use PLS is how to produce the portfolio-specific information needed to implement PLS. We have described two ways. In the first, a certain number of portfolio revaluations are used as an input to PLS. These portfolio revaluations can be thought of as "preprocessing" because they are needed to choose the PLS-based factors, in advance of the main VaR calculation. The number of portfolio revaluations must be large enough to enable PLS to identify the correct risk factors. We can expect that more preprocessing would increase the relevance of the PLS-based factors, thus increasing the accuracy of the VaR estimate. We must keep in mind, however, that the increased computational burden of the PLS-based method over the JZ method is almost entirely due to the preprocessing step. The second way to produce portfolio-specific information for PLS is to use already available information, such as deltas and gammas. Using deltas and gammas, we can construct a quadratic approximation to the portfolio's value, eliminating the preprocessing step.

Figure 7 shows the difference between the ways of producing portfolio-specific information for the "butterfly" portfolio. The two lines labeled "PLS 100" and "PLS 500" use preprocessing with 100 and 500 portfolio revaluations, respectively. As expected, the accuracy of the VaR estimates improves with more preprocessing. The PLS 100 VaR estimates always fall short of the true VaR, while the PLS 500 VaR estimates are quite close to the true VaR

Figure 8: Response of the portfolio of caps and floors to parallel shifts in the yield curve



for three factors, and are right on for four or more factors. The line labeled "PLS $\delta\gamma$" uses a delta-gamma approximation to bypass the preprocessing step. These VaR estimates are right on top of the PLS 500 estimates. These two methods choose nearly the same factors. This may not be surprising since the "butterfly" portfolio is made up of bonds whose value can be well captured by a quadratic approximation.

Our third portfolio also presents a risk measurement challenge. The basic portfolio consists of 50 written caps and floors.[16] Figure 8 shows the nonlinear dependence of this portfolio on the overall level of interest rates. This portfolio is designed to roughly approximate the portfolio of a dealer in interest rate options.[17] We will look at the unhedged caps-and-floors portfolio and several hedged portfolios. Portfolios that are hedged pose a particular problem to the JZ method since it concentrates on the first few PCA yield curve factors, which are likely to be the factors hedged against.

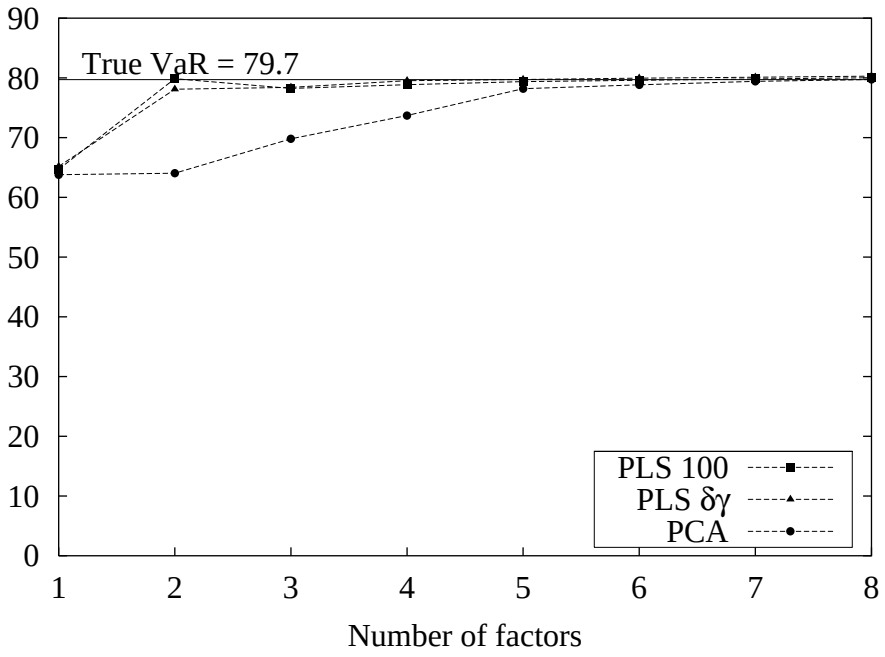
The estimated VaR for the unhedged portfolio of caps and floors is shown in Figure 9.[18]

[16]The portfolio has an equal number of caps and floors, with maturities distributed uniformly over 2–10 years, reference rates distributed uniformly over +/- 2 percent from the current short rate, and notional amounts distributed uniformly over the interval (10,100).

[17]The shape of Figure 8 is similar to Kambhu's (1998, Chart 1) estimate of the net interest rate exposure of U.S. dollar interest rate options dealers in March 1995.

[18]Yield volatilities are assumed constant at a flat 20 percent. In practice, additional risk factors would be needed to account for volatility risk.

Figure 9: Value at Risk for caps and floors portfolio, unhedged



It takes five PCA-based factors to accurately represent the risk of this portfolio, while only two PLS-based factors are needed. Both the "preprocessing" and delta-gamma techniques for picking the PLS factors do a good job on this portfolio.

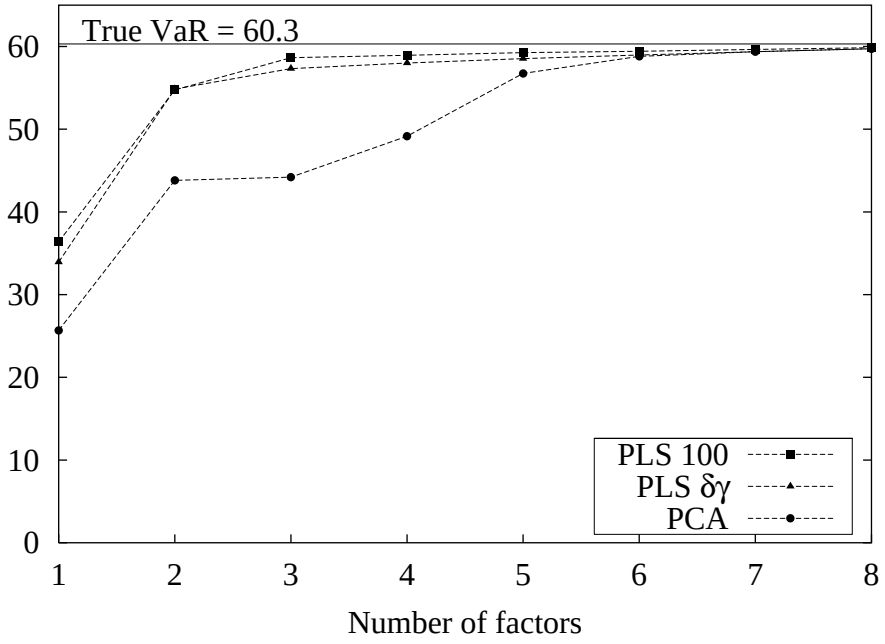
A derivatives dealer would likely try to hedge the exposure to yield curve shifts. One simple hedging strategy would be duration hedging. Duration hedging requires taking a position in zero-coupon bonds to neutralize the portfolio's exposure to parallel shifts in the yield curve. We assume for simplicity that the hedge position is taken in five-year zero-coupon bonds.

Estimates of VaR for the duration-hedged portfolio of caps and floors are shown in Figure 10. The gap between the PLS-based methods and the PCA-based method is wider, compared with Figure 9. This is what we expected, since the first PCA factor looks a lot like a parallel shift in the yield curve. Thanks to the duration hedging, the first PCA-based factor, where the JZ method's discrete pricing grid is finest, will be nearly irrelevant for the hedged portfolio's value.

Another hedging technique is to hedge against the principal components of yield curve fluctuations. Since the first three principal components capture over 90 percent of the variation in the yield curve, some risk managers advocate hedging against principal components.[19]

[19]See, for example, Hill and Vaysman (1998) and Golub and Tilman (1997).

Figure 10: Value at Risk for caps and floors portfolio, duration hedged



Obviously, this will pose problems for the JZ method.

Figures 11, 12, and 13 show the VaR estimates when the portfolio of caps and floors is hedged against the first principal component, the first two principal components, and the first three principal components, respectively. We see that the performance of the PCA-based JZ method gets worse and worse as the degree of hedging increases, just as we would expect. By contrast, the PLS-based factors do a reasonably good job of capturing the portfolio's value. In all three cases, a three- or four-factor PLS model comes quite close to the true VaR. For the more extensively hedged portfolios, a gap opens up between the PLS $\delta\gamma$ method and the PLS method with "preprocessing" to choose the factors. This suggests that more complicated portfolios are more likely to benefit from "preprocessing" to choose PLS factors.

To conclude this subsection, we have demonstrated that basing VaR measurement on risk factors derived using PCA will only be acceptable when the portfolio happens to be sensitive to the PCA factors. On the other hand, risk factors derived using PLS can do a good job of capturing the risk of an arbitrary portfolio. The JZ method will understate the true risk of a portfolio *unless* the risk manager can be certain that the portfolio is only sensitive to the PCA risk factors.

Figure 11: Value at Risk for caps and floors portfolio, hedged against the first principal component

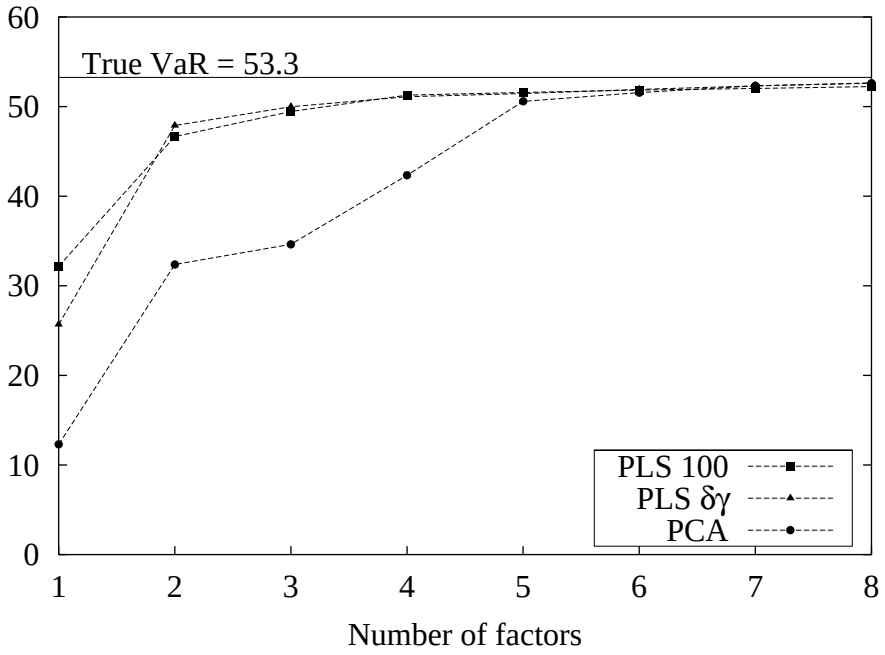


Figure 12: Value at Risk for caps and floors portfolio, hedged against the first two principal components

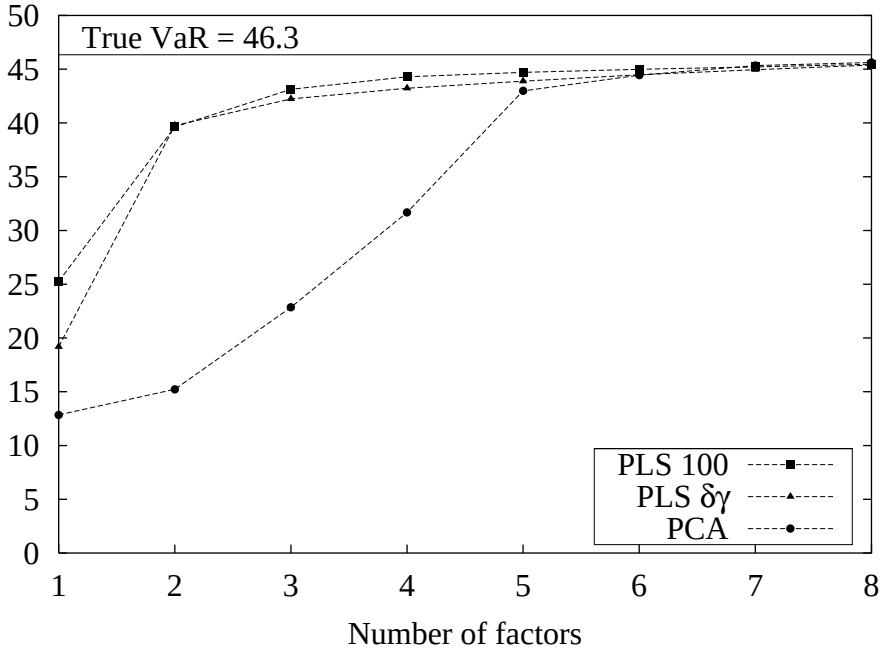
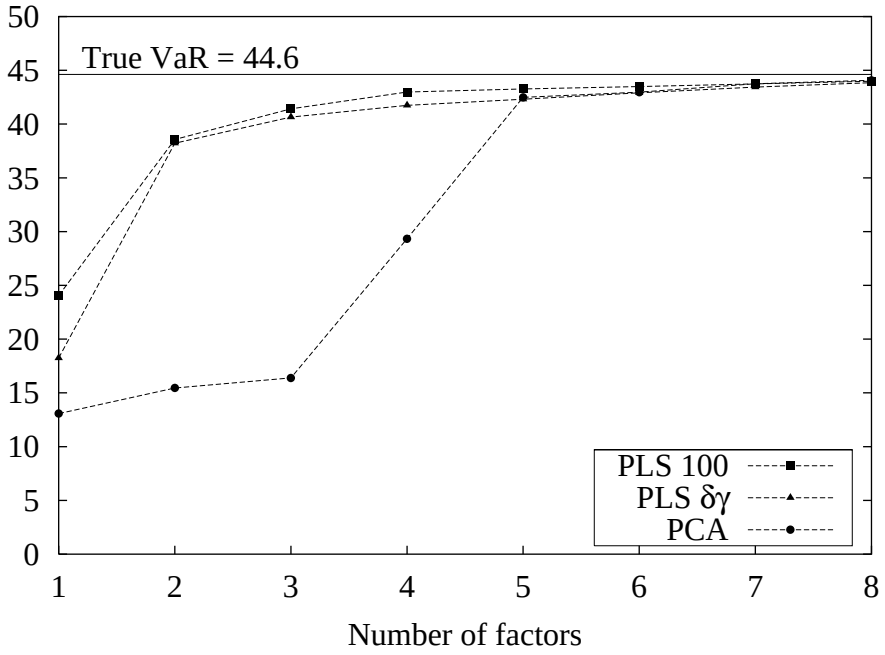


Figure 13: Value at Risk for caps and floors portfolio, hedged against the first three principal components



4.2 Using Grid Monte Carlo

The goal of the Grid Monte Carlo method is to generate risk measures that have the accuracy of Monte Carlo with full repricing, but at a lower computational cost. The computational costs of any grid method increase exponentially in the dimension of the risk factor grid. For this reason, we advocate applying the Grid Monte Carlo method to subportfolios whose common risks can be captured by a few risk factors. In this section we illustrate how to apply Grid Monte Carlo to a subportfolio. In our conclusion (section 5) we discuss how to integrate the results from Grid Monte Carlo applied at the subportfolio level into firm wide risk computations.

To apply the Grid Monte Carlo method at the subportfolio level requires four basic steps:

1. Choose the risk factors.

2. Create a grid of risk factor realizations. The subportfolio will be fully repriced at the nodes of the grid.

3. Choose an interpolation method to approximate the portfolio's value for factor realizations that do not fall on a node of the grid.

4. Generate factors and compute VaR as in the pseudo Monte Carlo method but use the grid for pricing.

The fourth step is straightforward, but the quality of the results from step 4 depends on the care used to implement steps 1 through 3. The most critical step is the first: choosing the risk factors. For a grid method to be computationally feasible, but produce good results, it is necessary to choose a small number of risk factors which capture the risk of the portfolio. We advocate choosing the factors using Partial Least Squares because it extracts and ranks factors based on their ability to explain variation in the portfolio.

To implement the second step, the number of nodes of the factor grid should be kept small to economize on computational costs, and the nodes should be chosen so that the pricing function approximation in the third step is reasonably accurate. This suggests that grid points should be chosen near discontinuities of the pricing function as well as near local maxima, minima, and inflection points. However, choosing the grid points optimally in this way requires significant prior knowledge about the behavior of the pricing function.

An alternative method of implementing the second step is to choose an initial set of grid points based on readily available prior information; additional grid points can then be added if the initial set of grid points appears to be insufficient according to some criterion. For example, the initial set of grid points could be constructed as in JZ where the most important factor has 7 realizations, the next most important factor has 5 realizations, and the least important factor has 3 realizations. Moreover, the placement of the initial grid points could be governed by the amount of nonlinearity in each dimension as measured by the portfolio's δ and Γ with respect to the factors. Given an initial choice of grid points and the pricing method in step 3, the adequacy of the pricing approximation in the neighborhood of different nodes of the grid can be cross-validated against the true price to identify whether additional nodes should be added in that neighborhood, or whether a different pricing approximation should be used.[20]

The third step can be implemented in a number of ways. The simplest way is linear interpolation (or extrapolation) from the prices at nearby nodes. We refer to this method as Grid Monte Carlo with Interpolation (Interp GridMC). An alternative interpolation method is to compute the price for a particular factor realization by using a Taylor series expansion

[20]If the PLS risk factors are chosen by fully repricing the portfolio, then those prices, and the risk factors associated with them can be used to cross-validate prices that are generated by using the grid for pricing. Depending on the method of pricing in step 3, it may also be possible to assess the performance of the grid by how well the price at one node on the grid can be computed based on information at a nearby node or nodes.

centered at the nearest node on the grid. We refer to this method as Delta-Gamma Grid Monte Carlo (DelGam GridMC) because we will implement it using a second order Taylor series.

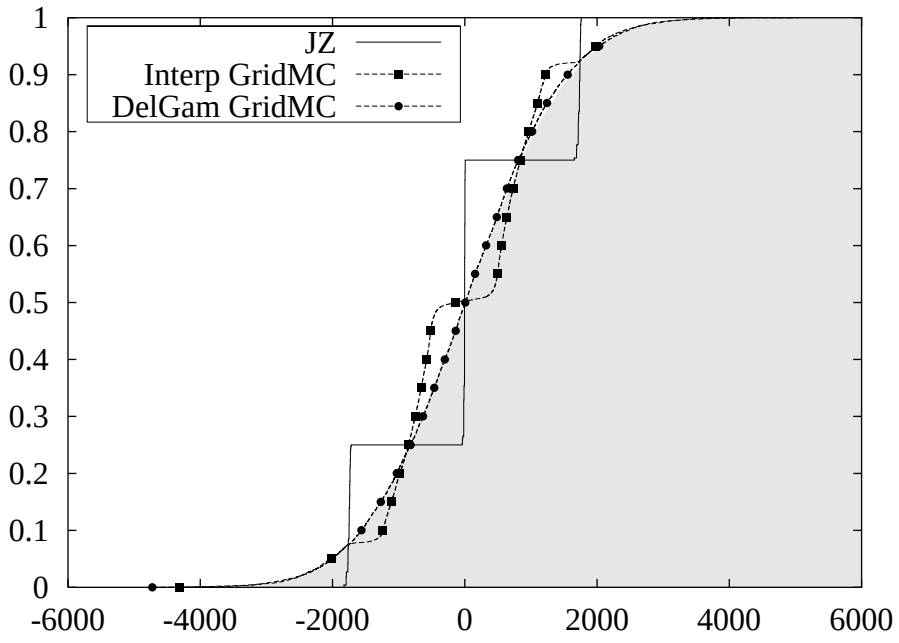
The Interp GridMC method and the DelGam GridMC method both compute prices by combining global information on prices (prices at the nodes of the grid) with a local approximation method such as linear interpolation or a Taylor series. Which of these methods work better will ultimately depend on the properties of the true pricing function. If the pricing function is nearly quadratic in the neighborhood between adjacent grid points, then the Taylor series expansion should perform well. However, if the derivatives of the pricing function are not well behaved, then a method based on linear extrapolation or interpolation may be preferable. Of course, many more sophisticated interpolation methods exist that could be used in this step.

Below, we consider two examples to illustrate the use of the Grid Monte Carlo method. Because our focus in this subsection is on errors resulting from discretizing the probability distribution of the risk factors, we assume here that the risk factors have been chosen well. For simplicity, we do this by simulating yield curves and instrument prices using only the first three PCA risk factors. Four estimators of VaR are contrasted in this section, JZ's method (JZ), both Grid Monte Carlo methods (Interp GridMC and DelGam GridMC), and Monte Carlo with full repricing. The Monte Carlo estimates used 10,000 random draws and the grid-based methods used the grid of factor realizations from Table 1.

Our first example is a \$1 million bond portfolio which has been constructed so that its δ is nonzero for the third risk factor, but zero for factors 1 and 2. The estimated CDFs of the change in portfolio value produced by the different VaR estimators are shown in Figure 14. From the example in Figure 4 we know that JZ's approach cannot fit the CDF of one-day changes in portfolio value for this case with this grid. In this special case, both Grid Monte Carlo approaches come much closer to fitting the true CDF than do JZ's estimator; but there are important differences between the grid-based estimators. In particular, the Delta-Gamma Grid Monte Carlo estimator of the CDF is nearly identical to the true CDF, but the Interpolation Grid Monte Carlo CDF departs significantly from the true CDF. Because the grid methods do not make errors from discretizing the probability distribution, the source of error in the Interpolation method is pure pricing error. These results show that the pricing errors made using linear interpolation can be much larger than those made using Delta-Gamma approximation.[21]

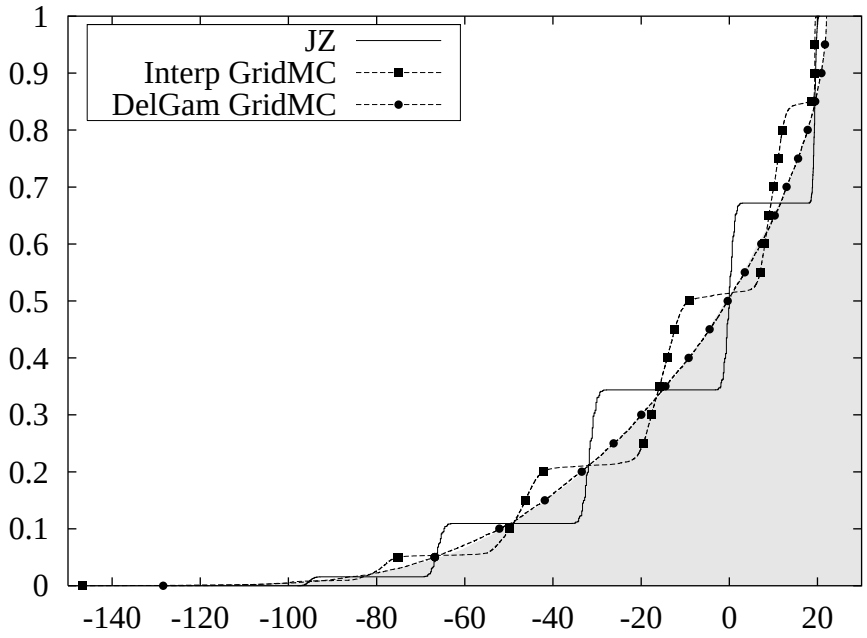
[21]This is not necessarily surprising. Provided the price function is differentiable, in local neighborhoods

Figure 14: Estimated CDFs for the Bond Portfolio



Notes: For a $1 million bond portfolio that is designed to mimic the third PCA factor, the figure presents three grid-based estimates of the cumulative distribution function: one using linear interpolation and Grid Monte Carlo (Interp GridMC), one using a local second-order Taylor series approximation and Grid Monte Carlo (DelGam GridMC), and one using the JZ discretization (JZ). The area under the "true" CDF (estimated by Monte Carlo with full repricing) is shaded for comparison.

Figure 15: Estimated CDFs for the Butterfly Position



Notes: For a butterfly option position with five days to expiry, the figure presents three grid-based estimates of the cumulative distribution function: one using linear interpolation and Grid Monte Carlo (Interp GridMC), one using a local second-order Taylor series approximation and Grid Monte Carlo (DelGam GridMC), and one using the JZ discretization (JZ). The area under the "true" CDF (estimated by Monte Carlo with full repricing) is shaded for comparison.

Our second example is a butterfly spread position which involves buying a put at each of two strike prices, and writing two puts at an intermediate strike price. More specifically, they are European options to sell a 10-year coupon bond with a face value of $1 million; and all three options have five days to expiration. The CDFs for the one day change in value of this position are presented in Figure 15. For this position, most of the probability mass of Jamshidian and Zhu's approximation to the CDF appears to cluster near five points. In this case the Interpolation GridMC estimator, while better, is not a major improvement on JZ. However, the Delta-Gamma GridMC method is much better in this case, nearly matching the true CDF.

VaR estimates at the 1 percent and 5 percent confidence levels for both of these examples

around the grid points the errors from using a delta-gamma approximation will be of a smaller order than from linear interpolation. If the grid points are close enough together delta-gamma approximation may be better than linear interpolation for most factor realizations.

Table 2: VaR Estimates for Grid Monte Carlo

A. Bond Portfolio

Confidence Level	JZ	Interp GridMC	DelGam GridMC	True
0.05	1785	2031	2037	2035
0.01	1827	2879	2887	2884

B. Butterfly Position

Confidence Level	JZ	Interp GridMC	DelGam GridMC	True
0.05	67	71	67	67
0.01	95	85	91	92

Notes: For the bond portfolio and butterfly option position described in section 4.2, the table presents VaR estimates using Jamshidian and Zhu's method (JZ), the Interpolation Grid Monte Carlo method (Interp GridMC), the Delta Gamma Grid Monte Carlo method (DelGam GridMC), and Monte Carlo with full repricing (True). The reported figures are averages obtained from 100 different simulations.

are shown in Table 2. The VaR estimates are averaged across 100 independent simulations to ensure that the results are not dominated by any particular Monte Carlo draw. As anticipated from Figure 14, the JZ method produces very poor estimates for the bond portfolio (panel A) while both Grid Monte Carlo methods produce very accurate VaR estimates.

The performance of the estimators for the butterfly portfolio (panel B) is somewhat different. Even though the CDF for the JZ method does not closely resemble the true CDF, the VaR estimates using the JZ method are quite accurate at the 5 percent confidence level, and pretty accurate at the 1 percent confidence level. The DelGam GridMC is better still, which is consistent with how closely its CDF matches the true CDF. The Interp GridMC performs the worst of all three methods. This suggests that at least in some cases, the pricing errors from using the Interp GridMC method with linear interpolation can be worse than the the probability discretization errors from using JZ's method.

These results suggest that Delta Gamma Grid Monte Carlo generates quite good estimates of the distribution of changes in portfolio value. It can significantly improve the performance of grid-based methods for computing VaR, such as that of JZ. Linear interpolation does not necessarily improve on JZ's method. More sophisticated interpolation methods, which use the properties of the local deltas and gammas, will likely dominate the performance of the two methods that are used here.

5 Conclusions

We have shown that grid-based methods of risk measurement can be improved by choosing risk factors according to their ability to explain the portfolio's value, and by avoiding discrete risk factor distributions which, while simple to use, can be a poor approximation to the true distribution of portfolio value. We conclude with our thoughts on how a grid-based VaR calculation for fixed-income portfolios can be intergrated into a firm-wide risk management process. We make the following recommendations:

Recommendation 1 *Within the class of grid-based methods, the best method for computing firm-wide VaR is Grid Monte Carlo.*

Recommendation 2 *Each trading desk should choose its own risk factors and grid using PLS.*

Recommendation 2 says that each trading desk should choose its own risk factors using PLS. The alternative would be for all trading desks to use a common set of risk factors. (This would be necessary if a JZ-style discretization were to be used to calculate firm-wide VaR.) We have shown that, if different desks take on different types of risk, a common set of risk factors may seriously understate some desks' risk. (This is true even if the common set of risk factors is chosen by PLS.) Once each trading desk chooses its own risk factors for risk measurement, using a JZ-style discretization to compute VaR is not possible. This, along with the potential for discretization errors to be large, leads to recommendation 1, to use Grid Monte Carlo to compute firm-wide VaR.

These recommendations imply a division of labor between the firm-wide risk manager and the desk-level risk managers. The firm-wide risk manager is responsible for generating Monte Carlo draws for the underlying yield curve factors, sending the yield curve draws to the trading desks, and aggregating the results. Yield curve draws can be generated in whatever way is thought most appropriate (parametric modeling, historical simulation, etc.).

The desk-level risk manager is responsible for modeling the response of the desk's portfolio to an arbitrary shock in the underlying yield curve factors. The most appropriate modeling technique will depend on the portfolio, of course. One technique would be to break the portfolio down into simple building blocks for which analytical pricing formulas are available. For many portfolios containing complex derivatives, this will not be possible. For these, our results suggest that Grid Monte Carlo with the factors chosen by PLS would be a reasonable modeling technique. Once the yield curve draws are received at the trading desk, they must

be mapped into the appropriate PLS risk factors. Then, the change in the value of the desk's portfolio can be estimated and transmitted back to the firm-wide risk manager.

References

Abken, Peter A., 2000. An empirical evaluation of VaR by scenario simulation. Office of the Comptroller of the Currency, E&PA Working Paper 2000-3, March 2000.

Chishti, Adnan, 1999. Simulation of fixed-income portfolios using grids. *Algo Research Quarterly* 2:2 (June), 41–50.

Frye, Jon, 1996. Principals of risk: Finding Value-at-Risk through factor-based interest rate scenarios. Working paper, NationsBanc-CRT (June).

Garthwaite, Paul H., 1994. An interpretation of partial least squares. *Journal of the American Statistical Association* 89:425, 122–127.

Golub, Bennett W. and Leo M. Tilman, 1997. Measuring yield curve risk using principal components analysis, Value at Risk, and key rate durations. *Journal of Portfolio Management* 23:4, 72–84.

Hill, Charles F. and Simon Vaysman, 1998. An approach to scenario hedging. *Journal of Portfolio Management* 24:2, 83–92.

Jamshidian, Farshid and Yu Zhu, 1997. Scenario simulation model: Theory and methodology. *Finance and Stochastics* 1:1.

Kambhu, John, 1998. Dealers' hedging of interest rate options in the U.S. dollar fixed-income market. *Federal Reserve Bank of New York Economic Policy Review* 4:2 (June), 35–57.

Pritsker, Matthew, 1997. Evaluating Value at Risk methodologies: Accuracy versus computational time. *Journal of Financial Services Research* 12:2/3, 201–242.

Stone, M. and R. J. Brooks, 1990. Continuum regression: Cross-validated Sequentially Constructed Prediction Embracing Ordinary Least Squares, Partial Least Squares and Principal Components Regression. *Journal of the Royal Statistical Society*, Ser. B, 52:2, 237–269.

A Partial Least Squares Algorithm

The algorithm we use for computing PLS factors is from Garthwaite (1994). It is summarized here for the convenience of interested readers.

We want to estimate the relationship between a sample of n yield curve shocks and the value of a portfolio. Let $x_1, \cdots, x_m$ be $n \times 1$ vectors of shocks to m points on the yield curve. Let y be the $n \times 1$ vector of the change in portfolio value under each yield curve shock.

Center the y and x_i around their means $\bar{y}$ and $\bar{x}_i$ as follows:

$$\begin{aligned} u_1 &= y - \bar{y} \\ v_{1i} &= x_i - \bar{x}_i \ \ \forall i = 1, \cdots, m \end{aligned}$$

To construct the first PLS factor, follow these steps.

1. One at a time, regress each of $v_{11}, \cdots, v_{1m}$ on u_1.
2. Call the estimated regression coefficients $\hat{\beta}_{1i} = v_{1i}' u_1 / (v_{1i}' v_{1i})$.
3. The first PLS factor, t_1, is constructed as the simple average of the predicted values from the m regressions:

$$t_1 = \frac{1}{m} \sum_{i=1}^{m} \hat{\beta}_{1i} v_{1i} \tag{1}$$

To construct the second PLS factor to be orthogonal to the first, we need to isolate the variation in y and $x_1, \cdots, x_m$ that is not in t_1. Define v_{2i} as the residual from a regression of v_{1i} on t_1. Define u_2 as the residual from a regression of u_1 on t_1. Using these newly constructed regressors, repeat steps 1–3 from the preceding paragraph. (Replace v_{1i} with v_{2i}, u_1 with u_2, β_{1i} with β_{2i}, and t_1 with t_2.) This defines the second PLS factor t_2.

Repeat these steps to construct the third and subsequent PLS factors.