Value at Risk for Interest Rate-Dependent Securities:
A Nonparametric Two-Dimensional Kernel Approach

Nusret Cakici

and

Kevin R. Foster

Economics Department
City College of New York
Convent Avenue at 138th Street
New York, NY 10031
Tel: (212) 650-6201
email: ncakici@yahoo.com, kfoster@ccny.cuny.edu

Value at Risk for Interest Rate-Dependent Securities:
A Nonparametric Two-Dimensional Kernel Approach

This paper shows how to calculate Value at Risk (VaR) for interest rate dependent securities, using an extension of the nonparametric estimator in Stanton [1997]. The method uses a two-dimensional kernel with adjustable bandwidth to model the risk as it changes with the level of interest rates. Since the variance, skewness, kurtosis, and higher moments of the distribution of interest rate changes are allowed to vary nonparametrically with the level of rates, the model produces dynamically changing boundaries of risk for VaR. These risk boundaries are based only on the data and are independent of a modeler's discretion. These VaR bounds are shown to be significantly better measures by a variety of criteria. The hit ratios using the model presented here are generally more accurate than a GARCH specification in backtesting.

Value at Risk for Interest Rate-Dependent Securities:
A Nonparametric Two-Dimensional Kernel Approach

VaR modelers face difficult tradeoffs in choosing between the two major VaR methodologies: historical/Monte Carlo simulations or parametric models.[1] The historical approach uses the recent history of the asset price (or, for derivatives, the price of the underlier) while the parametric approach imposes functional form assumptions upon the expected returns and uses history only to estimate the few governing parameters. Each model has its own limitations. The parametric model gains computational ease at the cost of unrealistic assumptions upon the forward distribution, often even assuming normal errors despite a long history of demonstrated abnormal tail behavior (Mandelbrot [1963]). The historical model suffers the "rewind" problem: there is no guarantee that the future will replay the past in the same way. Randomization (Monte Carlo) experiments mitigate this problem but at computational cost. Duffie and Pan [1997] and Linsmeier and Pearson [2000] give excellent overviews.

Interest rate modeling has allowed for a wide degree of nonparametric flexibility in how the forward distribution of rate changes depends on the level. Whether the classic Cox, Ingersoll, and Ross (CIR, 1985) specification that volatility increases with the square root of the level, or the nonparametric Aït-Sahalia [1996a] and Stanton [1997] models, it is widely believed that this relationship is important in accurately representing interest rate changes. A VaR calculated by ignoring this dependence of volatility on the level of rates would miss this essential feature. Parametric methods that explicitly model this dependence, such as CIR, assume normal errors (actually, non-central chi-squared in the discrete case, although this is built from Wiener processes) and so understate the actual kurtosis of the distribution, leading to inaccurate tail probability and VaR estimates.

Contemporary nonparametric methods for valuing derivatives, such as Stutzer [1996] and Stanton [1997], have made significant forays in option pricing but are only just

beginning to find prominence in VaR analyses. These can accurately reflect the thick tails, pronounced skewness, and excess kurtosis of financial asset price returns. The use of kernel density estimation can model this tail behavior in flexible yet workable ways. Stutzer provides a method of constructing a risk-neutral probability distribution using the maximal amount of information from the historical record of the underlier. Replacing the conventional log-normal assumption with this risk-neutral historical distribution allows a rich variety of effects on option values. Rather than squeezing the empirical variation into a single volatility parameter, Stutzer allows an entire dimension of variation to enter nonparametrically -- but only a single dimension, while interest-rate dependent securities need more.

While keeping the desirable nonparametric features of Stutzer and Stanton, the method demonstrated in this paper allows an extension of the historical record to two dimensions, to include the characteristic dependence of rate changes upon the present level of interest rates. This approach is not just a nonparametric stochastic volatility model but allows all of the higher moments to vary, so it may be termed a stochastic higher moments kernel model. The multidimensional model suitable for VaR analysis is termed the Historical Higher Moments VaR (HHM-VaR). This method allows a tremendously flexible, completely nonparametric dependence of the distribution of asset price changes upon the level of interest rates. This method uses a normal kernel with adjustable bandwidth (see Chapman and Pearson [2000] for the importance of bandwidth choice).

Between the twin perils of over-parameterization or historical "rewind" we show that this method is both feasible and desirable. Many competing measures of VaR suffer from additional "Model Risk" -- the possibility that the model is mis-specified and so the VaR bounds are incorrect. The HHM-VaR necessitates very few assumptions about the driving processes, requiring little prior information, which makes it an ideal "plug-in" when comparing to other models. In many applications, a completely nonparametric baseline model that makes very few assumptions about implied future rate paths can form an enlightening comparison.

An attractive feature of the HHM-VaR is that it does not assume symmetry in the error distributions. The asymmetry is determined entirely by the data on the underlying security. Any asymmetry in the empirical returns is important to a risk manager, since it means that long and short positions do not cancel each other out by a VaR risk measure. Our method can be extended to evaluate portfolios of bonds, although adding additional factors may exacerbate worries of over-fitting. While there is debate about whether the Aït-Sahalia and Stanton models go too far in giving the maximum voice to the data (see Chapman and Pearson), we believe that this HHM-VaR gives an interesting description of the risk that faces a particular portfolio. No portfolio manager should believe that there is a single true VaR somewhere, but rather a variety of different measures (based on particular statistical representations) that can be used to triangulate the true risk facing the position.

The performance of the HHM-VaR is tested against two alternate specifications: a simple rolling normal VaR and a restricted GARCH(1,1) (see Vlaar [2000]). The historical higher moments method of constructing VaR boundaries is indistinguishable from its theoretical value in nineteen of the thirty comparisons, while the GARCH is indistinguishable in just twelve. The HHM-VaR produces hit ratios that are generally more accurate than a GARCH specification, and even where the HHM-VaR is somewhat less accurate, it is never as far off as the GARCH can be. The worst miss for the HHM measure is 0.65 percentage points. The GARCH is at least that far away in eight of thirty subgroups and as far away as 1.67 percentage points.

Section 2 of the paper details the model of the joint distribution of the level and change of interest rates, with the historical higher moments kernel that generates the VaR. Section 3 explains the data used and displays the estimated VaR bounds. Section 4 presents summary statistics of how accurate those VaR boundaries are. Section 5 concludes and suggests avenues for further research.

2. Joint Distribution of r_t and Δr_{t+1}

The modeling of the dependence of the distribution of rate changes as a function of the level extends at least back to the classic Cox, Ingersoll, and Ross [CIR 1985] paper.

3

This model posits a Brownian motion process whose diffusion parameter (variance) is itself a function of the level, in addition to the standard dependence of the drift upon the level. The model, with a linear error-correction term for the drift and a diffusion proportional to the square root of the level, is:

$$dr = \boldsymbol{k}\left(r - \boldsymbol{q}\right)dt + \left(\boldsymbol{s}\sqrt{r}\right)dW \tag{1}$$

where $\boldsymbol{k}$ is the rapidity at which deviations from the long-term rate, $\boldsymbol{q}$, are corrected and $\boldsymbol{s}\sqrt{r}$ is the diffusion. CIR provide an exact solution so that the discrete-time difference of this continuous-time interest rate process is distributed as a non-central chi-square.

Aït-Sahalia's [1996a] paper made an important generalization by specifying the diffusion as depending nonparametrically upon the level of interest rates. However that specification kept the assumption of a linear error-correction factor, so that the rate process is given as:

$$dr = \boldsymbol{k}\left(r - \boldsymbol{q}\right)dt + \boldsymbol{s}\left(r\right)dW \tag{2}$$

Finally, Stanton [1997] gives a further generalization by allowing the drift process to also depend nonparametrically upon the level, so:

$$dr = \boldsymbol{m}(r)dt + \boldsymbol{s}\left(r\right)dW \tag{3}$$

so both $\boldsymbol{m}(r)$, the drift, and $\boldsymbol{s}\left(r\right)$, the diffusion, can be flexibly estimated. The model is criticized by Chapman and Pearson [2000] for making a low power estimation at very high levels of interest rates. This low power estimation will make the VaR bounds too tight at extreme values, so backtesting will reveal if this problem is important.

The empirical literature has run parallel to the theory, with various proposals to estimate or calibrate model parameters. Traditional methods of recovering the structure of interest rate changes from the historical record run regressions of the general form, $\Delta r_{t+1} = g(r_t, \boldsymbol{e}_t)$, where $\Delta r_{t+1} = r_{t+1} - r_t$ and the function $g(r_t, \boldsymbol{e}_t)$ may be as simple as an error correction model or may be expanded to include various GARCH specifications of the error term, lags of either r_t or e_t, or further complications (Chan, Karolyi, Longstaff, Sanders [1992]). A few approaches, such as that of Conley, Hansen, Luttmer, and Scheinkman [1997], address the identification issues forced by the discrete nature of the data collected.

The regression coefficients represent the central tendency only and do not allow for a heterogeneity of effects. Methods such as kernel regressions can allow a richer variety of effects, but these are still methods of summarizing the rich variation of the joint distribution, $f(r_t, \Delta r_{t+1})$, and compressing it into a few parameters. But is there a need for such compression? The joint distribution can be estimated directly and all of the variation can be employed, without the need for compaction, to simulate bond and option prices. Stutzer [1996] used this basic approach in the univariate case. We generalize to higher dimensions to estimate the joint distribution of the change in the interest rate, Δr_{t+1} and the rate itself, r_t. This joint distribution is estimated with a two-dimensional kernel.

A kernel estimate is basically like a histogram, but where a histogram stacks boxes to represent each data point, the kernel stacks "bumps" and a normal kernel uses Gaussian "bumps". The width of the "bump" (for the normal kernel, the variance) is controlled by the bandwidth choice, h. The two-dimensional kernel constructs the empirical probability density function $\hat{f}(r_t, \Delta r_{t+1})$, which will, under appropriate regularity conditions, converge to the true probability density function, $f(r_t, \Delta r_{t+1})$.

The kernel estimator will be appropriate, as long as the generating distribution is stationary, ergodic, and the data generating process satisfies some relatively weak assumptions of stochastic equicontinuity in order for a Uniform Law of Large Numbers result to obtain (Andrews [1994]). Stochastic equicontinuity of an empirical process, $H_T(\boldsymbol{q})$, is defined as

$\forall \boldsymbol{e} > 0$ and $\forall \boldsymbol{h} > 0$, $\exists \boldsymbol{d} > 0$ such that

$$\overline{\lim_{T \to \infty}} P\left[\sup_{\boldsymbol{q}, \boldsymbol{q}' \in \Theta, r(\boldsymbol{q}, \boldsymbol{q}') < \boldsymbol{d}} \left\| H_T(\boldsymbol{q}') - H_T(\boldsymbol{q}) \right\| > \boldsymbol{h} \right] < \boldsymbol{e} \ . \tag{4}$$

There is a tradeoff between the smoothness required of the function and the dimension of the random variable, here two. The kernel estimator at each point is constructed as

$$H_T(x) = \frac{1}{h^2 T} \sum_{i=1}^{T} K\left(\frac{x - X_i}{h} \right) \tag{5}$$

where K is the normal (Gaussian) kernel. A normal kernel does not limit the estimated p.d.f. to be normal. Since Chapman and Pearson have demonstrated the importance of bandwidth choice in estimating the drift and diffusion functions, we use an adjustable bandwidth kernel. This adjustable bandwidth estimator of the kernel density can be seen as a version of their recommended least squares cross-validation bandwidth choice for kernel regression functions. The adjustable bandwidth starts with a pilot estimate of the density. This pilot estimate uses the i.i.d. bandwidth,

$$h = \hat{s}T^{-1/6} \tag{6}$$

where $\hat{s}$ is the standard deviation of the sample and the two variates are pre-whitened to have the same variance. This bandwidth choice will be acceptable for a wide range of functional forms. Then the pilot estimate of the density is used to scale the bandwidth for the calculation of the adjustable-bandwidth kernel, so that areas of high density use a small bandwidth while areas of low density use a higher bandwidth choice. Following Silverman [1986], we define local bandwidth factors, l_i, by

$$l_i = \left[\tilde{f}(X_i)/g\right]^{-a} \tag{7}$$

where the pilot estimate is $\tilde{f}(X_i)$ g is the geometric mean,

$$g = \exp\left\{\tfrac{1}{T}\sum \tilde{f}(X_i)\right\} \tag{8}$$

and a is the sensitivity parameter, which here is 0.5.

Then the adaptive bandwidth kernel is

$$\hat{f}(x) = \sum_{i=1}^{T} \frac{1}{h^2 l_i^2 T} K\left(\frac{x - X_i}{h l_i}\right) \tag{9}$$

where K is again the normal (Gaussian) kernel. The kernel is estimated on a 500x500 grid and converted to a conditional cumulative distribution at each discretization of r_t. Each conditional density is allowed to be different, for different levels of the interest rate, r_t.

Calculating the Value at Risk interest rate is straightforward: the percentiles of the conditional kernel density are calculated at the gridpoints and the actual VaR is interpolated. Specifically, to find the Value at Risk interest rate, we calculate a value, $V(r_t)$, such that there is probability of a that the interest rate t days from now is greater than it:

$$\Pr\left\{r_{t+t} > V\left(r_t\right)\right\} = a \tag{10}$$

or for the other tail, $V'\left(r_t\right)$ such that

$$\Pr\left\{r_{t+t} < V'\left(r_t\right)\right\} = 1 - a . \tag{11}$$

We calculate results for the ten-day ($t = 10$) and one-day ($t = 1$) cases.

3. The Stochastic Kernel VaR Model

The data on interest rates incorporate a variety of maturities, downloaded from the Federal Reserve website. The one-month and three-month Eurodollar interest rates are each business day from January 4, 1971 to February 15, 2002.[2] The short Eurodollar rates were converted to a continuously compounded yield-to-maturity basis. The longer-maturity rate series are on constant maturity Treasury bonds, beginning January 2, 1962. The paper examines the one-year, five-year, and ten-year rates. With a kernel specification, we use the maximum amount of data that is available. Other models such as the GARCH seem to generate results that depend crucially upon the exact time period used, and therefore upon the discretion of the particular modeler. If a risk manager wants an analysis that is not so dependent upon these idiosyncrasies, then the kernel method may be better.

The first four moments of a kernel estimate of the one-day change in the short-term (one-month) rate are shown in Exhibit 1 along with 90% error bands. Exhibit 2 shows the moments of the ten-day changes. In both it is clear that while the variance increases with the level of the interest rate, the higher moments also change. Clearly the variance increases with the level of interest rates, with a suggestion of either a break or perhaps a parabolic shape, as found by Aït-Sahalia's [1996b] study of interest rate volatility. These changing moments do not appear to have a straightforward parametric form, suggesting that any simple model would miss important characterizations of the evolution of rate changes.[3] A simple VaR calculation based on normally distributed errors would have to assume that the higher moments are constant, even if it accounted for the changing variance, and so lead to important inaccuracies in the estimation of the tail probabilities. Exhibit 3 shows the moments of the ten-day changes for the ten-year rate[4]: the level of interest rates has an important influence on all the conditional

moments. The strong negative skew will influence the VaR bounds calculated below. Again, it is clear that a proper VaR model should take account of these effects.

The estimated kernel, at each discretization point of the interest rate level, is converted to a cumulative density and the appropriate percentiles are linearly interpolated. These percentiles represent the VaR bounds, either as levels of interest rates or as corresponding bond prices. Since the distribution of rate changes is allowed to vary depending on the level, the VaR bounds will also dynamically change depending on present conditions.

These dynamically changing VaR bounds, for a ten-day-ahead prediction period, are shown in Exhibit 4. This figure shows the one-percent and five-percent VaR bounds (in each direction, as well as the median) for the one-month Eurodollar and the ten-year Treasury (constant maturity) interest rates. The shorter rate percentile bounds are more sensitive to the level, showing a "waist" of narrowed volatility for rates from five to seven percent. For a one-day prediction period, the figures look substantially similar, although with tighter bounds (these figures, as well as the results for other maturities, are omitted but are available from the authors).

The asymmetry in the bounds is particularly important, but is difficult to see in Exhibit 4 and better views are shown in the next figures. These figures are intended to show the intuition of how the bounds depend on the level of interest rates. The asymmetry may be crucial in evaluating the risk of a portfolio since equal short and long positions will not carry equal levels of risk -- the skewness and kurtosis (and higher moments) of the forward distribution can change the risk. While a GARCH model can be specified that allows this asymmetry (see Campbell and Hentschel [1992] and Wu [2001]), the choice of how to specify the asymmetry adds further discretion to the GARCH modeler's risk analysis. The HHM-VaR handles this asymmetry naturally and flexibly, without undue dependence upon the modeler's choices. Rather the data determine the nature and extent of the dependence of skewness and kurtosis on the level of rates.[5]

Exhibit 5 shows the actual value at risk, for a bond with a given maturity and a current value of 100 (i.e. a face value of $100\exp\{rT\}$). This value at risk is calculated as:

$$VaR_a = \left(100\exp\{(r - \hat{r}_a)T\}\right) - 100 \tag{12}$$

where r is the present level of the interest rate, T is the time to maturity, and $\hat{r}_a$ is the critical a-percentile bound. Exhibits 5, 6, and 7 show the five-percent value at risk over a ten-day horizon for a bond with a $100 current value and maturity changing from one month to three months, one year, five, and ten years. (The figures for other percentiles of value at risk are similar.) In each figure, two values at risk are shown: the solid line shows the VaR for a long position (if interest rates rise) while the dashed line shows the VaR for a short position (if rates fall). With this simple model of the interest rate, more complex derivatives can be straightforwardly modeled.

The top panel of Exhibit 5 shows the short and long VaR on one-month Eurodollar bonds over a standard ten-day interval depending on today's interest rate (on the x-axis). The dashed line is the VaR for a short bond position, representing the chance that rates will fall; the solid line is the VaR for a long bond position. At low levels, the one-month leaves only a few pennies at risk however as rates rise this value at risk rises steeply so that at interest rates above 10% the VaR is at least four times higher. This should convince any risk manager that the dependence of forward rate changes upon the present level is important even for very short-maturity instruments and should be carefully modeled. At these short rates, there is little pronounced skew to push the bounds apart. The three-month rates, in the bottom panel of Exhibit 5, have a similar shape, although it more dramatically shows that not only does the VaR liftoff at rates above 9% but for rates from 4% to 6% the VaR may actually fall.

Exhibits 6 and 7 show the Treasury bond rates at longer maturities. The one-year is similar to the short instruments, while the longer maturities show a marked pivot of the long and short VaR. There is clearly more VaR from a short bond position when interest rates are high, while it is smaller when rates are low. The pivot point, for both the five-year and ten-year, is about 7%. For ten-year bonds, at high interest rates the Value at Risk of a short position is almost $2 more than the VaR of a long position. This difference is caused by the skewness in the distribution, which was clearly shown in

Exhibit 3. This skew is an important characterization of the temporal evolution of long-maturity interest rates yet a model based on normally distributed errors would overlook it entirely. The skew means that a portfolio with balanced short and long positions would not be riskless by the VaR criterion.

4.Evaluating Performance of Stochastic Kernel VaR Models

Exhibits 5, 6, and 7, showing the evolution of the Value at Risk as a function of the level of interest rates, give an overview of how these bounds are set, but they must be accurate predictors of the VaR over time as well. To test the performance of these HHM-VaR bounds, we backtest the ability of the HHM-VaR to realistically characterize the risks facing the bondholder. The performance of the HHM-VaR is tested against two primary alternate specifications: a simple rolling normal VaR and a restricted GARCH(1,1).

For the rolling normal VaR, the last T datapoints are used to imply the standard deviation of rate changes going forward if they were normally distributed. In this study we use $T=100$. This backward-looking window is obviously not a predictor that would be based on rational expectations, and of course the window of how many previous datapoints are used may change depending on a particular analyst's view of market conditions, but we intend the rolling backward-looking VaR to be a useful benchmark rather than the best possible measure.

The GARCH(p,q) model is based on the conventional specification of the evolution of the conditional variance, $\boldsymbol{s}_t^2$, of rate innovations, that

$$\boldsymbol{s}_t^2 = \boldsymbol{k} + \sum_{i=1}^{p} \boldsymbol{g}_i \boldsymbol{s}_{t-i}^2 + \sum_{j=1}^{q} \boldsymbol{a}_j \boldsymbol{e}_{t-j}^2 \tag{13}$$

As often in VaR modeling, our basic specifications have found little improvement whether the parameters are estimated or specified *a priori* to the RiskMetrics figures: $\boldsymbol{k} = 0$, $\boldsymbol{g} = 0.94$, $\boldsymbol{a} = 0.06$, and $p = q = 1$.[6] The restriction is intended to reduce the risk of over-fitting within sample. This GARCH specification, however, does not take into account the predictable effects of the level of interest rates upon the variance and higher moments. These effects are predictable and (in other strands of literature) commonly

accounted for. The innovation of this paper lies in showing a model that explicitly but flexibly takes this dependence seriously.

The typical analysis of the performance of a VaR measure counts the number of times the stated bound was hit. However, gauging the accuracy of VaR numbers is difficult without a well-specified loss function. While the "hit ratio," the number of times the actual loss is greater than the predicted VaR, is important, the costs of setting too high or too low a VaR are unclear. A model that tightened up the VaR without sacrificing the hit ratio would still be an improvement (see Christoffersen [2001] for a discussion of possible loss functions). The following analysis will assume that misses of the VaR in either direction are penalized.

The comparisons of hit rates for a ten-day window are in Exhibits 8-12, with each table showing the results for a different maturity. The hit rates for a one-day window are substantially similar and so are omitted. The tables show both the lower tail of rates that were lower than predicted and the upper tail of rates that were higher than predicted with bounds set for levels of a at 1%, 2.5%, and 5%. While an a of 1% is most common, we present the results for a broader range of categories, since we believe that a risk manager would not consider a VaR estimate that hit the 1% level in backtesting at the expense of inadequate power at other percentiles. So to check the robustness of the model, the higher percentiles are also shown. An important feature of the HHM-VaR is that it does not assume symmetry in the error distributions. Therefore this paper presents the tails separately instead of hiding behind composite or two-tailed tests. Portfolios of long and short positions in bonds and derivatives will generally necessitate an accurate idea of the risk of both up and down movements in rates.

Each pair of rows gives the percentage and the actual number of "hits" during backtesting, when the specified bound was breached. The first column shows the theoretical percent and number, which is what each method is trying to approximate. The second column gives the HHM-VaR hits, the third has the hits for the rolling normal VaR, and the fourth column gives the hits for the restricted GARCH(1,1)

specification.[7]   The rolling normal statistics are given to provide a baseline for how much information any model is adding, in order to better assess the marginal contribution.

An asterisk indicates if the model's VaR hit ratios are statistically indistinguishable at a 95% level (see Kupiec [1995] and Christoffersen [1998]) from the true value. The HHM-VaR measure is statistically indistinguishable from the true value in nineteen of the thirty maturity/risk cases while the GARCH only meets this standard in twelve. In six cases of the thirty, neither measure falls within a 95% confidence interval around the true value.

Exhibit 8 shows the estimates for the 10-day VaR of one-month Eurodollar rates. The one percent risk on the lower tail of interest rates should only be breached 73 times during the back-testing period (=1% of 7329). The HHM-VaR sets a too-conservative bound that is breached just 33 times which is just 0.45% of the periods, while the GARCH sets a closer bound that is breached 106 times, at 1.45%. The rolling normal is clearly worse: it breaches the bound 440 times, which is 4.57%. The one percent bound for the upper tail set by the HHM-VaR is broken a conservative 0.78% (57 times). This is not significantly different than the theoretical 1% value, unlike the 2.44% that the GARCH bound is broken (179 hits).

For the 2.5% level, the GARCH is a clear winner for the lower tail, setting boundaries that are hit 189 times compared with 183 "true" while the HHM is again conservative and sets bounds that are hit only 143 times (1.95%). On the other tail, however, the HHM-VaR is better: 208 hits (2.84%) are tallied rather than the GARCH's 295 hits (4.03%). At the five percent level, the HHM and GARCH bounds are essentially tied for estimating the lower bound but the upper bound is set significantly tighter by the HHM-VaR. If an analyst set a level of rates, above which only five percent of the ten-day windows would rise, it should be hit 366 times. The HHM-VaR hits such a level 396 times (5.40%, which is indistinguishable from 5%) while the GARCH sets a level that gets 491 hits (6.70%) -- clearly much worse.

This relative performance of the two chief measures holds for other VaR bounds and other maturities: since the GARCH specifies symmetric conditional errors, it cannot account for the skewness of rate changes. One boundary is typically set well by both models but the GARCH may miss the other boundary by as much as a full percentage point. In no test at any maturity/risk partition did the HHM-VaR miss by over 0.65 percentage points. As shown in the figures, since skewness and excess kurtosis are persistent characteristics of the data and are of varying intensity depending on the level of the interest rates, any quantification of the risk of a portfolio must take account of these effects. Since these effects are predictable consequences of the level of the interest rates, the GARCH model where heteroskedasticity is a function of lagged deviations must struggle to keep up.

Exhibit 9 shows the same comparison for the three-month Eurodollar rates. These show that even when the HHM-VaR misses, it is rarely far. The worst performance of the HHM-VaR of all the targets for the three-month bonds, where the one-percent lower tail should be hit 79 times but the conservative HHM bound is hit only 31 (0.39%), is 48 misses. The three-month GARCH model misses by at least 48 in three of six maturity/risk comparisons. In every comparison at each maturity, the HHM-VaR is either better or only slightly worse, while the GARCH is at times far from the true value. Sometimes the HHM-VaR and GARCH targets are on opposite sides of the "true" level, with one a bit too conservative and the other a bit too lax. This suggests that, while risk managers might not throw away GARCH-set VaR targets, the HHM-VaR may be another useful tool to gauge the riskiness of portfolios.

The comparisons for the one-year constant maturity Treasury bond interest rates are in Exhibit 10. The HHM measure is statistically indistinguishable from the true value at the higher risk levels, and at the one percent level the GARCH wins for the lower tail. In four of the six comparisons the HHM is so close to its theoretical value as to be statistically indistinguishable; in three of six the GARCH is very close. Since these longer-maturity interest rates are less volatile, there is a smaller difference between the GARCH and HHM-VaR performance.

In Exhibit 11 we find the evaluations for the five-year constant maturity Treasury bond rates. The two measures trade places for best at the one-percent level: the GARCH is better for the lower tail and the HHM-VaR for the upper. In all the remaining categories the HHM-VaR is indistinguishable from the theoretical value. The GARCH misses the lower tails of both the 2.5% and 5% as well.

The ten-year maturity comparisons are in Exhibit 12. For each risk level, the HHM is adequate for the upper tail but misses significantly in the lower tail. The GARCH is quite good for the 2.5% level and the upper tail of the five-percent. In both the 2.5% and 5%, the GARCH and HHM miss conservatively while the one-percent bound is breached too often with the GARCH but still set conservatively by the HHM-VaR.

To summarize the results of the back-testing: the historical higher moments method of constructing VaR boundaries is statistically indistinguishable from the theoretical value in nineteen of the thirty comparisons, while the GARCH is statistically indistinguishable in just twelve. When the GARCH misses, it can miss very wide of the mark. The worst miss for the HHM measure is 65 basis points. The GARCH is more than 65 basis points away in eight subgroups and as far away as 167 basis points. The HHM-VaR is particularly good in comparisons of the shorter maturities, which are more volatile. The two Eurodollar short maturities are the most flattering comparison: the HHM-VaR is not significantly different from the theoretical value in seven of twelve, while the GARCH is not significantly different three times. The three longer maturities show HHM wins in 12 of 18 while GARCH wins in 9.

5. Conclusion

This study shows the importance of constructing VaR bounds that vary dynamically depending on the level of the interest rate. This dependence, which is crucial for interest rate-dependent securities, has been modeled nonparametrically in papers such as Aït-Sahalia [1996a] and Stanton [1997]. The nonparametric method can accurately reflect the thick tails, pronounced skewness, and excess kurtosis of financial asset price returns. An adjustable bandwidth two-dimensional kernel is a flexible yet workable estimator of

the VaR. This allows all the higher moments to vary, not just the first two, so it is not stochastic volatility but stochastic higher moments with a kernel -- the HHM-VaR.

This study shows how to construct one-day and ten-day HHM-VaR bounds that depend upon the present level of interest rates. Since the variance, skewness, kurtosis, and higher moments of the distribution are allowed to vary nonparametrically with the level of rates, the HHM-VaR produces dynamically changing boundaries of risk. These VaR bounds are shown to be significantly better measures by a variety of criteria. The HHM-VaR produces hit ratios that are generally more accurate than a GARCH specification, and even where the HHM-VaR is somewhat less accurate, it is never as far off as the GARCH can be. The worst miss for the HHM measure is 0.65 percentage points. The GARCH is at least that far away in eight of thirty subgroups and as far away as 1.67 percentage points.

The methods of updating are very different between the GARCH and HHM-VaR approaches. The GARCH imposes a smooth relationship, where a sudden shock enters with a particular weight and begins a gradual process where the forward risk measures are changed. The HHM-VaR, however, doesn't impose such smoothness. A shock today that changes the level of the interest rate will immediately move the forward distribution to that implied by the new level, without any "memory" of previous history.

While a careful risk manager may not wish to entirely disregard GARCH-based VaR projections in favor of the HHM-VaR, the results of this study make a pellucid case for the use of kernel-based estimators of dynamically varying moments. The method of Barone-Adesi and Giannopoulos [2001] is another nonparametric attempt to take seriously this dependence upon past return volatility, but our approach for interest rates takes advantage of our knowledge that the distribution of interest rate changes is dependent upon the present level of the rate.

Complications remain, which are specific to modeling interest rates. When evaluating portfolios of bonds the analyst must attend to the co-movements along the yield curve, so further studies must examine these complexities. The basic HHM-VaR methodology

gives an interesting direction for yield-curve modeling. Further studies can be made to determine the utility of kernel methods in accurately describing the rate variation. Generalizing interest rate modeling to include an extreme value approach (Bali [2001]) also seems interesting and promising.

The historical higher moments Value at Risk (HHM-VaR) methodology can improve the estimation and prediction of interest rate percentiles for risk analysis. The use of a kernel estimator can accurately model the way in which the level of interest rates shifts the entire distribution of forward changes, not only variance but all the higher moments as well. This kernel estimator is ideally suited for many VaR applications where a "plug-in" estimator is desired that entails a minimal amount of prior information yet is still flexible and realistic, to eliminate the "Model Risk" of other measures. This kernel estimator is a portmanteau model that can incorporate a wide variety of other possible representations including jump and diffusion processes.

Bibliography

Aït-Sahalia, Yacine. "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64(3) (1996), 527-560.

_____. "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9(2) (1996), 385-426.

Andrews, Donald W.K. "Empirical Process Methods in Econometrics," *Cowles Foundation Discussion Paper* No. 1059, 1994.

Bali, Turan G. "An Extreme Value Approach to Estimating Volatility and Value at Risk," *Journal of Business* (2001).

Barone-Adesi, Giovanni, and Kostas Giannopoulos. "Non-parametric VaR Techniques: Myths and Realities," *Economic Notes*, 30(2) (2001), 167-81.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner. "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52 (1992), 5-59.

Campbell, John Y., and Ludger Hentschel. "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31 (1992), 281-318.

Chapman, David A., and Neil D. Pearson. "Is the Short Rate Drift Actually Nonlinear?" *Journal of Finance*, 55(1) (2000), 355-88.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices," *Journal of Financial Economics*, 61 (2001), 345-81.

Christoffersen, Peter F. "Evaluating Interval Forecasts," *International Economic Review*, 39(4) (1998), 841-62.

Christoffersen, Peter, Jinyong Hahn, and Atsushi Inoue. "Testing and Comparing Value-at-Risk Measures," *Journal of Empirical Finance*, 8 (2001), 325-42.

Conley, Timothy G., Lars Peter Hansen, Erzo G.J. Luttmer, and José A. Scheinkman. "Short-Term Interest Rates as Subordinated Diffusions," *The Review of Financial Studies*, 10(3) (1997), 525-77.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross. "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2) (1985), 385-408.

Duffie, Darrell, and Jun Pan. "An Overview of Value at Risk," *Journal of Derivatives*, 4(3) (1997), 7-49.

Kupiec, Paul H. "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, Winter (1995).

Linsmeier, Thomas J., and Neil D. Pearson. "Value at Risk," *Financial Analysts Journal*, 56(2) (2000), 47-67.

Longin, François M. "From Value at Risk to Stress Testing: The Extreme Value Approach," *Journal of Banking and Finance*, 24 (2000), 1097-1130.

Mandelbrot, Benoit. "The Variation of Certain Speculative Prices," *Journal of Business*, 36(4) (1963), 394-419.

Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

Stanton, Richard. "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, 52(5) (1997), 1973-2002.

Stutzer, Michael. "A Simple Nonparametric Approach to Derivative Security Valuation," *Journal of Finance*, 51 (1996), 1633-52.

Vlaar, Peter J.G. "Value at Risk for Dutch Bond Portfolios," *Journal of Banking and Finance*, 24 (2000), 1131-54.

Wu, Guojun. "Determinants of Asymmetric Volatility," *The Review of Financial Studies*, 14(3) (2001), 838-59.
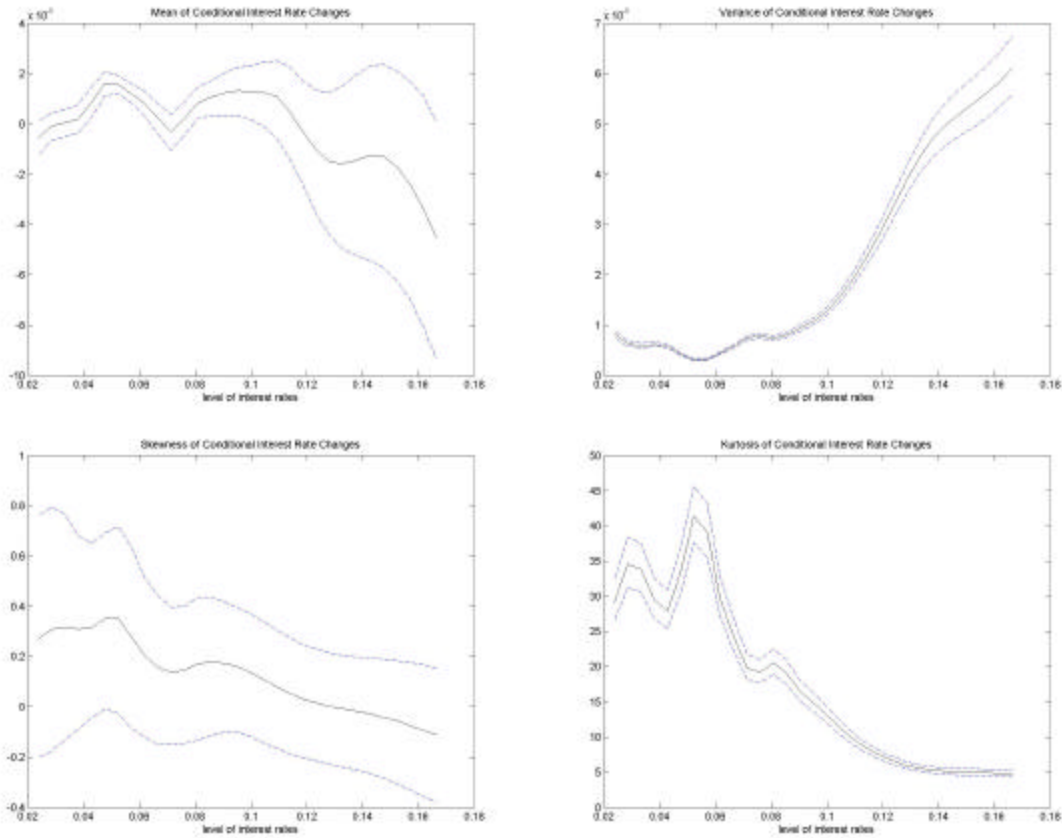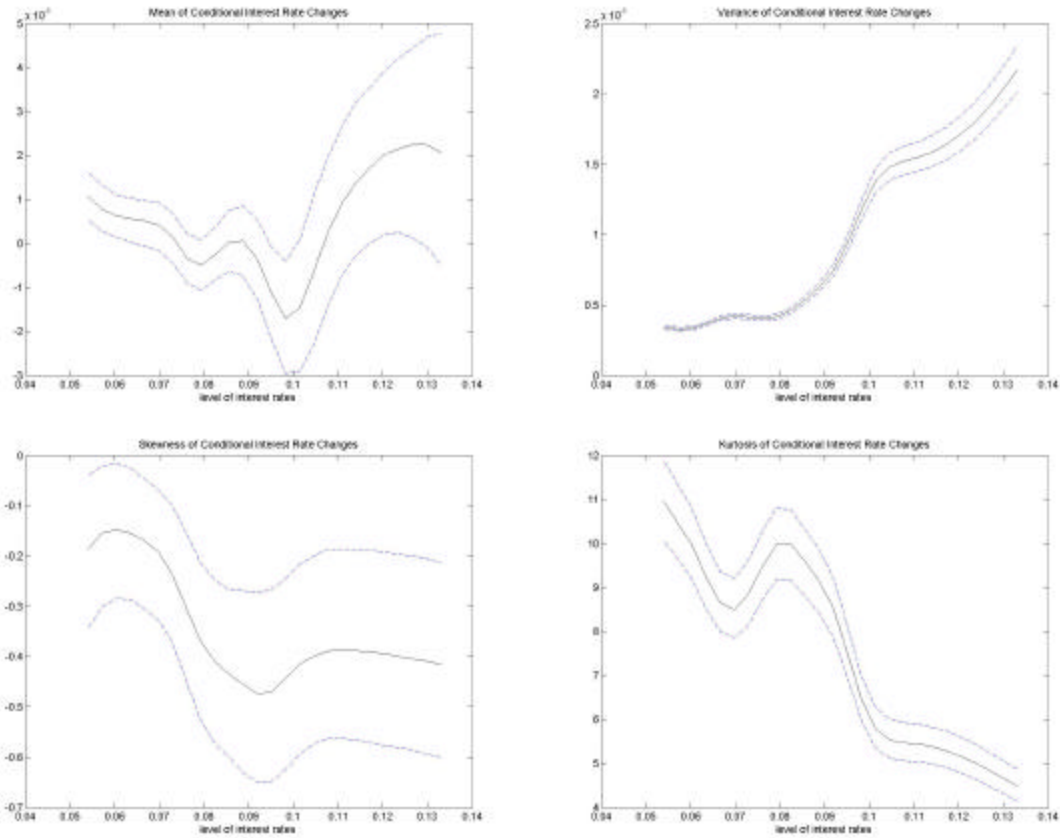
Exhibit 1

First Four Central Moments of One-Day Interest Rate Changes, Conditional on Level, For One-Month Eurodollar Rates.
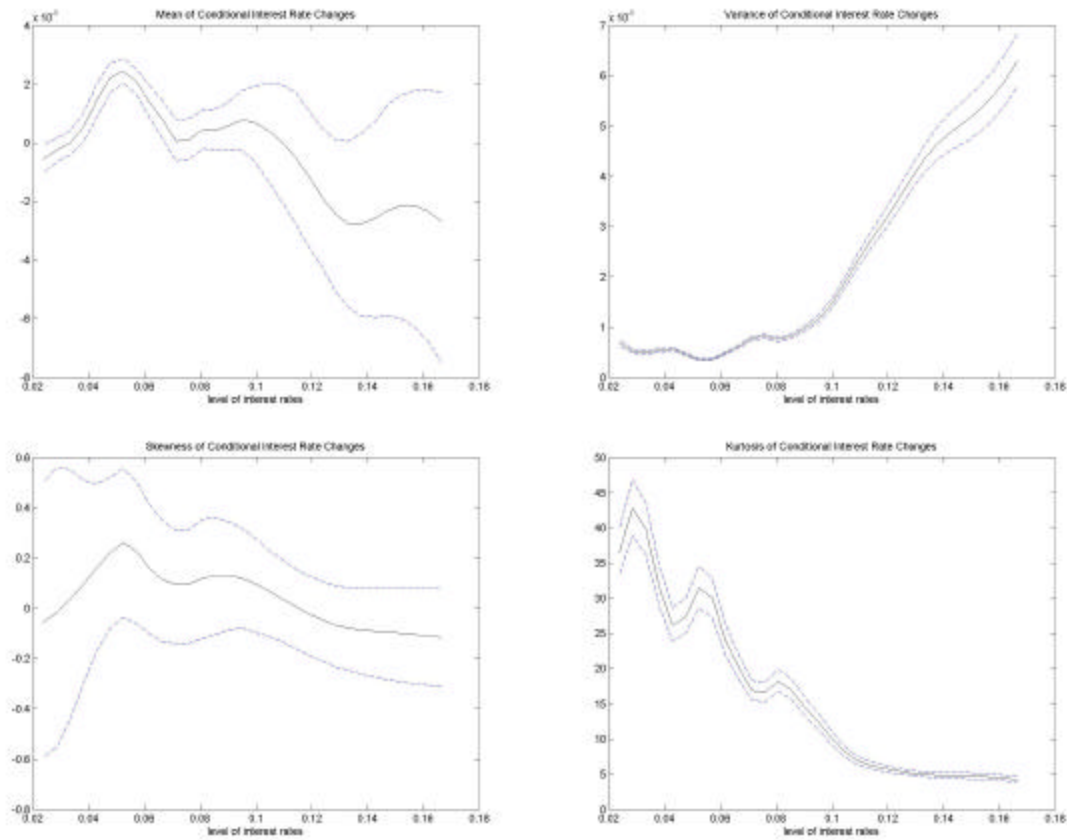
The kernels are estimated with a two-dimensional adjustable bandwidth normal kernel. Each plots the central tendency and 90% error bands.

Exhibit 2

First Four Central Moments of Ten-Day Interest Rate Changes, Conditional on Level, For One-Month Eurodollar Rates.

The kernels are estimated with a two-dimensional adjustable bandwidth normal kernel. Each plots the central tendency and 90% error bands.

Exhibit 3

First Four Central Moments of Ten-Day Interest Rate Changes, Conditional on Level, For Ten-Year Treasury Rates.
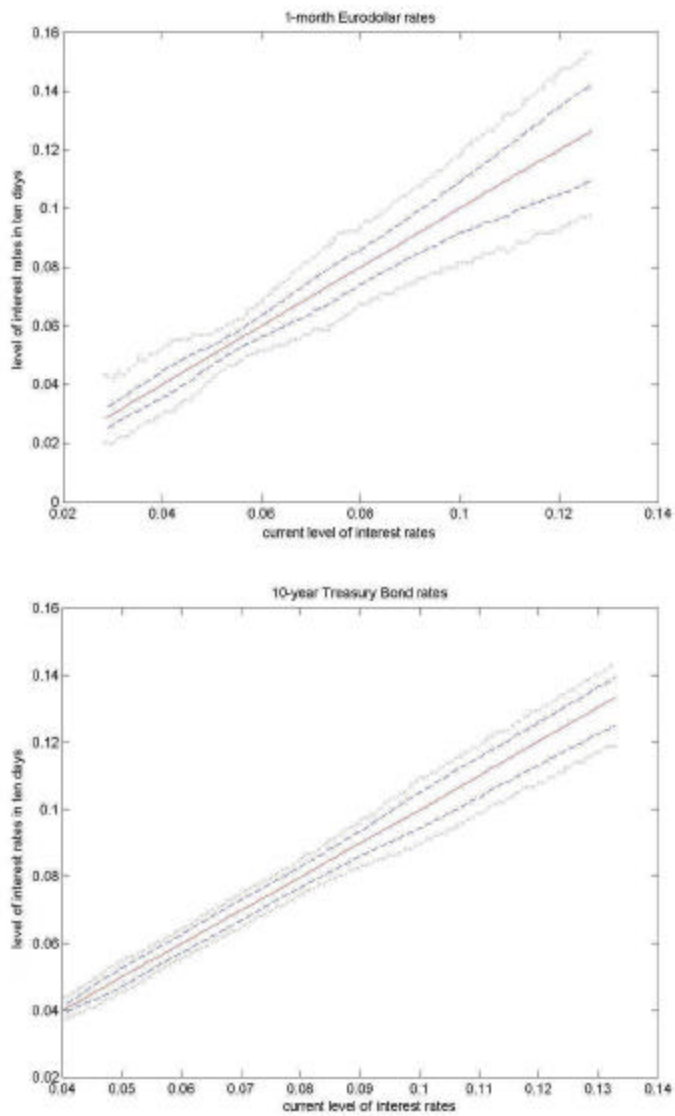

The kernels are estimated with a two-dimensional adjustable bandwidth normal kernel. Each plots the central tendency and 90% error bands.
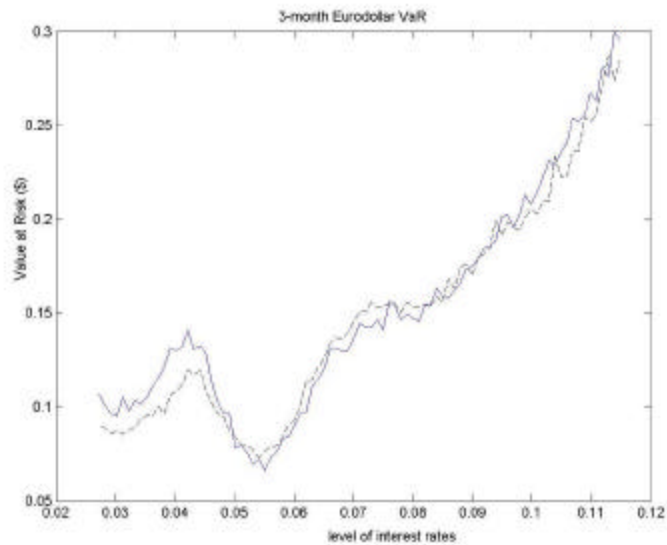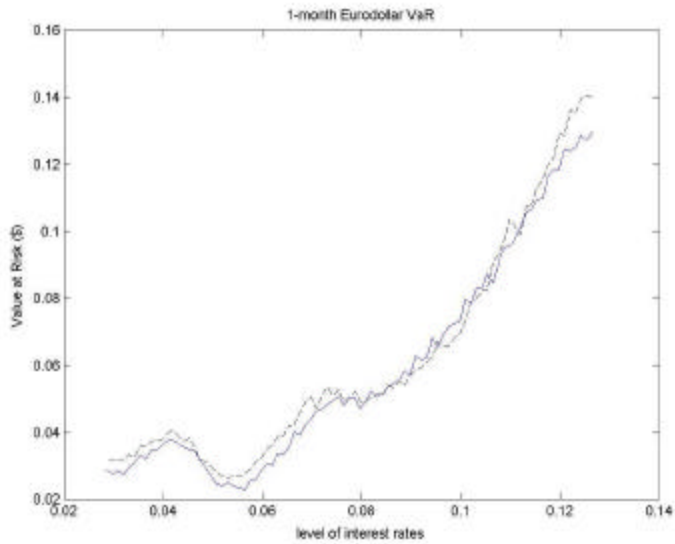
Exhibit 4

10-day Percentile Bounds as a Function of Initial Interest Rate

For 1-Month Eurodollar and Ten-Year Treasury

In each figure, the central line is the median, the dashed lines are 5% and 95%, and the dotted lines are 1% and 99%. The bounds are calculated from an adjustable bandwidth normal kernel.

Exhibit 5

5% Value at Risk of Bonds over 10-Day Horizon,

Depending on Level of Interest Rate

Each line gives the 5% Value at Risk (in dollars) on a bond with current value of $100 and of the specified maturity over a ten-day period. The solid line is the VaR from a long position (if interest rates rise); the dashed line is the VaR from a short position.
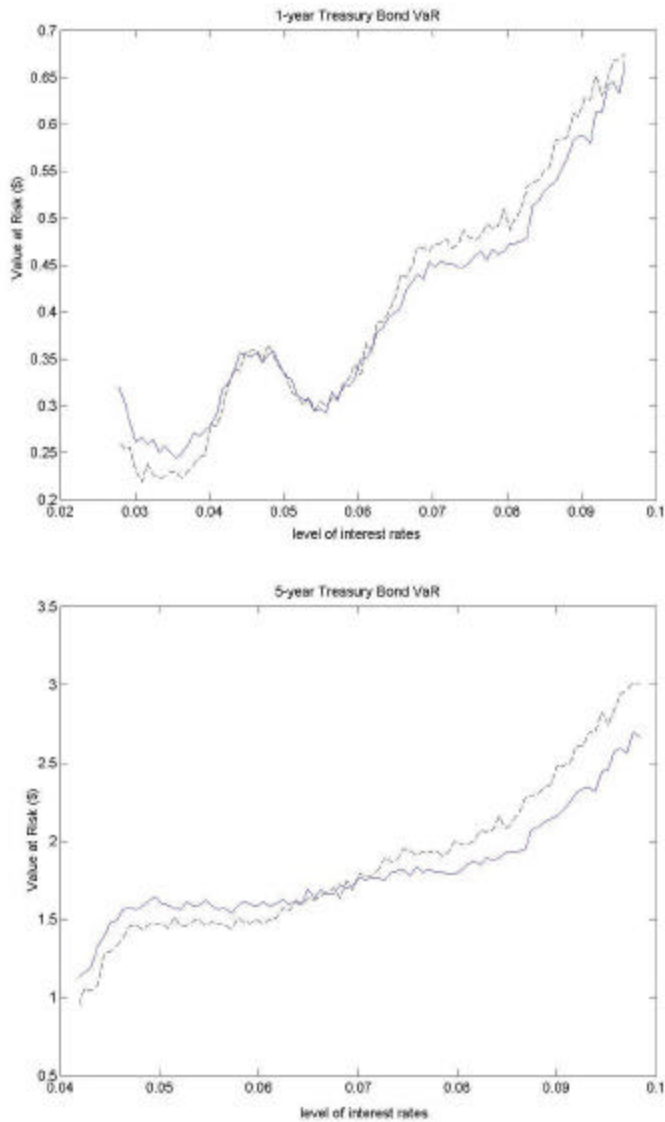
Exhibit 6

5% Value at Risk of Bonds over 10-Day Horizon,

Depending on Level of Interest Rate

Each line gives the 5% Value at Risk (in dollars) on a bond with current value of $100 and of the specified maturity over a ten-day period. The solid line is the VaR from a long position (if interest rates rise); the dashed line is the VaR from a short position.
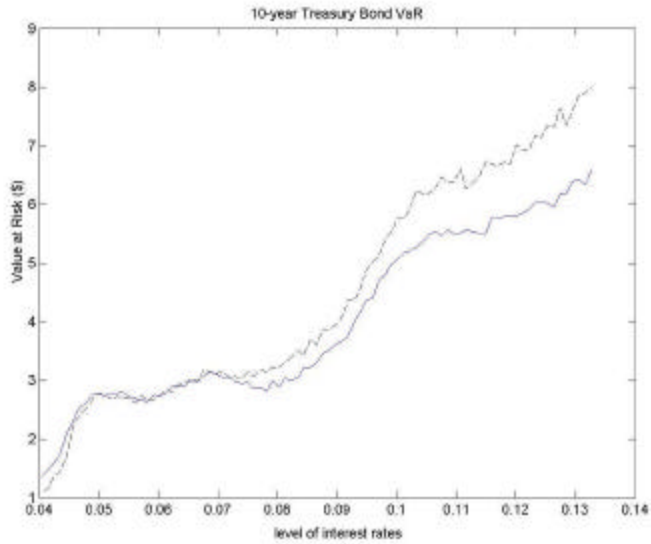
Exhibit 7

5% Value at Risk of Bonds over 10-Day Horizon,

Depending on Level of Interest Rate

Each line gives the 5% Value at Risk (in dollars) on a bond with current value of $100 and of the specified maturity over a ten-day period. The solid line is the VaR from a long position (if interest rates rise); the dashed line is the VaR from a short position.

Exhibit 8

Comparing Hit Ratios for Ten Day VaR, as given by Historical Higher Moments (HHM), Rolling Normal, and GARCH for One-Month Eurodollar Interest Rates

	Theoretical value	HHM	Rolling Normal	GARCH	tail
percent	**0.01**	0.004503	0.045713	0.014463	lower
number (of 7329)	**73**	33	335	106	
percent	**0.01**	0.007777	0.039479	0.024424	upper
number	**73**	57*	289	179	
percent	**0.025**	0.019512	0.061781	0.025788	lower
number	**183**	143	453	189*	
percent	**0.025**	0.02838	0.053747	0.040251	upper
number	**183**	208*	394	295	
percent	**0.05**	0.044481	0.083253	0.044344	lower
number	**366**	326	610	325	
percent	**0.05**	0.054032	0.073556	0.066994	upper
number	**366**	396*	539	491	

The first column (in bold) is the theoretical probability and number of times that the interest rate will be beyond the given VaR threshold. The remaining columns show the fraction and number of days when the rate was actually beyond, as calculated by the Historical Higher Moments (column three), the rolling normal (column four) and restricted GARCH(1,1) (fifth column). The last column indicates whether the comparison is for rates which stretch below the lower bound or those which extend above the upper bound in the tail of errors. An asterisk indicates if the particular specification was not significantly different (at 95%) from the theoretical value. Because the rolling normal doesn't use the first 100 observations, the hit numbers were re-scaled to be comparable.

Exhibit 9

Comparing Hit Ratios for Ten Day VaR, as given by Historical Higher Moments (HHM), Rolling Normal, and GARCH for Three-Month Eurodollar Interest Rates

	Theoretical value	HHM	Rolling Normal	GARCH	tail
percent	**0.01**	0.003913	0.038398	0.014641	lower
number (of 7923)	**79**	31	304	116	
percent	**0.01**	0.008709	0.037246	0.01666	upper
number	**79**	69*	295	132	
percent	**0.025**	0.020194	0.053373	0.025243	lower
number	**198**	160	423	200*	
percent	**0.025**	0.02802	0.05862	0.033068	upper
number	**198**	222*	464	262	
percent	**0.05**	0.046952	0.077051	0.05036	lower
number	**396**	372*	610	399*	
percent	**0.05**	0.053768	0.083835	0.062603	upper
number	**396**	426*	664	496	

The first column (in bold) is the theoretical probability and number of times that the interest rate will be beyond the given VaR threshold. The remaining columns show the fraction and number of days when the rate was actually beyond, as calculated by the Historical Higher Moments (column three), the rolling normal (column four) and restricted GARCH(1,1) (fifth column). The last column indicates whether the comparison is for rates which stretch below the lower bound or those which extend above the upper bound in the tail of errors. An asterisk indicates if the particular specification was not significantly different (at 95%) from the theoretical value. Because the rolling normal doesn't use the first 100 observations, the hit numbers were re-scaled to be comparable.

Exhibit 10

Comparing Hit Ratios for Ten Day VaR, as given by Historical Higher Moments (HHM), Rolling Normal, and GARCH for One-Year Treasury Bond (Constant Maturity) Interest Rates

	Theoretical value	HHM	Rolling Normal	GARCH	tail
percent	**0.01**	0.007605	0.034403	0.010908	lower
number (of 9993)	**100**	76	344	109*	
percent	**0.01**	0.007405	0.0427	0.013509	upper
number	**100**	74	427	135	
percent	**0.025**	0.022015	0.054235	0.021515	lower
number	**250**	220*	542	215	
percent	**0.025**	0.024417	0.063746	0.022816	upper
number	**250**	244*	637	228*	
percent	**0.05**	0.045832	0.079531	0.040528	lower
number	**500**	458*	795	405	
percent	**0.05**	0.050335	0.092381	0.045832	upper
number	**500**	503*	923	458*	

The first column (in bold) is the theoretical probability and number of times that the interest rate will be beyond the given VaR threshold. The remaining columns show the fraction and number of days when the rate was actually beyond, as calculated by the Historical Higher Moments (column three), the rolling normal (column four) and restricted GARCH(1,1) (fifth column). The last column indicates whether the comparison is for rates which stretch below the lower bound or those which extend above the upper bound in the tail of errors. An asterisk indicates if the particular specification was not significantly different (at 95%) from the theoretical value. Because the rolling normal doesn't use the first 100 observations, the hit numbers were re-scaled to be comparable.

Exhibit 11

Comparing Hit Ratios for Ten Day VaR, as given by Historical Higher Moments (HHM), Rolling Normal, and GARCH for Five-Year Treasury Bond (Constant Maturity) Interest Rates

	Theoretical value	HHM	Rolling Normal	GARCH	tail
percent	**0.01**	0.007405	0.035718	0.011108	lower
number (of 9993)	**100**	74	357	111*	
percent	**0.01**	0.009507	0.039158	0.01381	upper
number	**100**	95*	391	138	
percent	**0.025**	0.024117	0.054639	0.021415	lower
number	**250**	241*	546	214	
percent	**0.025**	0.024017	0.06243	0.025918	upper
number	**250**	240*	624	259*	
percent	**0.05**	0.048034	0.079429	0.041029	lower
number	**500**	480*	794	410	
percent	**0.05**	0.049335	0.09056	0.048234	upper
number	**500**	493*	905	482*	

The first column (in bold) is the theoretical probability and number of times that the interest rate will be beyond the given VaR threshold. The remaining columns show the fraction and number of days when the rate was actually beyond, as calculated by the Historical Higher Moments (column three), the rolling normal (column four) and restricted GARCH(1,1) (fifth column). The last column indicates whether the comparison is for rates which stretch below the lower bound or those which extend above the upper bound in the tail of errors. An asterisk indicates if the particular specification was not significantly different (at 95%) from the theoretical value. Because the rolling normal doesn't use the first 100 observations, the hit numbers were re-scaled to be comparable.

Exhibit 12

Comparing Hit Ratios for Ten Day VaR, as given by Historical Higher Moments (HHM), Rolling Normal, and GARCH for Ten-Year Treasury Bond (Constant Maturity) Interest Rates

	Theoretical value	HHM	Rolling Normal	GARCH	tail
percent	**0.01**	0.006401	0.031857	0.013303	lower
number (of 9998)	**100**	64	319	133	
percent	**0.01**	0.008602	0.033475	0.013503	upper
number	**100**	86*	335	135	
percent	**0.025**	0.020904	0.054814	0.023705	lower
number	**250**	209	548	237*	
percent	**0.025**	0.022104	0.051173	0.024905	upper
number	**250**	221*	512	249*	
percent	**0.05**	0.043509	0.080502	0.044909	lower
number	**500**	435	805	449	
percent	**0.05**	0.046309	0.0804	0.046809	upper
number	**500**	463*	804	468*	

The first column (in bold) is the theoretical probability and number of times that the interest rate will be beyond the given VaR threshold. The remaining columns show the fraction and number of days when the rate was actually beyond, as calculated by the Historical Higher Moments (column three), the rolling normal (column four) and restricted GARCH(1,1) (fifth column). The last column indicates whether the comparison is for rates which stretch below the lower bound or those which extend above the upper bound in the tail of errors. An asterisk indicates if the particular specification was not significantly different (at 95%) from the theoretical value. Because the rolling normal doesn't use the first 100 observations, the hit numbers were re-scaled to be comparable.

Endnotes

[1] The third, stress testing, is as much art as science and so is not addressed here, but see Longin [2000].

[2] Due to vagaries in the data collection in the early 1970s, the datasets have slightly different numbers of observations.

[3] Of course, there are direct implications for the form of the diffusion process, however in this exercise we wish to remain agnostic about the form of the jump component.

[4] The conditional moments for the other maturities are similar and therefore not reported here.

[5] Also note that, in applying this method to other price series, there is no necessary reason for this particular state variable, the level of interest rates. The higher moments of different price series may depend upon other state variables. This analysis is meant only to be illustrative of the broad technique.

[6] Certainly a determined econometrician could search to find some combination of parameters that improved VaR bounds (see Chen, Hong, and Stein [2001] for a fascinating endeavor), but part of the appeal of the historical higher moments kernel lies exactly in the fact that it does not need to be "tweaked". Estimating GARCH parameters that balance the twin risks of overfitting and poor back-testing is as much art as science. The kernel method does not require a strong prior belief and so can provide a baseline figure that is not idiosyncratic relative to the particular modeler's judgement. See Christoffersen et al. [2001] for a lucid overview.

[7] Several other specifications were tested. Rolling normal VaRs based on predictions of either the absolute level of the interest rates (rather than the change) as well as rolling normal VaRs based on the percent change were not much different from the rolling normal shown. The percent change can give a rough approximation to the notion that rate change volatilities depend on the level.