TESTING DISTRIBUTIONS OF STOCHASTICALLY GENERATED YIELD CURVES

BY

GARY G. VENTER

ABSTRACT

A method is introduced for testing the distribution of yield curves that are produced by asset scenario generators. The method is based on historical relationships in the conditional distributions of yield spreads given the short-term rate. As an illustration, this method is used to test a few selected models. To provide background, stochastic modeling for interest rates and fitting methods are briefly discussed.

KEYWORDS

Interest rates, financial modeling.

Property-liability insurers are looking to financial modeling to address the risk from assets, liabilities and current underwriting results, and often other sources as well. Models combine asset outcomes with liability development and underwriting returns to give a risk profile of the company. The asset models typically generate a large variety of scenarios, ideally each according to its probability of occurrence, which are applied to the asset portfolio to measure the distribution of asset returns.

Although a useful and general approach in theory, such a model might not capture the full range of economic outcomes, or it could over-weight the probabilities of some occurrences that are in fact not all that likely to happen. Thus a significant risk to this methodology is generating an unrealistic distribution of financial events. This paper looks at evaluating interest rate generators by testing, at selected time periods, the distributions of the yield curves they produce. This would help verify that the scenarios show up in the model somewhat in proportion to their actual probability of occurrence.

One approach to the modeling of interest rates is to first fit the dynamics of the short-term rate, and then use the probabilities of its future changes to build up the yield curve. The fit is evaluated primarily on how well it captures the short-term rate's dynamics. Some methods for evaluating the yield curves so created have been identified. For instance, Dai and Singleton (2000) evaluate a class of interest rate models called "affine models" by their impacts on the

yield curves they produce. Some of these criteria are specific to affine models, such as the correlations between factors, but some, like how well a model replicates pricing of selected derivatives, are more general and could be used to test any scenario generator.

Backus, Telmer and Wu (1999) expose several facts about the properties of US Treasury yield curves since 1970 and use these to test generated scenarios. Some of the properties they present characterize the shape of the yield curve, which is related to the yield spreads discussed here. They also find strongly autoregressive behavior for both the yields and the residuals of fits of the yields to simple autoregressive models, and they look at how well models can replicate derivative prices. Their tests are not particular to the affine class of models, so the scenarios generated by any model can be tested by the properties identified.

In a more actuarial vein, Cairns (2003) details some features of historical interest rate data to provide a basis for testing models for actuarial applications. Some of the features he identifies include:

- Short and long-term interest rates are highly correlated in the medium and long term
- They are still correlated, but to a lesser extent in the short term
- Short-term interest rates are more volatile
- Short and long-term interest rates have sustained periods of high and low rates.

From his review he recommends that models meet certain criteria, such as:

- The model should be arbitrage-free and show the features noted
- All interest rates should be positive and all rates should be able to take values arbitrarily close to zero (even if with small probability)
- Interest rates should be mean reverting.

The tests in these papers give a number of criteria for evaluating interest rate models. However there is still another criterion of use for actuarial applications, namely the distribution of the generated yield curves at key target time points, say at the ends of the next few years. Interest rate models in other areas of finance tend to be used to price derivatives, so they are evaluated on how well they can match such prices and the movements of interest rates. Insurer models are often focused on joint asset and loss risks, including the risks inherent in holding various loss/asset mixtures for a period of time, and so to actuaries realistic distributions of ending yield curves could be of more direct concern than option prices. Historical data and a methodology for testing models on such distributions are discussed below.

Other actuarial applications of interest rate generators include pricing of imbedded options and testing investment strategies. The former would need a generator that is good on option pricing, and the latter would be sensitive to the dynamics of yield movements. Thus all the cited tests could be important to actuarial modeling. The incorporation of asset scenarios in broader risk models is somewhat unique, however, in that an important issue is the distribution

across scenarios of the yield curves. Pricing of derivatives looks at the outcome of the derivative across multiple scenarios, but in the risk-neutral measure. The tests proposed below on distributions across scenarios are more specific to building insurer risk models.

1. Models of Interest Rates

The primary focus here is on arbitrage-free models of interest rates. The tests on the yield curve distributions, however, can be used on any model that generates yield curve scenarios. Interest rates are further assumed to be default free. Modeling default probabilities adds a degree of complexity that is not addressed here.

Modeling interest rates is a rapidly developing field. A standard textbook is Rebonato (2000) but a more contemporary treatment is provided by James and Webber (2002). They make the point that the purpose of this modeling, especially insofar as it is published, is not to predict interest rates but rather to describe distributional properties of interest rate movements. Typical applications include option pricing and risk management.

James and Webber discuss a number of different approaches to modeling interest rates, and some of these generate arbitrage-free interest rate scenarios. The method illustrated below models the short-term interest rate, denoted by r, directly, and uses the implied behavior of r, along with market considerations, to infer the behavior of longer-term rates. For these models, r is usually treated as a continuously fluctuating process, but jump components are being built into some models, e.g., see Zhou (2001), Das and Foresi (1996), Johannes (2003).

Another popular approach, following Heath, Jarrow and Morton (1992), models the forward rate curve rather than the short-term rate. The forward rate $f_t(T_1, T_2)$ is the continuously compounded rate at time t for borrowing at time T_1 and repaying at time T_2. Forward rate models are also called whole yield curve models, as modeling all the forward rates is equivalent to modeling the yield curve.

The most common financial models for continuous processes are based on Brownian motion. A Brownian motion has a simple definition in terms of the probabilities of outcomes over time: the change in r between time zero and time t is normally distributed with mean zero and variance $\sigma^2 t$ for some σ. In differential notation, the instantaneous change in r is expressed as $dr = \sigma dZ$. Here Z represents a standard Brownian motion (that is, $\sigma = 1$), and so its variance after a time period of length t is just t. If r also has a drift (i.e., a trend) of at during time t, the process could be expressed as $dr = a dt + \sigma dZ$.

Rebonato gives several examples of whole yield curve models. If you denote by $P(t,T)$ the price of a bond at time t that pays 1 at time T, and r_t is the short rate at time t, this kind of model can be expressed by:

$$dP(t,T) = r_t P(t,T)dt + v\big(t,T,P(t,T)\big)dZ, \text{ where } v \text{ is a volatility function} \quad (1)$$

By specifying v the current yield curve can be matched, and then its future dynamics can be generated. Rebonato shows that you can produce forward

rates with different volatilities if you express v using a deterministic function $u(t,T)$ of time by:

$$v(t,T,P) = P(t,T)\ln(P(t,T))u(t,T) - \int_t^T \ln(P(t,s))\partial u(t,s)/\partial s \, ds$$

The function u makes this somewhat complicated. Short rate models tend to be simpler. Cox, Ingersoll and Ross (1985) provide a model of the motion of the short-term rate that has been widely studied. In the CIR model, r follows the following process:

$$dr = a(b - r)dt + sr^{1/2}dZ \qquad (2)$$

Here b is the level of mean reversion. If r is above b, then the trend component is negative, and if r is below b it is positive. Thus the trend is always towards b. The speed of this mean reversion is expressed by a. Note that the volatility depends on r itself, so higher short-term rates would be associated with higher volatility. Also, if $r = 0$ there is no volatility, so the trend takes over. With $r = 0$ the trend would be positive, so r would move to a positive value. The mean reversion combined with rate-dependent volatility thus prevents negative interest rates.

If this model were discretized it could be written:

$r_t - r_{t-1} = a(b - r_{t-1}) + sr_{t-1}^{1/2}\varepsilon_t$, where ε_t, is a standard normal residual.

This is a fairly standard autoregressive model, so the CIR model can be considered a continuous analogue of an autoregressive model.

Some other models of the short rate differ from CIR only in the power of r in the dZ term. The Vasicek model takes the power to be zero. Another choice is taking a power of unity.

Multi-factor models of the short rate make some of the parameters of the stochastic process themselves stochastic. For instance the reverting mean or the volatility scalar could change over time. These models can overcome some of the limitations of the Gaussian assumptions. Both heavier tails and higher long-term autocorrelation of interest rates can be produced by multi-factor models than by simple Brownian motion.

For example, constant mean reversion is problematic. The short rate sometimes seems to gravitate towards a temporary mean for a while, then shift and revert towards some other. One way to account for this is to let the local mean reversion level b in equation (1) itself be stochastic. This can be done by adding a second stochastic equation to the model:

$$db = j(q - b)dt + wb^{1/2}dZ_1$$

Here Z_1 is a second, independent standard Brownian motion, and so b follows a mean reverting process gravitating towards q. Again different models can be produced by taking different powers for b in the stochastic term. Such two factor models are popular in actuarial literature. For instance, Hibbert, Mowbray and Turnbull (2001) use a two factor generalization of the Vasicek model with b and r both to the zero power, so they drop out of the dZ terms.

The volatility s can also be stochastic. Hull and White (1987) consider such a model.

Combining stochastic volatility and stochastic mean reversion, Andersen and Lund (1997) and (1998) explore the model:

$$dr = a(b - r)dt + sr^k dZ_1 \qquad k > 0$$
$$dln\ s^2 = c(p - ln\ s^2)dt + vdZ_2$$
$$db = j(q - b)dt + wb^{1/2}dZ_3$$

The AL model uses three independent standard Brownian motion processes, Z_1, Z_2, and Z_3. The volatility parameter s^2 now also varies over time, but via a mean reverting geometric Brownian motion process (i.e., Brownian motion on the log). In total there are eight parameters: a, c, j, k, p, q, v, and w and three varying factors r, b, and s. It is thus labeled a three-factor model. The power k on r in the stochastic term is a parameter that can be estimated.

Models of the short rate can be used to generate yield curves by their forecasted future behavior. However this requires an additional feature – the market price of risk. It turns out that the Brownian motion processes defining single or multi-factor short rate models can be adjusted for the market price of risk by changing their drift terms. Then bonds can be priced using the probabilities of the future evolution of the adjusted process, which is called the risk-neutral process. This is discussed in more detail below.

An important class of short rate models is the affine models. In any Brownian motion model of the short rate, the drift and squared-volatility of dr are functions of time and the short rate itself. Rebonato defines affine models as those where these functions (in the risk-neutral measure) are linear in r – but not necessarily in time. A primary result is that then the yield curve can be derived directly from the bond price function:

$$P(t,T) = exp[A(t,T) - B(t,T)r_t]\ \text{for deterministic functions } A \text{ and } B$$

Dai and Singleton study multi-factor affine models and identify a sub-class that is optimal in matching actual yield processes. AL is not affine, however, due to the log form of the diffusion equation for the volatility s^2.

Other processes besides Brownian motion can be used for the movements of r and for the other factors. Processes with jumps could be used, for example. The normally distributed tail of Brownian motion seems too light in practice, and heavier-tailed processes are being studied. For instance, Raible (2000) looks at modeling the yield curve with Levy processes. The autocorrelation of interest rates drops off more slowly than Brownian motion would support. This and similar time series issues have led some authors to look at fractal approaches, such as fractional Brownian motion. E.g., Lantsman and Major (2001) find that multi-factor Brownian motion models like AL can capture some but not all of the fractal-like features of interest rate series. In fact they find that in a long simulation AL displays fractal-like behavior, but actual US treasury rates follow a more complex multi-fractal pattern. Multi-fractal models are still being studied to find practical methods for arbitrage-free application.

2. Dynamics of Short-Term Rates – Fitting and Empirical Findings

Estimation of model parameters should be distinguished from calibration to current states. The permanent parameters of the models for real world rather than risk-neutral dynamics are estimated from historical data, whereas the variable factors are re-calibrated to current yield curves to capture the latest market conditions. A bond trader would usually calibrate to derivative prices, but matching the current yields might suffice for actuarial risk modeling.

Multi-factor Brownian motion models can be difficult to estimate. Some single-factor models, such as CIR, can be can be integrated out to form a time series, which can be estimated by maximum likelihood. In the case of CIR, the conditional distribution of the short rate at time $t + T$ given the rate at time t follows a non-central chi-squared distribution:

$f(r_{t+T}|r_t) = ce^{-u-v}(v/u)^{q/2}I_q(2(uv)^{1/2})$, where
$c = 2as^{-2}/(1 - e^{-aT})$, $q = -1 + 2abs^{-2}$, $u = cr_t e^{-aT}$, $v = cr_{t+T}$ and I_q can be expressed by $I_q(2z) = \sum_{k=0}^{\infty} z^{2k+q}/[k!(q+k)!]$ with factorial off integers defined by the gamma function

This is not usually possible for multi-factor models, where the volatility and other factors can change stochastically. Further, the short-term rate is observed, or is closely related to observed rates for very short terms, but the other factors, like the reverting mean and the volatility scalar, are not typically observed. Thus fitting techniques that match models to data will not be applicable for these factors.

A few fitting techniques have been developed for stochastic processes. The general topic of what these techniques are and how they work is beyond the scope of this paper, but one method that has been used successfully – the efficient method of moments (EMM) – is briefly discussed below. An alternative not discussed further here is quasi-maximum likelihood, which applies MLE to estimates of the unobserved factors, perhaps made by smoothing methods. EMM and a related method SMM, the simulated method of moments, are discussed in James and Webber (2002).

EMM is a special case of GMM, the generalized method of moments. A generalized moment is any quantity that can be averaged over a data set, such as $(3/x)\ln x$. GMM fits a model by matching a selection of generalized moments of the model to the data. In fact, MLE is a case of GMM as well, since it matches the scores of the log-likelihood function (the first partial derivatives of the log-likelihood function with respect to each model parameter) to the theoretical value of zero. EMM is a particular choice of generalized moments that has some favorable statistical properties when used to fit stochastic models. EMM for a particular data set starts by finding the best time series model, called the auxiliary model, that can be fit to that data. This might be a GARCH model, for example. The auxiliary model is fit by maximum likelihood, so the scores of that model will be zero when calculated over the data set.

The parameters of the multi-factor Brownian motion model are then found by an iterative search procedure. For a trial set of parameters, a long time series

is simulated from the model. The auxiliary model is then tested to see how well it fits this generated time series. The test is based on evaluating the score functions from the MLE estimates of the auxiliary model on the generated series. Usually the scores will not be zero, because the parameters of the auxiliary model were fit to the actual series not the generated series. But the parameters of the multi-factor model can be iterated until a weighted squared error function of the scores of the generated data is minimized. The score functions of the auxiliary model can thus be viewed as generalized moments, which are all zero when calculated on the actual data and are as close to zero as possible when calculated on the generated data from the optimized parameters.

The result of this technique is a parameterized stochastic model whose simulated values have all the same dynamics as the data, as far as anybody can tell by fitting time-series models to both. Then the modeled factors (starting values for the short rate, volatility, reverting mean, etc.) can be calibrated to current economic conditions for simulating future scenarios.

Anderson and Lund (1998) did an empirical study of short-term rate dynamics by EMM-fitting their above model to four decades of US Treasury notes, incorporating data from the 1950's through the 1990's. Their results provide an empirical background to evaluate other models as well.

AL estimate k as about 0.55, which supports the power of $\frac{1}{2}$ in the CIR model. In fact the AL model with this parameter is close to the CIR model at any instant of time, but makes the CIR parameters subject to change over time. Other models with $k = 0$ or $k > 1$ appear to be implausible given this result.

The period October 1979 – September 1982 had high rates and high volatility, and studies influenced by this period have found the power of $\frac{1}{2}$ on r too low, e.g., Chan, Karolyi, Longstaff, and Sanders 1992 who use data from 1964 through 1989. There has been some debate about whether or not to exclude 79-82 in fitting models. These results happened, so they can happen, but it was an unusual confluence of conditions not likely to be repeated. By taking a longer period which incorporates this interval AL do not exclude it but reduce its influence. Another approach by Bliss and Smith (1998), finds that 79-82 fits the definition of regime shift. With separate parameters for that period they find that models that have k between $\frac{1}{2}$ and 1 are supported for other periods.

One of the conclusions of Dai and Singleton is: "…consistent with the analysis of a central tendency factor in Andersen and Lund (1998) and Balduzzi et al. (1996), we find that the short rate tends to mean-revert relatively quickly to a factor that itself has a relatively slow rate of mean reversion to its own constant long-run mean." In the AL model, the speed of mean reversion of the short rate is given by the parameter a, and for the reverting mean by j. These are estimated as 2.72 and 0.13, respectively. The higher factor will make the drift towards the mean higher, which is interpreted as a greater speed of mean reversion. The relative sizes of these parameters is apparently consistent across a number of models.

All parameters in the AL model were statistically significant. This implies that dependence of the volatility on r is not enough to capture the changes in volatility of interest rates – stochastic volatility is also needed. There have been

periods of high volatility with low interest rates, for example. Thus the one and two-factor models without stochastic volatility appear to be insufficient to capture US treasury rate dynamics.

3. GENERATING YIELD CURVES

The modeled dynamics of the short rate can produce implied yield curves. This is done by modeling the prices of zero-coupon bonds with different maturities, from which the implied interest rates can be backed out. $P(T)$, the current (time 0) price of a bond paying $1 at maturity T, can be calculated as the risk adjusted discounted expected value of $1 using the continuously evolving interest rate r from the short-term model. Here "expected value" indicates that the discounted mean is calculated over all possible paths for r. This can be expressed as:

$$P(T) = E^*[exp(-\int_0^T r_t dt)],$$

where r_t is the interest rate at time t, the integral is over the time period 0 to T, and E^* is the risk-adjusted (risk-neutral) expected value of the discounted value over all paths r can take.

If E were not risk adjusted, the expectation that gives $P(T)$ could be approximated by simulating the r process to time T over small increments and then discounting back over each increment. The risk-adjusted expected value is obtained instead by using a risk-adjusted process to simulate the r's. This process is like the original process except that it is adjusted to generate higher r's over time. These higher rates usually produce an upward-sloping yield curve.

What is the risk adjusted process for r needed to so generate the yield curves? If the price at time t for a bond maturing at time T is a Brownian motion with drift u and volatility v, i.e.,

$$dP(t,T) = u(t,T)dt + v(t,T)dZ$$

then it can be shown (Vasicek (1977)) that arbitrage-free pricing requires the drift u to be a function of the risk-free rate r, the volatility v and a quantity λ called the market price of risk, with:

$$u(t,T) = rP(t,T) + \lambda(t,r)v(t,T)$$

Thus the value of the bond grows at the risk-free rate plus a mean zero stochastic term $v(t,T)dZ$ plus a risk premium given by the product of the bond's volatility with the market price of risk. The market price of risk $\lambda(t,r)$ does not depend on the maturity date T, but it could depend on the interest rate r and the current time t.

Thus the market price of risk in the bond price process is the link that specifies the risk-adjustment to the interest rate process for generating the bond prices as the discounted expected value. It turns out that to do this, only the drift of the interest rate process needs to be risk-adjusted, and this adjustment is to add the market price of risk times a function of the volatility of the

interest rate process. For instance, AL suggest using the following adjusted process to simulate the interest rates in the bond price calculation:

$$dr = a(b - r + \lambda_1 rs)dt + sr^k dZ_1 \quad k > 0$$
$$d\ln s^2 = c(p - \ln s^2)dt + v dZ_2$$
$$db = j(q - b + \lambda_3 b)dt + wb^{1/2} dZ_3$$

This adds risk-premium terms to the drift of the first and third equations but not the second, as AL feel there is little price effect of stochastic volatility. The risk-price factors λ_1 and λ_3 can be calibrated to the current yield curve along with r, s, and b. These factors do not depend on T, so are held constant throughout any simulated yield curve calculation, but they can change stochastically when a new yield curve is calculated from a new time 0.

In the AL model you have to simulate the dynamics of the risk-adjusted process to get the yield curves. However, in the case of the CIR model, a closed form solution exists which simplifies the calculation. The yield rate for a zero coupon bond of maturity T is given by:

$Y(T) = A(T) + rB(T)$ where:

$$A(T) = -2(ab/s^2 T)\ln C(T) - 2aby/s^2$$
$$B(T) = [1 - C(T)]/yT$$
$$C(T) = (1 + xye^{T/x} - xy)^{-1}$$
$$x = [(a - \lambda)^2 + 2s^2]^{-1/2}$$
$$y = (a - \lambda + 1/x)/2$$

Note that the only occurrence of r is in the Y equation, so Y is a linear function of r – but not of course of T. The linearity will come into play when we look at the distribution of Y across the generated scenarios. Since all the yield rates for different maturities are linear functions of r, they will also be linear functions of each other.

4. HISTORICAL DISTRIBUTIONS OF YIELD CURVES

To develop tests of distributions of yield curves, it is necessary to find some properties of these distributions which remain fairly constant over time. As it is difficult to describe properties of the distribution of the entire curve, which is multi-variate, the focus will be on the univariate distributions of selected yield spreads, i.e., the differences between yields.

For a property to test the models against, however, the historical distribution of a given yield spread is not necessarily all that germane. When short-term rates are high, the yield curve tends to get compressed or even inverted, so spreads get low or even negative, and when the short rates are low the spreads then to be higher. Thus the yield spread distributions vary as a function of the short rate. So the conditional distribution of yield spreads given the short rate could be expected to be more consistent than the unconditional distributions of yield spreads. This is supported in the data, although the conditional distributions themselves do change in certain ways over time.
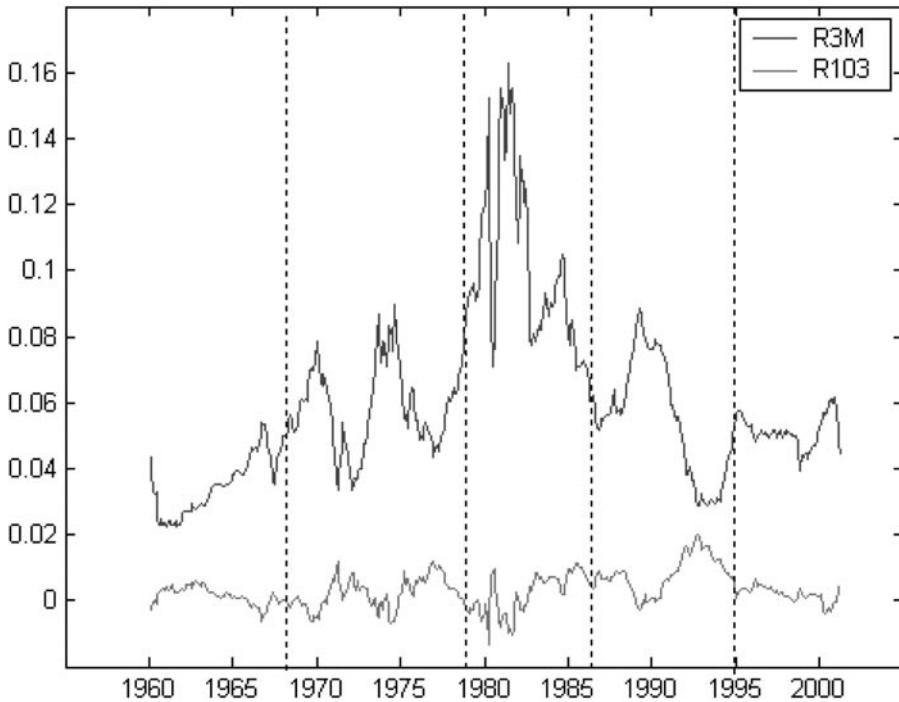
FIGURE 1: Three Month Rate (R3M) and Three Year to Ten Year Spread (R103)

Figure 1 displays a history of the US Treasury 3-month rate and the spread of the 3-year rate to the 10-year rate (10-year rate less 3-year rate). An inverse relationship is seen, as would be expected given the greater stability of the longer rates. However, the spread itself is more stable than the short rate. Thus the relationship between them will depend on whether the short rate is higher or lower. This period is divided into five sub-periods, which were selected to maintain somewhat consistent relationships between the spread and the short-term rate in each period. From the 60's to the early 80's, the short-term rates increased (sub-periods 1-3), then came back down after that (4 and 5).

Figure 2 shows the spread as a function of the three-month rate. Each sub-period shows a negative slope, with the slopes in the range of –0.2 to –0.3. For the entire period, there still seems to be a negative relationship between the short-term rate and the spread, but the slope is much flatter than for the individual periods. The different levels of the lines in the five periods suggests that interest rates follow somewhat different processes in each period. This supports the use of multi-factor or perhaps regime swapping models.

Within each period it appears that for a given short-term rate, the yield spreads tend to gather near the trend line at that rate, with a degree of dispersion around the line. The different yield spreads observed during each period could thus be regarded as random draws from a conditional distribution with
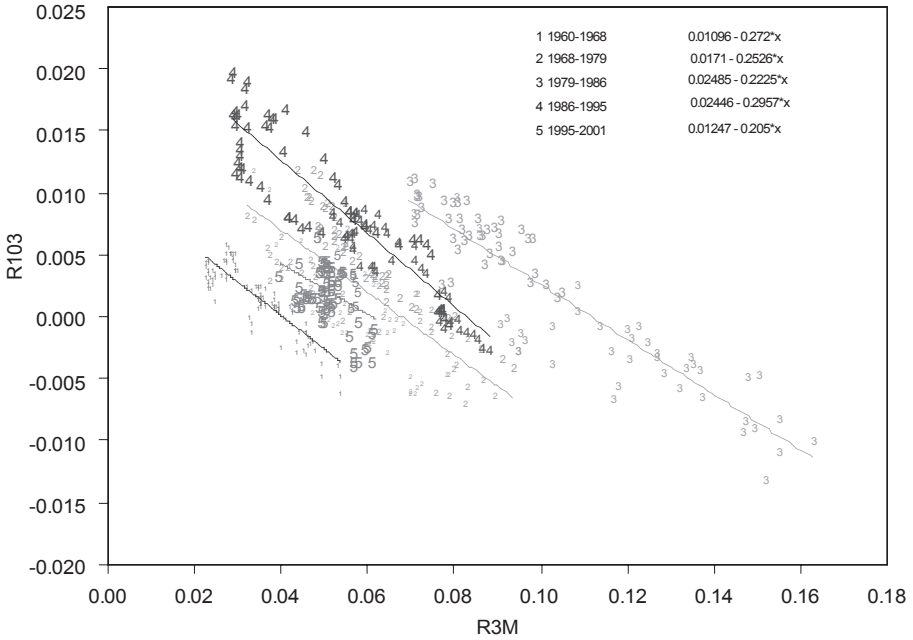
1 1960-1968	0.01096 - 0.272*x
2 1968-1979	0.0171 - 0.2526*x
3 1979-1986	0.02485 - 0.2225*x
4 1986-1995	0.02446 - 0.2957*x
5 1995-2001	0.01247 - 0.205*x

FIGURE 2: Three Year to Ten Year Spread (R103) vs. Three Month Rate (R3M)

1 1960-1968	0.005539 - 0.134*x
2 1968-1979	0.008277 - 0.1175*x
3 1979-1986	0.012 - 0.1055*x
4 1986-1995	0.01246 - 0.1564*x
5 1995-2001	0.003488 - 0.05258*x
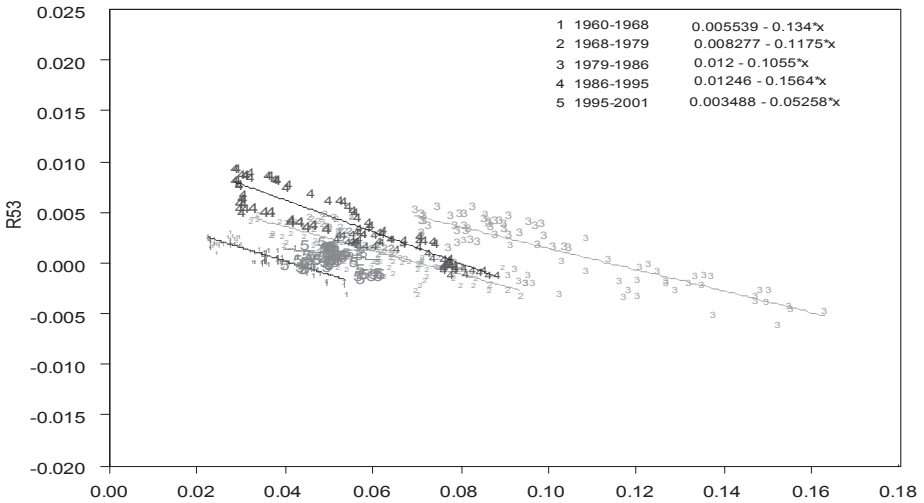
FIGURE 3: Three Year to Five Year Spread (R53) vs. Three Month Rate (R3M)

mean equal to a fixed linear function of the short rate. This behavior suggests that the distribution of yield spreads across scenarios of interest rates generated for a horizon of a few years should follow a similar distribution. The different short rates generated would have a range of yield spreads associated with them,
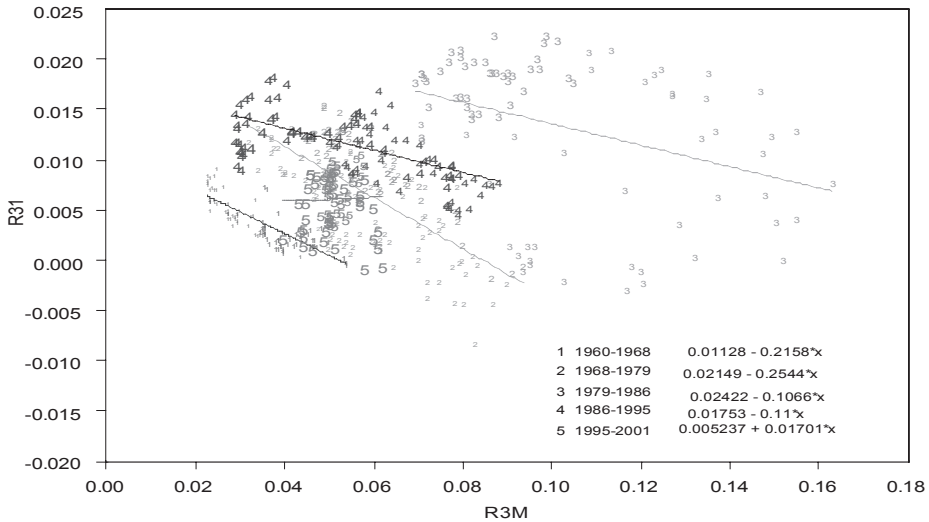
FIGURE 4: One Year to Three Year Spread (R31) vs. Three Month Rate (R3M)

but these would be centered around a linear function of the short rate, with a fair amount of spread around a slope in the range of –0.2 to –0.3. A model could be tested by graphing the scenarios it generated to see if they are generally consistent with this pattern. It would not be appropriate to use the conditional distribution from the entire period as a test of a scenario generator, since for a several-year period the steeper slopes as in the historical sub-periods would be more likely to prevail.

The three-year to five-year spreads in Figure 3 have slopes about half of those in the 10-3 case, except for the latest period, which is even flatter. The short-term rates in the last period have stayed in a fairly narrow range, however, making it harder to estimate the slope. In any case, it would seem that models producing a somewhat flatter slope in the near future would be consistent with the recent experience. The one to three-year spreads in Figure 4 are a little different: the trend was steeper than –0.2 in the 60's and 70's, near –0.11 in the 80's to mid-90's, and close to zero in the last period.

5. TESTING MODELS AGAINST HISTORICAL DISTRIBUTIONS

Within each of the five historical time periods the conditional distribution of yield spreads given the short rate appears to be a scattering around a negative-sloping line. This feature could be used to test the distribution of yield curves at a point in time across scenarios generated by a stochastic generator. Basically this test would be to see if the distribution of generated yield spreads follows a similar pattern. This is not a test of how individual scenarios progress over time, although it would be important to test that as well, using some of

the methods referenced above. This is more of an eyeball test, like a QQ plot, than a strict statistical test of rejection by probability level.

If long-term scenarios are needed, the scenarios could be tested to see if they produce different periods with similar slopes but different intercepts, as in the historical data. However the test proposed here is to look at a single time period and test the distribution across scenarios. The historical time series observations are not independent of each other, but the relationships identified above within each period have ignored the time dimension. Thus they can be thought of as a resampling of the historical data. It seems reasonable to require that the scenarios generated have a similar distribution of yield spreads as the resampled historical data. For a fairly short-term projection, the conditional distribution of the spreads given the short rate should be similar to the latest observed period. For longer projections there is the chance of moving into another regime, and so the spreads would be expected to look like the combination of a couple of periods, which would have a flatter slope.

In this section three models will be tested under a short-term projection as an illustration of this procedure. Two are extensions of the Anderson Lund specification for the short-term rate generator, differing in the treatment of the market price of risk. The third is the CIR model.

The market price of risk has, at all times, to be independent of the bond maturity to guarantee arbitrage-free dynamics. But it can change stochastically when generating scenarios for the yields at another time period. Allowing the market price of risk to change stochastically produces somewhat more variability among the yield curves generated. In one model, the constant lambda model, the two AL market price of risk parameters λ_1 and λ_3 are held constant across all simulations. In the variable lambda model, on the other hand, stochastic changes in λ_1 and λ_3 are generated from one period to the next. How best to do that is a subject of ongoing research. The variable lambda model tested here is one of many possible models of this type and has not been optimized for this test. In fact the testing below suggests that it probably introduces a bit too much variability into the market-price of risk.

The market price of risk parameters, as well as the current values of the three factors r, b, and s can be calibrated to the current yield curve to get starting values for scenario generation. For this example, a yield curve from May 2001 was used for calibration. Parameters were selected to generate a current yield curve that most closely matches the selected target curve. Then yield curves were simulated at various projected periods. For periods in the near future, the curves would not be expected to be too different from the current curves. But going out a few years produces a wider variety of yield curve scenarios. In this case the sets of curves generated for year end 2004 are used in the distributional tests. This seems like a long enough projection period to expect to see the kind of variability that exists within the five individual periods historically.

Models can be tested for the conditional distributions of all of the yield spreads. First examined is the three-year to five-year spread. Recall that the slope for this was about –0.05 in the latest period, but ranged from –0.11 to –0.16 in earlier segments. Figures 5 and 6 show the relationship for the simulated spreads
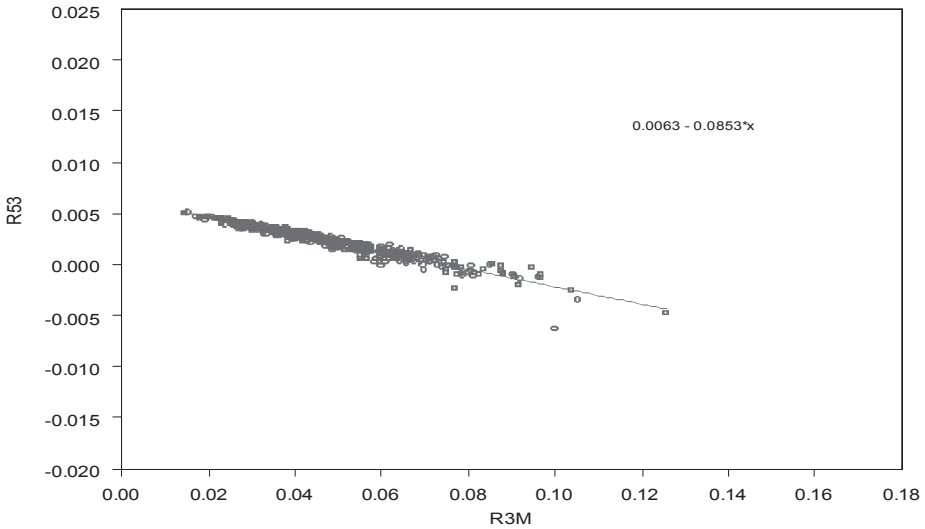
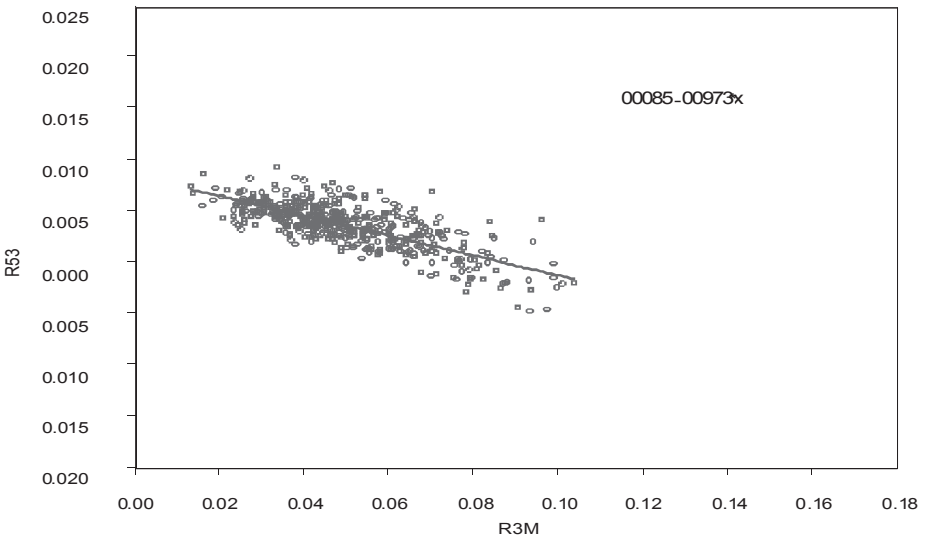FIGURE 5: Three to Five Year Spread (R53) from Constant Lambda Model

FIGURE 6: Three to Five Year Spread (R53) from Variable Lambda Model

under the two Anderson-Lund-based models. The constant lambda model shows a slope of about –0.09, vs. –0.1 for the variable lambda, which are both in the historical range. There is a difference apparent in the scattering around the trend line, with the constant lambda model showing little scattering, and the variable lambda showing a good deal more, which is more compatible with the historical data.

For the CIR model it was shown above that any yield spread is a linear function of the three-month rate. Although this model does have a fair amount of flexibility in determining the slope of that relationship, there will be no variability possible around the trend line. Graphically this would look narrow like the constant lambda case, only more so, with all points falling exactly on the trend line. This suggests that the CIR model will necessarily produce a restricted set of yield curve scenarios, and these will not have all the variability present in historical yield curves. Thus yield curve scenarios will not be present in proportion to their probability of occurring, contrary to the criteria established above for asset generators.

A more quantitative comparison of the conditional distributions can be made by looking at the slopes of the trend lines through the various yield spreads for each of the five time periods, as graphed in Figures 2-4, and the dispersion of the points around those lines, expressed as residual standard errors (mean squared distances from the trend line)$^{1/2}$. Table 1 summarizes the historical and modeled slopes and residual standard errors. For instance, R31 is the three-year to one-year spread graphed in Figure 4. The slopes tend to be negative, denoted by parentheses.

The slopes of both models are similar to those in period 5, which seems appropriate for a model calibrated to current yields. The shortest spreads have slopes near zero for period 5 and in both models. It is not clear why this has changed from historical patterns, but the models seem to be reflecting this trend to some degree. The standard errors for the constant lambda model tend to be lower than the historical variability around the trend, whereas this particular formulation of the variable lambda model produces perhaps too high standard errors in the longer spreads. This suggests that allowing somewhat less variability in the stochastic processes that generates the market prices of risk could lead to still more realistic models.

TABLE 1

Slopes and SE's	R10 3	R10 5	R5 3	R3 1
Period 1	(0.2720)	(0.1380)	(0.1340)	(0.2158)
Period 2	(0.2526)	(0.1351)	(0.1175)	(0.2544)
Period 3	(0.2225)	(0.1170)	(0.1055)	(0.1066)
Period 4	(0.2957)	(0.1393)	(0.1564)	(0.1100)
Period 5	(0.2050)	(0.1524)	(0.0526)	0.0170*
Constant λ	(0.2489)	(0.1635)	(0.0853)	(0.0721)
Variable λ	(0.2960)	(0.1987)	(0.0973)	(0.0615)
Period 1 se	0.0013	0.0009	0.0006	0.0013
Period 2 se	0.0031	0.0022	0.0013	0.0037
Period 3 se	0.0026	0.0013	0.0017	0.0070
Period 4 se	0.0022	0.0012	0.0012	0.0024
Period 5 se	0.0020	0.0013	0.0009	0.0028
Constant λ se	0.0008	0.0005	0.0004	0.0013
Variable λ se	0.0042	0.0028	0.0015	0.0021

* Not significantly different from zero.

6. Testing Residual Distributions

The test above looks at the slope of the conditional yield spreads and their degree of scattering around the trend line. It should also be possible to test the form of the distribution of these residuals. Unfortunately with the data available
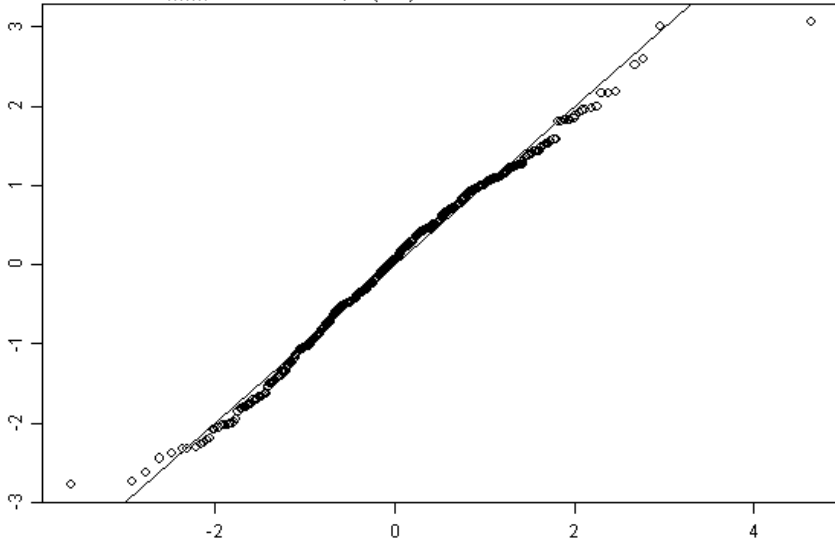

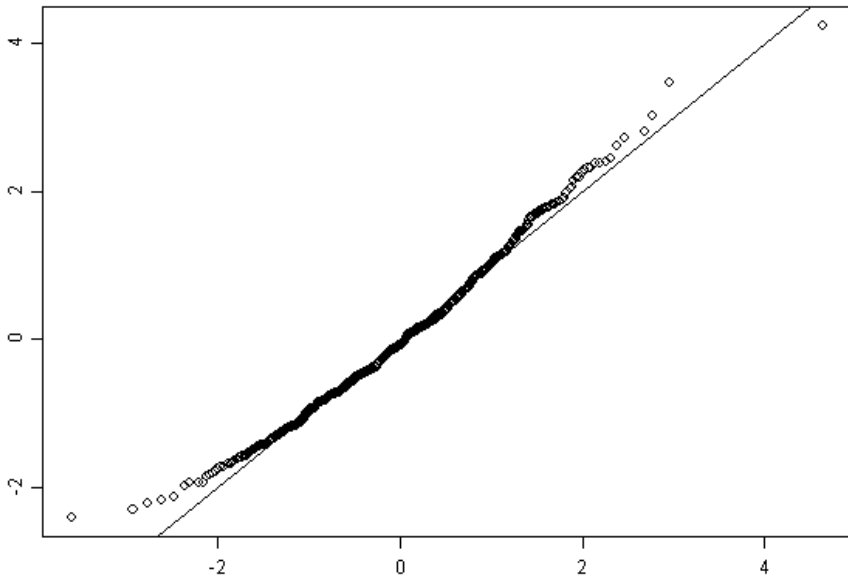FIGURE 7: 10-3 year spread historical residuals against t-33 fit

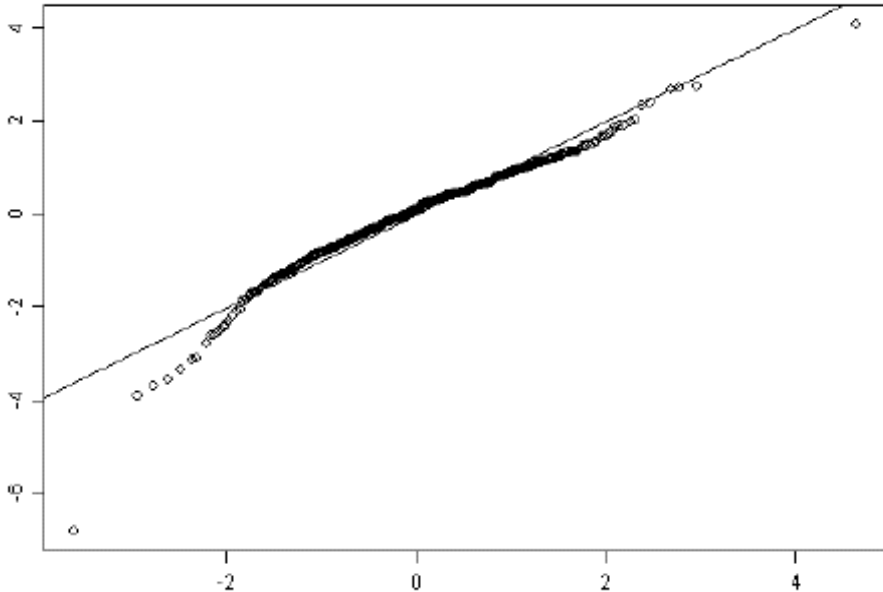FIGURE 8: 10-3 year spread variable lambda residuals against t-33 fit

FIGURE 9: 10-3 year spread constant lambda residuals against t-33 fit

it has not been possible to find a convincing fit for the historical distributions. Nonetheless, some comparisons can be made based on the deviations of the residuals from a selected distribution. Figure 7 is a QQ-plot of a t-distribution with 33 degrees of freedom fit to the residuals from the trend lines of the five sets of ten-year-to-three-year spreads. This is the closest to Gaussian of any of the yield spreads reviewed. For this calculation, all of the residuals around the five lines from Figure 2 were combined and fit to a single t-distribution. The fit looks reasonable, but there are some outliers in both tails that could be due to transition times between periods. Figures 8 and 9 show QQ-plots which graph the percentiles of the generated Anderson-Lund residuals against the percentiles of the fitted t's.

The 10-3 constant and variable lambda residuals look a lot like the data except in the left tail, where the constant lambda diverges oppositely. This test would be more convincing with a fitted distribution that more closely matched history in the tails. While both AL models compare well with the historical slopes, the variable lambda model looks more like the historical data in both the standard errors and the distribution of the residuals.

CONCLUSION

Empirical studies have found that models of the dynamics of the short-term rate need to incorporate a relatively fast mean reversion to a temporary mean which itself reverts slowly to the long-run mean, stochastic volatility, and mean

sensitive volatility proportional approximately to the mean raised to a power between $\frac{1}{2}$ and 1. Both short and long-term rates display a high degree of persistency, i.e., autocorrelations that decrease only very slowly as the period lengthens.

Testing the conditional distribution of yield spreads given the short-term rate appears to be a reasonable way to see if a model is generating a realistic distribution of yield curves. Within periods of a few years, these distributions for US Treasuries look similar to a t-distribution around a conditional mean which is itself a downward sloping linear function of the short-term rate. The level of these lines has varied historically between periods, but the slopes have been more consistent over time. While historical yield spreads can be viewed as samples from this distribution drawn sequentially, it is proposed here to use samples of generated curves at a singe point of time as a test of the reasonability of the generator. The unconditional distribution of generated yield spreads would not necessarily be comparable to the historical distribution, however, because different spreads are associated with different short-term rates, and a simulation at a particular time might not be generating a distribution of short-term rates that matches the entire historical record.

As with most tests of distributional issues, this one is not a formulaic system that gives a strict "yes/no" answer to a model's output. But it does provide a realm of reasonable results so an analyst can give an opinion of the "probably ok / probably not" type. For example, having no variability around the conditional trend line would seem to be too limiting. Slopes that are much steeper than historical would also seem unsuitable, as would distributions of residuals around the slopes that differ substantially from the t-distributions fit. Even though these tests are not strict, better results could be sought than those of any of the models tested.

An application issue is how much variability you should have for projection periods of different lengths. When projecting out four or five years, a conditional distribution similar to those of the historical sub-periods might be appropriate. However there is some chance of entering a new regime – i.e., changing periods – over that much time. In all five of the periods graphed, a shift to an adjacent period during a time frame would tend to flatten the conditional trend.

REFERENCES

ANDERSEN, T.G. and LUND, J. (1997) "Estimating continuous-time stochastic volatility models of the short-term interest rate", *Journal of Econometrics* **72**, 343-377.

ANDERSEN, T.G. and LUND, J. (1998) "Stochastic volatility and mean drift in the short term interest rate diffusion: Sources of steepness, level and curvature in the yield curve", Working paper, Northwestern University.

BACKUS, T. and WU (1999) "Design and Estimation of Affine Yield Models", Carnegie Mellon University, Graduate School of Industrial Administration working paper November. http://bertha.gsia.cmu.edu/files/papers/aff99.ps.

BLISS, R. and SMITH, D. (1998) "The Elasticity of Interest Rate Volatility: Chan, Karolyi, Longstaff, and Sanders Revisited", Federal Reserve Bank of Atlanta working paper 97-13a.

CAIRNS, A. (2004) "A family of term-structure models for long-term risk management and derivative pricing". To appear in *Mathematical Finance*. http://www.ma.hw.ac.uk/~andrewc/papers/ajgc30.pdf.

CHAN, K., KAROLYI, G., LONGSTAFF, F. and SANDERS, A. (1992) "An Empirical Investigation of Alternative Models of the Short-Term Interest Rate", *Journal of Finance* **47**, 1209-1227.

COX, J., INGERSOLL, J. and ROSS, S. (1985) "A Theory of the Term Structure of Interest Rates", *Econometrica* **53**, 385-408.

DAI, Q. and SINGLETON, K. (2000), "Specification Analysis of Affine Term Structure Models", *Journal of Finance* **55**, 1943-78.

DAS, S. and FORESI, S. (1996) "Exact Solutions for Bond and Options Prices with Systematic Jump Risk", *Review of Derivatives Research*, **1**, 7-24.

HEATH, D., JARROW R. and MORTON, A. (1992), "Bond Pricing and Term Structure of Interest Rates: A New Methodology", *Econometrica* **60(1)**, 77-105.

HIBBERT, MOWBRAY and TURNBULL (2001) "A Stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes", Technical Report, Barrie & Hibbert Limited June.

HULL, J. and WHITE, A. (1987) "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, **XLII(2)**, 281-300.

JAMES, J. and WEBBER, N. (2002) *Interest Rate Modelling*, Wiley.

JOHANNES, M. (2003) "The Statistical and Economic Role of Jumps in Interest Rates", *Journal of Finance*, to appear.

LANTSMAN, Y. and MAJOR, J. (2001) "Actuarial Applications of Multifractal Modeling Part II: Time Series Applications", Winter Forum, Casualty Actuarial Society 375-85.

RAIBLE, S. (2000) "Lévy Processes in Finance: Theory, Numerics, and Empirical Facts", Dissertation zur Erlangung des Doktorgrades der Mathematischen Fakultät der Albert-Ludwigs-Universität Freiburg, January.

REBONATO, R. (2000) *Interest Rate Option Models*, Second Edition, Wiley.

VASICEK, O. (1977) "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* **5**, 177-88.

ZHOU, H. (2001) "Jump-diffusion term structure and Ito conditional moment generator", Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.) 2001-28.

GARY G. VENTER
Guy Carpenter Instrat®
One Madison Avenue, New York
NY USA 10010
Tel: +1-917-937-3277
Fax: +1-917-937-3777
E-mail: gary.g.venter@guycarp.com