Advances in Modeling of Financial Series

Gary G. Venter, ASA, CERA, FCAS, MAAA
gary.venter@gmail.com

Abstract

There have been continual advances in the modeling of financial series but most are aimed at the pricing of derivatives. Different criteria are needed for development of scenarios for risk management. Some recent methods will be reviewed with an eye on risk-management applications, including using the simulated method of moments to parameterize multifactor models, fractional differencing and other methods to model series with persistent autocorrelation, and models to flatten out the volatility smile, such as jump-diffusion models. These methods will be illustrated with applications to inflation, interest rates, equity prices and exchange rates.

Keywords: Interest Rates; Inflation; Autocorrelation; Multifactor Models

Introduction

Most academic literature on modeling financial series emphasizes the pricing of ever-more-complex derivatives. However, if the purpose of the modeling is to generate scenarios for risk management, derivatives pricing is less important.[1] What is important is generating scenarios that are representative of what could happen. This is not always easy to judge, but matching past properties of the data is probably a good place to start.

One key property is persistence of shocks. If a jump in interest rates or inflation is quickly damped out, the financial consequences for insurers are quite different than if it tends to persist as a new pattern. For interest rates, having a variety of yield curves produced is necessary in order to match history. Distributional aspects, like heaviness of tails, are also important.

Parameterization of models is also different in the risk-management arena. For derivatives pricing, every parameter in a model can be used to calibrate to the current derivative prices. The model then becomes in essence an elaborate interpolation/extrapolation scheme. For risk management, the majority of parameters are intended to capture features of the process, and these would be fit to historical data. Some parameters, however, represent constantly changing values, and these would be calibrated to the latest data, such as option prices and yield curves.

For some models, maximum likelihood estimation (MLE) is not possible. If the features of historical data to match are clear, a convenient though ad hoc method of parameterization is the simulated method of moments. Basically a long series of data is simulated from the model with trial parameters, and the simulated data is checked to see how well it reproduces the desired features of the data (called "moments" even though the features being checked could be much more general than the usual concept of moments of distributions). The parameters are then refined to produce the best match possible. This is done by minimizing a selected function that includes weights for how important each property is considered to be. Studies have found that simulated method of moments can approach the efficiency of maximum likelihood, depending on the moments chosen.

The organization of the paper is to first review concepts of time series and stochastic process models, including fractional differencing, in Section 1. Sections 2 and 3 discuss Treasury interest

[1] The major exception in the insurance world is the management of the risks inherent in selling variable annuities, which utilizes derivative-trading strategies.

rates and inflation modeling, including estimation by simulated method of moments, and Section 4 addresses ways to correlate these. Section 5 gets into modeling of equity prices, including jump diffusions. Section 6 introduces some of the concepts of foreign exchange models. Section 7 brings in models of bonds with default and liquidity risk. Section 8 concludes.

1. Time Series and Stochastic Process Concepts

1.1 Lag Operator

If X is a time series, say inflation observed monthly from March 1980 to December 2009, the first lag of the series is the series of immediately previous observations, in this case Y, the inflation rates from February 1980 to November 2009. The second lag of X is the first lag of Y, here Z, the inflation rates from January 1980 to October 2009. Using L to denote the lag operator, we can write Y = LX and Z = LY. The lag operator basically works algebraically, so $Z = L^2X$ is the second lag.

The first difference of a series is the series minus its first lag, so for X this is X – Y, or (1 – L)X. The second difference is the first difference of the first difference, or (X – Y) – (Y – Z) = X + Z – 2Y = $X + L^2X - 2LX$. This can be written as $(1 - L)^2X$, continuing the algebraic treatment of L.

The first autocorrelation of a series is the correlation of X with LX; the second autocorrelation is the correlation of X with L^2X, etc. If only a few autocorrelations are statistically nonzero, then any shock in the process quickly washes out. However, if the autocorrelations are high for many lags, the effect of the shock persists. Thus, looking at the autocorrelations as a function of lag can indicate if shocks persist or fade away.

1.2 Models for Autocorrelation

A basic time series model is the AR(1), or first-order autoregressive process. This can be written:

$$r_{i+1} = a + br_i + s\varepsilon_{i+1} .$$

Here ε_{i+1} is a standard normal variate. The starting value is r_0. If b is assumed to be less than 1 in absolute value, the k^{th} autocorrelation of r is b^k. If say b = ½ the autocorrelations get small quickly, and any shocks fade out. However if b is just below 1, the autocorrelations decline slowly and shocks can persist for quite a while.

The expected value of r_t can be shown to be $r_0b^t + a(1 - b^t)/(1 - b)$ and so in the limit is $a/(1 - b)$. The variance of r_t is $s^2(1 - b^{2t+1})/(1 - b^2)$, which approaches and is limited by $s^2/(1 - b^2)$.

The simple random walk model is given by a=0, b=1. Here r is its initial value r_0 plus a series of random noise draws. A shock from a long time ago has the same effect as a recent shock, so does not fade out. In contrast, the effects of the random shocks in the AR(1) model decline by powers of b. For the random walk with t observations to date, the k^{th} autocorrelation is $(1 - k/t)^{1/2}$. It turns out that this function is hard to distinguish from that of an AR(1) model with b just below 1, so it is difficult to tell a random walk from such an AR(1) process. There are tests for this, but they are not particularly powerful, so there tends to be a fair amount of debate in the academic literature about whether a particular series of interest is a random walk or an AR(1) with high b. The expected value of r_t is just r_0, but its variance is s^2t, which grows without bound as time passes. Thus the probability of r_t being found in any particular fixed range goes to zero over time.

Often the first differences of a series can be modeled as an AR(1) process. For instance, a process with a trend would not have a finite mean or variance, but its differences might. Even a process like a random walk with an infinite ultimate variance has well-behaved first differences. If a process is growing proportionally, first differences of the log might work.

1.3 Problems with Persistent Autocorrelation

Persistent autocorrelation does not always start near 1. Some series might have autocorrelation around ½ at lag one but only slowly declining for later autocorrelations. This is a form of persistent autocorrelation but it is more difficult to model. One possibility is that it is the sum of processes, like an AR(1) with a low b plus either a random walk or an AR(1) with a high b. This gets into somewhat problematical modeling issues, since none of those process are likely to be observed separately.

For example, the U.S. monthly CPI inflation rate, seasonally adjusted from 1947 to August 2009, has the first 10 autocorrelations:

1	2	3	4	5	6	7	8	9	10
0.57	0.42	0.39	0.37	0.36	0.36	0.37	0.36	0.38	0.39

These do not drop consistently below 10 percent until about seven years, or 84 lags. This does not fit the pattern of either the AR(1) model or the random walk. A method for modeling processes like these is partial differencing. It turns out to be possible to define $(1 - L)^{1/2}$, or the half-difference, in such a way that if it is applied twice you get the first difference. More generally, for any real d, Newton's generalization of the binomial theorem gives:

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-L)^k = 1 - dL + \frac{d(d-1)}{2!} L^2 - \frac{d(d-1)(d-2)}{3!} L^3 + \cdots$$

Thus fractional differences can be defined as declining sums and differences of previous lags. The series of such differences has finite variance if $|d| < ½$, so an AR(1) model can apply to the differenced series. The autocorrelation at lag k is approximately ck^{2d-1}, where c is a constant that depends on d. This can be slowly declining. For the U.S. CPI inflation rate, d has been estimated as around 0.4 to 0.5 by various authors, with similar results in other countries. Discussion of this methodology and some extensions can be found in Baillie et al. (2002).

1.4 Multifactor Processes

Modeling inflation as a sum of simpler processes is discussed in the inflation section below. One popular approach is the double mean-reverting process. An AR(1) process is mean-reverting. Consider the change $r_{i+1} - r_i = a + (b-1)r_i$ plus a random mean zero term. If r_i is at the long-term mean $a/(1-b)$, the expected change is zero. Since $b - 1$ is negative, if r_i is above the long-term mean, the expected change will be less than this, and so negative, pulling the series toward the long-term mean. If r_i is below the mean, the pull will be upwards.

To make the mean reversion more explicit, sometimes this process is written as:

$r_{i+1} = r_i + c(m - r_i) + s\varepsilon_{i+1}$, where $c = 1 - b$ and m is the mean $a/(1-b)$. Then the expected change is readily seen as positive or negative depending on whether r_i is below or above m.

The double mean-reverting process is to also let m be an AR(1) process, say:

$m_{i+1} = m_i + h(\mu - m_i) + \sigma\eta_{i+1}$, and now

$r_{i+1} = r_i + c(m_i - r_i) + s\varepsilon_{i+1}$.

In this process, r reverts toward the temporary mean m, which itself reverts toward the long-term mean μ. Also it is useful to consider the process $q_i = r_i - m_i$. By subtracting the m from the r equation,

$q_{i+1} = q_i - cq_i - h(\mu - m_i) + s\varepsilon_{i+1} - \sigma\eta_{i+1}$.

The last three terms are all independent normal mean zero variates, so their sum is also normal mean zero. This shows that q is an AR(1) process with $a = 0$. Thus $r = m + q$ is a sum of two AR(1) processes, one with mean zero. Hence a slight generalization of the double mean-reverting process is just a sum of two AR(1) processes. This tends to be the more commonly used form recently.

Partial differencing has an advantage in simulating future scenarios in that it builds in the observed past lags, whereas double processes have unobserved terms that somehow have to be calibrated to start the simulation.

1.5 Brownian Motion

Brownian motion is a continuous version of the random walk. It is a continuous process whose change from one time to a time t periods later is normally distributed with mean zero and variance ts^2. Standard Brownian motion is usually denoted by B, W or Z and has s = 1. Over short periods it tends to be a very jumpy process, as the standard deviation $t^{½}$ is a lot greater than t for small t.

The random walk $r_{i+1} = r_i + s\varepsilon_{i+1}$ can be written $\Delta r_{i+1} = s\varepsilon_{i+1}$. The corresponding form for Brownian motion is $dr_t = sdW_t$. A deterministic time trend can be added as in $dr_t = adt + sdW_t$. This can be made mean-reverting by making the sign of the trend depend on whether the process is below or above the mean, e.g., $dr_t = a(m - r_t)dt + sdW_t$. If the log of a process follows a Brownian motion, the process is said to be a geometric Brownian motion and is written $d(\log r_t) = dr_t/r_t = sdW_t$.

The other commonly used stochastic process is the compound Poisson process. Here the number of events in time t is Poisson in λt, and each event size is an independent draw from a single distribution. If N(μ) denotes the number of events with Poisson mean μ and X_k is the k^{th} jump size, then the process can be written a $dr_t = d\left(\sum_{i=0}^{N(\lambda t)} X_i\right)$.

Both Brownian motion and the compound Poisson process are simulated by taking short periods to represent the instantaneous change dt. If t is measured in years, sometimes t = 1/252 is used to represent one trading day. The Brownian motion standard deviation then becomes $s/252^{½}$ for this period, so a mean zero normal draw with that standard deviation is used to represent dW_t.

2. Treasury Interest Rates

2.1 Basic Models

A popular way to model Treasury (assumed risk-free) interest rates is through short-rate models. These model the short-term rate of interest, usually the three-month rate due to data availability, and then use arbitrage-free concepts to build up the yield curve from the short rate.

The two most common short-rate models are the Vasicek model and the Cox, Ingersoll, Ross (CIR) model. The Vasicek model is a basic mean-reverting Brownian motion for the short-rate r:

$$dr_t = (b - ar_t)dt + s dW_t.$$

The CIR model adds a twist to this: the standard deviation, or volatility, is proportional to the square root of the interest rate:

$$dr_t = (b - ar_t)dt + s r_t^{1/2} dW_t.$$

There are three attractions to the CIR model: empirical work has found that higher rates are indeed associated with higher volatility; it produces heavier-tailed distributions of rates, which are more realistic; and making the random term zero when the short rate becomes zero makes it impossible to produce negative rates. However it comes with some costs: the distribution for a short interval is approximately normal, but over a longer period it is not. This requires simulating on short intervals and complicates the estimation of parameters, since the distribution is complicated.

2.2 Yield Curves

The popularity of these two models arises for another reason: the yield curves for both can be calculated in closed form. The calculation of the yield curve in the arbitrage-free framework is otherwise quite computationally intensive, so this is a significant advantage.

There are various ways to define arbitrage, but here it will be taken to mean a position built up from a net investment of zero, which thus has no chance of a loss, but which has a positive probability of a profit. Since Treasury bonds are assumed to be risk-free, and borrowing and lending them is possible, inconsistencies in the yield curve can easily lead to arbitrage opportunities. While such might exist for short periods in the real market, they do not tend to last very long, so should not be in the yield curves used for financial planning. It is also worth noting that a number of practices loosely referred to as arbitrage are not really such, as they have risks not being focused on, like buyouts not going through or liquidity problems arising.

The theory of arbitrage-free yield curves has produced a general calculation rule for bond prices, from which the yield curve can be extracted. The rule is that the bond price has to be the expected present value of the payments, discounted back along all possible paths of the short rate, from now to the time of maturity. However, these discount rates are the modeled short rates adjusted upward for the market price of risk. For the bond prices to be arbitrage-free, the adjustment has to be an increase in the deterministic trend term in the model and has to be related in specific ways to the volatility of the Brownian motion term. Also the same deterministic trend increase has to be used for all maturities. This higher trend tends to discount more for longer maturities, producing a lower bond price and so an upward-sloping yield curve. There could be model situations where the resulting curve is not always entirely upward-sloping, however, which sometimes happens in reality as well. The model with the added trend is called the risk-neutral process, as the price is the mean under this process, and the mean would be the price if the market were risk-neutral.

If the model is being simulated on a daily interval, you need many simulations that go out to 30 years in order to get that bond price, and if you want such yield curves for several points in time in many simulations, the calculations required can become extensive. That is why having a closed-form bond price is so useful. Even a somewhat messy formula for the bond price is worth tolerating.

There is a little flexibility in the trend adjustment. With market price of risk λ, the formulas below add a trend of $\lambda s r_t dt$ to the Vasicek and CIR models to get the risk-neutral short-rate process. Other models leave out the r_t factor. The bond price at time t for a bond maturing at time T with a payment of 1 and making no other payments, i.e., the price of a zero-coupon bond, is denoted as P(t,T).

In either model, let k = a – λs and q = b/k. Then $P(t,T) = A(t,T)/\exp[B(t,T)r_t]$. A and B differ in the two models.

For the Vasicek model:

$$B(t,T) = \left[1 - e^{k(t-T)}\right]/k$$

$$log[A(t,T)] = \left[q - \frac{s^2}{2k^2}\right][B(t,T) - T + t] - \frac{s^2}{4k}B(t,T)^2.$$

In the CIR model:

let $h = \sqrt{k^2 + 2s^2}$ and $C(t,T) = 2h + (k+h)\left(e^{h(T-t)} - 1\right)$. Then

$$B(t,T) = 2\frac{e^{h(T-t)} - 1}{C(t,T)}$$

$$A(t,T) = \left[\frac{2he^{(k+h)(T-t)/2}}{C(t,T)}\right]^{2kq/s^2}.$$

2.3 Interest Rates

It is often convenient to express the bond price as a single interest rate. If the rate is viewed as continuously compounding, then the payment of 1 at T can be expressed using the continuously compounded interest rate R(t,T) by $1 = P(t,T)\exp[R(t,T)(T - t)]$. Alternatively, an annually compounded rate Y(t,T) would give $1 = P(t,T)[1+Y(t,T)]^{T-t}$. Either rate can be expressed as a function of P(t,T) by backing it out of these formulas.

The drawback of the closed-form yield-curve formulas is that the entire yield curve is determined by the market price of risk and the short rate. This overly restricts the shapes of yield curves that can occur. This is a problem in pricing options, but also for risk management. If some yield curves are overrepresented in the scenario set, and others are missing entirely, the risk of possible outcomes of various positions would be misstated.

A way out of this problem is multifactor versions of the Vasicek and CIR models. It is possible to make double-mean-reverting processes for either form, but typically these days the interest rate is expressed as a sum of two partial interest rates, each following the same process but with different parameters. Thus you can have a two factor Vasicek process with the short rate the sum of two partial interest rates, each following a Vasicek process. The same is possible for the CIR process, but the CIR processes need to be independent. Fortunately it turns out that the bond prices and yield curves are still of closed form for the multifactor models. In fact, for the multifactor CIR model, the bond price is the product of the bond prices from the individual partial interest rates, and this makes the continuously compounded interest rates just the sum of those from the partial interest rates. The Vasicek model is similar, but can have an extra term for the correlation of the partial rates.

2.4 Some Empirical Findings

The interest rates generated from the Vasicek model are normally distributed. However, the $r^{1/2}$ factor on volatility in the CIR model destroys normality. The CIR rates end up being distributed as the sum of a series of gamma distributions. This makes MLE awkward at best. The normality of the Vasicek model also makes the two-factor Vasicek model tractable even if the factors are correlated. The two-factor CIR model requires independent factors in order to have closed-form yield curves. This ends up restricting the yield curve shapes.

Jagannathan et al. (2003) test CIR models with up to three factors. They show that MLE is possible using fast Fourier transforms to calculate the distribution function. They find that the three-factor model does capture enough of a variety of yield curve shapes, while the two-factor model does not, but that even the three-factor model is not sufficient for all options pricing, particularly for options sensitive to stochastic volatility.

Andersen and Lund (1998) fit a more complex generalization of the CIR model to U.S. Treasury rates. Their model can be expressed as:

$$dr_t = (b_t - ar_t)dt + s_t r_t^p dW_{rt}.$$

$$db_t = (m - cb_t)dt + hb_t^{1/2} dW_{bt}.$$

$$d\log s_t^2 = u(v - \tfrac{1}{2}\log s_t^2)dt + wdW_{st}.$$

The first two factors are a form of double-mean reversion. The third is stochastic volatility. They find that volatility does increase with interest rates, and the power $p = \frac{1}{2}$ on the rate is not unreasonable. But they also find that stochastic volatility is significant.

There is no simple formula for bond prices in their full model, so the grind-out simulation is the only alternative. This makes their model impractical for market pricing, but is not so bad for risk management. If you are going to update the model once a quarter, even taking a week to simulate it on a PC is not prohibitive. Venter (2004) tested this model for generating yield curves and found that it can produce a realistically wide variety of shapes, but that the market prices of risk need to be stochastic to get a historically reasonable distribution of shapes. The market price of risk for a given yield curve starting at time t has to be constant for all T, but it can change for different t and for different simulations. Andersen and Lund (1998) suggest that two market prices of risk are needed for this model, for r and b, but not for s.

There is ongoing work on modeling stochastic volatility. One of the simpler of these is Balduzzi et al. (1996), who reduce the yield curve calculation to the numerical solution of a system of two ordinary differential equations. Their model is:

$$dr_t = (b_t - ar_t)dt + s_t^{1/2} dW_{rt}.$$

$$db_t = (m - cb_t)dt + hdW_{bt}.$$

$$ds_t = u(v - s_t)dt + ws_t^{1/2} dW_{st}.$$

$$dW_{rt}W_{st} = \rho\, dt.$$

They start with a double-mean-reverting Vasicek framework, but add a stochastic volatility which follows a square-root process, to keep it positive, and then correlate it with the interest rate, which tends to make the volatility move with the interest rate as in CIR. In Andersen and Lund (1998), the values of both r and s affect the volatility, but here only s does. Also r and b can go negative. The advantage here is in the simpler calculation of the yield curve.

It turns out that ignoring stochastic volatility is not a problem for the prices of some options, but is for others. It might not be so bad to ignore it for most risk-management work. However some asset-liability hedging strategies used in life insurance may need tweaks in periods of high volatility, and if so this could be an issue. On the other hand, stochastic volatility increases the overall risk and probably affects the shape of the yield curve, so there is a case that it should always be included.

2.5 An Example

A three-factor CIR model is fit to U.S. three-month Treasury rates as an example. Data was taken from the St. Louis Federal Reserve Fred database at http://research.stlouisfed.org/fred2/categories/116. The interest rate r = x+y+z is the sum of three partial interest rates each following a CIR model:

$$dx_t = (b_1 - a_1x_t)dt + s_1x_t^{1/2}dW_{xt}.$$

$$dy_t = (b_2 - a_2y_t)dt + s_2y_t^{1/2}dW_{yt}.$$

$$dz_t = (b_3 - a_3z_t)dt + s_3z_t^{1/2}dW_{zt}.$$

There are nine parameters to estimate: 3 a's, 3 b's and 3 s's. While MLE is possible, it is complex, so an ad hoc procedure involving simulated moments is used for illustration. First some properties of the interest rates are identified that would be desirable for the model to mimic. Then starting parameters are selected and a long series simulated of x, y and z and so r. The series of r's are then tested to see how well they match the selected properties, and the parameters are iterated to produce the best match overall. This requires a nonlinear optimizer better than that which comes in common spreadsheets, but auxiliary packages like Poptools are available with such capabilities; or other software than spreadsheets can be used.

Gallant and Tauchen (1996) show that this method can approach MLE in efficiency if the right properties are selected. Their approach, called efficient method of moments, involves fitting an auxiliary model and then tweaking the parameters so that the simulated series fits as closely as possible to that model. However here a selected set of properties will be matched for illustration.

One thing selected to try to match is the autocorrelation structure. Here a question quickly arises as to what period to use for the empirical rates. There was a change in the Federal Reserve management of interest rates and monetary aggregates in the early 1980s, and some series show different behavior if viewed from 1983 on. The interest rates are highly autocorrelated in all periods, but a bit less so after 1983. Taking September 2009 as an ending point, comparing the short rate from June 1949 to that starting on January 1, 1983 shows a somewhat different autocorrelation pattern. Both series start at lag 1 at 98.6 percent correlation, but the 1983 series' autocorrelations drop faster, then pick up again later. At lag 70 the longer series has a 41 percent autocorrelation, versus 35 percent for beginning in 1983. The AR(1) value of 0.986^{70} = 37% is a bit between these but too far from either.

In the end, the longer series was selected as the basis to match. One reason was that several financial series recently have broken out of their 1983–2005 ranges. A series of 10,000 months was simulated from selected starting parameters (a's, b's and s's), and the sum of the absolute differences between the simulated and empirical autocorrelations for the first 180 lags was taken as one moment error to try to minimize by improving the parameters.

Other moments selected were the mean, standard deviation, CV (standard deviation/mean) and skewness of the interest rates. Also moments of the monthly changes in the interest rate and the absolute value of the monthly changes were selected. Finding parameters to match all of these would make the simulated series reasonably similar in statistical properties to the data.

Absolute errors instead of squared errors were minimized in an attempt to keep any outliers from having too much weight. Also the various moments were given different selected weights, reflecting a judgment on their importance. Tables 1 and 2 show the moments attempted to be matched, the weights given to each, the target and fitted values, and the fitted parameters:

Table 1. Moments to Match

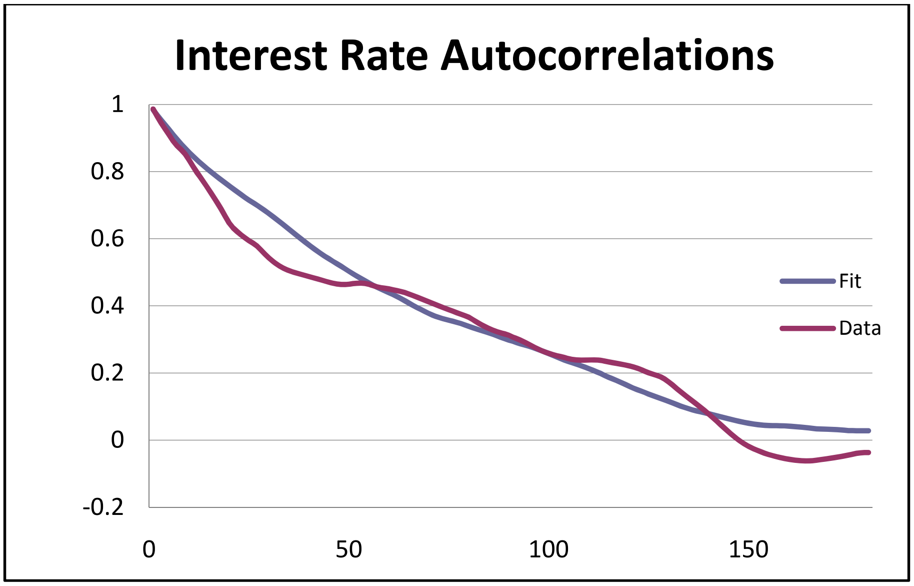
Weight	Moment	Target	Fitted
1	Each of 180 autocorrelations	see	graph
1000	Mean interest rate	4.76%	4.76%
1000	Std dev interest rate	2.89%	1.90%
10	CV interest rate	60.7%	40.0%
10	Skw interest rate	1.05	2.55
10	Std dev monthly change	0.43%	0.34%
0.1	Skw monthly change	-1.78	-0.25
100,000	Mean abs. monthly change	0.2345%	0.2345%
10,000	Std dev abs. monthly change	0.36%	0.24%
1	CV abs. monthly change	1.54	1.04
1	Skw abs. monthly change	5.17	2.36

Table 2. Fitted Parameters

a	b	s
0.49076	0.01301	0.00281
0.22609	0.00304	0.02854
0.01191	0.00008	0.00766

The fits do appear to respond to the selected weights, which probably are not ideal. The match of some moments was not great. The fitted and empirical autocorrelations are graphed in Figure 1. The AR(1) coefficient corresponds roughly to 1 – a, so the third partial interest rate, with a = 1.2%, has high autocorrelation. The reverting mean is b/a, so these are 2.65%, 1.34% and 0.66%, respectively. The second process is the most volatile, with the highest s.

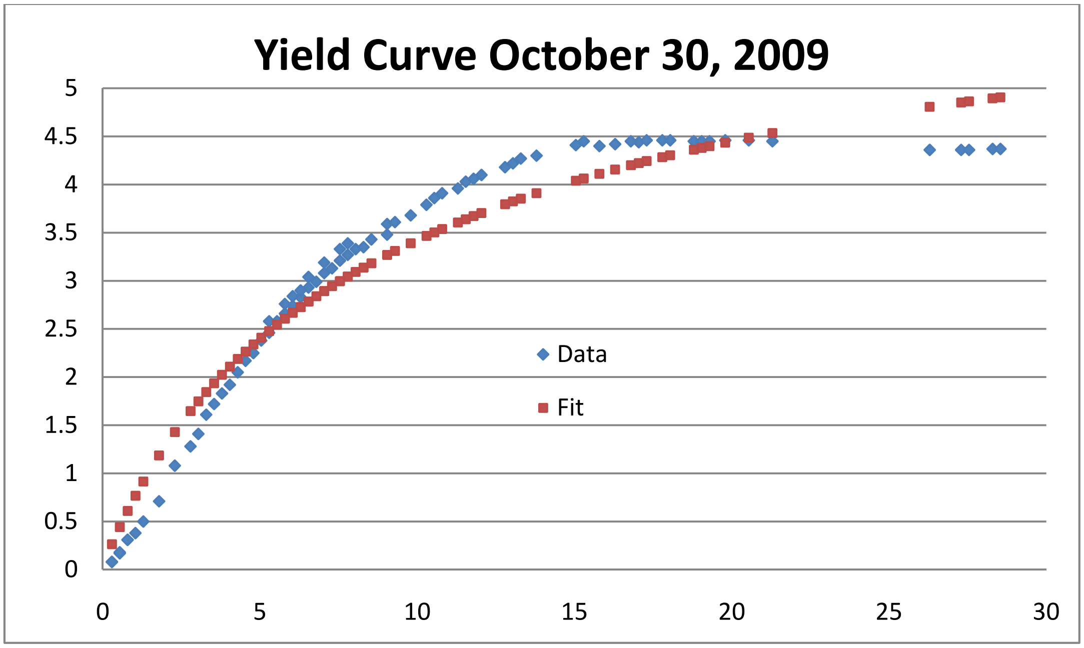
Figure 1. 180 Autocorrelations of Fitted and Empirical Monthly Short Rates



A simulation of future scenarios needs starting values of the three processes. These can be calibrated, along with the three market prices of risk, by trying to fit the latest observed yield curve using the bond-price formulas. One snag here is that most bonds are coupon-bearing, whereas the formulas are for zero-coupon. Each coupon payment can be viewed as a zero-coupon bond itself, and there are formulas for backing out a zero-coupon curve from the published bond prices. The data used here is taken from the zero-coupon curve published by the *Wall Street Journal* at http://online.wsj.com/mdc/public/page/2_3020-tstrips.html. It is particularly difficult to get a good fit at the time of this writing, as the short rate is 0.04 percent. The yield curves then basically come from the market prices of risk only, as all the partial short rates are virtually zero. This gives three parameters to fit to the curve instead of the usual six.

The best fit was found by setting the first rate to 0.04 percent with a market price of risk of 1.9, and all the other rates and prices of risk to zero. The fitted and empirical yield curves are shown in Figure 2. The fit is reasonable but not great. This turns out to be not atypical for short-rate models.

Figure 2. Fitted and Empirical Yield Curves as of Oct. 30, 2009



2.6 Improving the Fit of Yield Curves from Short-Rate Models

Any of the short-rate models discussed here can be made to fit exactly to the initial term structure, basically by making the b parameters deterministic time-dependent functions chosen to make the fit exact. This process traces back at least to Hull and White (1990) and so is often referred to in their names. The function needed can be computed directly from the bond prices P(t,T).

First, define the forward rate f(t,T) as $f(t,T) = -{}^{\partial logP(t,T)}/_{\partial T}$. The bond price is the expected discounted price of the payout of 1 discounted along all possible paths of the future risk-neutral rate. The forward rate is a representation of all the possible rates being discounted over at time T, so using it takes care of the expectation. In formulas, $P(t,T) = Ee^{-\int_t^T r_s ds} = e^{-\int_t^T f(t,s)ds}$. Let f(0,T) denote the forward rate curve for the market yields at time 0.

If we extend the basic Vasicek model, already adjusted for market price of risk, to:

$$dr_t = (b(t) - kr_t)dt + sdW_t,$$

it is possible to show that the market yields are reproduced by taking:

$$b(t) = \frac{\partial f(0,t)}{\partial dt} + kf(0,t) + \frac{s^2}{2k}(1 - e^{-2kt}) \, .$$

The bond price can still be written as P(t,T) = A(t,T)/exp[B(t,T)r_t]. Now A and B are

$$B(t,T) = \left[1 - e^{k(t-T)}\right]/k$$

$$A(t,T) = \frac{P(0,T)}{P(0,t)} exp\left\{B(t,T)f(0,t) - \frac{s^2}{4k}(1 - e^{-2kt})B(t,T)^2\right\}$$

Brigo and Mercurio (2007) show similar formulas for CIR and multifactor models, and is an excellent source for interest rate models in general.

If this is done so the yield curve matches the current curve exactly, then the market prices of risk and the partial interest rates would be calibrated to options prices.

Another way to improve the fit of multifactor short-rate models, introduced by Dai and Singleton (2000), is to add terms for interaction and correlations among the factors. They find a fairly general framework where the yield curve can be calculated in closed form once two ordinary differential equations are solved numerically. They call this an almost closed-form solution. This scheme allows, for instance, positive correlation among the CIR factors to be included. They test a number of models empirically, and find that a generalization along the lines they introduce of the Balduzzi et al. (1996) model is the best-fitting in this class. This model allows positive and negative correlations among the factors, which appears to provide more realistic fits to empirical data. Their best-fitting model just subtracts a multiple of dW_{bt} from the r diffusion and a multiple of dW_{rt} from the b diffusion, which introduces negative correlation between b and r.

2.7 Summary of Interest Rates

The study of interest-rate models is much more extensive, but the basic models with closed-form yield curves have been covered. As these do not appear to be capable of handling stochastic volatility, the trade-off of ignoring that versus analytical tractability has to be considered. Probably some effort to incorporate at least the simpler models with stochastic volatility would be worthwhile for insurer risk management.

Another issue often ignored is that the market price of risk, which has to be constant for every T at time t, probably changes over t. Studying this for a given model would involve calibrating yield curves at various historical points to get historical values for the market prices of risk, and then building a model for their evolution. This is an area worthy of more research.

3. Inflation

In the United States, the government management of inflation changed in 1983 along with that of interest rates, and there are differences in the properties of the series depending on which time frame is used, as can be seen in Table 3. That is a problem from the start for inflation modeling, and it is made no easier by the fact that in the last year or two the inflation rate has escaped from the range of its 1983 to 2005 values. Perhaps modelers should use scenarios from models fit to each period.

The Fred database is also a source of inflation data. The seasonally adjusted series used here is at:

http://research.stlouisfed.org/fred2/series/CPIAUCSL?cid=9.

The persistent autocorrelation of the longer series is more difficult to fit, as discussed before. Although partial differencing has been used to model this, it is also possible to use multifactor AR(1) processes somewhat analogously to what is done in interest-rate modeling. The double-mean-reverting process discussed earlier can quite easily model the 1983 to 2009 series, as perhaps could a simple AR(1) model, but it is not able to capture the features of the 1947 to 2009 series. However, a process consisting of three independent AR(1) processes for partial inflation rates is able to do so.

Table 3. Properties of U.S. Monthly Inflation Rate Seasonally Adjusted in Percentage Points

Moment	1947-9/2009	1983-9/2009	Fit
Mean inflation=100 log(CPI_{j+1}/CPI_j)	0.3068	0.2469	0.3068
Stdv inflation	0.3533	0.2566	0.3535
Mean Abs 1st difference of inflation	0.2188	0.1964	0.2188
Stdv Abs 1st difference	0.2423	0.1953	0.1654
# of autocorrelations above 5%	89	1	177
-Ratio max/min	1.15	0.81	1.24

The model will be denoted in the following form: Inflation = $q = q_1 + q_2 + q_3$

$$\Delta q_1(t) = u_1[v_1 - q_1(t)] + w_1\varepsilon_1$$

$$\Delta q_2(t) = u_2[v_2 - q_2(t)] + w_2\varepsilon_2$$

$$\Delta q_3(t) = u_3[v_3 - q_3(t)] + w_3\varepsilon_3.$$

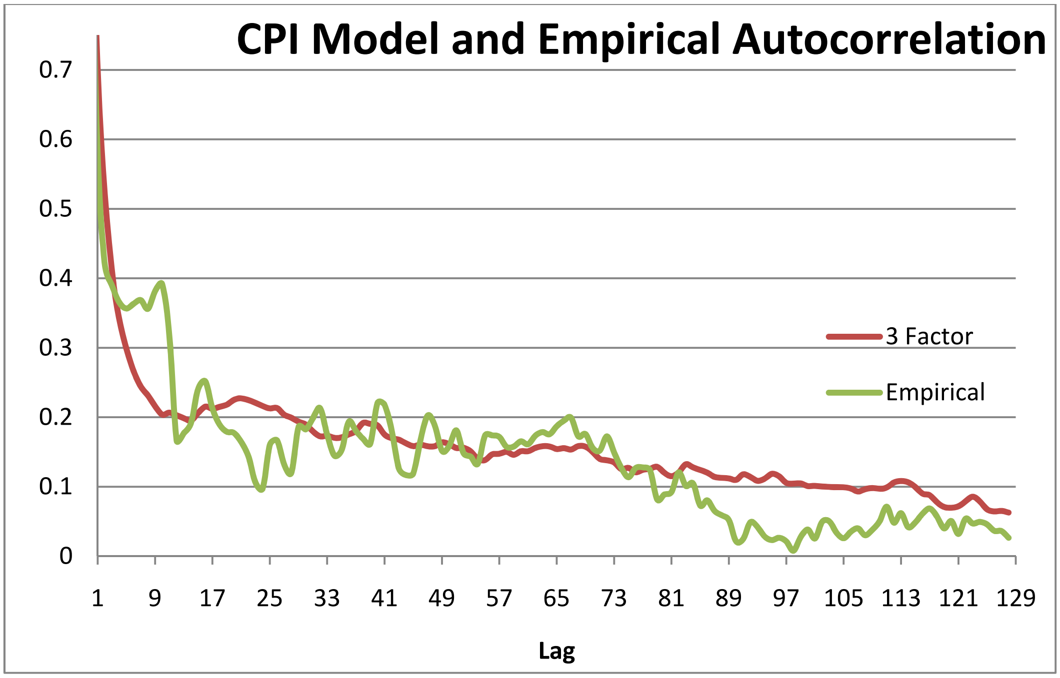
In this form, the reverting mean is v_j and the AR(1) b coefficient that determines autocorrelation is $1 - u_j$. The variance of a process is $w_j^2/u_j(2-u_j)$ The following parameters were fit by simulated method of moments:

Table 4. Parameters for Three-Factor Inflation Model

u	v	w
.681	.354	.008
.007	.022	.023
.384	-.021	.243

The second process has mean near zero but high autocorrelation and is somewhat volatile. The third process has a slightly negative mean and is quite volatile. The first process is most of the inflation and is fairly stable. The generalized moments fit better than those for the interest rates. The first 89 autocorrelations were targeted to match, but the resulting model maintained a higher level of autocorrelation than the data after that. If more autocorrelations were matched this might have been less pronounced. Figure 3 graphs the empirical and fitted autocorrelations.

Figure 3. Empirical and Fitted Autocorrelations CPI 1947 to September 2009



The double-mean-reverting process fits well for the 1983-onward period, with all general moments matching well under simulated method of moments. However, this period did not experience any high inflation or interest rates, and generated scenario sets should probably allow some possibility that these could occur, so closely matching the data from that period is probably not sufficient for risk management purposes.

Calibrating the unobserved factors in these models is a bit more arbitrary than with interest rates, as there is not a yield curve to calibrate to. Perhaps the temporary mean in the double-mean-reverting process could be set to an average of recent rates, or in the three-factor model the latest inflation rate could be subdivided in proportion to the means of the partial inflation rates.

4. Association of Interest Rates and Inflation

Interest rates and inflation tend to be high or low at the same time. This association is not exactly a dependency, as it is not clear that one series strictly drives the other. There are various ways to look at this association. The correlation coefficient between the three-month rate and inflation depends on the period selected. From 1947 to September 2009, it is 43 percent; but from 1983, it is 24 percent. For the 25 years from January 1983 to December 2007, missing the low rates of both series since then, it is even less, at 21 percent.

Besides the degree of correlation, there is an issue of where in the distribution the correlation takes place. For inflation and interest there appears to be strong tail association: both series tend to be unusually low or unusually high at the same time. One way to quantify this is the tail association functions L and R defined in Venter (2002). For z in (0,1), and two random variables X and Y, $L(z) = \Pr(F_X(X)<z|F_Y(Y)<z)$, and R(z) just reverses both inequalities. By applying the definition of conditional probability it can readily be seen that these definitions are symmetric in X and Y. These functions express the probabilities of both variables being small or large at the same time, as defined by probability levels.

Since $L(1) = R(0) = 1$, the main interest is in L for $z < ½$ and R for $z > ½$. The limits R(1) and L(0) are the tail concentration coefficients. It is not unusual for both of these to be zero, as for instance in the case of the bivariate normal distribution. If they are positive, that indicates that a fair degree of the correlation takes place in the more extreme cases.

Figure 4 graphs the empirical L and R functions for inflation and the short rate starting in 1947. While these tend to get jumpy near 0 and 1, values of the limits L and R around 10 percent to 20 percent seem reasonable. Some of the odd look of segments of the graph is due to ties in the interest rate ranks. Moving away from ½ the slope of L is flatter, suggesting more local correlation in the smaller losses. This could be a result of the fact that interest rates are more volatile at higher values.

Figure 5 shows the same functions but starting in 1983.

Figure 4. L and R Functions for CPI Inflation and Three-Month Rate from 1947

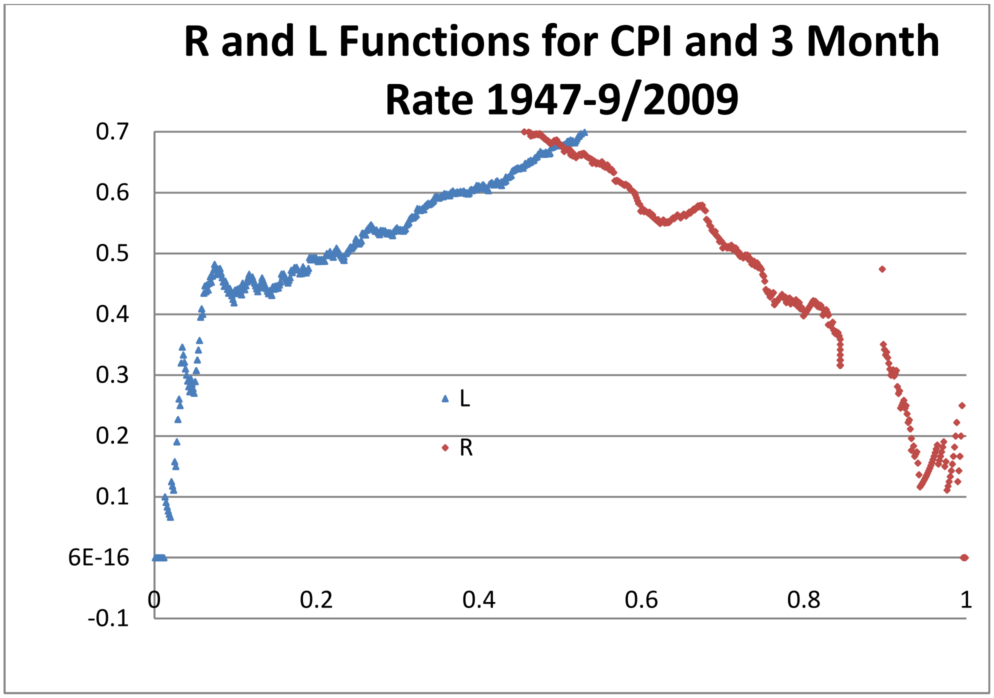
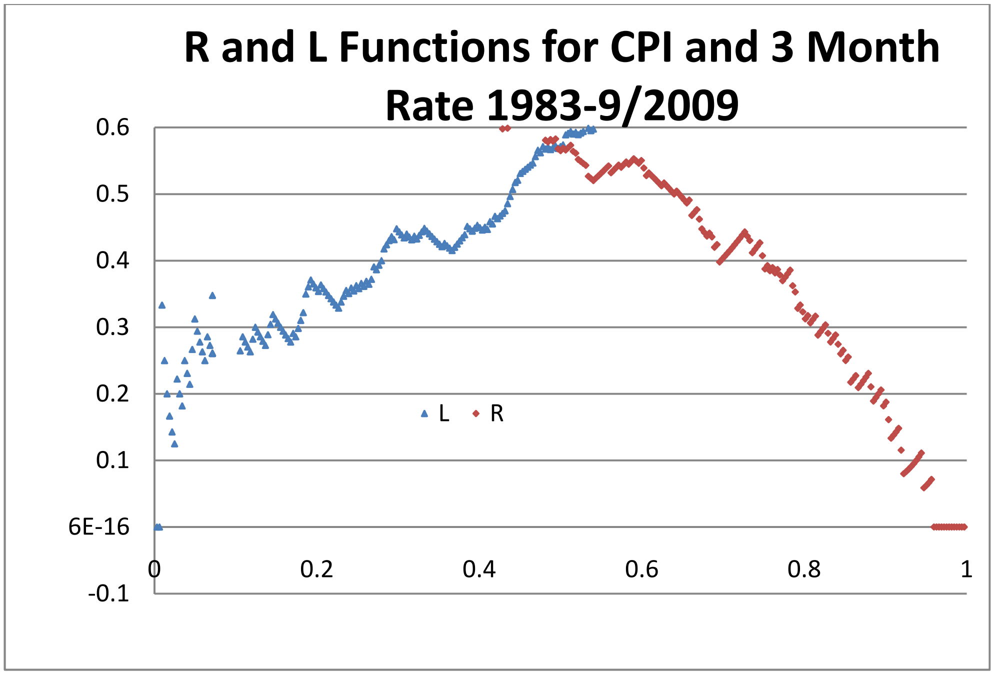


Figure 5. L and R Functions for CPI Inflation and Three-Month Rate from 1983



Here R seems headed straight to zero, while L seems to have a higher limit than before. In this period there was never a time of high inflation or high interest rates, which there was before 1983; this is what is creating the difference in the two R functions. Also a period of deflation with only moderately low interest rates in the late 1940s is excluded from the later graph, which strengthens the L values there.

Probably R and L values of 10 percent to 15 percent in the limits would be reasonable. These could be achieved with a t-copula with between 5 to 7 degrees of freedom, depending on the degree of overall correlation targeted.

If the three-factor CIR and AR(1) models are used for interest rates and inflation, the question arises about how association between the series can be simulated. Part of the problem is that the three CIR factors have to be independent of each other. Thus the correlation with inflation has to be done in such a way that it does not induce a correlation among the CIR factors.

One way to do this would be to correlate the largest CIR factor with each of the inflation factors. The inflation factors do not have to be independent, but if they are not, new parameters might be needed. This could readily be handled by simulated method of moments estimation. Another approach could correlate the largest CIR factor with the largest inflation factor, the middle with the middle, and the smallest with the smallest, measured by means. Selecting a target correlation of 33 percent between inflation and interest rates, a 92 percent correlation factor-to-factor was found to work.

To simulate the two correlated first factors, two correlated standard normal variates would be needed. Correlated standard normal variates can be generated in general by the Cholesky decomposition of the correlation matrix, but for just a single correlation ρ between two factors, this reduces to the following simple algorithm:

1. Generate two independent standard normal draws x_1 and x_2. E.g., in Excel, x_1 = normsinv(rand()) and similarly for x_2.

2. The correlated standard normal variates are then $y_1 = x_1$ and $y_2 = \rho x_1 + (1 - \rho^2)^{1/2} x_2$.

It is easy to see that the variance of y_2 is 1. This algorithm could be repeated to get correlated draws for the 2nd and 3rd factors that are independent of these.

This algorithm generates correlated bivariate normal variates, but the tail concentration coefficients are zero for these. A way to get nonzero tail concentration is to simulate from the t-copula, which takes a few more steps. Using n degrees of freedom these are:

3. Draw v from a chi-squared distribution with n degrees of freedom. In Excel this is v = 2*gammainv(rand(), n/2, 1).

4. Let $z_1 = y_1(n/v)^{½}$ and $z_2 = y_2(n/v)^{½}$ (same v). These are correlated bivariate t variates, but have to be converted to normally distributed variates. This is done by calculating the t probabilities then applying the inverse normal to get the percentiles.

5. The correlated variates are thus $w_1 = \Phi^{-1}T_n(z_1)$ and $w_2 = \Phi^{-1}T_n(z_2)$, where Φ is the standard normal distribution function and T_n is the t-distribution with n degrees of freedom. In Excel, $T_n(z)$ = ½ sign(z)betadist[$z^2/(n+z^2)$, ½, n/2] + ½ .

To get a tail concentration around 10 percent to 15 percent between two sums may require something different than 5 – 7 degrees of freedom for the individual factors, but the relatively high correlation factor-to-factor will in itself increase the factor-by-factor tail concentrations.

Inflation-indexed Treasury bonds present a more challenging modeling problem. The real and nominal yield curves and options prices are available and there are also options on the inflation index itself. One popular approach to modeling these processes is the JY model from Jarrow and Yildirim (2003). They look at the CPI as the exchange rate between the real and nominal economies. Their model starts with a formulation of the evolution of the real and nominal forward rates, but they provide a formulation for the short rates that can be expressed as Hull-White models.

Letting n_t and r_t denote the nominal and real interest rates and I_t the CPI, the processes can be expressed using a correlated trivariate Brownian motion (W_n, W_r, W_I), which has correlations $\rho_{n,e}$, $\rho_{n,I}$, and $\rho_{r,I}$ as follows:

$$dn_t = [b_n(t) - k_n n_t]dt + s_n dW_{n,t}$$

$$dr_t = [b_r(t) - \rho_{r,I} s_I s_r - k_r r_t]dt + s_r dW_{r,t}$$

$$dI_t/I_t = (n_t - r_t)dt + s_I dW_{I,t} \, .$$

Here both b_r and b_t follow the Hull-White formula, which letting x be n or r is:

$$b_x(t) = \frac{\partial f_x(0,t)}{\partial dt} + k_x f_x(0,t) + \frac{s_x^2}{2k_x}(1 - e^{-2k_x t}).$$

The CPI can now be simulated as a geometric Brownian motion with drift n – r, and n and r can be simulated as mean-reverting Brownian motions. These processes are overly simplistic for generating realistic scenarios, but at least they tie together the real and nominal rates with the CPI in an arbitrage-free environment, and match the latest actual yield curves. Options pricing has been worked out using these processes.

5. Equities

The starting point for modeling equity prices is geometric Brownian motion. This is the basis of the Black-Scholes options pricing formula, for example. With drift, the process for S can be written:

$$dS(t)/S(t) = \mu dt + \sigma dZ(t).$$

However, over time, evidence has accumulated to suggest that this process is not heavy enough in the tails. Merton (1976) already had proposed adding compound Poisson jumps to the log process. He suggested lognormally distributed jumps. More recently a series of papers, starting with Kou (2002) and including parameter estimation by Ramezani and Zeng (2007), work out options pricing with a heavier-tailed jump-size distribution and separate compound Poisson processes for upward and downward jumps. They find that there are more and larger downward than upward jumps.

Details of the power tail distributions they use are given below. If X_i is the i^{th} downward jump factor, and Y_j is the j^{th} upward jump factor, then $X_1 - 1$ and $Y_j - 1$ are added to the log process to reflect the jumps. If v and w are the Poisson parameters for downward and upward jumps, respectively, then the process is:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dZ(t) + d\left(\sum_{i=0}^{N(vt)} [X_i - 1] + \sum_{j=0}^{N(wt)} [Y_j - 1] \right).$$

The downward jump size distribution is defined for $0 < x < 1$ by $F(x) = x^{\alpha}$ and the upward jump size distribution is defined for $y > 1$ by $F(y) = 1 - y^{-\beta}$. The log of the jump or its negative is exponentially distributed in either case, so sometimes this process is called the double-exponential jump diffusion process (DEJD).

Ramezani and Zeng (2007) show that the distribution function of S(t) given S(0) is expressible in closed form, and so they are able to derive the likelihood function for equally spaced observations. They use daily values of the S&P 500 and Nasdaq indexes to fit the DEJD by MLE. Unfortunately the optimization is quite calculation intensive and requires substantial computational resources. They find, however, that the DEJD fits these series considerably better than do geometric Brownian motion or the lognormal jump diffusion model, even when they use a conservative (high) penalty for extra parameters. Even though difficult to estimate, the model is easy enough to simulate from.

The estimated parameters and their standard deviations are:

Table 5. DEJD Parameters

w	v	β	α	μ	σ
0.4640	0.5624	174.09	185.92	0.0007	0.0047
0.0714	0.0933	0.43	0.44	0.0000	0.0000

The relatively high values for β and α simulate jump factors $p^{1/\alpha}$ and $p^{-1/\beta}$ near 1 most of the time. Here t=1 is one trading day. The w and v parameters suggest there is about 1 jump per day on the average. The popularity of the DEJD is due to its goodness of fit and the fact that many options prices have been worked out for it. There are numerous other possible equity price models however.

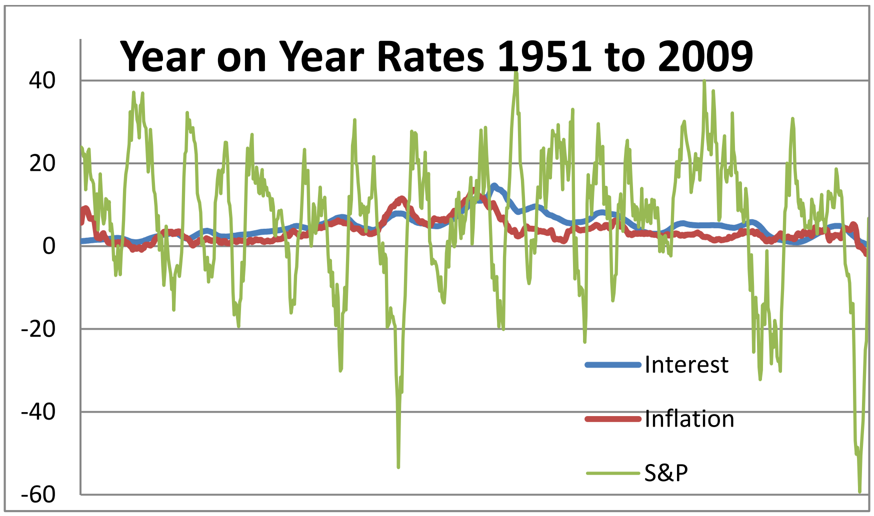
Associating equity prices with inflation or interest rates is difficult. In the early stages of a recession, stock prices, inflation and interest rates would all be expected to go down. But the stock market could start rising well before either of the others. Table 6 shows some correlation coefficients.

Table 6. S&P 500 Correlation Coefficients at 9/2009

S&P 500 Correlation with	From 1950	From 1983
Inflation	-6.2%	0.8%
Three-month Treasury rate	-3.1%	8.4%

Part of the difficulty is that equity prices are much more volatile than inflation or interest rates, which makes correlation coefficients unstable. Figure 6 graphs year-on-year changes for each series.

Figure 6. Year-on-Year Changes for Inflation, Interest and S&P 500



Perhaps meaningful associations could be established if the asset scenarios were imbedded in a more comprehensive model of the whole economy. Otherwise, making equities independent of inflation and interest does not seem a bad place to start.

6. Currency Exchange Rates

Currency exchange rates tend to be highly volatile, and explanatory variables have little predictive power, so the random element of any model is likely to dominate the generated scenarios. Also exchange rates have complex correlation patterns, which again highlights distributional issues. They have persistent autocorrelations as well.

The lag one autocorrelation of the US $ exchange rate to other major currencies, measured monthly, tends to be in the range 97 percent to 99 percent, and the subsequent lags fall off roughly as powers of lag one, as would an AR(1) process that is close to a random walk. A random walk or its continuous analogue, a driftless Brownian motion, is in fact often proposed as a model of exchange rates.

The theory of interest-rate parity holds that the difference in interest rates in the two economies should be the deterministic trend in the exchange-rate process. This idea, which can be traced back to at least Fisher (1930), means that a deposit made in either currency would have the same expected return in each currency. Empirical studies have consistently rejected this theory, however.

The form so stated is called uncovered interest parity (UIP). Covered interest parity is a similar idea, but it includes holding a position in the forward rates as well. Taking the US $ as the base currency, let $\Delta s_{t,k}$ denote the change from time t to time t+k of the log of the cost of the other currency in dollars. UIP says this is the interest rate difference in the k-year bond at time t, $i_{t,k} - i_{t,k}{}^*$ plus a random term. Testing this usually involves doing a regression:

$\Delta s_{t,k} = \alpha + \beta(i_{t,k} - i_{t,k}^*) + \varepsilon_{t,k}$ where UIP implies $\alpha = 0$, $\beta = 1$.

Froot and Thaler (1990) survey the literature on such tests and find that β is always less than 1, is usually less than 0, and averages –0.88. More recent studies and surveys have been in the same ballpark. They also consistently find a large error variance. So not only does UIP fail to predict the movement of the exchange rate, but it also fails to even predict the direction of movement.

An interesting exception is Chinn and Merideth (2001), who find the standard result for short-term rates, but find that UIP works well, although still with a high error variance, when using five- and 10-year interest rates to predict movements in the exchange rate over five- and 10-year periods.

What does this all suggest for modeling exchange rates? This is not so much a matter of fitting a model as constructing one that seems to make sense. Considerable judgment is required. Taking X as the exchange rate, a starting point is driftless geometric Brownian motion,

$dX_t/X_t = s\,dW_t$.

You could get s by an empirical study of changes in the rate. If you do not like the random walk aspect, you can introduce a slow drift towards X_0, which could be either the current rate or a consensus estimate of the future rate:

$dX_t/X_t = 0.1(X_0 - X_t)dt + s\,dW_t$, thinking of t measured in years.

Fitting an AR(1) model might give a good handle on the speed of mean reversion, here taken as

0.1. Perhaps an aspect of UIP can be included by adding in a slow drift based on the difference in the five-year rate:

$$dX_t/X_t = 0.1(X_0 - X_t)dt + 0.2(i_{t,5} - i_{t,5}^*)dt + sdW_t.$$

This requires a model of interest rates in each currency. These tend to be correlated, so that correlation should be reflected in modeling the interest rates before the exchange-rate model is applied. As we saw above, it is possible but a bit awkward to correlate multifactor models.

If more than one currency is being modeled, the associations among the exchange rates need to be considered as well. Even if Brownian motion is being used for each rate, a multivariate Brownian motion based on the multivariate normal distribution would not work, as there is a fair amount of tail concentration in exchange rates. Even if two rates are not particularly correlated usually, they tend to be in extreme changes. One source of extreme changes is something happening to the base economy, in which case all the exchange rate would be affected. Using t-copulas with normally distributed marginals would be one way to get in this effect.

Venter et al. (2007) look at modeling exchange rates via copulas. While the t-copula is reasonable for this, it has some weaknesses. In particular it strongly links the tail concentration index for different pairs of variates to their correlations. However, in the data, two fairly uncorrelated rates like the Swedish kroner and Canadian dollar can have comparable tail concentration to more highly correlated rates like the Swedish kroner and the yen. This is probably due to extreme movements that arise from US $ issues instead of individual currency issues.

Basically the t-copula does not have enough parameters to control for effects like this. Alternatives include the grouped t-copula of Daul et al. (2003) and the individuated t-copula of Venter et al. (2007). Getting the right copula for exchange rates is still an open issue.

With its high volatility, getting the distributional aspects of exchange rate scenario generation right is probably more important than modeling the drifts. The volatility, autocorrelations and correlations across currencies are all part of this. However some mean reversion and a degree of inclusion of UIP would probably improve the reasonableness of the scenarios.

7. Risky Bonds

Corporate and municipal bonds involve default risk, as well as liquidity and taxation issues. The pricing for these bonds includes all these issues, and perhaps others. A key method for pricing risky bonds was presented by Duffie and Singleton (1999). They show that all the models for Treasury bonds also apply to risky bonds if you use a risk-adjusted short-term rate. This adjustment includes default risk, but also could include other elements, such as liquidity premium.

If a particular class of risky bonds were being modeled in isolation, the same models as for Treasury bonds could be applied directly. This is the approach of Jagannathan et al. (2003), who apply it to LIBOR bonds. However, if scenarios are being generated that need to have consistent yield curves for Treasury and risky bonds, more work is needed.

Say r_t is the Treasury short rate, and the adjusted risky rate is $r_t + s_t$, where s_t is non-negative. What is needed is yield curves for both processes. Duffie and Singleton do this by making both r and s combinations of the same underlying factors.

For example, consider the square-root process used in the CIR models:

$$dx_t = (b - ax_t)dt + sx_t^{1/2}dW_t.$$

Multiplying this by a constant still leaves it a square-root process, just with different parameters. So if there are three factors x, y and z, setting

$$r_t = g_0 + g_1x_t + g_2y_t + g_3z_t \text{ and}$$

$$s_t = h_0 + h_1x_t + h_2y_t + h_3z_t$$

makes r still a sum of three independent square-root processes, and makes s and r + s such sums as well. Thus the bond price formulas for CIR processes can be used for both the Treasury and the risky bonds, possibly each with their own market prices of risk. The constant g_0 cancels out in the diffusions for dr and enters the bond price as a factor $\exp(-g_0)$. Since the partial bond prices multiply to the total bond price, g_0 can be spread in any way among the factors. Although arguably $g_0 = 0$, the same works for g_0+h_0 in the risky bond diffusion.

Higher yield spreads for risky bonds are sometimes seen with higher Treasury rates, but there could be cases where you would like to have negative correlation between two of the factors instead of independent factors. For instance, in a recession, the Treasury rate might be very low but s might be higher due to more default risk. This would make the risky bond rates more stable than the Treasury rates in these times, as higher spreads would to some extend offset the lower Treasury rates. Negative correlation was also found by Dai and Singleton (2000) to give more realistic yield curve scenarios in any case. Their generalization of the Balduzzi et al. (1996) three-factor model can apply to Treasury versus corporate rates, but the short rate is no longer the sum of the factors as it is for

CIR and Vasicek. Nonetheless, it is possible to combine the r and r+s processes into a single model form. Both Dai and Singleton (2000) and Duffie and Singleton (1999) discuss this.

In any case, the parameters can be estimated by an application of simulated method of moments to the Treasury and risky bonds simultaneously. Also different s parameters could be used for different classes of bonds, and these could even be nested to force a relativity among bond classes. That is, r + s + u might be the risk-adjusted rate for BB bonds when r + s is the rate for AA bonds, etc.

8. Concluding Remarks

Equities, while volatile, are reasonably modeled. Inflation is not too hard to model either, but the main question is what time period to use as a base. Foreign exchange rates are not very predictable, and models should recognize that and at least get the stochastic elements right, like tail concentration. Bond models are the most complex. Trying to keep a closed-form yield curve constrains the models, perhaps by more than is worthwhile. Interacting Treasuries, inflation-adjusted Treasuries and defaultable bonds, with complex yield-curve shapes for each, is a challenge.

Here ad hoc estimation by simulated method of moments has been used for illustration. In some ways this is a natural for scenario generation, in that what you really want are simulated scenarios to match the properties of the historical data. However there is always a question about the selection of properties to match and the accuracy of parameters that come from ad hoc methods.

MLE can be used whenever the distribution of incremental data can be formulated. Jagannathan et al. (2003) show that this can be done for CIR models, but it is involved. The efficient method of moments was used by both Andersen and Lund (1998) and Dai and Singleton (2000). This involved fairly complex auxiliary models in both cases. Both methods have the standard robustness issues of MLE. If the model or auxiliary model is misspecified, or there are outliers in the data, the estimation can be thrown off. For instance, if partial differencing is appropriate but not included in the set of auxiliary models, efficient method of moments might not get the properties of the data that would reproduce the historical series

Even a misspecified model can work for scenario generation. The question for this is not whether or not the model generated the data, but whether or not the model reproduces the key properties of the data. Even if MLE or efficient method of moments is used, testing the fit by matching of simulated properties could be informative.

References

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.

Baillie, R.T., Han, Y.W., and Kwon, T-G. 2002. "Further Long Memory Properties of Inflationary Shocks." Southern Economic Journal 68(3): 496–510.

Balduzzi, P., Das, R.S., Foresi, S., and Sundaram, R. 1996. "A Simple Approach to Three-Factor Affine Term Structure Models." The Journal of Fixed Income, December: 43–53.

Brigo, D., and Mercurio, F. 2007. *Interest Rate Models—Theory and Practice,* 2nd ed. Berlin, Heidelberg, New York: Springer.

Chinn, M., and Merideth, G. 2001. "Testing Uncovered Interest Parity at Short and Long Horizons." Manuscript at http://dipeco.economia.unimib.it/Persone/Colombo/finarm/Meredith-Cinn-UIP.pdf . Also see Chinn, M. 2006. "The (Partial) Rehabilitation of Interest Rate Parity in the Floating Rate Era: Longer Horizons, Alternative Expectations and Emerging Markets." Journal of International Money and Finance 25: 7–21.

Dai, Q., and Singleton, K. 2000. "Specification Analysis of Affine Term Structure Models." Journal of Finance 55(5): 1943–1978.

Daul, S., De Giorgi, E., Lindskog, F., and McNeil, A. 2003. "The Grouped t–Copula with an Application to Credit Risk." RISK 16: 73–76.

Duffie, D., and Singleton, K. 1999. "Modeling Term Structures of Defaultable Bonds." Review of Financial Studies 12(4): 687–720.

Fisher, I. 1930. *The Theory of Interest.* New York: Macmillan.

Froot, K.A., and Thaler, R.H. 1990. "Foreign Exchange." Journal of Economic Perspectives 4(3): 179–192.

Gallant, A. R. and G. E. Tauchen. 1996. "Which Moments to Match?" Econometric Theory 12: 657–681.

Hull, J., and White, A. 1990. "Pricing Interest Rate Derivative Securities." The Review of Financial Studies 3: 573–592.

Jagannathan, R., Kaplin, A., and Sun, S. 2003. "An Evaluation of Multi-Factor CIR Models Using LIBOR, Swap Rates and Cap and Swaption Prices." Journal of Econometrics 116: 113–146.

Jarrow, R., and Yildirim, Y. 2003. "Pricing Treasury Inflation Protected Securities and Related Derivatives Using an HJM Model." Journal of Financial and Quantitative Analysis 38(2): 337–359.

Kou, S. 2002. "A Jump Diffusion Model for Option Pricing." Management Science 48(8): 1086–1101.

Merton, R.C. 1976. "Option Pricing When Underlying Stock Returns are Discontinuous." Journal of Financial Economics 3: 224–244.

Ramezani, C.A., and Zeng, Y. 2007. "Maximum Likelihood Estimation of the Double Exponential Jump-Diffusion Process." Annals of Finance 3: 487–507.

Venter, G. 2002. "Tails of Copulas." Proceedings of the Casualty Actuarial Society 89: 68–113.

Venter, G. 2004. "Testing Distributions of Stochastically Generated Yield Curves." ASTIN Bulletin 34(1): 229–247.

Venter, G., Barnett, J., Kreps, R., and Major, J. 2007. "Multivariate Copulas for Financial Modeling." Variance 1(1): 103–119.