CHAPTER 1

Interest Rate Models

Oren Cheyette, Ph.D.
Vice President
Fixed Income Research
BARRA, Inc.

An interest rate model is a probabilistic description of the future evolution of interest rates. Based on today's information, future interest rates are uncertain: An interest rate model is a characterization of that uncertainty. Quantitative analysis of securities with rate dependent cash flows requires application of such a model in order to find the present value of the uncertainty. Since virtually all financial instruments other than default- and option-free bonds have interest rate sensitive cash flows, this matters to most fixed-income portfolio managers and actuaries, as well as to traders and users of interest rate derivatives.

For financial instrument valuation and risk estimation one wants to use only models that are arbitrage free and matched to the currently observed term structure of interest rates. "Arbitrage free" means just that if one values the same cash flows in two different ways, one should get the same result. For example, a 10-year bond putable at par by the holder in 5 years can also be viewed as a 5-year bond with an option of the holder to extend the maturity for another 5 years. An arbitrage-free model will produce the same value for the structure viewed either way. This is also known as the *law of one price*. The term structure matching condition means that when a default-free straight bond is valued according to the model, the result should be the same as if the bond's cash flows are simply discounted according to the current default-free term structure. A model that fails to satisfy either of these conditions cannot be trusted for general problems, though it may be usable in some limited context.

For equity derivatives, lognormality of prices (leading to the Black-Scholes formula for calls and puts) is the standard starting point for option calculations. In the fixed-income market, unfortunately, there is no equally natural and simple assumption. Wall Street dealers routinely use a multiplicity of models based on widely varying assumptions in different markets. For example, an options desk most likely uses a version of the Black formula to value interest rate caps and floors, implying an approximately lognormal distribution of interest rates. A few feet away, the mortgage desk may use a normal interest rate model to evaluate their passthrough and CMO durations. And on the next floor, actuaries may use variants of both types of models to analyze their annuities and insurance policies.

It may seem that one's major concern in choosing an interest rate model should be the accuracy with which it represents the empirical volatility of the term structure of rates, and its ability to fit market prices of vanilla derivatives such as at-the-money caps and swaptions. These are clearly important criteria, but they are not decisive. The first criterion is hard to pin down, depending strongly on what historical period one chooses to examine. The second criterion is easy to satisfy for most commonly used models, by the simple (though unappealing) expedient of permitting predicted future volatility to be time dependent. So, while important, this concern doesn't really do much to narrow the choices.

A critical issue in selecting an interest rate model is, instead, ease of application. For some models it is difficult or impossible to provide efficient valuation algorithms for all financial instruments of interest to a typical investor. Given that one would like to analyze all financial instruments using the same underlying assumptions, this is a significant problem. At the same time, one would prefer not to stray too far from economic reasonableness—such as by using the Black-Scholes formula to value callable bonds. These considerations lead to a fairly narrow menu of choices among the known interest rate models.

The organization of this chapter is as follows. In the next section I provide a (brief) discussion of the principles of valuation algorithms. This will give a context for many of the points made in the third section, which provides an overview of the various characteristics that differentiate interest rate models. Finally, in the fourth section I describe the empirical evidence on interest rate dynamics and provide a quantitative comparison of a family of models that closely match those in common use. I have tried to emphasize those issues that are primarily of interest for application of the models in practical settings. There is little point in having the theoretically ideal model if it can't actually be implemented as part of a valuation algorithm.

VALUATION

Valuation algorithms for rate dependent contingent claims are usually based on a risk neutral formula, which states that the present value of an uncertain cash flow at time T is given by the average over all interest rate scenarios of the scenario cash flow divided by the scenario value at time T of a money market investment of $1 today.[1] More formally, the value of a security is given by the expectation (average) over interest rate scenarios

$$P = E\left[\sum_i \frac{C_i}{M_i}\right] \tag{1}$$

where C_i is the security's cash flows and M_i is the money market account value at time t_i in each scenario, calculated by assuming continual reinvestment at the prevailing short rate.

The probability weights used in the average are chosen so that the expected rate of return on any security over the next instant is the same, namely the short rate. These are the so-called "risk neutral" probability weights: They would be the true weights if investors were indifferent to bearing interest rate risk. In that case, investors would demand no excess return relative to a (riskless) money market account in order to hold risky positions—hence equation (1).

It is important to emphasize that the valuation formula is not dependent on any *assumption* of risk neutrality. Financial instruments are valued by equation (1) *as if* the market were indifferent to interest rate risk *and* the correct discount factor for a future cash flow were the inverse of the money market return. Both statements are false for the real world, but the errors are offsetting: A valuation formula based on probabilities implying a nonzero market price of interest rate risk and the corresponding scenario discount factors would give the same value.

There are two approaches to computing the average in equation (1): by direct brute force evaluation, or indirectly by solving a related differential equation. The brute force method is usually called the Monte Carlo method. It consists of generating a large number of possible interest rate scenarios based on the interest rate model, computing the cash flows and money market values in each one, and averaging. Properly speaking, only path generation based on random numbers is a Monte Carlo method. There are other scenario methods—e.g., complete sampling of a tree—that do not depend on the use of random numbers.

[1] The money market account is the *numeraire*.

Given sufficient computer resources, the scenario method can tackle essentially any type of financial instrument.[2]

A variety of schemes are known for choosing scenario sample paths efficiently, but none of them are even remotely as fast and accurate as the second technique. In certain cases (discussed in more detail in the next section) the average in equation (1) obeys a partial differential equation—like the one derived by Black and Scholes for equity options—for which there exist fast and accurate numerical solution methods, or in special cases even analytical solutions. This happens only for interest rate models of a particular type, and then only for certain security types, such as caps, floors, swaptions, and options on bonds. For securities such as mortgage passthroughs, CMOs, index amortizing swaps, and for some insurance policies and annuities, simulation methods are the only alternative.

MODEL TAXONOMY

The last two decades have seen the development of a tremendous profusion of models for valuation of interest rate sensitive financial instruments. In order to better understand these models, it is helpful to recognize a number of features that characterize and distinguish them. These are features of particular relevance to practitioners wishing to implement valuation algorithms, as they render some models completely unsuitable for certain types of financial instruments.[3] The following subsections enumerate some of the major dimensions of variation among the different models.

One- versus Multi-Factor

In many cases, the value of an interest rate contingent claim depends, effectively, on the prices of many underlying assets. For example, while the payoff of a caplet depends only on the reset date value of a zero coupon bond maturing at the payment date (valued based on, say, 3-month LIBOR), the payoff to an option on a coupon bond depends on the exercise date values of all of the bond's remaining interest and principal payments. Valuation of such an option is in principle an inherently multidimensional problem.

Fortunately, in practice these values are highly correlated. The degree of correlation can be quantified by examining the covariance matrix of

[2] This is true even for American options. For a review see P. Boyle, M. Broadie, and P. Glasserman, "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control* (1997), pp. 1267–1322.

[3] There is, unfortunately, a version of Murphy's law applicable to interest rate models, which states that the computational tractability of a model is inversely proportional to its economic realism.

changes in spot rates of different maturities. A principal component analysis of the covariance matrix decomposes the motion of the spot curve into independent (uncorrelated) components. The largest principal component describes a common shift of all interest rates in the same direction. The next leading components are a twist, with short rates moving one way and long rates the other, and a "butterfly" motion, with short and long rates moving one way, and intermediate rates the other. Based on analysis of weekly data from the Federal Reserve H15 series of benchmark Treasury yields from 1983 through 1995, the shift component accounts for 84% of the total variance of spot rates, while twist and butterfly account for 11% and 4%, leaving about 1% for all remaining principal components.

The shift factor alone explains a large fraction of the overall movement of spot rates. As a result, valuation can be reduced to a one factor problem in many instances with little loss of accuracy. Only securities whose payoffs are primarily sensitive to the shape of the spot curve rather than its overall level (such as dual index floaters, which depend on the difference between a long and a short rate) will not be modeled well with this approach.

In principle it is straightforward to move from a one-factor model to a multi-factor one. In practice, though, implementations of multi-factor valuation models can be complicated and slow, and require estimation of many more volatility and correlation parameters than are needed for one-factor models, so there may be some benefit to using a one-factor model when possible. The remainder of this chapter will focus on one-factor models.[4]

Exogenous versus Endogenous Term Structure

The first interest rate models were not constructed so as to fit an arbitrary initial term structure. Instead, with a view towards analytical simplicity, the Vasicek[5] and Cox-Ingersoll-Ross[6] (CIR) models contain a few constant parameters that define an endogenously specified term structure. That is, the initial spot curve is given by an analytical formula in terms of the model parameters. These are sometimes also called "equilibrium" models, as they posit yield curves derived from an assumption of

[4] For an exposition of two-factor models, see D.F. Babbel and C.B. Merrill, *Valuation of Interest Sensitive Financial Instruments* (New Hope, PA: Frank J. Fabozzi Associates and Society of Actuaries, 1996).

[5] O. Vasicek, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* (November 1977).

[6] J.C. Cox, J.E. Ingersoll Jr., and S.A. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica* (March 1985).

economic equilibrium based on a given market price of risk and other parameters governing collective expectations.

For *dynamically* reasonable choices of the parameters—values that give plausible long-run interest rate distributions and option prices—the term structures achievable in these models have far too little curvature to accurately represent typical empirical spot rate curves. This is because the mean reversion parameter, governing the rate at which the short rate reverts towards the long-run mean, also governs the volatility of long-term rates relative to the volatility of the short rate—the "term structure of volatility." To achieve the observed level of long-rate volatility (or to price options on long-term securities well) requires that there be relatively little mean reversion, but this implies low curvature yield curves. This problem can be partially solved by moving to a multi-factor framework—but at a significant cost as discussed earlier. These models are therefore not particularly useful as the basis for valuation algorithms—they simply have too few degrees of freedom to faithfully represent real markets.

To be used for valuation, a model must be calibrated to the initial spot rate curve. That is, the model structure must accommodate an exogenously determined spot rate curve, typically given by fitting to bond prices, or sometimes to futures prices and swap rates. All models in common use are of this type.

There is a "trick" invented by Dybvig that converts an endogenous model to a calibrated exogenous one.[7] The trick can be viewed as splitting the nominal interest rate into two parts: the stochastic part modeled endogenously, and a non-stochastic drift term, which compensates for the mismatch of the endogenous term structure and the observed one. (BARRA has used this technique to calibrate the CIR model in its older fixed-income analytics.) The price of this method is that the volatility function is no longer a simple function of the nominal interest rate.

Short Rate versus Yield Curve

The risk neutral valuation formula requires that one know the sequence of short rates for each scenario, so an interest rate model must provide this information. For this reason, many interest rate models are simply models of the stochastic evolution of the short rate. A second reason for the desirability of such models is that they have the *Markov property*, meaning that the evolution of the short rate at each instant depends only on its current value—not on how it got there. The practical significance of this is that, as alluded to in the previous section, the valuation prob-

[7] P. Dybvig, "Bond and Bond Option Pricing Based on the Current Term Structure," in M. A. H. Dempster and S. Pliska (eds.), *Mathematics of Derivative Securities* (Cambridge, U.K.: Cambridge University Press, 1997).

lem for many types of financial instruments can be reduced to solving a partial differential equation, for which there exist efficient analytical and numerical techniques. To be amenable to this calculation technique, a financial instrument's cash flow at time t must depend only on the state of affairs at that time, not on how the evolution occurred prior to t, or it must be equivalent to a portfolio of such securities (for example, a callable bond is a position long a straight bond and short a call option).

Short-rate models have two parts. One specifies the average rate of change ("drift") of the short rate at each instant; the other specifies the instantaneous volatility of the short rate. The conventional notation for this is

$$dr(t) = \mu(r, t)dt + \sigma(r, t)dz(t) \tag{2}$$

The left-hand side of this equation is the change in the short rate over the next instant. The first term on the right is the drift multiplied by the size of the time step. The second is the volatility multiplied by a normally distributed random increment. For most models, the drift component must be determined through a numerical technique to match the initial spot rate curve, while for a small number of models there exists an analytical relationship. In general, there exists a no-arbitrage relationship linking the initial forward rate curve, the volatility $\sigma(r,t)$, the market price of interest rate risk, and the drift term $\mu(r,t)$. However, since typically one must solve for the drift numerically, this relationship plays no role in model construction. Differences between models arise from different dependences of the drift and volatility terms on the short rate.

For financial instruments whose cash flows don't depend on the interest rate history, the expectation formula (1) for present value obeys the Feynman-Kac equation

$$\frac{1}{2}\sigma^2 P_{rr} + (\mu - \lambda)P_r + P_t - rP + c = 0 \tag{3}$$

where, for example, P_r denotes the partial derivative of P with respect to r, c is the payment rate of the financial instrument, and λ, which can be time and rate dependent, is the market price of interest rate risk.

The terms in this equation can be understood as follows. In the absence of uncertainty ($\sigma = 0$), the equation involves four terms. The last three assert that the value of the security increases at the risk-free rate (rP), and decreases by the amount of any payments (c). The term $(\mu - \lambda)P_r$ accounts for change in value due to the change in the term structure with time, as rates move up the forward curve. In the absence of uncertainty it is easy to express ($\mu - \lambda$) in terms of the initial forward rates. In the presence of uncertainty this term depends on the volatility as well, and we also have the first term, which is the main source of the complexity of valuation models.

The Vasicek and CIR models are models of the short rate. Both have the same form for the drift term, namely a tendency for the short rate to rise when it is below the long-term mean, and fall when it is above. That is, the short-rate drift has the form $\mu = \kappa(\theta - r)$, where r is the short rate and κ and θ are the mean reversion and long-term rate constants. The two models differ in the rate dependence of the volatility: it is constant (when expressed as points per year) in the Vasicek model, and proportional to the square root of the short rate in the CIR model.

The Dybvig-adjusted Vasicek model is the mean reverting generalization of the Ho-Lee model,[8] also known as the mean reverting Gaussian (MRG) model or the Hull-White model.[9] The MRG model has particularly simple analytical expressions for values of many assets—in particular, bonds and European options on bonds. Like the original Vasicek model, it permits the occurrence of negative interest rates with positive probability. However, for typical initial spot curves and volatility parameters, the probability of negative rates is quite small.

Other popular models of this type are the Black-Derman-Toy (BDT)[10] and Black-Karasinski[11] (BK) models, in which the volatility is proportional to the short rate, so that the ratio of volatility to rate level is constant. For these models, unlike the MRG and Dybvig-adjusted CIR models, the drift term is not simple. These models require numerical fitting to the initial interest rate and volatility term structures. The drift term is therefore not known analytically. In the BDT model, the short-rate volatility is also linked to the mean reversion strength (which is also generally time dependent) in such a way that—in the usual situation where long rates are less volatile than the short rate—the short-rate volatility decreases in the future. This feature is undesirable: One doesn't want to link the observation that the long end of the curve has relatively low volatility to a forecast that in the future the short rate will

---

[8] T.S.Y. Ho and S.B. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* (December 1986); and, J. Hull and A. White, "Pricing Interest Rate Derivative Securities," *The Review of Financial Studies*, 3:4 (1990).

[9] This model was also derived in F. Jamshidian, "The One-Factor Gaussian Interest Rate Model: Theory and Implementation," Merrill Lynch working paper, 1988.

[10] F. Black, E. Derman and W. Toy, "A One Factor Model of Interest Rates and its Application to Treasury Bond Options," *Financial Analysts Journal* (January/February 1990).

[11] F. Black and P. Karasinski, "Bond and Option Prices when Short Rates are Lognormal," *Financial Analysts Journal* (July/August 1992).

become less volatile. This problem motivated the development of the BK model in which mean reversion and volatility are delinked.

All of these models are explicit models of the short rate alone. It happens that in the Vasicek and CIR models (with or without the Dybvig adjustment) it is possible to express the entire forward curve as a function of the current short rate through fairly simple analytical formulas. This is not possible in the BDT and BK models, or generally in other models of short-rate dynamics, other than by highly inefficient numerical techniques. Indeed, it is possible to show that the only short-rate models consistent with an arbitrary initial term structure for which one can find the whole forward curve analytically are in a class that includes the MRG and Dybvig-adjusted CIR models as special cases, namely where the short-rate volatility has the form[12]

$$\sigma(r, t) = \sqrt{\sigma_1(t) + \sigma_2(t)r}.$$

While valuation of certain assets (e.g., callable bonds) does not require knowledge of longer rates, there are broad asset classes that do. For example, mortgage prepayment models are typically driven off a long-term Treasury par yield, such as the 10-year rate. Therefore a generic short-rate model such as BDT or BK is unsuitable if one seeks to analyze a variety of assets in a common interest rate framework.

An alternative approach to interest rate modeling is to specify the dynamics of the entire term structure. The volatility of the term structure is then given by some specified function, which most generally could be a function of time, maturity, and spot rates. A special case of this approach (in a discrete time framework) is the Ho-Lee model mentioned earlier, for which the term structure of volatility is a parallel shift of the spot rate curve, whose magnitude is independent of time and the level of rates. A completely general continuous time, multi-factor framework for constructing such models was given by Heath, Jarrow, and Morton (HJM).[13]

It is sometimes said that all interest rate models are HJM models. This is technically true: In principle, every arbitrage-free model of the term structure can be described in their framework. In practice, however, it is impossible to do this analytically for most short-rate Markov models. The only ones for which it is possible are those in the MRG-CIR family described

[12] A. Jeffrey, "Single Factor Heath-Jarrow-Morton Term Structure Models Based on Markov Spot Interest Rate Dynamics," *Journal of Financial and Quantitative Analysis*, 30:4 (December 1995).

[13] D. Heath, R. Jarrow, and A. Morton, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60:1 (January 1992).

earlier. The BDT and BK models, for instance, cannot be translated to the HJM framework other than by impracticable numerical means. To put a model in HJM form, one must know the term structure of volatility at all times, and this is generally not possible for short-rate Markov models.

If feasible, the HJM approach is clearly very attractive, since one knows now not just the short rate but also all longer rates as well. In addition, HJM models are very "natural," in the sense that the basic inputs to the model are the initial term structure of interest rates and a term structure of interest rate volatility for each independent motion of the yield curve.

The reason for the qualification in the last paragraph is that a generic HJM model requires keeping track of a potentially enormous amount of information. The HJM framework imposes no structure other than the requirement of no-arbitrage on the dynamics of the term structure. Each forward rate of fixed maturity evolves separately, so that one must keep track of each one separately. Since there are an infinite number of distinct forward rates, this can be difficult. This difficulty occurs even in a one factor HJM model, for which there is only one source of random movement of the term structure. A general HJM model does not have the Markov property that leads to valuation formulas expressed as solutions to partial differential equations. This makes it impossible to accurately value interest rate options without using huge amounts of computer time, since one is forced to use simulation methods.

In practice, a simulation algorithm breaks the evolution of the term structure up into discrete time steps, so one need keep track of and simulate only forward rates for the finite set of simulation times. Still, this can be a large number (e.g., 360 or more for a mortgage passthrough), and this computational burden, combined with the inefficiency of simulation methods, has prevented general HJM models from coming into more widespread use.

Some applications require simulation methods because the assets' structures (e.g., mortgage-backed securities) are not compatible with differential equation methods. For applications where one is solely interested in modeling such assets, there exists a class of HJM models that significantly simplify the forward rate calculations.[14] The simplest version of such models, the "two state Markov model," permits an arbitrary dependence of short-rate volatility on both time and the level of interest rates, while the ratio of forward-rate volatility to short-rate volatility is solely a function of term. That is, the volatility of $f(t,T)$, the term T forward rate at time t takes the form

[14] O. Cheyette, "Term Structure Dynamics and Mortgage Valuation," *Journal of Fixed Income* (March 1992). The two state Markov model was also described in P. Ritchken and L. Sankarasubramanian, "Volatility Structure of Forward Rates and the Dynamics of the Term Structure," *Mathematical Finance*, 5(1) (1995), pp. 55–72.

$$\sigma_f(r, t, T) = \sigma(r, t)e^{-\int_t^T k(u)du} \tag{4}$$

where $\sigma(r,t) = \sigma_f(r,t,t)$ is the short-rate volatility and $k(t)$ determines the mean reversion rate or equivalently, the rate of decrease of forward rate volatility with term. The evolution of all forward rates in this model can be described in terms of two state variables: the short rate (or any other forward or spot rate), and the slope of the forward curve at the origin. The second variable can be expressed in terms of the total variance experienced by a forward rate of fixed maturity by the time it has become the short rate. The stochastic evolution equations for the two state variables can be written as

$$\begin{aligned} d\tilde{r}(t) &= (V(t) - k(t)\tilde{r})dt + \sigma(r, t)dz(t) \\ \frac{\partial V}{\partial t} &= \sigma^2(r, t) - 2k(t)V(t) \end{aligned} \tag{5}$$

where $\tilde{r}(t) \equiv r(t) - f(0, t)$ is the deviation of the short rate from the initial forward rate curve. The state variable $V(t)$ has initial value $V(0)=0$; its evolution equation is non-stochastic and can be integrated to give

$$V(t) = \int_0^t \sigma_f^2(r, s, t)ds = \int_0^t \sigma^2(r, s)e^{-2\int_s^t k(u)du} ds \tag{6}$$

In terms of these state variables, the forward curve is given by

$$f(t, T) = f(0, T) + \phi(t, T)\left(\tilde{r} + V(t)\int_t^T \phi(t, s)ds\right) \tag{7}$$

where

$$\phi(t, T) = \sigma_f(r, t, T)/\sigma_f(r, t, t) = e^{-\int_t^T k(s)ds}$$

is a deterministic function.

Instead of having to keep track of hundreds of forward rates, one need only model the evolution of the two state variables. Path independent asset prices also obey a partial differential equation in this model, so it appears possible, at least in principle, to use more efficient numerical methods. The equation, analogous to equation (3), is

$$\frac{1}{2}\sigma^2 P_{\tilde{r}\tilde{r}} + (V - k\tilde{r})P_{\tilde{r}} + (\sigma^2 - 2kV)P_V + P_t - rP + c = 0. \tag{8}$$

Unlike equation (3), for which one must use the equation itself applied to bonds to solve for the coefficient μ–λ, here the coefficient functions are all known in terms of the initial data: the short-rate volatility and the initial forward curve. This simplification has come at the price of adding a dimension, as we now have to contend also with a term involving the first derivative with respect to *V*, and so the equation is much more difficult to solve efficiently by standard techniques.

In the special case where $\sigma(r,t)$ is independent of *r*, this model is the MRG model mentioned earlier. In this case, *V* is a deterministic function of *t*, so the P_V term disappears from equation (8), leaving a two-dimensional equation that has analytical solutions for European options on bonds, and straightforward numerical techniques for valuing American bond options. Since bond prices are lognormally distributed in this model, it should be no surprise that the formula for options on pure discount bounds (PDB's) looks much like the Black-Scholes formula. The value of a call with strike price *K*, exercise date *t* on a PDB maturing at time *T* is given by

$$C = P(T)N(h_1) - KP(t)N(h_2), \tag{9}$$

where

$$h_1 = \frac{k}{(1 - e^{-k(T-t)})\sqrt{V(t)}} \ln\frac{P(T)}{KP(t)} + \frac{\sqrt{V(t)}(1 - e^{-k(T-t)})}{2k},$$

$$h_2 = h_1 - \frac{\sqrt{V(t)}(1 - e^{-k(T-t)})}{k},$$

$N(x)$ is the Gaussian distribution, and $P(t)$ and $P(T)$ are prices of PDB's maturing at *t* and *T*. (The put value can be obtained by put-call parity.) Options on coupon bonds can be valued by adding up a portfolio of options on PDBs, one for each coupon or principal payment after the exercise date, with strike prices such that they are all at-the-money at the same value of the short rate. The Dybvig-adjusted CIR model has similar formulas for bond options, involving the non-central χ^2 distribution instead of the Gaussian one.

If $\sigma(r,t)$ depends on r, the model becomes similar to some other standard models. For example, $\sigma(r,t)=a\sqrt{r}$ has the same rate dependence as the CIR model, while choosing $\sigma(r,t)=br$ gives a model similar to BK, though in each case the drift and term structure of volatility are different.

Unless one has some short- or long-term view on trends in short-rate volatility, it is most natural to choose $\sigma(r,t)$ to be time independent, and similarly $k(u)$ to be constant. This is equivalent to saying that the shape of the volatility term structure—though not necessarily its magnitude—should be constant over time. (Otherwise, as in the BDT model, one is imposing an undesirable linkage between today's shape of the forward rate volatility curve and future volatility curves.) In that case, the term structure of forward-rate volatility is exponentially decreasing with maturity, and the integrals in equations (6) and (7) can be computed, giving for the forward curve

$$f(t,T) = f(0,T) + e^{-k(T-t)}\left(\tilde{r} + V(t)\,\frac{1-e^{-k(T-t)}}{k}\right). \tag{10}$$

Finally, if the volatility is assumed rate independent as well, the integral expression for $V(t)$ can be evaluated to give

$$V(t) = \sigma^2\,\frac{1-e^{-2kt}}{2k}\,, \tag{11}$$

and we obtain the forward curves of the MRG model.

Empirically, neither the historical volatility nor the implied volatility falls off so neatly. Instead, volatility typically increases with term out to between 1 and 3 years, then drops off. The two state Markov model cannot accommodate this behavior, except by imposing a forecast of increasing then decreasing short-rate volatility, or a short run of negative mean reversion. There is, however, an extension of the model that permits modeling of humped or other more complicated volatility curves, at the cost of introducing additional state variables.[15] With five state variables, for example, it is possible to model the dominant volatility term structure of the U.S. Treasury spot curve very accurately.

[15] O. Cheyette, "Markov Representation of the HJM Model," working paper, 1995.

EMPIRICAL AND NUMERICAL CONSIDERATIONS

Given the profusion of models, it is reasonable to ask whether there are empirical or other considerations that can help motivate a choice of one model for applications. One might take the view that one should use whichever model is most convenient for the particular problem at hand—e.g., BDT or BK for bonds with embedded options, Black model for caps and floors, a two-state Markov model for mortgages, and a ten-state, two-factor Markov-HJM model for dual index amortizing floaters. The obvious problem with this approach is that it can't be used to find hedging relationships or relative value between financial instruments valued according to the different models. I take as a given, then, that we seek models that can be used effectively for valuation of most types of financial instruments with minimum compromise of financial reasonableness. The choice will likely depend on how many and what kinds of assets one needs to value. A trader of vanilla options may be less concerned about cross-market consistency issues than a manager of portfolios of callable bonds and mortgage-backed securities.

The major empirical consideration—and one that has produced a large amount of inconclusive research—is the assumed dependence of volatility on the level of interest rates. Different researchers have reported various evidence that volatility is best explained (1) as a power of the short rate[16] ($\sigma \propto r^{\gamma}$)—with γ so large that models with this volatility have rates running off to infinity with high probability ("explosions"), (2) by a GARCH model with very long (possibly infinite) persistence,[17] (3) by some combination of GARCH with a power law dependence on rates,[18] (4) by none of the above.[19] All of this work has been in the context of short-rate Markov models.

Here I will present some fairly straightforward evidence in favor of choice (4) based on analysis of movements of the whole term structure of spot rates, rather than just short rates, from U.S. Treasury yields over the period 1977 to early 1996.

The result is that the market appears to be well described by "eras" with very different rate dependences of volatility, possibly coinciding with periods of different Federal Reserve policies. Since all the models in

[16] K.C. Chan, G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, "An Empirical Comparson of Alternative Models of the Short Rate," *Journal of Finance* 47:3 (1992).
[17] See R.J. Brenner, R.H. Harjes, and K.F. Kroner, "Another Look at Alternative Models of the Short-Term Interest Rate," University of Arizona working paper (1993), and references therein.
[18] Ibid.
[19] Y.Aït-Sahalia, "Testing Continuous Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9:2 (1996).

common use have a power law dependence of volatility on rates, I attempted to determine the best fit to the exponent (γ) relating the two. My purpose here is not so much to provide another entrant in this already crowded field, but rather to suggest that there may be no simple answer to the empirical question. No model with constant parameters seems to do a very good job. A surprising result, given the degree to which the market for interest rate derivatives has exploded and the widespread use of lognormal models, is that the period since 1987 is best modeled by a nearly *normal* model of interest rate volatility.

The data used in the analysis consisted of spot rate curves derived from the Federal Reserve H15 series of weekly average benchmark yields. The benchmark yields are given as semiannually compounded yields of hypothetical par bonds with fixed maturities ranging from 3 months to 30 years, derived by interpolation from actively traded issues. The data cover the period from early 1977, when a 30-year bond was first issued, through March of 1996. The spot curves are represented as continuous, piecewise linear functions, constructed by a root finding procedure to exactly match the given yields, assumed to be yields of par bonds. (This is similar to the conventional bootstrapping method.) The two data points surrounding the 1987 crash were excluded: The short and intermediate markets moved by around ten standard deviations during the crash, and this extreme event would have had a significant skewing effect on the analysis.

A parsimonious representation of the spot curve dynamics is given by the two-state Markov model with constant mean reversion k and volatility that is time independent and proportional to a power of the short rate: $\sigma = \beta r^{\gamma}$. In this case, the term structure of spot rate volatility, given by integrating equation (4), is

$$\sigma(r_t)v(T) = \beta r_t^{\gamma} \frac{1 - e^{-kT}}{kT} \tag{12}$$

where T is the maturity and r_t is the time t short rate. The time t weekly change in the spot rate curve is then given by the change due to the passage of time ("rolling up the forward curve") plus a random change of the form $v(T)x_t$, where for each t, x_t, is an independent normal random variable with distribution $N(\mu, \sigma(r_t)\sqrt{52})$. (The systematic drift μ of x_t, over time was assumed to be independent of time and the rate level.) The parameters β, γ, and k are estimated as follows. First, using an initial guess for γ, k is estimated by a maximum likelihood fit of the maturity dependence of $v(T)$ to the spot curve changes. Then, using this value of k, another maximum likelihood fit is applied to fit the variance of x_t to the power law model of $\sigma(r_t)$.

The procedure is then iterated to improve the estimates of k and γ (although it turns out that the best fit of k is quite insensitive to the value of γ, and vice versa).

One advantage of looking at the entire term structure is that we avoid modeling just idiosyncratic behavior of the short end, e.g., that it is largely determined by the Federal Reserve. An additional feature of this analysis is proper accounting for the effect of the "arbitrage-free drift"—namely, the systematic change of interest rates due purely to the shape of the forward curve at the start of each period. Prior analyses have typically involved fitting to endogenous short-rate models with constant parameters not calibrated to each period's term structure. The present approach mitigates a fundamental problem of prior research in the context of one-factor models, namely that interest rate dynamics are poorly described by a single factor. By reinitializing the drift parameters at the start of each sample period and studying the volatility of changes to a well-defined term structure factor, the effects of additional factors are excluded from the analysis.

The results for the different time periods are shown in Exhibit 1.1. (The exhibit doesn't include the best fit values of β, which are not relevant to the empirical issue at hand.) The error estimates reported in the exhibit are derived by a bootstrap Monte Carlo procedure that constructs artificial data sets by random sampling of the original set with replacement and applies the same analysis to them.[20] It is apparent that the different subperiods are well described by very different exponents and mean reversion. The different periods were chosen to include or exclude the monetarist policy "experiment" under Volcker of the late 1970s and early 1980s, and also to sample just the Greenspan era. For the period since 1987, the best fit exponent of 0.19 is significantly different from zero at the 95% confidence level, but not at the 99% level. However, the best fit value is well below the threshold of 0.5 required to guarantee positivity of interest rates, with 99% confidence. There appears to be weak sensitivity of volatility to the rate level, but much less than is implied by a number of models in widespread use—in particular, BDT, BK, and CIR.

The estimates for the mean reversion parameter k can be understood through the connection of mean reversion to the term structure of volatility. Large values of k imply large fluctuations in short rates compared to long rates, since longer rates reflect the expectation that changes in short rates will not persist forever. The early 1980s saw just such a phenomenon, with the yield curve becoming very steeply inverted for a brief period. Since then, the volatility of the short rate (in absolute terms of points per year) has been only slightly higher than that of long-term rates.

[20] B.J. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap* (New York: Chapman & Hall, 1993).

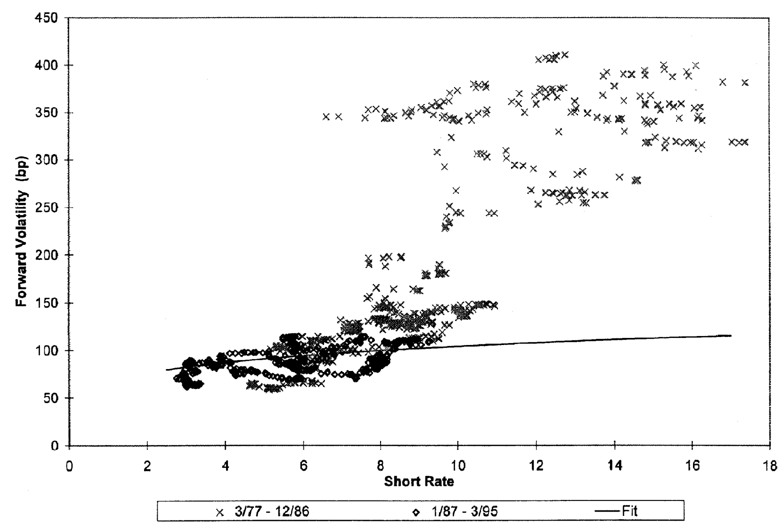
1-Cheyette Page 19 Thursday, August 29, 2002 9:58 AM

EXHIBIT 1.1 Parameter Estimates for the Two-State Markov Model with Power Law Volatility over Various Sample Periods*

Sample Period	Exponent (γ)	Mean Reversion (k)	Comments
3/1/77–3/29/96	1.04 ± 0.07	0.054 ± 0.007	Full data set
3/1/77–1/1/87	1.6 ± 0.10	0.10 ± 0.020	Pre-Greenspan
3/1/77–1/1/83	1.72 ± 0.15	0.22 ± 0.040	"Monetarist" policy
1/1/83–3/29/96	0.45 ± 0.07	0.019 ± 0.005	Post high-rate period
1/1/87–3/29/96	0.19 ± 0.09	0.016 ± 0.004	Greenspan

* The uncertainties are one standard deviation estimates based on bootstrap Monte Carlo resampling.

EXHIBIT 1.2 52-Week Volatility of Term Structure Changes Plotted Against the 3-Month Spot Rate at the Start of the Period



The x's are periods starting 3/77 through 12/86. The diamonds are periods starting 1/87 through 3/95. The data points are based on the best fit k for the period 1/87–3/96, as described in the text. The solid curve shows the best fit to a power law model. The best fit parameters are β=91 bp, γ=0.19. (This is *not* a fit to the points shown here, which are provided solely to give a visual feel for the data.)

Exhibit 1.2 gives a graphical representation of the data. There is clear evidence that the simple power law model is not a good fit and that the data display regime shifts. The exhibit shows the volatility of the factor in equation (12) using the value of k appropriate to the period January 1987–March 1996 (the "Greenspan era"). The vertical coordinate of each dot represents the volatility of the factor over a 52-week period; the horizontal coordinate shows the 3-month spot rate (a proxy for the short rate) at the start of the 52-week period. (Note that the maximum likelihood estimation is not based on the data points shown, but on the individual weekly changes.) The dots are broken into two sets: The *x*'s are for start dates prior to January 1987, the diamonds for later dates. Divided in this way, the data suggest fairly strongly that volatility has been nearly independent of interest rates since 1987—a time during which the short rate has ranged from around 3% to over 9%.

From an empirical perspective, then, no simple choice of model works well. Among the simple models of volatility, the MRG model most closely matches the recent behavior of U.S. Treasury term structure.

There is an issue of financial plausibility here, as well as an empirical one. Some models permit interest rates to become negative, which is undesirable, though how big a problem this is isn't obvious. The class of simple models that provably have positive interest rates without suffering from explosions and match the initial term structure is quite small. The BDT and BK models satisfy these conditions, but don't provide information about future yield curves as needed for the mortgage problem. The Dybvig-adjusted CIR model also satisfies the conditions, but is somewhat hard to work with. There is a lognormal HJM model that avoids negative rates, but it is analytically intractable and suffers from explosions.[21] The lognormal version of the two-state Markov model also suffers from explosions, though, as with the lognormal HJM model, these can be eliminated by capping the volatility at some large value.

It is therefore worth asking whether the empirical question is important. It might turn out to be unimportant in the sense that, properly compared, models that differ only in their assumed dependence of volatility on rates actually give similar answers for option values.

The trick in comparing models is to be sure that the comparisons are truly "apples to apples," by matching term structures of volatility. It is easy to imagine getting different results valuing the same option using the MRG, CIR, and BK models, even though the initial volatilities are set equal—not because of different assumptions about the dependence of volatility on rates, but because the long-term volatilities are different in the three models even when the short-rate volatilities are the same. There are a number of published papers claiming to demonstrate dramatic differences between models, but which actually demonstrate just that the models have been calibrated differently.[22]

[21] Heath, Jarrow, and Morton, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation."

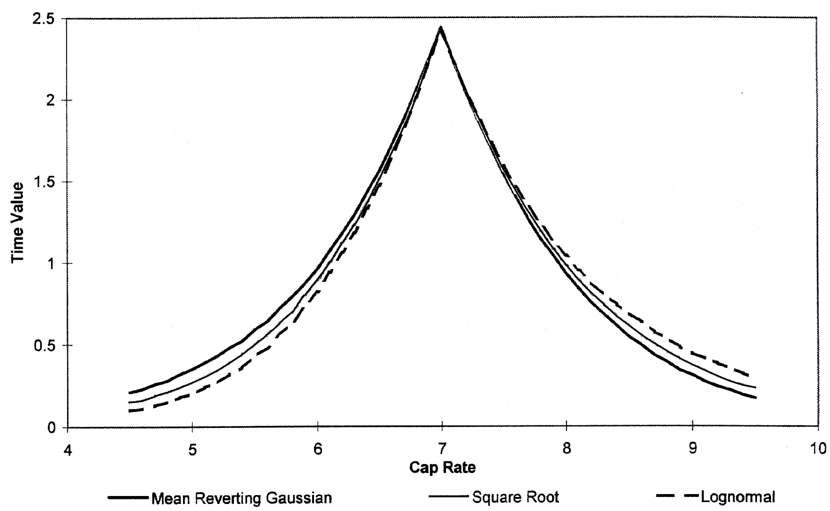
[22] For a recent example, see M. Uhrig and U. Walter, "A New Numerical Approach to Fitting the Initial Yield Curve," *Journal of Fixed Income* (March 1996).

The two-state Markov framework provides a convenient means to compare different choices for the dependence of volatility on rates while holding the initial term structure of volatility fixed. Choosing different forms for $\sigma(r)$ while setting k to a constant in expression (4) gives exactly this comparison. We can value options using these different assumptions and compare time values. (Intrinsic value—the value of the option when the volatility is zero—is of course the same in all models.) To be precise, we set $\sigma(r, t) = \sigma_0(r/r_0)\gamma$, where σ_0 is the initial annualized volatility of the short rate in absolute terms (e.g., 100 bp/year) and r_0 is the initial short rate. Choosing the exponent $\gamma = \{0, 0.5, 1\}$ then gives the MRG model, a square root volatility model (not CIR), and a lognormal model (not BK), respectively.

The results can be summarized by saying that a derivatives trader probably cares about the choice of exponent γ, but a fixed-income portfolio manager probably doesn't. The reason is that the differences in time value are small, except when the time value itself is small—for deep in- or out-of-the-money options. A derivatives trader may be required to price a deep out-of-the-money option, and would get very different results across models, having calibrated them using at-the-money options. A portfolio manager, on the other hand, has option positions embedded in bonds, mortgage-backed securities, etc., whose time value is a small fraction of total portfolio value. So differences that show up only for deep in- or out-of-the-money options are of little consequence. Moreover, a deep out-of-the-money option has small option delta, so small differences in valuation have little effect on measures of portfolio interest rate risk. An in-the-money option can be viewed as a position in the underlying asset plus an out-of-the-money option, so the same reasoning applies.

Exhibit 1.3 shows the results of one such comparison for a 5-year quarterly pay cap, with a flat initial term structure and modestly decreasing term structure of volatility. The time value for all three values of γ peaks at the same value for an at-the-money cap. Caps with higher strike rates have the largest time value in the lognormal model, because the volatility is increasing for rate moves in the direction that make them valuable. Understanding the behavior for lower strike caps requires using put-call parity: An in-the-money cap can be viewed as paying fixed in a rate swap and owning a floor. The swap has no time value, and the floor has only time value (since it is out-of-the money). The floor's time value is greatest for the MRG model, because it gives the largest volatility for rate moves in the direction that make it valuable. In each case, the square root model gives values intermediate between the MRG and lognormal models, for obvious reasons. At the extremes, 250 bp in or out of the money, time values differ by as much as a factor of 2 between the MRG and lognormal models. At these extremes, though, the time value is only a tenth of its value for the at-the-money cap.

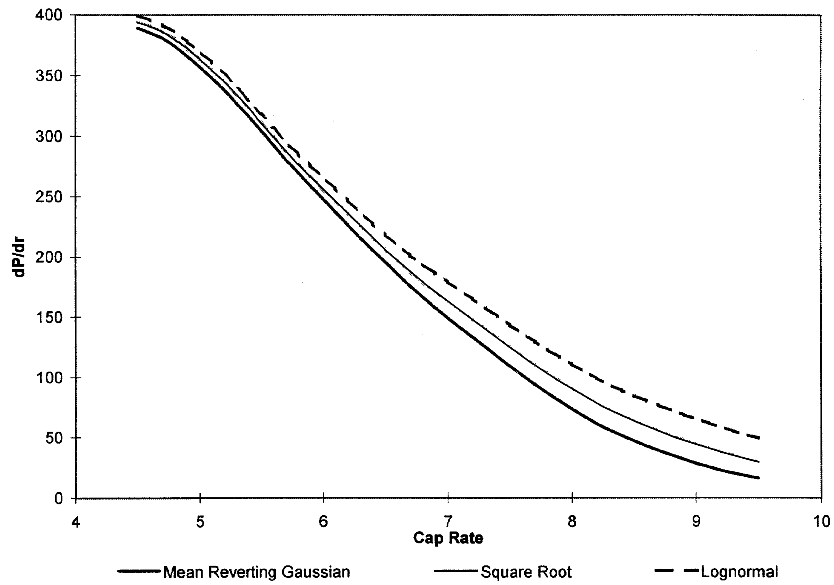
EXHIBIT 1.3 **Time Values for Five-Year Quarterly Pay Caps for Gaussian, Square Root, and Lognormal Two-State Markov Models with Identical Initial Term Structure of Volatility and a 7% Flat Initial Yield Curve***



* The model parameters (described in the text) are σ_0=100 bp/yr., k=0.02/yr., equivalent to an initial short-rate volatility of 14.8%, and a 10-year yield volatility of 13.6%.

If the initial term structure is not flat, the model differences can be larger. For example, if the term structure is positively sloped, then the model prices match up for an in-the-money rather than at-the-money cap. Using the same parameters as for Exhibit 1.3, but using the actual Treasury term structure as of 5/13/96 instead of a flat 7% curve, the time values differ at the peak by about 20%—about half a point—between the MRG and lognormal models. Interestingly, as shown in Exhibit 1.4, even though the time values can be rather different, the option deltas are rather close for the three models. (The deltas are even closer in the flat term structure case.) In this example, if a 9.5% cap were embedded in a floating-rate note priced around par, the effective duration attributable to the cap according to the lognormal model would be 0.49 year, while according to the MRG model it would be 0.17 year. The difference shrinks as the rate gets closer to the cap. This ⅓ year difference isn't trivial, but it's also not large compared to the effect of other modeling assumptions, such as the overall level of volatility or, if mortgages are involved, prepayment expectations.

EXHIBIT 1.4 Sensitivity of Cap Value to Change in Rate Level as a Function of Cap Rate*



* The cap structure and model parameters are the same as used for Exhibit 1.3, except that the initial term structure is the (positively sloped) U.S. Treasury curve as of 5/13/96. The short rate volatility is 19.9% and the ten-year yield volatility is 14.9%.

These are just two numerical examples, but it is easy to see how different variations would affect these results. An inverted term structure would make the MRG model time value largest at the peak and the lognormal model value the smallest. Holding σ_0 constant, higher initial interest rates would yield smaller valuation differences across models since there would be less variation of volatility around the mean. Larger values of the mean reversion k would also produce smaller differences between models, since the short-rate distribution would be tighter around the mean.

Finally, there is the question raised earlier as to whether one should be concerned about the possibility of negative interest rates in some models. From a practical standpoint, this is an issue only if it leads to a significant contribution to pricing from negative rates. One simple way to test this is to look at pricing of a call struck at par for a zero coupon bond. Exhibit 1.5 shows such a test for the MRG model. For reasonable parameter choices (here taken to be σ_0=100 bp/year, k = 0.02/year, or

20% volatility of a 5% short rate), the call values are quite modest, especially compared to those of a call on a par bond, which gives a feel for the time value of at-the-money options over the same period. The worst case is a call on the longest maturity zero-coupon bond which, with a flat 5% yield curve, is priced at 0.60. This is just 5% of the value of a par call on a 30-year par bond. Using the actual May 1996 yield curve, all the option values—other than on the 30-year zero—are negligible. For the 30-year zero the call is worth just 1% of the value of the call on a 30-year par bond. In October 1993, the U.S. Treasury market had the lowest short rate since 1963, and the lowest 10-year rate since 1967. Using that yield curve as a worst case, the zero coupon bond call values are only very slightly higher than the May 1996 values, and still effectively negligible for practical purposes.

Again, it is easy to see how these results change with different assumptions. An inverted curve makes negative rates likelier, so increases the value of a par call on a zero-coupon bond. (On the other hand, inverted curves at low interest rate levels are rare.) Conversely, a positive slope to the curve makes negative rates less likely, decreasing the call value. Holding σ_0 constant, lower interest rates produce larger call values. Increasing k produces smaller call values. The only circumstances that are really problematic for the MRG model are flat or inverted yield curves at very low rate levels, with relatively high volatility.

EXHIBIT 1.5 Valuation of a Continuous Par Call on Zero Coupon and Par Bonds of Various Maturities in the MRG Model

Model parameters are:

$$\sigma_0 = 100 \text{ bp/year}$$
$$k = 0.02\text{/year}$$

The value of the call on the zero coupon bond should be zero in every case, assuming non-negative interest rates.

	5% Flat Curve		7% Flat Curve		5/96 U.S. Tsy. Yields		10/93 U.S. Tsy. Yields	
Term	Zero Cpn.	Par Bond	Zero Cpn.	Par Bond	Zero Cpn.	Par Bond	Zero Cpn.	Par Bond
3-year	<0.01	0.96	<0.01	0.93	<0.01	0.65	<0.01	0.62
5-year	<0.01	1.93	<0.01	1.83	<0.01	1.43	<0.01	1.27
10-year	0.06	4.54	<0.01	4.07	<0.01	3.47	0.02	3.06
30-year	0.60	11.55	0.10	8.85	0.08	7.86	0.09	7.26

CONCLUSIONS

For portfolio analysis applications, the mean reverting Gaussian model has much to recommend it. For this model, it is easy to implement valuation algorithms for both path independent financial instruments such as bond options, and path dependent financial instruments such as CMOs and annuities. It is one of the simplest models in which it is possible to follow the evolution of the entire yield curve (à la HJM), making it especially useful for valuing assets like mortgage-backed securities whose cash flows depend on longer term rates. The oft raised bogeyman of negative interest rates proves to have little consequence for option pricing, since negative rates occur with very low probability for reasonable values of the model parameters and initial term structure.

Option values are somewhat (though not very) sensitive to the assumed dependence of volatility on the level of rates. The empirical evidence on this relationship is far from clear, with the data (at least in the United States) showing evidence of eras, possibly associated with central bank policy. The numerical evidence shows that, for a sloped term structure, different power law relationships give modestly different at-the-money option time values, and larger relative differences for deep in- or out-of-the-money options. These differences are unlikely to be significant to fixed-income portfolio managers, but are probably a concern for derivatives traders.