Lattice Models for Pricing American Interest Rate Claims

Anlong Li, Peter Ritchken, L. Sankarasubramanian

Journal of Finance, Volume 50, Issue 2 (Jun., 1995), 719-737.

THE JOURNAL OF FINANCE • VOL. L, NO. 2 • JUNE 1995

Lattice Models for Pricing American Interest Rate Claims

ANLONG LI, PETER RITCHKEN, and
L. SANKARASUBRAMANIAN*

ABSTRACT

This article establishes efficient lattice algorithms for pricing American interest-sensitive claims in the Heath, Jarrow, and Morton paradigm, under the assumption that the volatility structure of forward rates is restricted to a class that permits a Markovian representation of the term structure. The class of volatilities that permits this representation is quite large and imposes no severe restrictions on the structure for the spot rate volatility. The algorithm exploits the Markovian property of the term structure and permits the efficient computation of all types of interest rate claims. Specific examples are provided.

THIS ARTICLE DEVELOPS A FAMILY of models for pricing American options on interest rate derivatives. The models belong to the class of Heath, Jarrow, and Morton (HJM) (1992) models in that their only requirements are the initial yield curve and the volatility structure for all forward rates. Without severely restricting the structure of these volatilities, HJM showed that the evolution of the term structure could depend on the entire path taken by the term structure since it was initialized. Since the evolution of the term structure may not be Markovian with respect to a finite dimensioned state space, difficulties arise in implementing models for pricing interest rate derivatives. For example, if simple lattice procedures are used, then the paths may not reconnect and information may have to be manipulated over an exploding tree.[1] This article investigates ways to resolve these difficulties when additional assumptions are imposed on the permissible class of volatility structures.

Under very specific volatility restrictions, the path dependence can be completely removed. In this case the instantaneous spot rate can be shown to be the appropriate single state variable that contains all information relevant

*Li is at the First National Bank of Chicago. Ritchken is at the Weatherhead School of Management, Case Western Reserve University. Sankarasubramanian is at the Debt and Equity Markets Group, Merrill Lynch. We thank the participants of the Fourth Annual Derivative Securities Conference, held at Cornell University in 1994, for helpful comments. We also appreciate the helpful suggestions of an anonymous referee and the editor, René Stulz.

[1]The path dependence is made apparent in a related article by HJM (1990) where they use a bionomial approximation to provide an alternative derivation of their results. Due to the path dependence, the lattice grows exponentially with the number of time periods.

for pricing derivatives.[2] For this class of volatilities, closed form solutions are available for European contracts, and efficient computational schemes exist for American contracts.[3] Unfortunately, under the HJM paradigm, the total removal of the path dependence can only be achieved if the volatility structure for forward rates belongs to certain deterministic classes. All these classes imply that interest rates have normal distributions, a property that has been criticized, not only because interest rates can go negative with positive probability, but also because there is little empirical support for models of spot interest rates having variances that are independent of the level of rates.[4]

Recently, Ritchken and Sankarasubramanian (RS) (1995) have identified necessary and sufficient conditions on volatility structures that do not completely remove the path dependence, but rather capture it by a single sufficient statistic.[5] For volatility structures in their class, the evolution of the term structure can be made Markovian with respect to two state variables. That is, given these two variables, the entire term structure can be recovered. The class of volatilities that permit this representation is quite large, and includes, as a special case, the deterministic volatility structures that eliminate all forms of path dependence. Interestingly, the class imposes no particular restrictions on the structure of the spot rate volatility, and hence permits volatilities to fluctuate according to the level of the spot rate. To date, however, under the HJM approach, no analytical models of claims have been obtained for any nondeterministic structure of volatilities, including those in the RS class. Simulation procedures have been used to establish prices of European claims, and the analysis of American claims has largely been restricted to pricing short term claims, where implementing exploding lattice procedures may be computationally feasible.[6]

To our knowledge, there are no efficient HJM algorithms available that can accurately price most types of long-term American contracts, even those as simple as long term semiannual coupon callable bonds, when the volatility structure of forward rates is not deterministic. This article develops efficient

[2]Hull and White (1993c) and Caverhill (1994) provide necessary and sufficient conditions on the volatility structures that ensure the spot rate is the sole-state variable. The most popular models ion this class are referred to as generalized Vasicek models. Examples include Jamshidian (1989), Hull and White (1990), and HJM (1992).

[3]Examples of analytical solutions include Jamshidian (1993a), Amin and Jarrow (1991), HJM (1992), and Ritchken and Sankarasubramanian (1992, 1993).

[4]For example, see Chan, Karolyi, Longstaff and Sanders (1992) and Flesaker (1992).

[5]The sufficiency condition was also independently derived by Cheyette (1992), who applied the resulting Markovian model of the term structure to price mortgages.

[6]Heath *et al.* (1993) have used simple nonrecombining lattices. Unfortunately, the convergence behavior of these algorithms have not been investigated since the partition of the time period cannot be made very fine. Indeed, lattices with as few as 20 time periods would require searching millions of nodes. For time horizons less than 5 years, Amin and Bodurtha (1995) study discrete time path dependent lattices, and conclude 10 partitions may suffice for shorter term contracts. Amin and Morton (1994) use seven partitions in their empirical study on alternative volatility structures for short term contracts.

lattice procedures for pricing European and American claims for the class of volatility structures developed by RS. The resulting algorithms permit all types of claims to be efficiently computed, regardless of their maturity.

The article proceeds as follows. In Section I the basic results of HJM as well as RS are reviewed, with special attention placed on implementation issues. Section II then establishes a lattice approximation for implementing the RS model. Section III provides results that illustrate the convergence of the prices on the lattice. Section IV compares our HJM models to those developed under alternative paradigms. In particular we compare our models to the variants of the Black, Derman, and Toy (1990) models often used by professional trading houses. Section V summarizes the article.

I. Path Dependence and Volatility Structures

Let $f(t,T)$ be the forward rate at date t for instantaneous and riskless borrowing or lending at date T. The evolution of forward rates of every maturity T is given by a diffusion process of the form

$$df(t,T) = \mu_f(t,T)\,dt + \sigma_f(t,T)\,dw(t), \qquad f(0,T) \text{ given } T > t. \tag{1}$$

Here $\mu_f(t,T)$ and $\sigma_f(t,T)$ are the drift and volatility parameters that could depend on the level of the term structure itself. The instantaneous spot interest rate, $r(t)$ is given by $f(t,t)$. By definition, the price at date t of a pure discount bond with maturity date, T, is given by

$$P(t,T) = e^{-\int_t^T f(t,s)\,ds}. \tag{2}$$

The volatility function, $\sigma_f(t,T)$, holds the key to the pricing analysis, and can be chosen quite arbitrarily. In fact, its selection completely determines the price of all claims, since for each choice, the drift term is uniquely determined under the equivalent martingale measure by the no arbitrage condition:

$$\mu_f(t,T) = \sigma_f(t,T)\int_t^T \sigma_f(t,s)\,ds. \tag{3}$$

Let $g(0)$ represent the date 0 value of an European claim having a terminal payout at date s that is fully determined by yields drawn from the yield-curve at that time. HJM have shown that the fair price of the claim can be represented by

$$g(0) = \mathbf{E}_0[e^{-\int_0^s r(t)\,dt} g(s)]. \tag{4}$$

The expectation is computed under the process described by equation (1) with the drift term restriction imposed in equation (3).

In general, forward rates of all maturities cannot be represented as functions of a small number of state variables that have evolutions governed by Markovian processes. RS identify the class of volatility structures that permit the term structure to be represented by a two-state Markovian model. This class is characterized by

$$\sigma_f(t,T) = \sigma_f(t,t)k(t,T) \tag{5}$$

with

$$k(t,T) = e^{-\int_t^T \kappa(x)\,dx}.$$

Here $\sigma_f(t,t)$ is the volatility of the spot interest rate at date t, which could depend on the entire set of term structures to that date and $\kappa(x)$ is some exogeneously provided deterministic function.

If the volatility structure is of the form in equation (5), then bond prices at date t can be expressed in terms of price information at date 0, the spot rate at date t, $r(t)$, and a second statistic, $\phi(t)$, which represents the accumulated variance for the forward rate up to date t. Specifically,

$$P(t,T) = \left(\frac{P(0,T)}{P(0,t)}\right)e^{-\beta(t,T)(r(t)-f(0,t))-\frac{1}{2}\beta^2(t,T)\phi(t)} \tag{6a}$$

where

$$\beta(t,T) = \int_t^T k(t,u)\,du \tag{6b}$$

and

$$\begin{aligned}\phi(t) &= \int_0^t \sigma_f^2(u,t)\,du \\ &= \int_0^t \sigma_f^2(u,u)k^2(u,t)\,du.\end{aligned} \tag{6c}$$

Moreover, the price of an European interest rate claim, as given by equation (4), can be computed as

$$g(0) = \mathbf{E}_{r,\phi}[e^{-\int_0^s r(t)\,dt}g(s)] \tag{7}$$

where the expectation is taken under the risk-neutralized process

$$dr(t) = \mu(r,\phi,t)\,dt + \sigma_f(t,t)\,dw(t) \tag{8a}$$

$$d\phi(t) = \left(\sigma_f^2(t,t) - 2\kappa(t)\phi(t)\right)dt \tag{8b}$$

with

$$\mu(r,\phi,t) = \kappa(t)[f(0,t) - r(t)] + \phi(t) + \frac{d}{dt}f(0,t). \tag{8c}$$

The class of volatility structures described by equation (5) is quite large since no explicit restrictions are imposed, apart from boundedness, on the structure of spot rate volatilities, $\phi_f(t,t)$. In particular, the volatility could depend on both state variables, $r(t)$, and $\phi(t)$. That is,

$$\sigma_f(t,t) = \sigma(r(t),\phi(t),t). \tag{9}$$

As an example, consider the family of models generated by assuming that the volatility of the spot rate has constant elasticity,

$$\sigma_f(t,t) = \sigma[r(t)]^{\gamma}; \qquad \gamma \geq 0. \tag{10}$$

Interest rate models with this specification have been studied quite extensively in financial economics. For example, with $\gamma = 0$, we obtain the generalized Vasicek (1977) model, while with $\gamma = 0.5$, we obtain the square root structure considered by Cox, Ingersoll, and Ross (1985). More recently, Chan *et al.* (1992) test a wide variety of models resulting from this spot rate volatility specification and the additional requirement that the drift term of the spot rate process be linear in the level of the spot rate. They conclude that inelastic models with γ set to zero lead to term structure dynamics that appear inconsistent with observed data, and that models of spot rate behavior should permit the volatility to fluctuate according to the level of the rate. The family of RS models generated with the spot rate volatility assumption (9) share this property with the models tested by Chan *et al.* (1992).

In the RS family of interest rate models, the volatilities of all forward rates, $\sigma_f(t,T)$, are related to the volatility of the spot rate, $\sigma_f(t,t)$ through the exogenously specified parameter set, $\{\kappa(x)|x \geq 0\}$. This set can be made as parsimonious as desired. In particular, if we set $\kappa(z) = \kappa$, we obtain models that capture the notion that distant forward rates are less volatile than near forward rates. Cursory empirical evidence reported by Heath *et al.* (1993), however, indicates that the term structure of volatilities of forward rates may actually be humped, first increasing and then decreasing. To avoid the pricing errors that would result from an improperly specified initial volatility structure, many models, including those of Black, Derman, and Toy (1990), Hull and White (1990) and Black and Karasinski (1991) initialize not only the term structure but also the initial set of volatilities for all forward rates to a given set of values. In the RS class of models, this is achieved by appropriately choosing the set of $\kappa(\cdot)$s. For example, if $\kappa(x)$ were negative for small values of x and positive for large values, this would lead to the humped structure observed by Heath *et al.* (1993).

Of course the volatility restriction in equation (5) does preclude certain volatility structures that may have some intuitive appeal. For example, while the restriction does permit forward rate volatilities to fluctuate according to the level of any *single* spot rate, it does not permit volatilities of different forward rates to fluctuate according to *different* rates. Hence, a model in which volatilities vary according to the level of their forward rates would not be permissible.[7]

In the next section we describe a lattice procedure that can be used to efficiently implement the HJM paradigm for the RS volatility structures.

[7]Such volatility structures have been considered by Amin and Morton (1994).

II. Lattice Approximation of the Two-State Variable Process

The lattice approach follows a two step procedure. First, we transform the interest rate process in equation (8) to a form that has constant volatility. Second, we establish a path-reconnecting lattice approximation for the transformed process. To achieve the first step, we follow Nelson and Ramaswamy (1990) and consider the following transformation:

$$Y(t) = \int \frac{1}{\sigma[r(t), \phi(t), t]} dr(t) \tag{11}$$

where, as before, $\sigma[r(t), \phi(t), t]$ denotes the volatility of the spot rate at date t conditional on the values of $r(t)$ and $\phi(t)$. Further, let $r(t) = h(Y(t))$ be the inverse function. Then

$$dY(t) = m(Y, \phi, t)\, dt + dw(t) \tag{12a}$$

$$d\phi(t) = (\sigma^2[r(t), \phi(t), t] - 2\kappa(t)\phi(t))\, dt \tag{12b}$$

where

$$m(Y, \phi, t) = \frac{\partial Y(t)}{\partial t} + \mu(r, \phi, t)\frac{\partial Y(t)}{\partial r(t)} + \frac{1}{2}\sigma^2[r(t), \phi(t), t]\frac{\partial^2 Y(t)}{\partial r(t)^2}.$$

To illustrate the transform, first consider a *square root* model, where $\gamma = 0.5$ in equation (10) and $\kappa(t) = \kappa$. Then, for $r(t) > 0$, $Y(t) = 2\sqrt{r(t)}/\sigma$, and the specific structures for equations (12a) and (12b) are given by

$$\sigma^2[r(t), \phi(t), t] = \frac{\sigma^4[Y(t)]^2}{4}$$

$$m(Y, \phi, t) = \frac{1}{Y(t)}\left[\frac{\kappa}{2}\left(\frac{4v(\phi, t)}{\sigma^2} - [Y(t)]^2\right) - \frac{1}{2}\right]$$

where

$$v(\phi, t) = f(0, t) + \frac{d}{dt}f(0, t) + \frac{\phi(t)}{\kappa}.$$

As a second example, consider the *proportional* model where $\gamma = 1$. Then, for $r(t) > 0$, $Y(t) = \ln[r(t)/\sigma]$ and the specific structure for equation (12) is

$$\sigma^2[r(t), \phi(t), t] = \sigma e^{2\sigma Y(t)}$$

$$m(Y, \phi, t) = \frac{1}{\sigma}\left[v(Y, \phi, t) - \frac{1}{2}\sigma^2\right]$$

where

$$v(Y, \phi, t) = \left\{\kappa[f(0, t) - e^{\sigma Y(t)}] + \phi(t) + \frac{d}{dt}f(0, t)\right\}e^{-\sigma Y(t)}.$$

Once the transformed process with constant volatility is obtained, a lattice approximation can be established. As usual, this procedure begins by partitioning the interval of interest into subintervals of length Δt. Assume that at the beginning of some time increment the approximating variables are y^a

and ϕ^a. In the next time increment the variables move to either (y^{a+}, ϕ^{a+}) or to (y^{a-}, ϕ^{a-}) where,

$$y^{a+} = y^a + (J+1)\sqrt{\Delta t} \geq y^a + m(y^a, \phi^a, t)\Delta t \geq y^a + (J-1)\sqrt{\Delta t} = y^{a-}.$$

Here, J is an integer chosen so that the two terms, y^{a-} and y^{a+}, bracket the expected value of the interest rate in the next time increment.[8] Choosing J so that the expected value brackets the two successor points for the state variables ensures that a probability value for the jump can be found that is strictly between 0 and 1. Let $p = p(y^a, \phi^a)$ be the probability. Then,

$$p(y^{a+} - y^a) + (1-p)(y^{a-} - y^a) = m(y^a, \phi^a, t)\Delta t.$$

or

$$p = \frac{m(y^a, \phi^a, t)\Delta t + (y^a - y^{a-})}{(y^{a+} - y^{a-})}. \tag{13}$$

This method ensures that locally the means of the interest rates match the true drift and the variances of the approximated process converge to the true variance as the partition is refined.

Since the process for ϕ is locally deterministic (see equation (8b)), the values ϕ^{a+} and ϕ^{a-} are equal and fully determined by the current state variables (y^a, ϕ^a). Let ϕ^{a*} represent their common value. From equations (8b) and (9) we have

$$\phi^{a+} = \phi^{a-} = \phi^{a*} = \phi^a + [\sigma^2(h[y^a], \phi^a, t) - 2\kappa(t)\phi^a]\Delta t.$$

The value, ϕ^{a*}, is completely determined from its predecessor, (y^a, ϕ^a). Hence, after n steps in the lattice, the total number of distinct ϕ values at each node will equal the total number of unique paths leading to that node. Rather than keep track of all these values, we identify the two paths from the origin to each node that yields the maximum and minimum values for the state variable, and then partition the resulting interval into a finite number of points at which information will be computed. Let $\overline{\phi^a}$ and $\underline{\phi^a}$ represent the maximum and minimum values of the ϕ variable at this node.[9] Assume the interval $[\underline{\phi^a}, \overline{\phi^a}]$ is partitioned into m equidistant points with $\phi^a(k)$, $k = 1, 2, \ldots, m$ representing the kth point and

$$\underline{\phi^a} = \phi^a(1) < \phi^a(2) < \cdots < \phi^a(m) = \overline{\phi^a}.$$

Figure 1 provides an illustration of a lattice constructed over the first three periods. In this example the initial term structure is flat at 4 percent throughout and the volatility structure is of the form in equation (5) with

$$\sigma_f(t, T) = \sigma[r(t)]^\gamma e^{-\kappa(T-t)}. \tag{14}$$

[8]Formally, the integer J is chosen as follows. Let $Z = \text{int}[m(y^a, \phi^a, t)\sqrt{\Delta t}]$. Then

$$J = \begin{cases} \delta Z, & \text{if } Z \text{ is even;} \\ \delta|Z| + 1 & \text{otherwise.} \end{cases}$$

where $\delta = \text{Sign}(Z)$.

[9]Clearly, these two values will vary for each node, and only on the edges of the lattice will the two values be equal.

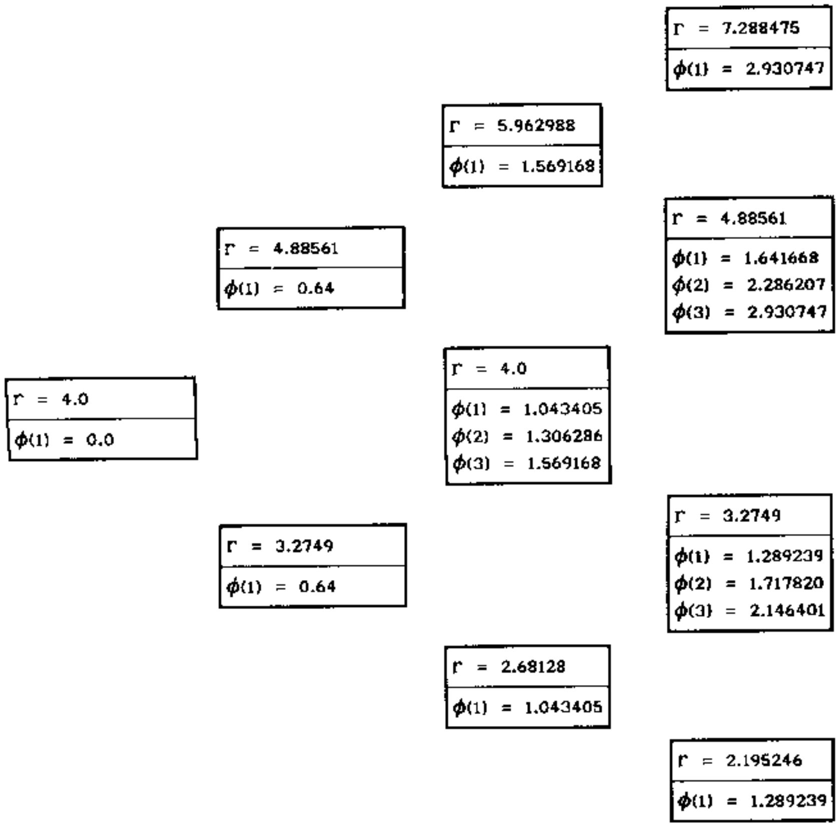

Figure 1. Construction of the lattice. The initial term structure is flat at 4 percent. The volatility structure is given by equation (14) with $\kappa = 2$ percent, $\sigma = 20$ percent, and $\gamma = 1.0$. The time partition is one year, and other than along the edges, three partitions are used for the state variable, ϕ. The lattice shows the interest rate at each node, together with the ϕ variables. The interest rate values are in percentage units, while the ϕ values are in percentage squared units. The lattice is constructed in a forward direction using the transform in equation (12). As an example, in the first period, $r = 0.04$ and hence $y = \ln(r)/\sigma = -16.094$. With $\Delta t = 1$, $y + = y_0 + 1 = -15.094$. Hence $r^* = \exp(\sigma y^+) = 4.88561$ percent. The initial value of ϕ is 0. Using equation (12b) leads to an updated value of ϕ, of 0.64. Since ϕ is locally deterministic, this value is used in the up and in the down node. In this example, $J = 1$ at all nodes. That is, there are no multiple jumps. In a three-period example, the maximum number of paths to any node is 3. All $\phi(2)$ values on the lattice are obtained by taking the average of the smallest value, $\phi(1)$, and the largest value, $\phi(3)$. In this example, after 3 periods there are 8 distinct states.

In this example, $\sigma = 20$ percent, $\kappa = 2$ percent, $\gamma = 1$, the time increment is one year, and the number of ϕ values at each node is restricted to no more than three. Figure 1 shows the evolution of the spot interest rate, r on the lattice as well as all the ϕ values at each node.

Once the lattice is established, claims can be priced using backward recursion. In particular, at the expiration date, for each (y^a, ϕ^a) location, the boundary values can be computed by first constructing the term structure using equation (6a) and then establishing the exercise value. Let $g_n(y^a, \phi^a)$ be the value of the claim in the nth period conditional on the state variables being (y^a, ϕ^a). Now consider the general backward recursion. Assume the values for the claim have been determined for all states in the lattice at the $(i+1)$st period, and that claim prices in period i are to be computed. Given the state is (y^a, ϕ^a), we know the successor nodes and the associated probabilities. The value of the claim at the ith time period, $g_i(y^a, \phi^a)$ say, is given by

$$g_i(y^a, \phi^a) = [\, p g_{i+1}(y^a, \phi^{a*}) + (1-p) g_{i+1}(y^{a-}, \phi^{a*})] e^{-r_i^a \Delta t}. \quad (15)$$

Actually, since ϕ^{a*} is completely determined by (y^a, ϕ^a), the values of the claims at the successor nodes may not be available. However, by construction, prices of the claims in the next time period at nodes (y^{a+}, ϕ_+^{a*}) and (y^{a+}, ϕ_-^{a*}) will be available where $\phi_-^{a*} \le \phi^{a*} \le \phi_+^{a*}$. In this case the price of the claim at (y^{a+}, ϕ^{a*}) can be obtained by linearly interpolating between the prices $g_{i+1}(y^{a+}, \phi_+ ua^*)$ and $g_{i+1}(y^{a+}, \phi_-^{a*})$. Similarly, linear interpolation may be necessary to obtain the prices at (y^{a-}, ϕ^{a*}).[10] If the claim is American, then the above recursion has to be modified. In particular, the price given in equation (15) has to be compared to the exercise value of the claim, with the higher of the two values recorded.

Figure 2 provides a step-by-step illustration of the pricing process for a three-year European call option that gives the holder the right to purchase a five-year discount bond for the current forward price. The values of the underlying bond for each set of state variables in the third year are presented. These values are computed using equation (6a). The boundary values for the claim are shown below the bond prices. Once the boundary values are obtained at the expiration date, equation (15) is used recursively to eventually obtain the theoretical price. At each stage of the valuation, the probability values for up-jumps are computed using equation (13) and are shown in parentheses in Figure 2.

Of course, the algorithm converges to the continuous time limit only when the time and ϕ partitions are made arbitrarily fine. The actual rate of convergence for a variety of interest rate claims are examined next.

III. Convergence Properties

The computational effort involved in pricing a stock option on a binomial lattice grows linearly with the number of time partitions. Our algorithm shares this property. In particular, if the number of ϕ values at each node is m, then our algorithm requires approximately m times more effort than the usual bionomial stock option model with the same number of time partitions.

[10] The idea of carrying a vector at each node of the lattice and using linear interpolation is not new. Examples include Hull and White (1993a) and Ritchken, Sankarasubramanian, and Vijh (1993), who price path-dependent contracts on underlying Geometric Wiener processes.

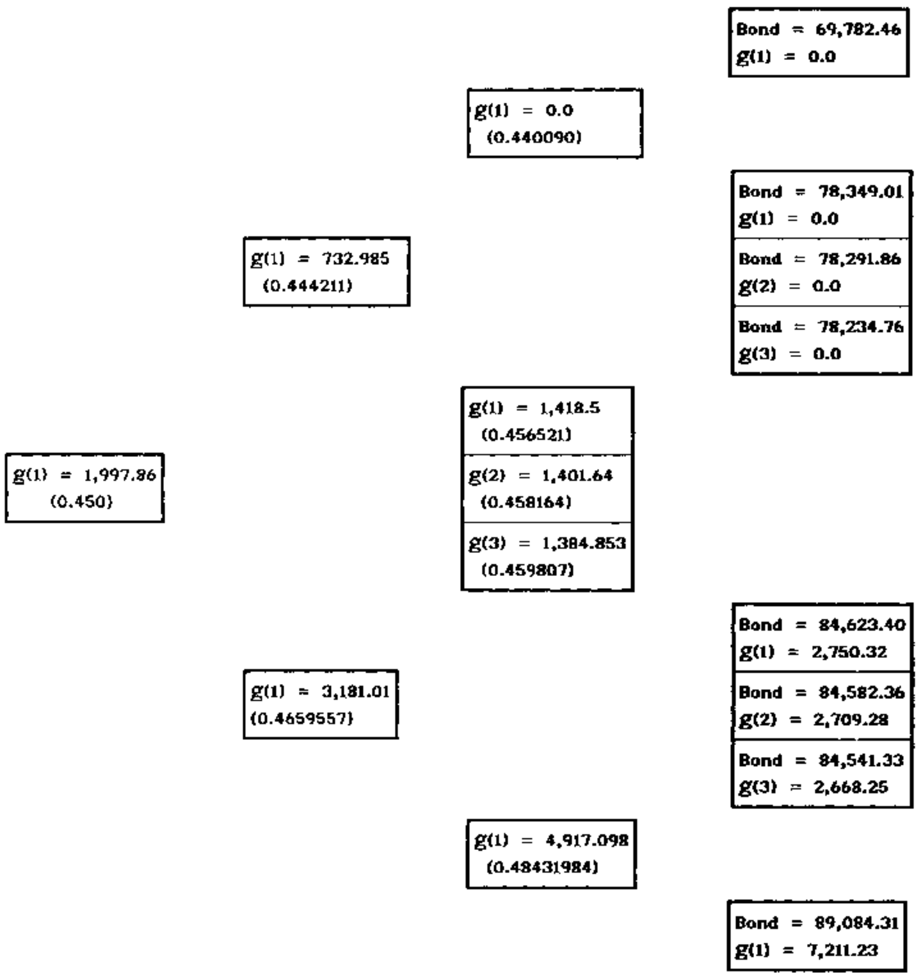

Figure 2. Construction of option prices. Consider an option that after 3 years provides the holder with the right to buy a five year discount bond with face value, \$100,000. The strike price is set at the current forward price of \$81,873.07. The probabilities of up-jumps from each state are computed using equaiton (13) and are shown in parentheses. Given r and ϕ, the value of y can be computed ($y = \ln(r)/\sigma$), and the formula is then used. For example, the probability of a first jump up is 0.450. At the expiration date, the value of the underlying bond in each state is shown. The price of the bond is computed using equation (8b). The terminal value of the call option on the bond is shown below the bond price. Given the probabilities at each node, the value of the claim in each state can then be computed using equation (15). The fair value of the option on this lattice is \$1,997.86.

In contrast, the computational effort involved in pricing claims using the general implementation of HJM doubles with the number of partitions.[11]

To illustrate the convergence behavior of prices produced by the models, we consider three applications. The first involves pricing short-dated claims on long-term bonds. The second involves pricing long-dated claims, namely long-term callable coupon bonds. The third application illustrates how complex interest rate exotics can be priced. In particular, we examine adjustable rate preferred stocks, which are complex interest rate derivatives, offered in perpetuity, and have dividends that are reset periodically at a fixed spread above, at, or below the highest of several points on the Treasury yield curve.

A. Short-Dated Options on Long-Term Bonds

Consider the problem of pricing a one-year at-the-money American put option on a 30-year discount bond. In this application, our algorithm only requires the lattice to be built over the lifetime of the claim, rather than over the lifetime of the underlying instrument. In particular, since the entire term structure can be constructed at the end of the year, the prices of the long-term bond in each state can be obtained, and the boundary values for the option can therefore easily be computed.[12]

The initial term structure was assumed flat at 5 percent, and the volatility structure for forward rates was taken to be of the form in equation (14). Figure 3 shows the convergence rate for at-the-money American put prices for different time partitions over the year and for different ϕ partitions, for $\kappa = 2$ percent, $\gamma = 1.0$, and $\sigma = 10$ percent.

Notice that as the number of ϕ-values increases, the option prices converge as the time partition increases. Figure 3 shows the convergence rate when the number of ϕ values at each node is 2, 5, 10, 25, 50, 100, and 200. The results are almost identical for all ϕ partitions that exceed 25. With 25 partitions, the number of computations for our algorithm is about 25 times larger than those required in a binomial stock option model with the same number of time partitions. Moreover, with regard to the time partitions, convergence is quite rapid, with 50 time partitions providing sufficient accuracy. The analysis for in- and out-the-money American put options yielded similar convergence properties.

Since no analytical solution exists for American put options, the true limiting price cannot be measured, and so to confirm that the lattice prices do converge to the true value, we compare the results of European put prices computed on the lattice to those generated by computer simulation with

[11] More formally, the complexity of the bionomial stock option model with n time steps is of the order n^2. The computational complexity of our algorithm is also n^2, while the general HJM implementation has order 2^n.

[12] Many term structure constrained models do not have this property and require a lattice to be built over the lifetime of the underlying security. As a result, the partitions often have to be quite large or time varying. Examples include Black, Derman, and Toy (1990), Black and Karasinski (1991), and Hull and White (1993b).

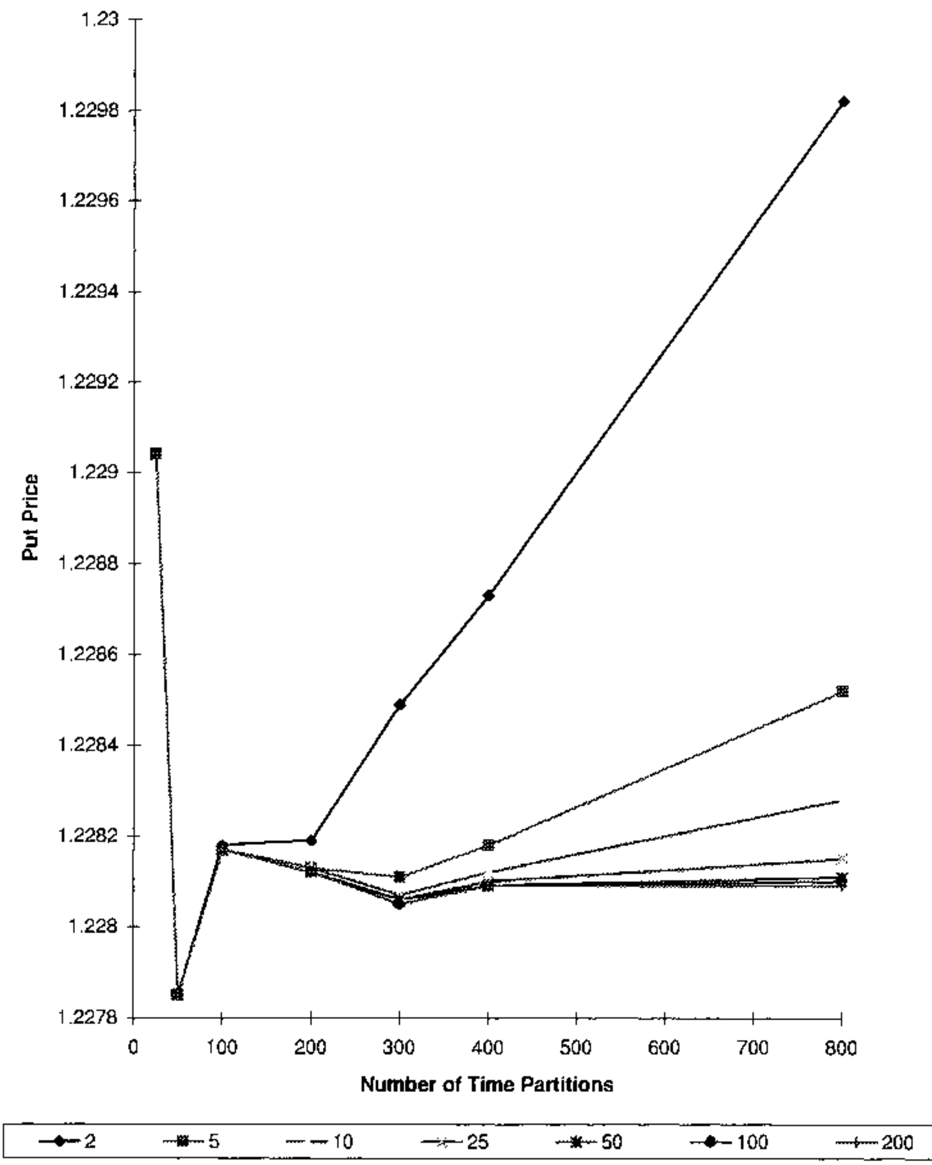

Figure 3. Convergence of American put prices. Figure 3 shows the rate of convergence of the price of a one year at the money American put option price. The initial term structure is flat at 5 percent, and the volatility structure is given by equation (14) with $\kappa = 2$ percent, $\sigma = 10$ percent, and $\gamma = 1$. The face value of the underlying bond is \$100, due in 31 years' time. The strike price of \$21.3130 is set at the forward price for the bond in one year's time. The graph indicates the very rapid convergence of prices as the ϕ-partition is refined. Indeed, for 25 or more values of ϕ at each node, the resulting prices are indistinguishable. ϕ represents the accumulated variance of the forward rate. The rate of convergence for different time partitions is also seen to be very rapid.

Table I

Sensitivity of Callable Bond Prices to the Partition for the ϕ Variable

The volatility structure is given by equation (14), with $\gamma = 1$, $\kappa = 2$ percent, and $\sigma = 10$ percent per year. The theoretical noncallable price represents the price of the Southern Bell Bond if it were not callable. The theoretical price for the actual bond is shown in the third column. The final column reports the theoretical value of the call feature. Notice that as the partition of ϕ values increases, the prices converge. ϕ represents the accumulated variance of the forward rate. A partition of 25-values appears satisfactory.

Number of ϕ Values	Noncallable Price (% of Par)	Callable Price (% of Par)	Option Value (% of Par)
2	104.55	101.44	3.11
5	104.55	101.49	3.06
10	104.55	101.50	3.05
25	104.55	101.51	3.04
50	104.55	101.51	3.04
100	104.55	101.51	3.04
200	104.55	101.51	3.04

control variates. For extremely large simulations, the differences in prices produced by the two methods are negligible.[13]

B. Long-Dated Derivatives

Long-term options on bonds most commonly arise as call features embedded in callable coupon bonds. To illustrate the pricing process, consider the Southern Bell 8 3/4 percent bond due on 9/1/2024. The bond has a maturity of 30 years and is callable, with a call schedule that begins at 104.20 in the first year, and then declines to 100 after 20 years. The term structure is initialized to the structure that existed on August 31, 1994. The volatility structure is taken to be of the form in equation (14) with $\kappa = 2$ percent, $\gamma = 1$, and $\sigma = 10$ percent. The 30-year time period is broken into 360 time periods. Table I shows the convergence of prices allowing for different numbers of ϕ values at each node. As with short-term derivatives, convergence is quite rapid and is accomplished with about 25 ϕ values.

This problem could not be solved using the usual path-dependent implementation of the HJM model. Specifically, with just two partitions a year, the path dependent lattice would have 2^{60} terminal nodes, making computation impossible. To date, the only efficient HJM models that can solve this problem require deterministic volatility structures. The example illustrates how long-dated contracts can be efficiently priced in the HJM paradigm, when volatilities are not deterministic.

[13] For a discussion on the simulation approach for this problem see Ritchken and Sankarasubramanian (1995).

Table II shows the sensitivity of the Southern Bell callable bond price to the volatility parameter, σ. For σ near 11 percent per year, the theoretical price aligns up with the actual market price for that date.

C. Adjustable Rate Preferred Stocks

Adjustable rate preferred stocks (ARPS) are complex interest rate derivatives that typically pay quarterly dividends, the size of which are linked to the best performing of the 3-month Treasury rate, the 10-year Treasury rate, and the 30-year Treasury rate. Usually a lifetime cap and floor are provided, as well as callable features. The contracts are usually offered in perpetuity. As an example, consider the Bank of America Corporation's adjustable rate preferred stock. The issue consists of 6 million shares at $50 each. The shares are currently callable at par. The payments at each quarter are linked to the Treasury curve as follows:

$$\text{Max}[3 \text{ mo}, 10 \text{ yr}, \text{Avg}\,[18, 19, 20, 21, 22 \text{ yrs}]] - 200 \text{ basis points}.$$

The payment has a floor of 6.5 percent and a cap of 14.5 percent. The yield curve is initialized to the structure that existed on August 31, 1994. The volatility structure for this model is given by equation (14) with $\kappa = 2$ percent, $\sigma = 10$ percent, and $\gamma = 1$.

To value such a claim calls for a lattice that has quarterly time steps at the very least. Further, at each node on the lattice, we must know the yield curve for at least the next 22 years. To value this perpetuity, we construct a 100-year quarterly lattice. At the end of 100 years, the value of the ARP is determined at each node under the assumption that the yield curve from then

Table II

Sensitivity of Prices to the Volatility

The table shows the sensitivity of the price of the Southern Bell callable coupon bond to changes in the volatility parameter. The second column indicates the theoretical price of the bond if it were not callable. The value remains constant regardless of the volatility parameter since the initial term structure is fixed. The value of the call feature is isolated in the last column. The actual price of this bond is 101.47 percent of par. This value implies a volatility close to 12 percent.

Volatility (Annual %)	Bond Price (Noncallable)	Bond Price (Callable)	Option Price
0	104.55	104.02	0.53
2.5	104.55	104.02	0.53
5.0	104.55	103.50	1.05
7.5	104.55	102.51	2.04
10.0	104.55	101.51	3.04
12.5	104.55	100.18	4.37
15.0	104.55	99.12	5.43
17.5	104.55	98.14	6.41
20.0	104.55	97.44	7.11

Table III

Sensitivity of ARPS Prices to the Partition for the ϕ Variable

The volatility structure is given by equation (14), with $\gamma = 1$, $\kappa = 2$ percent, and $\sigma = 10$ percent per year. The Noncallable Price represents the theoretical price of the Bank of America adjustable rate preferred stock if it were not callable. The theoretical price for the actual contract is shown in the third column. The final column reports the theoretical value of the all feature. Notice that as the partition of ϕ values increases, the prices converge. ϕ represents the accumulated variance of the forward rate. A partition of 25-values appears satisfactory.

Number of ϕ Values	Noncallable Price (% of Par)	Callable Price (% of Par)	Option Value (% of Par)
2	92.28	91.54	0.74
5	92.13	91.50	0.73
10	92.16	91.45	0.71
25	92.09	91.40	0.69
50	92.08	91.39	0.69
100	92.08	91.39	0.69
200	92.08	91.39	0.69

on is deterministic. Given the boundary values, the usual backward recursion could be used to price the preferred.[14] Table III shows the convergence of prices for different partition sizes for ϕ. The results again are similar to the other examples, in that 25 partitions appear to be sufficient. The sensitivity of prices to changes in the volatility parameter are shown in Table IV. The actual price of the ARP is 92.3 percent of par, a value that implies a volatility of about 12 percent.

Owing to the frequency of payouts over the lifetime of the security, it would be extremely difficult to value this ARP using the path-dependent implementation of HJM.

In all the above examples, the convergence rate of prices is quite rapid and obtains with about 25 ϕ values. In practice then, the computational effort for our implementation of the HJM models, when the volatility structure belongs to the RS class, is approximately 25 times greater than that required from a binomial stock option model with the same number of time partitions.

IV. Comparison with Alternative Models

The reconnecting lattice models developed here can be adapted to price claims when spot rates have quite general volatility structures as given in equation (5). However, when spot rate volatilities are specified as in equation (10) with $\gamma = 1$, our lattice provides an alternative to the Black, Derman, and Toy (BDT) (1990)) model. Both lattices have the ability to produce prices consistent with an observable term structure. In addition, both allow the spot

[14] The sensitivity of the price to changes in the date beyond which we take the term structure to be conditionally deterministic was investigated. The prices remain virtually unchanged for periods beyond 50 years.

Table IV

Sensitivity of Prices to Volatility Parameter

The table shows the sensitivity of the price of the Bank of America Corporation adjustable rate preferred stock to changes in the volatility parameter. The second column indicates the theoretical price of the preferred if it did not have a call feature. The value of the call feature is isolated in the last column. The actual price of this ARP is 92.37 percent of par. This value implies a volatility of about 12 percent.

Volatility (Annual %)	ARP (Noncallable)	ARP (Callable)	Option Price
0	85.70	85.70	0.00
2.5	86.62	86.61	0.01
5.0	88.15	88.10	0.05
7.5	90.09	89.81	0.28
10.0	92.09	91.40	0.69
12.5	93.61	92.55	1.06
15.0	96.35	94.07	2.28
17.5	97.49	94.60	2.89
20.0	98.96	95.08	3.88

rate volatility to be proportional to the level of the spot interest rate. The BDT lattice, however, *assumes* that the spot interest rate is the sole state variable that determines prices. In contrast, our lattice *derives* the appropriate state variables under the HJM paradigm. Moreover, in our paradigm, the structure of forward rate volatilities can be maintained over time. This is in contrast to the BDT lattice where the structure for future forward rate volatilities is not well understood, and indeed is fully determined by the specification of the initial term structure. That is, the structure of the future volatilities of forward rates cannot be decoupled from the initial specification of the term structure. Actually, in the BDT paradigm, the future level of the spot rate volatility, as well as future forward rate volatilities, can only be curtailed if more initial conditions are specified and more time varying functions introduced into the dynamics.[15]

For relatively simple contracts, such as options on specific bonds, the BDT models are quite efficient. Indeed, once the lattice of spot interest rates are computed, the algorithm is as rapid as the binomial stock option model, with an equivalent number of time steps. However, if the BDT model is used to price the one-year option on a thirty-year bond, the lattice would have to extend out 30 years, and the overall computational requirements may exceed that of our algorithm.

When the underlying contract is relatively complex and, at each node of the lattice, requires information from across the term structure, then the computational effort of BDT might equal or exceed the effort required in our lattice

[15] A good example of this approach is Hull and White (1993b). In their analysis, three time-varying functions are required in the stochastic process of the state variables in order to fit a variety of required initial conditions.

procedures. As an example, if the Bank of America's adjustable rate preferred stock is to be priced using BDT, then, with the underlying lattice constructed over quarters, *at least* $22 \times 4 = 88$ values of the term structure need to be carried at each node. This is in contrast to our lattice, where given the two state variables, the relevant points on the term structure can be immediately computed. In addition, the construction of the BDT lattice would have to go out at least 22 years further in time relative to our model. For this example, their overall computational requirements will exceed those of our algorithm.

In the examples presented in this article, other than the initial term structure, all that is required is estimates of two parameters, namely the volatility term, σ, and the mean reversion function $\kappa(t)$. For $\kappa(t) = \kappa$, the two parameters can be chosen so as to match the prices of two specific caps or so as to minimize some measure of error produced in calibrating the model to a term structure of cap prices or initial forward rate volatilities. Of course our lattice can match more initial conditions by requiring the parameter κ to be time varying. Clearly, the greater the number of parameters introduced into the analysis, the greater the number of prices that can be fitted. As usual, the cost of an overparameterized model usually takes the form of poor out-of-sample performance. Hence, if the model is to be used for hedging, where the intertemporal dynamics are important, fewer parameters may be desirable. Since the dynamics of the assumed state variable in the BDT paradigm are completely determined by the initial conditions, it is not possible to significantly reduce the number of parameters. It is therefore not surprising that the hedging performance of such models has been found to be quite poor. It remains for future empirical work to establish whether the RS models with simple volatility structures as in equation (14) perform well. In particular, the stability of estimates of κ and σ over time may provide good indicators of the viability of this approach.

V. Conclusion

This article establishes a lattice for pricing European and American interest rate claims using the HJM paradigm. To avoid the path dependence that arises in their general approach, attention is focused on a class of volatilities for forward rates that made the dynamics of the term structure Markovian with respect to two state variables. The class of volatilities that accomplishes this is large and imposes no severe restrictions on the volatility of spot rates. The performance of the algorithm is investigated. For European options, the rate of convergence to the limiting price is established, where the limiting price is obtained using computer simulation. The convergence behavior for American contracts is also examined. The examples illustrate the rapid rate of convergence of prices that can be obtained for HJM models in spite of the fact that volatilities are not deterministic. Indeed, the examples are selected to highlight the fact that the models can be established to price long-dated claims under a richer class of volatility assumptions than could previously be accomplished in the HJM paradigm. Finally, comparisons between the proportional volatility model in the RS class and the term structure constrained models of BDT are made.

It remains for future empirical work to establish how well models with these volatility restrictions perform, and to identify whether the addition of a few maturity varying parameters are necessary. If empirical studies show no support for any volatility structures in the RS class, then, implementing the HJM paradigm for long term claims and nondeterministic volatilities will continue to remain an unresolved problem. If, on the other hand, empirical support is found for volatility structures in our class, then this algorithm should be useful for pricing and hedging purposes. Either way, empirical research in this area has important consequences.

REFERENCES

Amin, K., and J. Bodurtha, 1995, Discrete time valuation of American options with stochastic interest rates, *Review of Financial Studies* 8, 193–234.

Amin, K., and R. Jarrow, 1991, Pricing foreign currency options under stochastic interest rates, *Journal of International Money and Finance* 10, 310–330.

Amin, K., and A. Morton, 1994, Implied volatility functions in arbitrage free term structure models, *Journal of Financial Economics* 35, 141–180.

Black, F., E. Derman, and W. Toy, 1990, A one factor model; of interest rates and its application to treasury bond options, *Financial Analysts Journal* 46, 33–39.

Black F., and P. Karasinski, 1991, Bond and option pricing when short rates are lognormal, *Financial Analysts Journal* 47, 52–59.

Caverhill A., 1994, When is the short rate markovian?, *Mathematical Finance* 4, 305–312.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders, 1992, An empirical comparison of alternative models of the term structure of interest rates, *Journal of Finance* 47, 1209–1228.

Cheyette O., 1992, Term structure dynamics and mortgage valuation, *Journal of Fixed Income*, March, 28–41.

Cox, J., J. Ingersoll, and S. Ross, 1985, A theory of the term structure of interest rate, *Econometrica* 53, 385–407.

Heath, D., R. Jarrow, and A. Morton, 1990, Bond pricing and the term structure of interest rates: A discrete time approximation, *Journal of Financial and Quantitative Analysis* 25, 419–440.

———, 1992, Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica* 60, 77–105.

Heath, D., R. Jarrow, A. Morton, and M. Spindel, 1993, Easier done than said, *Risk Magazine* 5, 77–80.

Hull, J., and A. White, 190, Pricing interest derivative securities, *Review of Financial Studies* 3, 573–592.

———, 1993a, Efficient procedures for valuing European and American path dependent options, *The Journal of Derivatives* 1, 21–32.

———, 1993b, One factor interest rate models and the valuation of interest rate derivative securities, *Journal of Financial and Quantitative Analysis* 28, 235–254.

———, 1993c, Bond option pricing on a model for the evolution of bond prices, *Advances in Options and Futures Research* 6, 1–13.

Jamshidian, F., 1989, An exact bond option formula, *Journal of Finance* 44, 205–09.

Nelson, D., and K. Ramaswamy, 1990, Simple bionomial processes as diffusion approximations in financial models, *Review of Financial Studies* 3, 393–430.

Ritchken, P., and L. Sankarasubramanian, 1992, Pricing the quality option in treasury bond futures, *Mathematical Finance* 2, 197–214.

———, 1993, Averaging and deferred payment yield agreements, *The Journal of Futures Markets* 13, 23–41.

———, 1995, Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance* 5, 55–72.

Ritchken, P., L. Sankarasubramanian, and A. Vijh, 1993, Valuation of path dependent options on the average, *Management Science* 39, 1202–13.

Vasicek, O., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.