

## AN IMPROVED FINITE DIFFERENCE APPROACH TO FITTING THE INITIAL TERM STRUCTURE

Kenneth R. Vetzal  
Centre for Advanced Studies in Finance  
University of Waterloo  
WATERLOO ON CANADA N2L 3G1

Telephone: 519 888 4567 ext. 6518  
Facsimile: 519 888 7562  
E-mail: kvetzal@uwaterloo.ca

### Abstract

This paper describes an alternative discretization strategy for mean-reverting models which permits the use of implicit finite difference methods along lines developed for explicit methods by Hull and White (1990a). This provides significant benefits in terms of stability, convergence, and flexibility. A technique for fitting the initial term structure is developed which is similar to that proposed by Uhrig and Walter (1996), but much more efficient computationally. The approach is illustrated with applications to three common single factor models and two factor stochastic volatility extensions of these models.

**Keywords:** Term structure of interest rates, Interest rate contingent claims, Bond options, Finite difference methods

**Acknowledgements:** Financial support for this research was provided by the Social Sciences and Humanities Research Council of Canada. Thanks to Phelim Boyle, Peter Forsyth, Bruce Simpson, and Robert Zvan for many helpful discussions. Chris Gmach provided excellent research assistance. Thanks are also due to Peter Forsyth for kindly providing access to the WATSIT sparse matrix software library.

## 1 Introduction

Two main approaches have emerged in the modern literature on the pricing of interest rate contingent claims. Each is intended to ensure that the model being used is consistent with an observed initial yield curve. The first strategy involves specifying the evolution of some small number of points on the yield curve (typically just one — the instantaneous risk free rate) and incorporating time-dependent parameters into the model in order to meet the consistency requirement. Among the early contributors to this literature were Hull and White (1990b), Black, Derman, and Toy (1990), Jamshidian (1991), and Black and Karasinski (1991).

The second line of attack was pioneered by Heath, Jarrow, and Morton (1992). It takes the entire term structure as a model input and specifies its subsequent evolution over time in an arbitrage-free manner. This is automatically consistent with today's yield curve, so there is no need to augment the model with time-dependent parameters. The price to be paid for this convenience is that the model typically ends up being non-Markovian and is then relatively hard to implement, especially for American-style claims.

If a time-dependent parameter approach is to be used, an efficient method must be available to determine these parameters from the observed term structure. As pointed out by Uhrig and Walter (1996), many of the early papers in this area feature techniques which are applicable only to a specific interest rate model. Perhaps the first authors to attempt to address this were Hull and White (1993), in the context of a trinomial lattice which was built upon the framework described in Hull and White (1990a). This is essentially an explicit finite difference method for solving a partial differential equation (PDE). As is well-known, such methods are prone to stability problems unless certain conditions are satisfied. These conditions restrict the size of a time step relative to the spacing between points on the discrete grid. It can be difficult to satisfy them unless the model is first transformed to one with constant standard deviation. Moreover, the Hull and White method requires that the model be mean-reverting after the transformation, which may not be true for some parameter values. Similarly, it can be hard to vary time step sizes so as to match cash flow dates. The entire approach is inherently rather inflexible, though Hull and White have managed to address many of these and related issues in a series of recent papers (Hull and White (1994a, 1994b, 1996)).

One promising alternative is to use an implicit finite difference technique. Such methods are unconditionally stable, and therefore quite flexible. For example, there are no obstacles to changing the time step size in order to hit a cash flow date exactly. There is no need to transform the model to one with constant standard deviation. As a result, some cases can be handled using this type of approach which cannot be fit into the Hull and White framework. Such an approach has been advocated in a recent paper by Uhrig and Walter (1996), who describe an "inverted implicit finite difference" method and apply it to several common single factor models.

The contributions of the present paper build on the ideas of both Hull and White (1990a) and Uhrig and Walter (1996). As will be discussed in greater detail below, one of the advantages of the Hull and White explicit method is that it avoids the specification of boundary conditions at the end points of the spatial grid. This can be a desirable feature in models with mean-reversion.

This paper demonstrates that this can effectively also be accomplished for implicit methods, through a slight change in the discretization strategy. As noted above, this offers superior stability, convergence, and flexibility properties. This alternative strategy can then be combined with the Uhrig and Walter approach to fit the initial term structure. Significant computational efficiency can be gained by using the discretization strategy proposed here in conjunction with a different choice for the time-dependent parameter than that advocated by Uhrig and Walter. The approach is demonstrated with applications to three common single factor term structure models as well as three bivariate stochastic volatility extensions of these models.

The outline of the paper is as follows. Section 2 describes in detail the alternative discretization strategy. Section 3 provides an example which shows the superior accuracy of this approach as well as some potential speed improvements. Section 4 describes one way how to calibrate models to fit an initial term structure, and it is followed by Section 5 which presents illustrative applications. Section 6 concludes with a brief summary.

## 2 A Discretization Strategy for Mean-Reverting Models

This section presents a discretization scheme for PDE models of contingent claims valuation where the underlying factor(s) are mean-reverting. In such situations, the state variables have a tendency to move back towards some central value as time moves forward. It is important to remember, though, that time actually moves backwards when one solves a PDE to find the price of a contingent claim (i.e. the pricing algorithm works back from the terminal payoff to the current time). In this context, the state variable(s) are not tending towards a central value, but rather spreading out. In other words, the flow of information runs out towards the boundaries of the computational domain. Consequently, it is desirable to design pricing algorithms in such a way that any boundary conditions imposed on the edges of the computational domain arise naturally from the problem at hand. The well-known Hull and White (1990a) trinomial lattice technique accomplishes this by avoiding any specification of boundary conditions (except of course for the terminal payoff). Unfortunately, it is an explicit type of method and is therefore prone to stability problems. The objective in this section is to circumvent these concerns by designing an implicit scheme with similar properties in terms of the boundary conditions. The method used is a simple two point upstream technique. Such methods are commonly used in other applications of numerical solutions of PDEs (e.g. computational fluid dynamics), but they appear to be relatively unknown in the finance literature.<sup>1</sup>

The easiest way to understand the scheme relies on knowledge of the relationship between the Hull and White (1990a) trinomial lattice technique and standard finite difference approaches, and so a brief review of those methods is presented first. The well-known one factor Vasicek (1977) term structure model is used throughout the section for illustrative purposes. In this model the instantaneous risk free interest rate  $r$  evolves according to the mean-reverting process:

$$dr = \kappa(\theta - r) dt + \sigma dB \quad (1)$$

where  $B$  is a standard Brownian motion and  $\kappa$ ,  $\theta$ , and  $\sigma$  are parameters (respectively the speed of adjustment, reversion level, and volatility). Assuming for simplicity that the market price of interest rate risk is zero, the value of an interest rate contingent claim  $u$  may be found by solving the following PDE:

$$\frac{1}{2}\sigma^2 u_{rr} + [\kappa(\theta - r)]u_r + u_t - ru = 0 \quad (2)$$

subject to an appropriate terminal boundary condition. For example, the price of a pure discount bond paying \$1 at maturity date  $T$ , denoted by  $u(r, t, T)$ , is found by solving (2) subject to  $u(r, T, T) = 1$ . In order to solve (2) numerically, a discrete grid can be constructed for values of  $r$  over some interval  $[r_{\min}, r_{\max}]$ . Assume that this grid has  $M$  evenly spaced points  $r_1, \dots, r_M$  and denote the distance between successive points by  $\Delta r$ . Let the value of  $u$  at the  $i$ -th grid point at time step  $n$  be  $u_i^n$ . In a centrally-weighted explicit scheme the derivatives in (2) are approximated by:

$$\begin{aligned} u_{rr} &= \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{(\Delta r)^2} \\ u_r &= \frac{u_{i+1}^{n+1} - u_{i-1}^{n+1}}{2\Delta r} \\ u_t &= \frac{u_i^{n+1} - u_i^n}{\Delta t} \end{aligned} \quad (3)$$

where  $\Delta t$  is the length of the time step between  $n$  and  $n + 1$ . It is common in finance (though not standard in the PDE literature) to use  $r_i u_i^n$  for the  $ru$  term in (2). Substitution into (2) and rearrangement yields:

$$u_i^n (1 + r\Delta t) = p_{i,i-1} u_{i-1}^{n+1} + p_{i,i} u_i^{n+1} + p_{i,i+1} u_{i+1}^{n+1} \quad (4)$$

where

$$\begin{aligned} p_{i,i-1} &= \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{\kappa(\theta - r_i) \Delta t}{\Delta r} \right] \\ p_{i,i} &= 1 - \frac{\sigma^2 \Delta t}{(\Delta r)^2} \\ p_{i,i+1} &= \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{\kappa(\theta - r_i) \Delta t}{\Delta r} \right] \end{aligned} \quad (5)$$

Note that the  $p$ 's sum to one. Provided that they are all positive, the scheme has the familiar interpretation of risk-neutral probabilities and a trinomial lattice (where, over the next time interval,  $r$  can either remain constant or go up or down by  $\Delta r$ ).<sup>2</sup> As noted by Hull and White (1990a, p. 92), positivity of the  $p$ 's is a sufficient condition for the stability of the explicit method.

Alternatively, the spatial derivatives  $u_{rr}$  and  $u_r$  can be approximated at time step  $n$  rather than  $n + 1$ . This yields an implicit scheme:

$$c_{i,i-1} u_{i-1}^n + c_{i,i} u_i^n + c_{i,i+1} u_{i+1}^n = u_i^{n+1} \quad (6)$$

where

$$\begin{aligned}
 c_{i,i-1} &= -\frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{\kappa(\theta - r_i) \Delta t}{\Delta r} \right] \\
 c_{i,i} &= 1 + \frac{\sigma^2 \Delta t}{(\Delta r)^2} + r_i \Delta t \\
 c_{i,i+1} &= -\frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{\kappa(\theta - r_i) \Delta t}{\Delta r} \right]
 \end{aligned} \tag{7}$$

Such schemes are unconditionally stable (there is no requirement that the  $c$ 's be positive), but require the solution of a system of linear equations as well as some boundary information at  $r_{\min}$  and  $r_{\max}$ . In the present one dimensional context, the equation system is tridiagonal and so it can be solved easily and quickly.

In models which feature mean-reversion, explicit schemes run into the difficulty that it is not easy to ensure that the  $p$ 's are all positive. In the Vasicek (1977) model, if  $r$  is low enough, the upward drift will be strong enough that  $p_{i,i-1}$  (the probability of  $r$  decreasing over the next time interval) becomes negative. Similarly, for sufficiently high  $r$ , the downward drift will cause  $p_{i,i+1}$  to be negative. This fact led Hull and White (1990a) to propose a modified trinomial lattice method.

The basic idea of Hull and White (1990a) can be described as follows.<sup>3</sup> Consider first moderate values of  $r$  such that the grid point closest to the expected value of  $r$  after another time step is the current grid point (i.e. the absolute value of the expected change in  $r$  over  $\Delta t$  is less than  $\Delta r/2$ ). In such cases construct a trinomial lattice by solving:

$$\begin{aligned}
 \hat{p}_{i,i-1}(-\Delta r) + \hat{p}_{i,i}(0) + \hat{p}_{i,i+1}(\Delta r) &= \kappa(\theta - r_i) \Delta t \\
 \hat{p}_{i,i-1}(-\Delta r)^2 + \hat{p}_{i,i}(0)^2 + \hat{p}_{i,i+1}(\Delta r)^2 &= \sigma^2 \Delta t + [\kappa(\theta - r_i) \Delta t]^2 \\
 \hat{p}_{i,i-1} + \hat{p}_{i,i} + \hat{p}_{i,i+1} &= 1
 \end{aligned} \tag{8}$$

for the  $\hat{p}$ 's. The idea is simply to find probabilities to match the first and second moments of the change in  $r$  over  $\Delta t$ . These probabilities are denoted by  $\hat{p}$  since they are not the same as the  $p$ 's for the standard explicit method in equation (5). However, if the second term on the right hand side of the middle equation in (8) is dropped, then it is straightforward to verify that the solution to (8) is equal to the  $p$ 's given in (5). This amounts to approximating the second moment by the variance. The error involved is of order  $(\Delta t)^2$ . It should be emphasized that this is only an error in the sense of matching the second moment in a lattice, not in the context of a formal discretization of the PDE. The  $\hat{p}$ 's which solve (8) are used provided that they are all positive. When  $r$  gets too low for this to be the case, say at value  $r_1$ , modify the branching in the lattice so that  $r$  either remains constant, goes up by  $\Delta r$ , or goes up by  $2\Delta r$ . Note that this truncates the computational domain since  $r$  will not go below  $r_1$ . Form a set of equations analogous to (8) to find probabilities to match the first two moments:

$$\hat{p}_{1,1}(0) + \hat{p}_{1,2}(\Delta r) + \hat{p}_{1,3}(2\Delta r) = \kappa(\theta - r_1) \Delta t$$

$$\begin{aligned}\hat{p}_{1,1}(0)^2 + \hat{p}_{1,2}(\Delta r)^2 + \hat{p}_{1,3}(2\Delta r)^2 &= \sigma^2 \Delta t + [\kappa(\theta - r_1) \Delta t]^2 \\ \hat{p}_{1,1} + \hat{p}_{1,2} + \hat{p}_{1,3} &= 1\end{aligned}\quad (9)$$

Similarly, when  $r$  becomes too high for the probabilities under the regular branching process to be positive, say at value  $r_M$ , modify the lattice so that  $r$  either remains at  $r_M$ , decreases by  $\Delta r$ , or decreases by  $2\Delta r$ . Again, this limits the range of computation since  $r$  is constrained to not increase beyond  $r_M$  and generates the following set of equations which can be used to determine probabilities:

$$\begin{aligned}\hat{p}_{M,M-2}(-2\Delta r) + \hat{p}_{M,M-1}(-\Delta r) + \hat{p}_{M,M}(0) &= \kappa(\theta - r_M) \Delta t \\ \hat{p}_{M,M-2}(-2\Delta r)^2 + \hat{p}_{M,M-1}(-\Delta r)^2 + \hat{p}_{M,M}(0)^2 &= \sigma^2 \Delta t + [\kappa(\theta - r_M) \Delta t]^2 \\ \hat{p}_{M,M-2} + \hat{p}_{M,M-1} + \hat{p}_{M,M} &= 1\end{aligned}\quad (10)$$

To summarize, the Hull and White (1990a) method is a trinomial lattice where (risk-neutral) probabilities are determined so as to match the first two moments of the change in  $r$  over time interval  $\Delta t$ . Owing to mean-reversion, the branching in the lattice is modified at  $r_1$  and  $r_M$ , truncating the computational domain. This method is popular, fast, and generally quite accurate. It has been extended by Hull and White (1990a, 1993, 1994a, 1994b, 1996) in a series of papers dealing with issues such as an additional state variable and matching initial yield curves and term structures of interest rate volatilities through incorporating time-dependent parameters. One potential problem in the single factor case is that the underlying state variable is required to have constant standard deviation. Models where this is not the case (such as the Cox, Ingersoll, and Ross (1985) square root specification) must be transformed in order to apply the Hull and White technique. However, as pointed out by Tian (1994), depending on the parameter values the transformed process may or may not be mean-reverting. If it is not, then the Hull and White method will not converge. Another potential difficulty is that applying the technique in the case of a two factor model requires that the factors be orthogonalized and that they be mean-reverting after the transformation. Again, this may not be the case. Furthermore, some models such as the EGARCH stochastic volatility specifications examined by Andersen and Lund (1996) and Vetzal (1997) cannot be transformed in a way such that the Hull and White method can be applied.

Finally, and most importantly for present purposes, Hull and White's approach is strictly an explicit type of method. They do not provide an implicit version, in part because in some situations there may be no obvious boundary conditions to apply at the end points  $r_1$  and  $r_M$ . Indeed, Hull and White claim that explicit methods are superior on these grounds since "the explicit finite difference method has the advantage that it can require the specification of fewer boundary conditions than the implicit method" (1990a, p. 92). Furthermore, "errors are introduced by the redundant boundary conditions in implicit methods" (p. 99). On one level, these comments are correct. Equation (2) is formally a parabolic equation and technically does require the specification of boundary conditions at the end points of the grid if an implicit method is used. However, as the absolute value of  $r$  gets large, the magnitude of what is referred to in the

PDE literature as the convection coefficient,  $\kappa(\theta - r)$ , swamps that of the diffusion coefficient, the constant  $\sigma^2/2$ . In other words, for extremely high or low values of  $r$  equation (2) is said to be *convection-dominated*. Although formally parabolic, it behaves numerically as if it were hyperbolic. Consequently, it is not only unnecessary to impose boundary conditions at the end points of the grid, in fact it is desirable to avoid specifying them. However, to accomplish this in the case of an implicit method requires a different discretization strategy.

Consider the following alternative. Construct a grid going from  $r_1$  to  $r_M$  and use a standard method for the interior points  $r_2, \dots, r_{M-1}$ . At the end points, proceed as follows. For  $r_1$ , take two Taylor's expansions around  $u(r_1)$ , one for  $u(r_1 + \Delta r)$  and one for  $u(r_1 + 2\Delta r)$ . Ignoring higher than second order terms, this gives:

$$\begin{aligned} u(r_1 + \Delta r) &= u(r_1) + u_r \Delta r + u_{rr} \frac{(\Delta r)^2}{2} \\ u(r_1 + 2\Delta r) &= u(r_1) + u_r 2\Delta r + u_{rr} \frac{(2\Delta r)^2}{2} \end{aligned} \quad (11)$$

These can be rearranged to yield:

$$\begin{aligned} u_{rr} &= \frac{u(r_1 + 2\Delta r) - 2u(r_1 + \Delta r) + u(r_1)}{(\Delta r)^2} \\ u_r &= \frac{-u(r_1 + 2\Delta r) + 4u(r_1 + \Delta r) - 3u(r_1)}{2\Delta r} \end{aligned} \quad (12)$$

Similarly, at  $r_M$  take two second order expansions around  $u(r_M)$ , one for  $u(r_M - \Delta r)$  and one for  $u(r_M - 2\Delta r)$ :

$$\begin{aligned} u(r_M - \Delta r) &= u(r_M) - u_r \Delta r + u_{rr} \frac{(\Delta r)^2}{2} \\ u(r_M - 2\Delta r) &= u(r_M) - u_r 2\Delta r + u_{rr} \frac{(2\Delta r)^2}{2} \end{aligned} \quad (13)$$

Rearranging gives:

$$\begin{aligned} u_{rr} &= \frac{u(r_M - 2\Delta r) - 2u(r_M - \Delta r) + u(r_M)}{(\Delta r)^2} \\ u_r &= \frac{-u(r_M - 2\Delta r) + 4u(r_M - \Delta r) + 3u(r_M)}{2\Delta r} \end{aligned} \quad (14)$$

This implies that an explicit method can be constructed in the following way. For the interior grid points  $r_2, \dots, r_{M-1}$  the standard explicit scheme given by (4) and (5) applies. At the minimum value of  $r$ , take the discrete analogues of the spatial derivatives in (11) at time step  $n + 1$  and substitute into the PDE (2) to obtain:

$$u_1^n (1 + \tau_1 \Delta t) = p_{1,1} u_1^{n+1} + p_{1,2} u_2^{n+1} + p_{1,3} u_3^{n+1} \quad (15)$$

where

$$\begin{aligned}
 p_{1,1} &= 1 + \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{3\kappa(\theta - r_1) \Delta t}{\Delta r} \right] \\
 p_{1,2} &= -\frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{2\kappa(\theta - r_1) \Delta t}{\Delta r} \\
 p_{1,3} &= \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{\kappa(\theta - r_1) \Delta t}{\Delta r} \right]
 \end{aligned} \tag{16}$$

Note that these  $p$ 's sum to one. Furthermore, recall the relationship for the interior grid points between the probabilities for the standard explicit method in (5) and the Hull and White probabilities generated by (8): they would be identical if the term in the second equation of (8) which is of order  $(\Delta t)^2$  was ignored. In fact, the same relationship holds here. The probabilities given by (16) are the solution to (9) when the term of order  $(\Delta t)^2$  is dropped from the right hand side of the second equation.

Similarly, at the maximum value of  $r$  substitute the discrete analogues of (13) at time step  $n + 1$  for the spatial derivatives in (2) to find:

$$u_M^n (1 + \tau_M \Delta t) = p_{M,M-2} u_{M-2}^{n+1} + p_{M,M-1} u_{M-1}^{n+1} + p_{M,M} u_M^{n+1} \tag{17}$$

where

$$\begin{aligned}
 p_{M,M-2} &= \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{3\kappa(\theta - r_M) \Delta t}{\Delta r} \right] \\
 p_{M,M-1} &= -\frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{2\kappa(\theta - r_M) \Delta t}{\Delta r} \\
 p_{M,M} &= 1 + \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{3\kappa(\theta - r_M) \Delta t}{\Delta r} \right]
 \end{aligned} \tag{18}$$

As before, these  $p$ 's sum to one and are also the solution to the corresponding Hull and White equations (10) when the term of order  $(\Delta t)^2$  is dropped.

It is clear that this proposed explicit method is quite similar to that of Hull and White, becoming identical to it as  $\Delta t \rightarrow 0$ . It is constructed so that the relationship between the standard explicit method and the Hull and White trinomial lattice for the interior grid points also holds at each of the end points. Moreover, the Taylor's series approach permits the construction of an implicit scheme along similar lines, merely by evaluating the derivatives in (12) and (14) at time step  $n$  rather than at  $n + 1$  as for the explicit case. In particular, at the minimum value of  $r$ :

$$c_{1,1} u_1^n + c_{1,2} u_2^n + c_{1,3} u_3^n = u_1^{n+1} \tag{19}$$

where

$$c_{1,1} = 1 + r_1 \Delta t - \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{3\kappa(\theta - r_1) \Delta t}{\Delta r} \right]$$

$$\begin{aligned}
c_{1,2} &= \frac{\sigma^2 \Delta t}{(\Delta r)^2} - \frac{2\kappa(\theta - r_1) \Delta t}{\Delta r} \\
c_{1,3} &= -\frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{\kappa(\theta - r_1) \Delta t}{\Delta r} \right]
\end{aligned} \tag{20}$$

Notice that this discretization strategy avoids the requirement of imposing a specific boundary condition at  $r_1$ . The standard implicit scheme given by (6) and (7) applies for the interior points  $r_2, \dots, r_{M-1}$ , and at the maximum value of  $r$ :

$$c_{M,M-2}u_{M-2}^n + c_{M,M-1}u_{M-1}^n + c_{M,M}u_M^n = u_M^{n+1} \tag{21}$$

where

$$\begin{aligned}
c_{M,M-2} &= -\frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{\kappa(\theta - r_M) \Delta t}{\Delta r} \right] \\
c_{M,M-1} &= \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{2\kappa(\theta - r_M) \Delta t}{\Delta r} \\
c_{M,M} &= 1 + r_M \Delta t - \frac{1}{2} \left[ \frac{\sigma^2 \Delta t}{(\Delta r)^2} + \frac{3\kappa(\theta - r_M) \Delta t}{\Delta r} \right]
\end{aligned} \tag{22}$$

Again, no specific boundary condition is required at the end point. It is also worth noting that although this implicit scheme is not tridiagonal, it can easily be converted to a tridiagonal form by taking appropriate linear combinations of the equations at  $r_1$  and  $r_2$  and of the equations at  $r_{M-1}$  and  $r_M$ . As a result, the equation system can be solved very quickly. Several further features of this proposed scheme deserve emphasis:

- Equally weighted combinations of the explicit and implicit versions may be taken to form a Crank-Nicolson scheme. This has the advantage of being second order accurate in time. Consequently, much larger time steps can be taken without significant loss of accuracy.
- The scheme is very flexible with regard to grid construction. There is no requirement that the coefficients be positive for either an implicit or Crank-Nicolson version, so the end points of the grid may be selected in advance in these cases, rather than determined as a function of parameters as in the Hull and White technique. Furthermore, it is easy to adapt the method to an unevenly spaced grid. This means that more grid points can be placed for greater accuracy in areas of particular interest (such as regions where options are near-the-money). Also, grid points may be placed exactly on nodes for greater accuracy in valuing barrier types of claims.<sup>4</sup> By contrast, the Hull and White technique is nowhere near as flexible. Being an explicit method, there is limited scope for modifying the spatial grid without also changing the time step.
- The implicit and Crank-Nicolson versions of the scheme are also flexible with regard to choice of time step, so this can be varied in such a way as to match cash flow dates exactly

or to provide greater accuracy near the maturity date of an option. Again, this is hard to do in the Hull and White method without also changing the spatial grid.

- The scheme can be easily extended to accommodate more than one mean-reverting factor.
- There is no requirement (except perhaps in the explicit case) for the underlying factor to be transformed to a new state variable with constant standard deviation, or in the case of multiple factors, for the factors to be orthogonalized. Consequently, the method can be used in cases where the Hull and White scheme cannot, such as situations where the factor is not mean-reverting after the transformation and the EGARCH stochastic volatility models alluded to previously.<sup>5</sup> It should be emphasized, however, that if the model is not transformed to one with constant standard deviation, then the PDE may not be convection-dominated for large values of the factor and the use of this discretization scheme needs to be justified on other grounds. This will be discussed further below in Section 5.

To summarize, the method outlined here shares with the Hull and White technique the important feature of avoiding the specification of boundary conditions for models with mean-reversion (and constant standard deviation). However, the method proposed here has a number of additional desirable properties. Most of these derive from the ability to construct an implicit and/or Crank-Nicolson scheme, which removes many of the grid constraints required for an explicit method. It is also more general in that it can be applied in some situations where the Hull and White method cannot be.

### 3 A Numerical Example

This section provides a comparison of the performance of the Hull and White (1990a) method with explicit and Crank-Nicolson versions of the scheme proposed in this paper. Again, the single factor Vasicek (1977) model is used for illustrative purposes. The following parameter values are used:  $\kappa = 1.2$ ,  $\theta = .08$ , and  $\sigma = .05$ .<sup>6</sup> The market price of interest rate risk is assumed to be zero and the current value of  $r$  is 8%. The interest rate grid spacing  $\Delta r$  is fixed at 100 basis points in all cases. Using Hull and White's recommended value of  $\Delta t = (\Delta r)^2 / 3\sigma^2$  implies 75 time steps per year. The model is solved for the price of a pure discount bond with a face value of \$100 out to a horizon of  $T = 30$  years.

In the cases of the Hull and White and explicit methods, the requirement of keeping the probabilities positive determines the range of the interest rate grid. For the Hull and White method, the values of  $r_{\min}$  and  $r_{\max}$  turn out to be -24% and 40% respectively. For the explicit method described here, the grid runs from -12% to 28%. As noted previously, there is no requirement for the coefficients to be positive in the Crank-Nicolson scheme, so the end points of the grid can potentially be set equal to any reasonable values. For reasons of comparability, here the grid is set to be the same as for the explicit scheme. To demonstrate the advantage of being able to take much larger time steps with the Crank-Nicolson method, results are reported for two different

Model	Time steps per year	$r_{\min}$	$r_{\max}$	Execution time (in seconds)	Maximum Absolute Error	
					Discount Function	Final Grid Values
HW	75	-24%	40%	0.48	2.228e-02	1.752e-02
Explicit	75	-12%	28%	0.41	1.946e-02	1.165e-02
CN	75	-12%	28%	0.48	9.011e-06	2.828e-05
CN-q	4	-12%	28%	0.04	1.204e-03	8.766e-04

cases: 75 time steps per year and 4 time steps per year. The algorithms were coded in Fortran-77 and run on a DEC Alpha 2100 workstation.

Table 1 summarizes the results. The first four columns of the table identify the model ("HW" for Hull and White, "Explicit" for the explicit scheme, "CN" for the Crank-Nicolson scheme with 75 time steps per year, and "CN-q" for the Crank-Nicolson scheme with quarterly time steps), the number of time steps, and the range of the interest rate grid. The fifth column contains the amount of CPU seconds required. The last two columns each present measures of the maximum absolute error from comparing the numerical solution with the analytic solution provided by Vasicek (1977). The first of these, labelled "Discount Function", is computed over each time step during the 30 year period at the grid point where  $r = 8\%$ . The second, "Final Grid Values", is calculated for each value of  $r$  on the grid at  $T = 30$  years.

All of the methods are quite fast and accurate, but it is the relative comparison across models which is of most interest here. The Hull and White model produces a maximum absolute error of slightly over 2 cents (given the \$100 face value) for the discount function and a little under 2 cents for the final grid values, taking less than 1/2 second of computer time. The explicit method presented here proves to be slightly faster (not surprising given the narrower grid) and also a bit more accurate. The Crank-Nicolson method is dramatically superior. With 75 time steps per year, it is just as fast (on a narrower grid) as Hull and White's method, but much more accurate.<sup>7</sup> With quarterly time steps, it is about 10 times faster than the other methods and yet retains significantly better accuracy than either the explicit or the Hull and White method. For instance, the maximum absolute error (for either the discount function or the final grid values) is almost 20 times smaller than for the Hull and White technique. Figure 1 plots the errors in computing both the discount function and the final grid values and reinforces these conclusions, showing that they apply uniformly across the errors, not just for the largest errors.

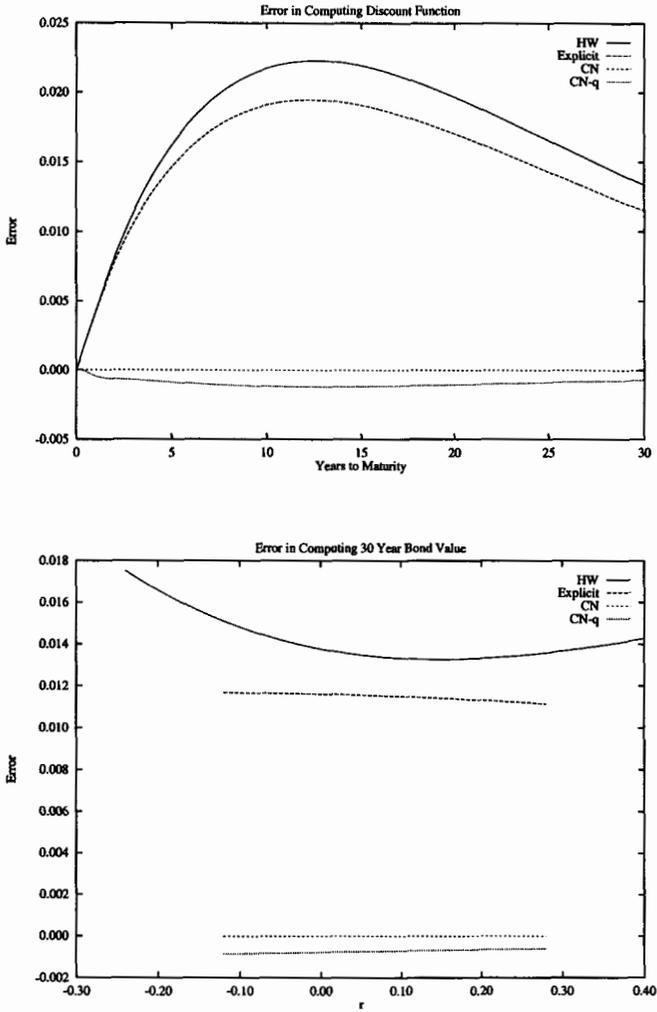


Figure 1: Errors in Computing Discount Function and Final Grid Values

## 4 Fitting The Initial Term Structure

This section describes a method for calibrating an interest rate model to a given term structure by using a time-dependent parameter. This has been done in various ways by authors such as Black, Derman, and Toy (1990), Jamshidian (1991), Black and Karasinski (1991), and Hull and White (1993, 1994a, 1994b, 1996). All of these papers are written in the context of tree-based methods and describe techniques which are inapplicable for either a Crank-Nicolson or implicit finite difference approach. Uhrig and Walter (1996) describe a technique for an implicit method, and the strategy taken here is in many ways similar to theirs. There are, however, some significant differences, and these will be described in detail below.

Before going further, it should be emphasized that any approach involving time-dependent parameters is actually only necessary in the case of American-style claims. European-style securities can be easily handled using an adjustment proposed by Dybvig (1989). As Dybvig's method will form the basis of the approach to be outlined below for American-style securities, it merits a summary here. Dybvig's main result is contained in his Theorem 1, which is reproduced below with minor notational changes:

**Theorem 1 (Dybvig)** *Let  $r^a$  be any interest rate process, let  $r^b$  be an interest rate process which is a function of time alone, and let the interest rate process  $r^c$  be defined by  $r^c = r^a + r^b$ . Let the current (time  $t$ ) price of a pure discount bond paying \$1 maturing at date  $s$  for processes  $a$  and  $b$  be denoted by  $D^a(r_t^a, t, s)$  and  $D^b(r_t^b, t, s)$ . Then the discount bond price for process  $c$  is given by  $D^c(r_t^c, t, s) = D^a(r_t^a, t, s) \times D^b(r_t^b, t, s)$ . Moreover, if interest rates follow the  $c$  process, consider a claim paying at  $s$  some function  $f[r_s^c, D^c(r_s^c, s, T_1), \dots, D^c(r_s^c, s, T_n)]$  of the spot rate and various bond prices at  $s$ . At time  $t$ , this claim has the same value as a claim paying  $D^b(r_t^b, t, s) f[r_s^a + r_s^b, D^a(r_s^a, s, T_1)D^b(r_s^b, s, T_1), \dots, D^a(r_s^a, s, T_n)D^b(r_s^b, s, T_n)]$  under process  $a$ .*

Proof: See Dybvig (1989).

To clarify some details, consider the example of a European call option on a coupon bond. In this case:

$$f[r_s^c, D^c(r_s^c, s, T_1), \dots, D^c(r_s^c, s, T_n)] \\ = \max \left[ \sum_{i=1}^{n-1} CD^c(r_s^c, s, T_i) + (C + F)D^c(r_s^c, s, T_n) - K, 0 \right]$$

where there are  $n$  coupon payments of amount  $C$  paid at times  $T_1, \dots, T_n$  after the option expiry date  $s$ , the face value of the bond is  $F$ , and the strike price of the option is  $K$ . Then:

$$D^b(r_t^b, t, s) f[r_s^a + r_s^b, D^a(r_s^a, s, T_1)D^b(r_s^b, s, T_1), \dots, D^a(r_s^a, s, T_n)D^b(r_s^b, s, T_n)] \\ = D^b(r_t^b, t, s) \\ \times \max \left[ \sum_{i=1}^{n-1} CD^a(r_s^a, s, T_i)D^b(r_s^b, s, T_i) + (C + F)D^a(r_s^a, s, T_n)D^b(r_s^b, s, T_n) - K, 0 \right]$$

which can be rewritten as:

$$D^b(r_t^b, t, s) D^b(r_s^b, s, T_n) \times \max \left[ \sum_{i=1}^{n-1} C D^a(r_s^a, s, T_i) \frac{D^b(r_s^b, s, T_i)}{D^b(r_s^b, s, T_n)} + (C + F) D^a(r_s^a, s, T_n) - \frac{K}{D^b(r_s^b, s, T_n)}, 0 \right]$$

Since the  $r^b$  process is deterministic,  $D^b(r_t^b, t, T) = D^b(r_t^b, t, s) D^b(r_s^b, s, T)$  for any  $T > s$ . Then the expression above simplifies to:

$$D^b(r_t^b, t, T_n) \times \max \left[ \sum_{i=1}^{n-1} C D^a(r_s^a, s, T_i) \frac{D^b(r_t^b, t, T_i)}{D^b(r_t^b, t, T_n)} + (C + F) D^a(r_s^a, s, T_n) - K \frac{D^b(r_t^b, t, s)}{D^b(r_t^b, t, T_n)}, 0 \right]$$

This means that the price of the option can be computed in a manner consistent with the observed initial term structure using the following simple procedure. First, calculate the discount function  $D^a(r_t^a, t, T)$  for the interest rate model under consideration. This just involves computing the prices of pure discount bonds out to the maturity of the coupon bond underlying the option. Next, determine the  $D^b(r_t^b, t, T)$  function by dividing the observed initial discount function  $D^c(r_t^c, t, T)$  by the model discount function  $D^a(r_t^a, t, T)$  for each  $T$ . Third, calculate the option price after multiplying the strike price by  $D^b(r^b, t, s)/D^b(r^b, t, T_n)$  and the coupon payments by  $D^b(r^b, t, T_i)/D^b(r^b, t, T_n)$ . Finally, multiply the computed option value by  $D^b(r^b, t, T_n)$ . This type of method provides an easy way to price European-style options under any interest rate model so that they are consistent with a given initial term structure.

American-style claims are somewhat more difficult to handle. For expositional simplicity, consider the following single factor model for the spot risk free rate:

$$dr = \kappa(\theta - r) dt + \sigma r^\beta dB \quad (23)$$

which implies this PDE for contingent claim valuation:

$$\frac{1}{2} \sigma^2 r^{2\beta} u_{rr} + [\kappa(\theta - r) - \phi(r) \sigma r^\beta] u_r + u_t - ru = 0 \quad (24)$$

where  $\phi(r)$  is the market price of interest rate risk. This is a well-known formulation, nesting the models of Vasicek (1977) when  $\beta = 0$ , Cox, Ingersoll, and Ross (1985) when  $\beta = 1/2$ , and Courtadon (1982) when  $\beta = 1$ . A common approach for pricing claims consistent with an initial yield curve is to make one of the parameters in (24) a function of time. The problem is to determine the time-dependent parameter in such a way that the model reproduces a given term structure when used to price discount bonds. In the previous literature, the parameter chosen has been either the reversion level  $\theta$  (see for example Hull and White (1996)) or some component of the market price of risk function (Uhrig and Walter (1996)). Here a different choice is made which turns out to have some significant potential advantages.

Recall from above that the basic idea underlying Dybvig's (1989) approach for European claims is to augment the interest rate model with an additional interest rate process  $r^b$  which is a function only of time. The price of a pure discount bond paying \$1 at  $T$  under the  $r^b$  process is  $\exp\left[-\int_t^T r^b(s)ds\right]$ . This is just the solution of the ordinary differential equation  $u_t - r^b(t)u = 0$  with boundary condition  $u(T) = 1$ . The form of this equation suggests the following modification of (24):

$$\frac{1}{2}\sigma^2 r^{2\beta} u_{rr} + [\kappa(\theta - r) - \phi(r)\sigma r^\beta] u_r + u_t - [r + r^b(t)] u = 0 \tag{25}$$

This formulation has two attractive properties. First, given that the  $r^b(t)$  process has been chosen so as to match the initial yield curve, European option values found by solving (25) are the same (up to numerical error) as those calculated using Dybvig's method. This is not true when  $\beta \neq 0$  if either the reversion level or the market price of risk is specified as a time-dependent parameter.<sup>8</sup> Second, it often turns out to be much quicker in the PDE context to determine the  $r^b(t)$  process than to find either a time-dependent reversion level or price of risk. The reason for this will be explained in detail below.

In general, in a Crank-Nicolson scheme with time-dependent parameters the solution at some time step  $n$  (denoted by  $u^n$ ) is moved to the preceding time step by solving a set of linear equations of the form  $A_n u^{n-1} = B_n u^n$  or  $u^{n-1} = A_n^{-1} B_n u^n$  where  $A$  and  $B$  are square matrices of size  $M$  and the subscript  $n$  is used to indicate that their elements may vary with the time-dependent parameter.<sup>9</sup> Similarly,

$$\begin{aligned} u^{n-2} &= A_{n-1}^{-1} B_{n-1} u^{n-1} \\ &= A_{n-1}^{-1} B_{n-1} A_n^{-1} B_n u^n \end{aligned} \tag{26}$$

It turns out that for models of the form of (25) the matrices on the right hand side of (26) commute under a fairly general set of circumstances. In other words,

$$u^{n-2} = A_{n-1}^{-1} B_{n-1} A_n^{-1} B_n u^n = A_n^{-1} B_n A_{n-1}^{-1} B_{n-1} u^n$$

This fact is the key as to why the method proposed here can be much faster than the Uhrig and Walter (1996) technique. The basic reason why the matrices might commute is that it is possible to find discretizations of (25) for which the terms involving  $r^b$  appear only on the diagonal of each row of each matrix. By contrast, if any coefficient of  $u_r$  (e.g. the reversion level or the market price of risk) changes over time, then off-diagonal elements in the matrices will change. More specifically, as  $r^b$  changes over time, for certain discretization schemes  $A_n$  and  $A_{n-1}$  (and  $B_n$  and  $B_{n-1}$ ) retain exactly the same structure, differing only by some constant on their diagonals, which under some conditions leads to the commuting property. This is summarized in the following proposition:

**Proposition 1** Assume that any one of  $A_n$ ,  $A_{n-1}$ ,  $B_n$ , or  $B_{n-1}$  has  $M$  linearly independent eigenvectors and that  $A_n$  and  $A_{n-1}$  and  $B_n$  and  $B_{n-1}$  differ only by a constant on their diagonals. Then

$$A_{n-1}^{-1}B_{n-1}A_n^{-1}B_n = A_n^{-1}B_nA_{n-1}^{-1}B_{n-1}$$

Proof: See Appendix.

The condition in the proposition that one of the four matrices has  $M$  linearly independent eigenvectors is sufficient, but not necessary. Also, it can be extremely difficult to verify in practice.<sup>10</sup> For these reasons, it is probably better to view the proposition as suggesting that the commuting property might hold and to simply check directly whether it does for the PDE discretization under consideration. Even if the matrices do not commute, it is still of course possible to use a standard forward induction approach such as that of Uhrig and Walter (1996).

To see why large increases in speed can be achieved if the matrices do commute, consider this illustrative example. Suppose the prices of one, two, and three period pure discount bonds are to be matched. The Uhrig and Walter (1996) forward induction procedure works as follows. First, determine the parameter at initial time zero such that the first period bond price is matched. That is, use some appropriate numerical search technique to find a parameter such that the element of  $A_0^{-1}B_0\vec{1}$  corresponding to the initial value of  $r$  generates the appropriate bond price, where  $\vec{1}$  is a column vector of 1's corresponding to the terminal payoff of the bond in each state. For the two period bond, it is necessary to find the value of the time-dependent parameter in  $A_1$  and  $B_1$  such that the appropriate element of  $A_0^{-1}B_0A_1^{-1}B_1\vec{1}$  satisfies the required condition, having stored the previously calculated value in  $A_0$  and  $B_0$ . Similarly, for the three period bond the problem is to find the parameter in  $A_2$  and  $B_2$  such that the appropriate element of  $A_0^{-1}B_0A_1^{-1}B_1A_2^{-1}B_2\vec{1}$  gives the correct price. Notice that for each time step, the model must be solved right back to the initial time, using the previously generated parameter values along the way. In this example, the model must be solved for one, two, and then three time steps, giving a total of six time steps.

However, if the matrices commute, the problem can be solved much more directly. For the two period bond,  $A_0^{-1}B_0A_1^{-1}B_1\vec{1} = A_1^{-1}B_1A_0^{-1}B_0\vec{1}$ . This means that the parameter value can be found by simply using the solution vector  $A_0^{-1}B_0\vec{1}$  from the preceding time step. This avoids having to solve the model all the way back to the initial time. Similarly, the solution vector for the three period bond is  $A_2^{-1}B_2$  times the solution vector for the two period bond, and the time-dependent parameter can again be found by using the vector of bond prices from the previous step. The three period matching exercise requires solving the model over only three time steps rather than six as is the case if the matrices do not commute. In general, for an application with  $N$  time steps, the method described by Uhrig and Walter (1996) will effectively require solving the model over  $\sum_{i=1}^N i = N(N+1)/2$  time steps rather than over  $N$  steps when the matrices commute. To get a feel for the magnitudes involved, suppose weekly bond prices are to be matched for a decade. This would require 520 time steps for commuting matrices and 135,460 time steps if the matrices do not commute.

Given this enormous potential improvement, it is natural to ask what kind of discretization schemes will produce commuting matrices. Without being able to provide a definitive answer,

it does seem that the key factor is that the discretization produces matrices which differ only by a constant on the diagonals. In every case tested where this has been true, the matrices have commuted. The issue which really must be addressed is boundary conditions, as any reasonable discretization for interior grid points will place the  $r^b(t)$  term only on the diagonal of the corresponding row in each matrix. The type of discretization scheme proposed in Section 2 works. It effectively sidesteps the issue of having to specify boundary conditions, and the criterion is satisfied because the coefficients of the  $ru$  term in the PDE show up only on the diagonals of the matrices for both  $r_{\min}$  and  $r_{\max}$ .<sup>11</sup>

What about other discretization schemes which require the specification of boundary conditions at the end points of the grid? Suppose, for example, a model of the form of equation (25) is being used with  $\beta \neq 0$ . Given mean-reversion,  $r$  will be prevented from going below zero, but there may not be an obvious boundary condition to apply at  $r_{\min} = 0$ . One possibility is to simply substitute  $r = 0$  into (25) to get  $\kappa\theta u_r + u_r - r^b(t)u = 0$ . The  $u_r$  term could be evaluated using a one-sided difference and the  $r^b(t)$  term will show up on the diagonal. For the condition at  $r_{\max}$ , a standard procedure would be to note that  $u = 0$  if  $r = +\infty$  and impose this value of  $u$  at some large  $r_{\max}$ . This effectively puts a 1 on the diagonal of the last row of the matrix and would appear to violate the requirement that the  $r^b(t)$  term appear there. However, it can be put there anyway because if  $u_M = 0$ , then  $[1 + r^b(t)]u_M = 0$ . This does suggest, though, that boundary conditions of the type where  $u$  is specified to be some non-zero value will not work, and this is indeed the case.<sup>12</sup> It should be emphasized, however, that these kinds of boundary conditions can easily be avoided for term structure models with mean-reverting factors by using the methods described in Section 2.

## 5 Numerical Results

This section provides some numerical results for bond option prices for some one and two factor models using the methods described above for discretizing the associated PDE and ensuring that the models are consistent with a given initial term structure. Six models are used. The first three are single factor models corresponding to the choices of  $\beta = 0$ ,  $\beta = 1/2$ , and  $\beta = 1$  in (23) and (25). As noted previously, these values of  $\beta$  generate respectively the Vasicek (1977), Cox, Ingersoll, and Ross (1985), and Courtadon (1982) models. The remaining models are EGARCH stochastic volatility extensions of these univariate models of the type investigated by Melino and Turnbull (1990). Such extensions have been shown to provide a much improved fit to historical short term interest rate data by Andersen and Lund (1996) and Vetzal (1997). They are given by the following bivariate specification:

$$\begin{bmatrix} dr \\ d \ln \sigma \end{bmatrix} = \begin{bmatrix} \kappa(\theta - r) \\ \alpha + \delta \ln \sigma \end{bmatrix} dt + \begin{bmatrix} \sigma r^\beta dB_t^{(1)} \\ \gamma \left( \rho dB_t^{(1)} + \sqrt{1 - \rho^2} dB_t^{(2)} \right) \end{bmatrix} \quad (27)$$

where  $\kappa$ ,  $\theta$ ,  $\alpha$ ,  $\delta$ ,  $\gamma$ , and  $\rho$  are parameters and  $B^{(1)}$  and  $B^{(2)}$  are standard Brownian motions. The PDE for contingent claims prices is:

$$\frac{1}{2} \left[ \sigma^2 r^{2\beta} u_{rr} + 2\rho\gamma\sigma r^\beta u_{rv} + \gamma^2 u_{vv} \right] + [\kappa(\theta - r)] u_r + [\alpha + \delta v] u_v + u_t - [r + r^b(t)] u = 0 \quad (28)$$

where  $v \equiv \ln \sigma$  and the market prices of risk have been set to zero both for simplicity and to facilitate comparisons across the models.<sup>13</sup> Several earlier papers have provided similar comparisons, particularly for the single factor models. Examples include the papers of Buser, Hendershott, and Sanders (1990), Hull and White (1993), Uhrig and Walter (1996), and Vetzal (1997). The general conclusion of these studies is that the models tend to produce quite similar values for European bond options, except possibly for deep out-of-the-money options, provided that they are constrained to have similar levels of volatility and that they are consistent with the initial term structure. An issue to be investigated in this section is the extent to which similar conclusions apply to American-style options.

Before proceeding further, it is necessary at this stage to discuss some technical issues regarding boundary conditions which arise when  $\beta \neq 0$ . The justification for the discretization scheme presented in Section 2 relies on the PDE being convection-dominated for extremely high or low values of  $r$ , but this may not be the case if  $\beta \neq 0$  for large positive values of  $r$ .<sup>14</sup> As a result, it is technically necessary to specify some boundary condition when  $r$  is large. Note that the  $-ru$  term in the PDE is a source of exponential decay, driving the solution to zero for large positive values of  $r$ . This also causes all derivatives with respect to  $r$  to approach zero, so the boundary condition requirement can be satisfied by imposing that either  $u$  itself or some derivative of  $u$  with respect to  $r$  equals zero. An obvious candidate is  $u_r = 0$ . This condition arises naturally in the numerical solution when the discretization strategy outlined in Section 2 is used due to the dominance of the exponential decay term.<sup>15</sup> This justifies the use of the scheme, even if the PDE is not convection-dominated for large positive values of  $r$ .<sup>16</sup>

The parameter values to be used are presented in Table 2.<sup>17</sup> The omitted values of  $\kappa$  are used to control for the level of bond price volatility. Specifically, the value of  $\kappa = 1.5$  for the single factor  $\beta = 1/2$  case is used for that model. The values of  $\kappa$  for the other models are found by using a standard numerical search procedure to match the price of a European option for a particular strike price to that produced by the one factor  $\beta = 1/2$  model. Note that  $\kappa$  is not specified as a time-dependent parameter, but rather is set to be a different constant depending on the option under consideration.<sup>18</sup> Specifying  $\kappa$  to be a function of time has some undesirable consequences. From a strictly numerical point of view, the problem is that it destroys the matrix commuting property. More fundamentally, making volatility time-dependent can cause significant pricing errors when the model is applied to claims which depend on the future level of volatility (e.g. American-style options).<sup>19</sup>

Before presenting option pricing results, it is useful to provide some indication of the computing time required for the methods used. For European-style options, one strategy is to simply use the adjustment proposed by Dybvig (1989). Another possibility (which can also be used for American-style options) is to compute the values of the time-dependent parameter  $r^b(t)$  in the

Table 2			
Parameter Values			
Single Factor Models			
$\beta$	$\kappa$	$\theta$	$\sigma$
0	—	.08	.05
1/2	1.5	.08	.15
1	—	.08	.50

Two Factor Models						
$\beta$	$\kappa$	$\theta$	$\alpha$	$\delta$	$\gamma$	$\rho$
0	—	.08	-10.5	-1.8	1.5	0.2
1/2	—	.08	-9.8	-2.2	1.5	-0.2
1	—	.08	-9.5	-3.0	1.5	-0.5

PDE and then solve the model using these values. Table 3 provides some results. In each case the programs were coded in Fortran-77 and run on a DEC Alpha 2100 computer. The sample problem involved pricing a five year option on a ten year coupon bond using weekly time steps. The Crank-Nicolson method was used for the discretization strategy proposed in Section 2. The grids used contained 31 points for  $r$ , and in the two factor case, 21 points for  $v$ . The outstanding feature of the table is the difference in the time required for computing the time-dependent parameter using the method of forward induction as opposed to the commuting matrices method proposed in this paper. It should be emphasized that both methods produce the same sequence of parameter values. For the single factor model, the commuting matrices technique takes about 1/3 of a second but the forward induction method takes around a minute. Even more dramatically, in the two factor case it takes about a minute and a half using the first method but about seven-and-a-half hours using the second method. The other feature of the table is that Dybvig's method is somewhat faster at producing option values than solving the model using the time-dependent parameter sequence. This is not surprising as the latter technique requires changing the matrices involved at each time step whereas the former does not (though it does require solving the model twice, first to generate the required adjustments for term structure consistency and second to compute the option value). As a final comment, it should be pointed out that there are a couple of reasons why the execution times for the two factor model are much slower than for the one factor case. The most fundamental reason is the well-known fact that the computational effort required to solve a PDE is exponential in the number of dimensions. Second, the sparse matrix software used here is designed for matrices of a general sparsity pattern and does not exploit the

Table 3		
Sample Execution Timing Results		
5 Year Option on 10 Year Coupon Bond, Weekly Time Steps		
CPU Seconds on DEC Alpha 2100		
Task	One Factor Model (31 grid points)	Two Factor Model (31 × 21 grid points)
Compute European option value using Dybvig's method	0.07	30.64
Generate time-dependent parameter sequence (commuting matrices)	0.35	84.38
Generate time-dependent parameter sequence (forward induction)	57.96	27, 183.43
Compute option value using time-dependent parameter sequence	0.13	33.48

particular structure of the matrices to speed up computation. By contrast, the one factor model produces tridiagonal matrices which can be solved very quickly.<sup>20</sup>

Table 4 presents some sample results for European call options. Two cases are considered: a one year option on a two year bond and a five year option on a ten year bond. The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The initial term structure is flat at 8%. The initial level of  $v$  is set to its reversion level of  $-\alpha/\delta$  for each of the stochastic volatility models. The Crank-Nicolson method was used on evenly spaced grids with weekly time steps in every case. The interest rate reversion speed parameter  $\kappa$  was adjusted across the models so that they all produced the same option value for a strike price of \$100. Analytic solutions are available for the single factor models where  $\beta \neq 1$  and are provided in parentheses.<sup>21</sup> These suggest that the pricing errors for the numerical method are fairly small in those particular cases. Broadly speaking, the results correspond to those reported in the previous literature cited above. There is not much difference in the option prices across the various models, except (in percentage terms) for deep out-of-the-money options. The stochastic volatility models produce somewhat higher values in these cases. Option prices generally decline with  $\beta$ , though some exceptions to this are apparent for options which are deeper in-the-money (more so for the longer maturity case).

Option	Strike	One Factor			Two Factor		
		$\beta = 0$	$\beta = 1/2$	$\beta = 1$	$\beta = 0$	$\beta = 1/2$	$\beta = 1$
1 year option on 2 year bond	98	1.7307 (1.7451)	1.7553 (1.7660)	1.7830	1.7876	1.7654	1.7444
	99	0.9349 (0.9447)	0.9637 (0.9706)	0.9893	0.9771	0.9552	0.9428
	100	0.3751 (0.3778)	0.3751 (0.3774)	0.3751	0.3751	0.3751	0.3751
	101	0.1005 (0.1010)	0.0757 (0.0772)	0.0559	0.1173	0.1169	0.1076
	102	0.0161 (0.0168)	0.0028 (0.0043)	0.0004	0.0454	0.0369	0.0238
5 year option on 10 year bond	98	0.9946 (0.9870)	1.0304 (1.0259)	1.0710	1.0407	1.0187	1.0097
	99	0.5177 (0.5103)	0.5492 (0.5408)	0.5768	0.5337	0.5245	0.5256
	100	0.2103 (0.2069)	0.2103 (0.2102)	0.2103	0.2103	0.2103	0.2103
	101	0.0616 (0.0628)	0.0511 (0.0491)	0.0419	0.0910	0.0833	0.0692
	102	0.0133 (0.0138)	0.0031 (0.0046)	0.0002	0.0432	0.0301	0.0190

The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The term structure is flat at 8%. For the two factor models, the initial value of  $v = -\alpha/\delta$ . Available analytic solutions are in parentheses.

To approximately gauge the effect of the initial level of  $r$ , Figure 2 plots the option values for the single factor  $\beta = 1/2$  model for the two cases with a strike price of \$100 across a range of values of  $r$ . Of course, these numbers are not economically correct in that they have all been adjusted so as to be consistent with a particular initial yield curve at  $r = 8\%$ . However, it is still interesting to observe how the large range in values for the short maturity case (from 2 cents to 87 cents) is compressed to only about 4 cents in the longer maturity case. This pattern is exactly what one would expect given mean-reversion — as maturity increases, the initial value becomes much less important. Nonetheless, it is perhaps surprising how quickly this effect occurs.

The two panels of Figure 3 illustrate the same type of phenomenon in the two factor case. The spread of values for the short maturity case is about \$1.65, dropping to around 5 cents for the other option. Interestingly, the effect of the initial level of  $v$  is almost non-existent in this latter

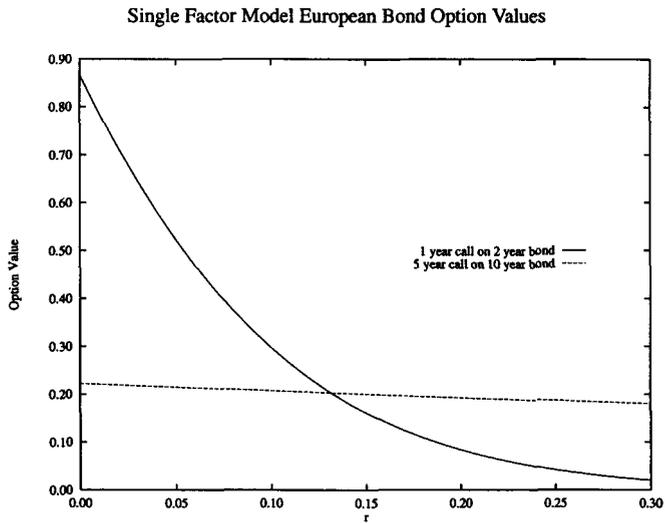
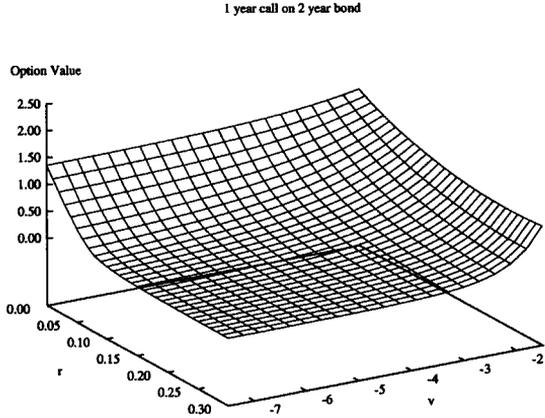


Figure 2: Single factor model ( $\beta = 1/2$ ) European bond option values. The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The strike price is \$100. Option values are computed so that they are consistent with an initial term structure which is flat at 8%.

## Two Factor Model European Bond Option Values



**Figure 3:** Two factor model ( $\beta = 1/2$ ) European bond option values. The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The strike price is \$100. Option values are computed so that they are consistent with an initial term structure which is flat at 8%.

Option	Strike	One Factor			Two Factor		
		$\beta = 0$	$\beta = 1/2$	$\beta = 1$	$\beta = 0$	$\beta = 1/2$	$\beta = 1$
1 year option on 2 year bond	100	3.5528	3.5535	3.5621	3.5764	3.5663	3.5560
	101	2.6319	2.6394	2.6570	2.6725	2.6563	2.6409
	102	1.7319	1.7553	1.7830	1.7905	1.7678	1.7451
	103	0.9359	0.9637	0.9894	0.9799	0.9574	0.9435
	104	0.3759	0.3752	0.3751	0.3776	0.3770	0.3758
	105	0.1008	0.0757	0.0559	0.1193	0.1182	0.1081
106	0.0162	0.0028	0.0004	0.0469	0.0376	0.0239	
5 year option on 10 year bond	100	3.5153	3.5165	3.5359	3.4440	3.4627	3.4691
	101	2.7135	2.7107	2.7247	2.6386	2.6577	2.6576
	102	1.9330	1.9302	1.9406	1.8787	1.8998	1.8959
	103	1.2107	1.1884	1.1897	1.1889	1.2082	1.2044
	104	0.6120	0.5706	0.5361	0.6482	0.6539	0.6246
	105	0.2200	0.1538	0.1057	0.3432	0.3078	0.2538
106	0.0510	0.0106	0.0030	0.1862	0.1225	0.0729	

The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The term structure is flat at 8%. For the two factor models, the initial value of  $v = -\alpha/\delta$ .

case. Clearly, slower rates of mean-reversion in  $v$  would imply greater effects, but relatively fast adjustment speeds in  $v$  seem to be characteristic of historical data.<sup>22</sup> This suggests that the effects of the initial level of  $v$  on long term European option values are likely to be quite small, unless the market price of volatility risk is such that the reversion rate of volatility is much slower under the equivalent martingale measure.

Table 5 presents results for American call options.<sup>23</sup> Note that the values of  $\kappa$  for the various models are kept the same as for the corresponding European option case — the models are still forced to produce the same price for a European call with a strike price of \$100. Comparing prices with those in Table 4, it can be seen that the early exercise premia are huge, but the same general pattern prevails.<sup>24</sup> The models all produce quite similar values for in-the-money options, but there are now somewhat larger discrepancies for out-of-the-money options. This is particularly true for the longer maturity options. The effects of  $\beta$  are most apparent for out-of-the-money options (the values of which decline with  $\beta$ ), again especially for the longer maturity case. The stochastic volatility models typically produce higher values than their single factor counterparts for options which are well out-of-the-money, but lower values for options which are relatively deeper into-the-money. There are, however, exceptions to this pattern for the shorter maturity cases when  $\beta \neq 1$ .

Table 6 repeats the exercise for American put options. There are a couple of noteworthy

Option	Strike	One Factor			Two Factor		
		$\beta = 0$	$\beta = 1/2$	$\beta = 1$	$\beta = 0$	$\beta = 1/2$	$\beta = 1$
1 year option	95	0.0094	0.0205	0.0384	0.0410	0.0267	0.0129
	96	0.1938	0.1993	0.2122	0.1465	0.1420	0.1412
	97	1.0647	1.0641	1.0638	1.0622	1.0629	1.0633
on 2 year bond	98	2.0602	2.0602	2.0601	2.0601	2.0602	2.0602
	92	0.0140	0.0710	0.1854	0.2064	0.1333	0.0600
5 year option	93	0.0956	0.2004	0.3457	0.3307	0.2413	0.1479
	94	0.3550	0.4870	0.6265	0.5457	0.4523	0.3638
on 10 year bond	95	0.8698	0.9655	1.0678	0.9270	0.8694	0.8131
	96	1.5800	1.6279	1.6881	1.5423	1.5191	1.4800
	97	2.3900	2.4101	2.4431	2.3152	2.3102	2.2773
	98	3.2604	3.2671	3.2882	3.1515	3.1549	3.1272
	99	4.1516	4.1481	4.1619	4.0179	4.0262	4.0022
	100	5.0602	5.0454	5.0519	4.9150	4.9211	4.9070
	101	5.9763	5.9537	5.9531	5.8813	5.8867	5.8860
	102	6.9311	6.9171	6.9134	6.8723	6.8773	6.8779

The underlying bond has a face value of \$100 and pays an 8% semi-annual coupon. The term structure is flat at 8%. For the two factor models, the initial value of  $v = -\alpha/\delta$ .

differences here. First, for out-of-the-money options, the values for the one factor models are increasing with  $\beta$ , whereas the prices for the two factor models are declining with  $\beta$ . Second, there is a very small range of values over which the models generate any noticeable differences in the shorter maturity case, but quite a large range in the longer maturity case. For all strike prices above \$98, the models all produce essentially identical prices for the one year option on the two year bond. By contrast, in the case of the five year option on the ten year bond the price differences between the corresponding one and two factor models is around 15 cents for a strike price of \$100, and although it is not shown in the table, there are still differences of a cent or two between the models even at a strike price of \$106. The two factor models generate higher values than their single factor counterparts for deep out-of-the-money options and lower values for relatively in-the-money options, except when  $\beta = 1$  in which case lower values are produced over the entire range of strike prices shown in the table.

To briefly summarize, while the results found here are more suggestive than conclusive, they are in line with the previous literature for European-style options. Constraining different models to match a given initial term structure and to be similar in terms of bond price volatility effectively forces them to generate quite similar option values, with the most significant differences occurring for deep out-of-the-money options. The results here show that similar conclusions

apply in the case of American call options, though the effects are somewhat larger. American put options also fit this general pattern, with even bigger price differences being observed for the longer maturity cases. The impact of stochastic volatility is quite complicated and depends on  $\beta$ . Generally speaking, the stochastic volatility models produce higher values than the corresponding single factor models for out-of-the-money options and lower values for options which are relatively deeper into-the-money. Among the exceptions to this, however, are the  $\beta = 0$  case for European calls where the stochastic volatility model generates higher prices for both in and out-of-the-money options. Also, the  $\beta = 1$  case for American puts produces lower values across the board.

## 6 Summary

The first contribution in this paper is the two point upstream discretization strategy for mean-reverting models. This improves on the well-known Hull and White (1990a) trinomial lattice in several ways. By permitting the construction of an implicit (and hence also a Crank-Nicolson) scheme, superior stability, grid spacing flexibility, and time step flexibility are attained. Unlike the Hull and White method, it does not require the use of any transformations. It is therefore more general because it can be applied in situations where the Hull and White technique cannot be. The general applicability of the method should be emphasized — it can be used for any model with mean-reverting factors provided that the PDE either becomes convection-dominated for extreme values of the factors or that exponential decay drives the solution to zero as the factors get large. In addition to the specifications dealt with in this paper, examples would include all of the general factor models described in Section 2 of Duffie and Kan (1996), such as the models of Longstaff and Schwartz (1992) and Chen and Scott (1992).

The second contribution is a demonstration of how to find European option values consistent with an initial term structure along the lines of Dybvig (1989), but in a PDE context which also permits the valuation of American-style claims in the same setting. Again, the generality of the method is a significant strength. It can be applied to any Markovian interest rate model. Most importantly, the two point upstream discretization scheme results in the commuting matrix property, which provides huge efficiency gains in computer time.

Finally, as a demonstration of the methods some numerical results are provided for three common single factor models and two factor stochastic volatility analogues of them. These results reproduce the general pattern observed in the previous literature for European options, but also show somewhat larger differences across models for American options, especially for longer maturity cases.

Future related research topics might include examining possible efficiency gains from using unevenly spaced grids and a exploration of time step selection schemes to take advantage of the flexibility offered in that area. The generality of the methods described here may also prove useful in empirical tests of alternative term structure models.

## Notes

<sup>1</sup>Willmott, Dewynne, and Howison (1993, section 22.4.1) use a two point upstream method in a PDE for Asian options. Zvan, Forsyth, and Vetzal (1996) explore the use of a more complicated variation known as a flux limiter, again in the context of Asian options.

<sup>2</sup>In other words, prices are computed by working backwards from final payoffs, discounting expected future values under the risk-neutral measure.

<sup>3</sup>The scheme is presented in a different form from that originally provided by Hull and White in order to facilitate comparison with the method to be proposed here.

<sup>4</sup>The importance of doing this has been emphasized in the context of equity options by Cheuk and Vorst (1996), among others.

<sup>5</sup>Some transformations may still be desirable on grounds of computational efficiency, but they are not required in order for the method to work.

<sup>6</sup>These values are approximately equal to (annualized) parameter estimates based on historical U.S. Treasury bill data reported by Vetzal (1997).

<sup>7</sup>There is a negligible upward bias in the computing time for both the Hull and White method and the explicit method in that bond prices are solved for over the entire interest rate grid for the full time horizon rather than back only to the initial value of  $r$  in the cone-shaped region generated by the trinomial lattice for the first few steps in the tree. For the Hull and White method, using the full grid over the entire time horizon requires evaluating the bond price at 65 interest rate grid points  $\times$  2,250 time steps or 146,250 points. In a trinomial lattice, there are a total of  $\sum_{i=0}^N (2i + 1) = (N + 1)^2$  points after  $N$  time steps. Here it takes 32 time steps to reach both  $r_{\min}$  and  $r_{\max}$ , so the total number of points required would be  $33^2 + (2,250 - 32) \times 65 = 145,259$  points. By comparison, the explicit method used here requires 2,250 time steps  $\times$  41 interest rate grid values or 92,250 points and is only about .07 seconds faster than the Hull and White method with 146,250 points.

<sup>8</sup>The reason is that making  $\theta$  or  $\phi$  time-dependent involves changing the mean level of  $r$  under the equivalent martingale measure, which in turn affects the variance in the model since it is a function of  $r$  if  $\beta \neq 0$ .

<sup>9</sup>The discussion here is in the context of a Crank-Nicolson scheme, but it applies to either a fully explicit method if  $A_n$  is an identity matrix for all  $n$  or to a fully implicit method if  $B_n$  is an identity matrix for all  $n$ .

<sup>10</sup>It is worth noting that this property is occasionally used in matrix stability analysis. For example, Geske and Shastri (1985, p. 56) state that it holds "in general" and proceed to derive some standard stability results.

<sup>11</sup>This remains true for the cases of unevenly spaced grids and extensions to two factor mean-reverting

models using the same type of discretization.

<sup>12</sup>One example could be the value at  $r_{\min}$  in the Vasicek (1977) model. Since there is nothing to prevent  $r$  from going negative in the model, one possible boundary condition is  $u = +\infty$  if  $r = -\infty$ , which could be imposed by setting  $u$  equal to some large positive value for some large negative value of  $r$ .

<sup>13</sup>Similarly, when solving (25) the market price of interest rate risk  $\phi(r)$  will be set to zero throughout this section.

<sup>14</sup>When  $\beta \neq 0$ , the lower boundary for  $r$  is zero. At this point the diffusion coefficient  $\sigma^2 r^{2\beta} / 2$  will vanish, and so the PDE is convection-dominated at and near the lower boundary.

<sup>15</sup>This is true for any value of  $\beta$ , including 0. In fact, this provides another perspective on the discretization scheme in that case. Instead of viewing the scheme as avoiding the specification of boundary conditions, it can alternately be seen as imposing the natural boundary condition  $u_{,r} = 0$  when  $r$  gets large. Of course, there is nothing unique about this discretization scheme in this regard. Since the boundary condition develops naturally in the numerical solution, any reasonable scheme will satisfy it.

<sup>16</sup>Note that for the stochastic volatility specifications  $v$  is mean-reverting with constant standard deviation, so the PDE is convection-dominated with respect to  $v$  for extremely high or low values of  $v$ .

<sup>17</sup>These are approximately equal to annualized values of the parameter estimates reported in Vetzal (1997).

<sup>18</sup>A more satisfactory approach in actual empirical applications would be to choose  $\kappa$  to minimize model pricing errors over a sample of options.

<sup>19</sup>For a good discussion of this issue, see Hull and White (1996), pp. 32–34.

<sup>20</sup>As a point of comparison, using the general purpose sparse matrix routines on the one factor model approximately triples the required computing time.

<sup>21</sup>These are based on the formulas for discount bond options provided by Jamshidian (1989) for the model with  $\beta = 0$  and Cox, Ingersoll, and Ross (1985) for the model with  $\beta = 1$ , converted to coupon bond option formula using the well-known method of Jamshidian (1989), and adjusted using Dybvig's (1989) technique to ensure consistency with the initial term structure.

<sup>22</sup>The estimate of  $\delta$  used here for the  $\beta = 1/2$  case implies that the half-life of a deviation from the mean level of  $v$  is about 0.3 years. This is based on a data sample running from 1980 through 1990. Using a much longer sample (1954–1995), Andersen and Lund (1996) report an estimate of close to 0.65 years, still not very long relative to the maturities of many interest rate derivative securities.

<sup>23</sup>Note that strike prices must be adjusted to reflect accrued interest when evaluating the early exercise decision. For particular details, see for example Longstaff (1993).

<sup>24</sup>Longstaff (1993) reports early exercise premia of similar magnitude.

## Appendix - Proof of Proposition 1

By Theorem 6.4.6 of Ortega (1987), any two  $M \times M$  matrices with  $M$  linearly independent common eigenvectors commute under multiplication. By assumption, one of  $A_{n-1}$ ,  $B_{n-1}$ ,  $A_n$ , or  $B_n$  has  $M$  linearly independent eigenvectors, so proving the proposition requires only showing that  $A_{n-1}^{-1}$ ,  $B_{n-1}$ ,  $A_n^{-1}$ , and  $B_n$  all have the same eigenvectors.

Since  $B_n$  and  $B_{n-1}$  only differ by some constant  $\gamma$  on the diagonal,  $B_{n-1} = B_n + \gamma I$  where  $I$  is an  $M \times M$  identity matrix. Any vector  $x$  which satisfies  $[\lambda I - B_n]x = 0$  also satisfies  $[(\lambda - \gamma)I - B_n]x = 0$ , which establishes that  $B_n$  and  $B_{n-1}$  have eigenvalues which differ by  $\gamma$  but a common set of eigenvectors.  $A_n$  and  $A_{n-1}$  also differ only by a constant on the diagonal and therefore also share a common set of eigenvectors. Since the eigenvectors of a matrix are the same as for its inverse,  $A_n^{-1}$  and  $A_{n-1}^{-1}$  have the same set of eigenvectors.

The last step is to show that the eigenvectors of  $A_n$  and  $B_n$  are the same. For a Crank-Nicolson scheme, it is always possible to write  $A_n = I - C_n$  and  $B_n = I + C_n$  for some matrix  $C_n$ . The eigenvectors  $x$  of  $A_n$  satisfy:

$$[\lambda I - (I - C_n)]x = 0$$

Multiplying by  $-1$  gives:

$$[(1 - \lambda)I - C_n]x = 0$$

Substitution of  $\lambda = 2 - \delta$  yields:

$$[(\delta - 1)I - C_n]x = 0$$

or

$$[\delta I - (I + C_n)]x = 0$$

which establishes that  $A_n$  and  $B_n$  have the same eigenvectors and completes the proof.

## References

- Andersen, T.G. and J. Lund (1996). Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics*, forthcoming.
- 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* (January-February), 33–39.
- Black, F., and P. Karasinski (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal* (July-August), 52–59.
- Buser, S.A., P.H. Hendershott, and A.B. Sanders (1990). Determinants of the value of call options on default-free bonds. *Journal of Business* 63 (Supplement), S33–S50.
- Chen, R.-R. and L.O. Scott (1992). Pricing interest rate options in a two-factor Cox-Ingersoll-Ross model of the term structure. *The Review of Financial Studies* 5, 613–636.
- Cheuk, T.H.F. and T.C.F. Vorst (1996). Complex barrier options. *Journal of Derivatives* 4 (Fall), 8–22.
- Courtadon, G. (1982). The pricing of options on default-free bonds. *Journal of Financial and Quantitative Analysis* 17, 75–100.
- Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* 53, 385–407.
- Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance* 6, 379–406.
- Dybvig, P.H. (1989). Bond and bond option pricing based on the current term structure. Working paper, Washington University in St. Louis.
- Geske, R. and K. Shastri (1985). Valuation by approximation: A comparison of alternative option valuation techniques. *Journal of Financial and Quantitative Analysis* 20, 45–71.
- Heath, D., R.A. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica* 60, 77–105.
- Hull, J. and A. White (1990a). Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis* 25, 87–100.
- Hull, J. and A. White (1990b). Pricing interest rate derivative securities. *The Review of Financial Studies* 3, 573–592.
- Hull, J. and A. White (1993). One-factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis* 28, 235–254.
- Hull, J. and A. White (1994a). Numerical procedures for implementing term structure models I: Single-factor models. *Journal of Derivatives* 2 (Fall), 7–16.
- Hull, J. and A. White (1994b). Numerical procedures for implementing term structure models II: Two-factor models. *Journal of Derivatives* 2 (Winter), 37–48.

- Hull, J. and A. White (1996). Using Hull-White interest rate trees. *Journal of Derivatives* 3 (Spring), 26–36.
- Jamshidian, F. (1989). An exact bond option formula. *Journal of Finance* 44, 205–209.
- Jamshidian, F. (1991). Forward induction and construction of yield curve diffusion models. *Journal of Fixed Income* 1 (June), 62–74.
- Longstaff, F.A. (1993). The valuation of options on coupon bonds. *Journal of Banking and Finance* 17, 27–42.
- Longstaff, F.A. and E.S. Schwartz (1992). Interest rate volatility and the term structure: A two factor general equilibrium model. *Journal of Finance* 47, 1259–1282.
- Melino, A. and S.M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* 45, 239–265.
- Ortega, J.M. (1987). *Matrix Theory*. Plenum Press, New York.
- Tian, Y. (1994). A re-examination of lattice procedures for interest-rate contingent claims. *Advances in Futures and Options Research* 7, 87–111.
- Uhrig, M. and U. Walter (1996). A new numerical approach for fitting the initial yield curve. *Journal of Fixed Income* 5 (March), 82–90.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177–188.
- Vetzal, K.R. (1997). Stochastic volatility, movements in short term interest rates, and bond option values. *Journal of Banking and Finance*, forthcoming.
- Willmott, P., J. Dewynne, and S. Howison (1993). *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, Oxford, UK.
- Zvan, R., P.A. Forsyth, and K.R. Vetzal (1996). Robust numerical methods for PDE models of Asian options. Technical Report CS-96-28, Department of Computer Science, University of Waterloo.

