

# Lattice methods for no-arbitrage pricing of interest rate securities

Toby Daghish\*

May 20, 2010

## Abstract

We explore calibration of single factor no-arbitrage short rate models to yield and volatility information. We note that the calculation of Arrow-Debreu prices for interest rate securities is analogous to solving the Kolmogorov Forward Equation. This insight allows us to implement implicit methods, which exhibit more rapid convergence than explicit methods. We develop an algorithm for calibrating a model to match both yield and volatility curves, which is general across single factor short rate models, and also across finite difference techniques. Numerical examples confirm that our approach vastly improves computation times for derivative pricing.

The use of short-rate no-arbitrage models for pricing interest rate securities has a strong popularity amongst academics and practitioners. Being able to calibrate a model to zero coupon bond prices

---

\*Senior Lecturer, Victoria University of Wellington. PO Box 600, Wellington, New Zealand 6140. E-mail: toby.daghish@vuw.ac.nz. Phone: 64-4-4635451. The author thanks Graeme Guthrie, Alan White, Mairead de Roiste, Bethanna Jackson, Vladimir Petkov and Leigh Roberts for their comments, as well as participants at the New Zealand Finance Colloquium 2010.

ensures that pricing of derivative securities is consistent with observed interest rates, rather than being based on some best fit, as would be the case for equilibrium interest rate models.

A number of such no-arbitrage models exist. The best known of these models are Ho and Lee [1986], Black, Derman, and Toy [1990], Hull and White [1990] and Black and Karasinski [1991]. While closed form solutions are available in some cases, for most implementations, the models are set up as binomial or trinomial trees, which are calibrated in order to match information on the yield curve. Hull and White [1993] show that their trinomial tree-building technique can in fact be applied to the other models in this family, and so presents a general numerical technique for no-arbitrage bond pricing.

Uhrig and Walter [1996] and Vetzal [1998] make a case that using implicit methods could improve upon the performance of trinomial models. Uhrig and Walter [1996] calibrate a time varying market price of risk to match the yield curve. Their approach entails full valuation of bonds of progressively longer maturities leading to excessive computational costs as the number of time steps increases. Vetzal [1998] avoids this shortfall by instead building in the current yield curve as a deterministic addition to his stochastic interest rate process. While this is tractable, it is unclear how Vetzal's model could be efficiently calibrated to match volatility term structures, unlike the Hull-White approach.

Our contribution is to provide an extension of the Hull and White [1993] approach which allows the use of implicit methods. We demonstrate how the Kolmogorov Forward Equation can be solved to obtain Arrow-Debreu prices for an interest rate model. We then demonstrate how we can use this result to derive an algorithm, similar to that used in binomial and trinomial models, to choose drift terms to match the existing term structure through forward induction. Next we extend our methodology to having two parameters in the drift term (a mean reversion rate as well as a level to which interest rates mean revert) and show how we can calibrate these to a term structure of

volatilities. Our resulting algorithm is considerably faster than existing methodologies (having second order convergence in time rather than first order, while still requiring only first order computational effort). This is important for many fixed income pricing problems where securities can have lives of up to thirty years, and therefore computational burdens are sizable. Further, if a user wishes to calibrate their model to option prices (as discussed in Hull and White [2001]) many evaluations of option prices may be necessary as part of the parameter selection process.

The layout of the remainder of the paper is as follows: section 1 describes the Arrow-Debreu pricing approach, describes the Hull-White approach to tree-building, and presents our more general lattice building technique. Section 2 describes calibration of one parameter to match the yield curve. Section 3 generalises this to the two parameter case. Section 4 shows some numerical examples of our method, compared to trinomial trees and a backward equation methods. Lastly section 5 concludes.

## 1 Finding Arrow-Debreu prices

An Arrow-Debreu security is conventionally defined as a security which pays off \$1 in a particular state at a particular time. In continuous time, this can be replaced by a state price density, specifying a continuous function, which can be used as an integrand with security payoffs to determine a security's price. This density incorporates the risk-neutral probabilities of reaching a particular point, but also the discounting which is necessary to reach this point. This state price density is sometimes also referred to as a Green's function since this is the term given to a kernel which can be used to find solutions to a differential equation for arbitrary boundary conditions.

Suppose that we have a short rate process under the risk-neutral probability measure given by:

$$dr = \mu(r, t)dt + \sigma(r, t)dW \quad (1)$$

where  $r$  is the short rate and  $\mu$  and  $\sigma$  are arbitrary functions.  $t$  measures time and  $dW$  is the innovation of a standard Brownian motion. It can be shown that the price of a contingent claim ( $f$ ) will solve the partial differential equation:

$$\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial r} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial r^2} = rf \quad (2)$$

Pricing of a particular claim is then generally a case of solving (2) subject to appropriate boundary conditions. Equation (2) is an example of a Backward Kolmogorov Equation, coupled with discounting (the term on the right hand side). The Backward Kolmogorov Equation allows us to evaluate expectations.

Clearly one way in which the state price density could be formed would be to solve (2) for a range of different boundary conditions. However, if our objective is to simply find all state price densities, an easier approach is to use the Kolmogorov *Forward* Equation, as described in Jamshidian [1991], to find the state price density, working forward from time 0. The Kolmogorov Forward Equation (including discounting) for the state price density ( $Q$ ), assuming that the short rate follows the process (1), is given by:

$$\frac{\partial Q}{\partial t} + \left( \mu - \frac{\partial \sigma^2}{\partial r} \right) \frac{\partial Q}{\partial r} + \left( \frac{\partial \mu}{\partial r} - \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2} \right) Q - \frac{1}{2} \sigma^2 \frac{\partial^2 Q}{\partial r^2} = -rQ \quad (3)$$

Note that this equation differs slightly from (2) in that the diffusion term now has a negative co-

efficient, and that an additional  $Q$  term appears on the left hand side. Given an initial state price density ( $Q(r, 0)$ ), we could solve (3) forward to find the state price density at any future date.

## 1.1 Trinomial approach

Hull and White [1993] tackle the evaluation of Arrow-Debreu prices by building an explicit approximation of (2). For brevity, we focus on interior points of the tree. For these points, the transition probabilities take the form:

$$\begin{aligned} p_1(r, t) &= \frac{\sigma^2 \Delta t}{2\Delta r^2} + \frac{\mu \Delta t^2}{2\Delta r^2} + \frac{\mu \Delta t}{2\Delta r} \\ p_2(r, t) &= 1 - \frac{\sigma^2 \Delta t}{\Delta r^2} - \frac{\mu \Delta t^2}{\Delta r^2} \\ p_3(r, t) &= \frac{\sigma^2 \Delta t}{2\Delta r^2} + \frac{\mu \Delta t^2}{2\Delta r^2} - \frac{\mu \Delta t}{2\Delta r} \end{aligned}$$

where  $p_1$  is the probability of moving from  $r$  (at time  $t$ ) to  $r + \Delta r$  (at time  $t + \Delta t$ ),  $p_2$  is the probability of  $r$  not changing and  $p_3$  is the probability of  $r$  moving to  $r - \Delta r$ .  $\Delta t$  is the timestep of the tree. Applying a Taylor expansion to  $f(r, t)$  we can show that:

$$\begin{aligned} f(r, t - \Delta t) &= e^{-r\Delta t} [p_1(r, t - \Delta t)f(r + \Delta r, t) + p_2(r, t - \Delta t)f(r, t) + p_3(r, t - \Delta t)f(r - \Delta r, t)] \\ \Rightarrow 0 &= \Delta t \left[ rf - \mu \frac{\partial f}{\partial r} - \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial r^2} - \frac{\partial f}{\partial t} \right] + O(\Delta r^2 \Delta t, \Delta t^2) \end{aligned}$$

so that the relationship implies:

$$\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial r} + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial r^2} = rf + O(\Delta t, \Delta r^2)$$

and the trinomial tree approximates (2) with second order accuracy in  $\Delta r$  and first order accuracy in  $\Delta t$ .

Denoting Arrow-Debreu prices for the tree as  $Q(r, t)$ , the Hull-White algorithm initialises  $Q(r_0, 0) = 1$  (where  $r_0$  is the initial short rate) and then calculates the subsequent nodes in the tree as:

$$\begin{aligned} Q(r, t + \Delta t) &= p_3(r + \Delta r, t)Q(r + \Delta r, t)e^{-(r+\Delta r)\Delta t} + p_2(r, t)Q(r, t)e^{-r\Delta t} \\ &\quad + p_1(r - \Delta r, t)Q(r - \Delta r, t)e^{-(r-\Delta r)\Delta t} \end{aligned}$$

Taking a two dimensional Taylor series around  $Q(r, t)$  and exploiting the fact that:

$$\begin{aligned} \frac{\sigma^2(r + \Delta r, t) - \sigma^2(r - \Delta r, t)}{2\Delta r} &= \frac{\partial \sigma^2}{\partial r}(r, t) + \frac{1}{6} \frac{\partial^3 \sigma^2}{\partial r^3}(r, t)(\Delta r)^2 + O(\Delta r^3) \\ \frac{\sigma^2(r + \Delta r, t) - 2\sigma^2(r, t) + \sigma^2(r - \Delta r, t)}{\Delta r^2} &= \frac{\partial^2 \sigma^2}{\partial r^2}(r, t) + \frac{1}{12} \frac{\partial^4 \sigma^2}{\partial r^4}(r, t)(\Delta r)^2 + O(\Delta r^3) \\ \frac{\sigma^2(r + \Delta r, t) + \sigma^2(r - \Delta r, t)}{2} &= \sigma^2(r, t) + \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2}(r, t)(\Delta r)^2 + O(\Delta r^3) \end{aligned}$$

with similar relationships holding for  $\mu$  and  $\mu^2$ , we obtain

$$\begin{aligned} Q(r, t + \Delta t) &= p_3(r + \Delta r, t)Q(r + \Delta r, t)e^{-(r+\Delta r)\Delta t} + p_2(r, t)Q(r, t)e^{-r\Delta t} \\ &\quad + p_1(r - \Delta r, t)Q(r - \Delta r, t)e^{-(r-\Delta r)\Delta t} \\ \Rightarrow \frac{\partial Q}{\partial t} \Delta t &= \Delta t \left[ Q \left( \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2} - \frac{\partial \mu}{\partial r} - r \right) + \frac{\partial Q}{\partial r} \left( \frac{\partial \sigma^2}{\partial r} - \mu \right) + \frac{\partial^2 Q}{\partial r^2} \frac{\sigma^2}{2} \right] + O(\Delta t^2, \Delta t \Delta r^2) \end{aligned}$$

so that

$$\frac{\partial Q}{\partial t} + Q \left( \frac{\partial \mu}{\partial r} - \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2} \right) + \frac{\partial Q}{\partial r} \left( \mu - \frac{\partial \sigma^2}{\partial r} \right) - \frac{\partial^2 Q}{\partial r^2} \frac{\sigma^2}{2} = -rQ + O(\Delta t, \Delta r^2)$$

and so the Hull-White approach to evaluating  $Q$  approximates (3) with similar accuracy to its approximation of (2).

### 1.1.1 Calibrating the yield curve

Hull and White [1994] finesse the approach of Hull and White [1993] for the specific case:

$$dr = (\theta(t) - \lambda r) dt + \sigma dW \tag{4}$$

where  $\theta(t)$  is an arbitrary function, chosen to match zero coupon bond prices from the model with observed yields.  $\lambda$  and  $\sigma$  are constants. Hull and White [1994] suggest that an appropriate tree-building technique is to set up a tree for

$$dr = -\lambda r dt + \sigma dW \tag{5}$$

They then solve for the Arrow-Debreu prices, at each step augmenting the interest rate by an amount  $\alpha(t)$  so that the price of that time's zero coupon bond is consistent with the yield curve.<sup>1</sup> Mathematically,

$$\begin{aligned} Q(r, t + \Delta t) &= p_3(r + \Delta r, t)Q(r + \Delta r, t)e^{-(r+\Delta r+\alpha(t))\Delta t} + p_2(r, t)Q(r, t)e^{-(r+\alpha(t))\Delta t} \\ &\quad + p_1(r - \Delta r, t)Q(r - \Delta r, t)e^{-(r-\Delta r+\alpha(t))\Delta t} \end{aligned} \tag{6}$$

---

<sup>1</sup>Their paper shows that the relationship between  $\alpha(t)$  and  $\theta(t)$  is  $(\theta(t) - \lambda\alpha(t))\Delta t = \alpha(t) - \alpha(t - \Delta t)$ .

If we define the price of a zero coupon bond maturing at time  $t + \Delta t$ , conditional on  $\alpha(t)$  as being  $P(t + \Delta t, \alpha(t))$ , then:

$$P(t + \Delta t, \alpha(t)) = \sum_j Q(r_j, t) e^{-(\alpha(t) + r_j)\Delta t} = e^{-\alpha(t)\Delta t} \sum_j Q(r_j, t) e^{-r_j\Delta t}$$

Hence knowing the level of  $Q(r, t)$  for the previous time step and knowing the price of a zero coupon bond maturing at time  $t + \Delta t$  is sufficient information to back out  $\alpha(t)$ . Once  $\alpha(t)$  is known, the Arrow-Debreu prices can be updated using (6). Once  $\alpha(t)$  is known for each time step, the finished tree can be used to price derivatives.

### 1.1.2 Calibrating volatility

To fit not only the yield curve, but also the volatility curve, it becomes necessary to introduce two functions of time into the equation. Following Hull and White [1993], we consider:

$$dr = (\theta(t) - \phi(t)r)dt + \sigma r^\beta dW \tag{7}$$

Here  $\theta(t)$  and  $\phi(t)$  are arbitrary functions of time, chosen to now match the yield curve *and* observed yield volatilities.  $\beta$  and  $\sigma$  are constants. Intuitively, if the mean-reversion speed ( $\phi(t)$ ) increases in  $t$ , it will damp later volatilities, while if it declines, later volatilities will be higher.

As noted in Hull and White [1993] this technique generalises to transformations of  $r$  so that (for example) the Black-Karasinski model can be implemented in a similar fashion. We refer readers to Hull and White [1993] for complete treatment of their approach. Their algorithm begins by using a binomial step where the entire yield curve is shifted up or down by a sufficient amount to match the

volatility curve specified. The algorithm then chooses  $\phi(t)$  and  $\theta(t)$  at each step so that the value of a zero coupon bond maturing at that time, conditional on the initial binomial step having been up, or having been down, are consistent with the zero coupon bond price observed at time zero.

## 1.2 Alternative solutions

The explicit method is not the only method that could be used to solve either (2) or (3). We can construct alternative finite difference approximations. For the backward equation, we define our finite difference method as:

$$f(r, t) - w\Delta t\mathcal{L}_{\mathcal{B}}f(r, t) = f(r, t + \Delta t) + (1 - w)\Delta t\mathcal{L}_{\mathcal{B}}f(r, t + \Delta t)$$

where  $\mathcal{L}_{\mathcal{B}}$  is the operator:

$$\mathcal{L}_{\mathcal{B}}f(r, t) = \mu(r, t)\frac{f(r + \Delta r, t) - f(r - \Delta r, t)}{2\Delta r} + \frac{\sigma(r, t)^2}{2}\frac{f(r + \Delta r, t) - 2f(r, t) + f(r - \Delta r, t)}{\Delta r^2} - rf(r, t)$$

Our goal is to evaluate  $f(r, t)$  given information regarding  $f(r, t + \Delta t)$ . For  $w = 0$ , we obtain an explicit relationship, where a single point at time  $t$  is calculated using three points at time  $t + \Delta t$ , as is the case in a trinomial tree. If  $w = 1$ , we obtain an implicit relationship, where three points at time  $t$  are related to a single point at time  $t + \Delta t$ . For  $w = 0.5$ , we obtain the Crank-Nicholson approach (see Crank and Nicholson [1947]), which relates three points at time  $t$  to three points at time  $t + \Delta t$ . These latter two approaches result in a system of equations with one equation for each spatial point at the new time step, which are solved to find  $f(r, t)$ . To deal with points on the boundary, following

Vetzal [1998], we can use the operators:

$$\begin{aligned}\mathcal{L}_{\mathcal{BL}}f(r, t) &= \mu(r, t) \frac{-f(r + 2\Delta r, t) + 4f(r + \Delta r, t) - 3f(r, t)}{2\Delta r} \\ &\quad + \frac{\sigma(r, t)^2}{2} \frac{f(r + 2\Delta r, t) - 2f(r + \Delta r, t) + f(r, t)}{\Delta r^2} - rf(r, t) \\ \mathcal{L}_{\mathcal{BU}}f(r, t) &= \mu(r, t) \frac{3f(r, t) - 4f(r - \Delta r, t) + f(r - 2\Delta r, t)}{2\Delta r} \\ &\quad + \frac{\sigma(r, t)^2}{2} \frac{f(r, t) - 2f(r - \Delta r, t) + f(r - 2\Delta r, t)}{\Delta r^2} - rf(r, t)\end{aligned}$$

where  $\mathcal{L}_{\mathcal{BL}}$  is applied at the lower boundary and  $\mathcal{L}_{\mathcal{BU}}$  is applied at the upper boundary. The resulting system of equations in  $f(r, t)$  is tridiagonal, except for the first and last equations. It is a simple matter to eliminate the non-tridiagonal elements prior to solving the system.<sup>2</sup>

We can develop a similar approximation for solution of (3):

$$Q(r, t + \Delta t) + w\Delta t\mathcal{L}_{\mathcal{F}}Q(r, t + \Delta t) = Q(r, t) - (1 - w)\Delta t\mathcal{L}_{\mathcal{F}}Q(r, t)$$

where  $\mathcal{L}_{\mathcal{F}}$  is the operator:

$$\begin{aligned}\mathcal{L}_{\mathcal{F}}Q(r, t) &= \left( \mu(r, t) - \frac{\partial\sigma^2}{\partial r}(r, t) \right) \frac{Q(r + \Delta r, t) - Q(r - \Delta r, t)}{2\Delta r} \\ &\quad - \frac{\sigma(r, t)^2}{2} \frac{Q(r + \Delta r, t) - 2Q(r, t) + Q(r - \Delta r, t)}{\Delta r^2} \\ &\quad + \left( \frac{\partial\mu}{\partial r}(r, t) - \frac{1}{2} \frac{\partial^2\sigma^2}{\partial r^2}(r, t) + r \right) Q(r, t)\end{aligned}$$

---

<sup>2</sup>Further discussion of finite difference techniques can be found in many textbooks. A good reference for this is Tavella and Randall [2000]. Reduction to tridiagonality is discussed in Vetzal [1998] and Cairns [2004].

with boundary conditions:

$$\begin{aligned}
\mathcal{L}_{\mathcal{FL}}Q(r, t) &= \left( \mu(r, t) - \frac{\partial \sigma^2}{\partial r}(r, t) \right) \frac{-Q(r + 2\Delta r, t) + 4Q(r + \Delta r, t) - 3Q(r, t)}{2\Delta r} \\
&\quad - \frac{\sigma(r, t)^2}{2} \frac{Q(r + 2\Delta r, t) - 2Q(r + \Delta r, t) + Q(r, t)}{\Delta r^2} \\
&\quad + \left( \frac{\partial \mu}{\partial r}(r, t) - \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2}(r, t) + r \right) Q(r, t) \\
\mathcal{L}_{\mathcal{FU}}Q(r, t) &= \left( \mu(r, t) - \frac{\partial \sigma^2}{\partial r}(r, t) \right) \frac{3Q(r, t) - 4Q(r - \Delta r, t) + Q(r - 2\Delta r, t)}{2\Delta r} \\
&\quad - \frac{\sigma(r, t)^2}{2} \frac{Q(r, t) - 2Q(r - \Delta r, t) + Q(r - 2\Delta r, t)}{\Delta r^2} \\
&\quad + \left( \frac{\partial \mu}{\partial r}(r, t) - \frac{1}{2} \frac{\partial^2 \sigma^2}{\partial r^2}(r, t) + r \right) Q(r, t)
\end{aligned}$$

Note that in this situation, we wish to solve for  $Q(r, t + \Delta t)$  given  $Q(r, t)$  (forward induction). As is the case for the Hull-White implementation, our initial condition is that  $Q(r_0, 0) = 1$  and  $Q(r, 0) = 0$  for  $r \neq r_0$ . Also similarly to the Hull-White implementation, we can find the price of a zero coupon bond maturing at time  $T$  as  $P(T) = \sum_j Q(r_j, T)$ . Again  $w = 0$  will result in the explicit method,  $w = 1$  will result in the implicit method, and  $w = 0.5$  will give the Crank-Nicholson method.

Of the three approaches, the Crank-Nicholson approach is generally taken to be the superior, since it has second order accuracy in the time dimension as well as the space dimension. Our applications in section 4 will make use of the Crank-Nicholson method.

### 1.2.1 Stability conditions for the Crank-Nicholson method

For a differential equation (being solved forward in time) which takes the form

$$\frac{\partial Q}{\partial t} = a_0 Q + a_1 \frac{\partial Q}{\partial r} + a_2 \frac{\partial^2 Q}{\partial r^2}$$

the Crank-Nicholson stability condition is equivalent to the condition necessary for the matrix which must be inverted for each time step to be diagonally dominant.<sup>3</sup> Simplifying this diagonal dominance condition, we obtain:

$$1 + \frac{\Delta t}{(\Delta r)^2} a_2 > \frac{1}{2} \left( a_0 \Delta t + \frac{\Delta t}{\Delta r} |a_1| - 2 \right)$$

Intuitively this condition states that the diffusion part of the equation ( $a_2$ ) is dominant over the advection part ( $a_1$ ). Since a user generally wants to fix the proportion between  $\Delta t$  and  $\Delta r$ , the term on the left hand side increases as finer discretisations are considered. The term on the right hand side shrinks as  $\Delta t$  declines. Hence stability is only ever an issue when a very coarse discretisation is used, or if the  $a_1$  term becomes very large.

For our numerical examples (see section 4) stability issues are only encountered when a time step of one year or more is used.

## 2 Calibrating to the yield curve

We focus here on the Hull-White model (4) and demonstrate how we can calibrate  $\theta(t)$  using the techniques outlined in section 1.2. For a general finite difference technique, we could write, in matrix notation, that

$$BQ_{t+\Delta t} = AQ_t$$

where  $Q_t$  is a vector containing the level of  $Q$  for each level of  $r$  at time  $t$ . The matrices  $A$  and  $B$  describe the system of equations that must be solved to find  $Q_{t+\Delta t}$  given  $Q_t$ . We will assume that these have (as was the case for the tree approach) been set up to be consistent with (5). Our goal,

---

<sup>3</sup>A similar condition can be derived for a backward equation, by making the substitution  $\tau = -t$ .

for each time step, is to find a level of  $\alpha(t)$  (see section 1.1.1) which results in the model correctly pricing a time  $t$  zero coupon bond.

Increasing the discount rate will simply affect the right hand side of (3), so we can adjust our updating to incorporate  $\alpha$  as:

$$(B + w\Delta t\alpha(t)I)Q_{t+\Delta t} = (A - (1 - w)\Delta t\alpha(t)I)Q_t \quad (8)$$

where  $w$  takes on value 1 for a fully implicit method, 0.5 for the Crank Nicholson method (discussed previously) or 0 for an explicit method. The matrix  $I$  is an identity matrix. Note that for the explicit method, the matrix  $B$  will be an identity matrix, and so there is a straightforward relationship between choice of  $\alpha(t)$  and the level of  $Q_{t+\Delta t}$ . Hence it is easy to solve for  $\alpha(t)$ .

For an implicit method, matters are complicated by the appearance of  $\alpha(t)$  on the left hand side of equation (8), creating an implicit relationship between choice of  $\alpha(t)$  and the level of  $Q_{t+\Delta t}$ . This requires us to solve the problem iteratively.

Consider that we guess that the level of  $\alpha(t)$  is  $\hat{\alpha}(t)$ , which has corresponding level of  $Q_{t+\Delta t}$  of  $\hat{Q}_{t+\Delta t}$ . Now consider perturbing  $\hat{\alpha}(t)$  by  $\Delta\alpha$ . Then the left hand side of (8) will change by

$$w\Delta t\Delta\alpha\hat{Q}_{t+\Delta t} + (B + w\Delta t\hat{\alpha}(t)I)\Delta Q_{t+\Delta t}$$

where  $\Delta Q_{t+\Delta t}$  is the change in time  $t + \Delta t$  Arrow-Debreu prices caused by the change in  $\hat{\alpha}(t)$ .

Meanwhile, the right hand side of the equation will change by:

$$-(1 - w)\Delta t\Delta\alpha Q_t$$

Equating the changes, and solving for  $\Delta Q_{t+\Delta t}$ , we obtain:

$$\Delta Q_{t+\Delta t} = -\Delta t \Delta \alpha (B + w \Delta t \hat{\alpha}(t) I)^{-1} ((1-w)Q_t + w\hat{Q}_{t+\Delta t}) \quad (9)$$

Furthermore, we know that the (given) price of a time  $t + \Delta t$  bond ( $P(t + \Delta t)$ ) should equal the sum of the Arrow-Debreu prices at time  $t + \Delta t$ . Hence we would like:

$$\mathbb{1}^T \Delta Q_{t+\Delta t} = P(t + \Delta t) - \mathbb{1}^T \hat{Q}_{t+\Delta t} \quad (10)$$

where  $\mathbb{1}$  is a vector of ones of the same length as  $Q_t$ . Substituting (9) into (10) and solving for  $\Delta \alpha$ , we obtain:

$$\Delta \alpha = \frac{-(P(t + \Delta t) - \mathbb{1}^T \hat{Q}_{t+\Delta t})}{\Delta t \mathbb{1}^T (B + w \Delta t \hat{\alpha}(t) I)^{-1} ((1-w)Q_t + w\hat{Q}_{t+\Delta t})}$$

Solving for  $\alpha(t)$  thus follows the following algorithm:

1. Solve equation (8) to obtain a preliminary level for  $\hat{Q}_{t+\Delta t}$  given a guess of  $\hat{\alpha}(t)$ .
2. Update  $\hat{\alpha}(t)$  to:

$$\hat{\alpha}(t) + \frac{\mathbb{1}^T \hat{Q}_{t+\Delta t} - P(t + \Delta t)}{\mathbb{1}^T R}$$

where the vector  $R$  solves

$$(B + w \Delta t \hat{\alpha}(t) I) R = \Delta t \left[ (1-w)Q_t + w\hat{Q}_{t+\Delta t} \right]$$

3. Repeat the iterations until  $\hat{\alpha}(t)$  converges.

In practice, given a reasonable guess of  $\hat{\alpha}(t)$  (such as  $\alpha(t - \Delta t)$ ) the algorithm converges to a high level

of precision within one or two iterations. Note that this algorithm is not peculiar to the particular form of volatility term used for the short rate – it could be implemented for a geometric or square root style process.

Further note that once  $\alpha(t)$  has been calibrated, a more conventional backward algorithm can be employed to actually price securities using  $\alpha(t)$ . This allows for American features to be easily dealt with.

## 2.1 Application to backward equation methods

An implicit method can be calibrated using solely the backward equation, as outlined in Uhrig and Walter [1996] and Vetzal [1998]. This requires repeated pricing of a time  $t$  bond, iterating over  $\hat{\alpha}(t + \Delta t)$  until this bond is priced correctly. In general, this would require  $t/\Delta t$  solutions of the backward finite difference equations for each evaluation of the time  $t$  bond's price. However, Vetzal [1998] notes that when evaluating bond prices, provided the only adjustment to the equations solved at each time step is a change in the diagonal elements (as is the case here with time varying  $\alpha(t)$ ), the application of time steps commutes. As a result, denoting the vector of zero coupon bond prices for maturity  $t$  (for each grid level of  $r$ ) as  $U(t)$ ,  $\hat{U}(t + \Delta t)$  may be evaluated by solving

$$(\bar{B} + wI\hat{\alpha}(t)\Delta t)\hat{U}(t + \Delta t) = (\bar{A} - (1 - w)I\hat{\alpha}(t)\Delta t)U(t)$$

where  $\bar{A}$  and  $\bar{B}$  are the (backward equation) matrices implied by the finite difference technique for  $\alpha(t) = 0$ . A similar reasoning to that used above for the forward equation method can be used to

show that the optimal update to  $\hat{\alpha}(t)$  is

$$\Delta\alpha(t) = (\hat{U}(t + \Delta t, r_0) - P(t + \Delta t)) / \bar{R}(r_0)$$

where  $\bar{R}$  solves

$$(\bar{b} + wI\hat{\alpha}(t)\Delta t)\bar{R} = (1 - w)U(t) + w\hat{U}(t + \Delta t)$$

Here  $\hat{U}(t, r_0)$  and  $\bar{R}(r_0)$  refer to the elements in the vector  $\hat{U}(t)$  and  $\bar{R}$  which correspond to  $r = r_0$ .

Since  $U(t)$  is known, updating  $\hat{\alpha}(t + \Delta t)$  requires only two matrix inversions, as is the case of the forward equation algorithm.

### 3 Calibrating to the volatility curve

We now turn our attention to fitting the more general model

$$dr = (\theta(t) - \phi(t)r)dt + \sigma(r, t)dW$$

using the finite difference techniques described in section 1.2. Here  $\theta(t)$  and  $\phi(t)$  must be selected in order to match not only the yield curve observed, but also a term structure of volatilities.

Our algorithm here is similar in spirit to our approach for matching the yield curve alone. Firstly, we note that the volatility of an interest rate for time horizon  $t$  can be written in terms of the short rate volatility ( $\sigma(r, t)$ ):

$$V(t) = \frac{-\frac{\partial P(t)}{\partial r}\sigma(r_0, 0)}{P(t)t}$$

Here we define volatility in absolute terms (the diffusion term of the process for the interest rate in

question). Adjusting to be a proportional volatility simply requires dividing  $V(t)$  by the interest rate for time horizon  $t$ :  $R(t) \equiv -\log(P(t))/t$ .

We know that  $P(t)$  is equal to the sum of  $Q(r, t)$  across all values of  $r$  at time  $t$ , so we may write:

$$V(t) \simeq -\frac{\sigma(r_0, 0)}{2t\Delta r} \frac{\mathbb{1}^T(Q_t^u - Q_t^d)}{\mathbb{1}^T Q_t} \quad (11)$$

where  $Q_t^u$  is the vector of Arrow-Debreu prices consistent with the time zero interest rate being equal to  $r_0 + \Delta r$  and  $Q_t^d$  is the vector of Arrow-Debreu prices at time  $t$  consistent with the time zero interest rate being equal to  $r_0 - \Delta r$ . Note that  $Q^u$  and  $Q^d$  both satisfy (3) but with boundary conditions such that  $Q^u(r, 0)$  is zero except when  $r = r_0 + \Delta r$  (where it equals one) and  $Q^d(r, 0)$  is zero except when  $r = r_0 - \Delta r$  (where it similarly equals one).

Our conditions we require to hold are thus (11) and the familiar

$$\mathbb{1}^T Q_t = P(t)$$

with both equations required to hold for each time step.

Our algorithm proceeds similarly to that in section 2. At each time step, we will iteratively find  $\theta(t)$  and  $\phi(t)$ , given some initial guess  $\hat{\theta}(t)$  and  $\hat{\phi}(t)$ . We will assume that the updating equations, including  $\phi(t)$  and  $\theta(t)$  are given by:

$$(B + w\hat{\theta}(t)D_1 + w\hat{\phi}(t)D_2)\hat{Q}_{t+\Delta t} = (A - (1-w)\hat{\theta}(t)D_1 - (1-w)\hat{\phi}(t)D_2)Q_t \quad (12)$$

where  $D_1$  and  $D_2$  are matrices representing the terms in the finite difference equations which have  $\theta$

and  $\phi$  as a coefficient, respectively (see section 1.2). We will also define:

$$Z(\theta(t), \phi(t)) \equiv B + w\theta(t)D_1 + w\phi(t)D_2$$

Now consider the effect of changing  $\hat{\theta}$  by  $\Delta\theta$  and  $\hat{\phi}$  by  $\Delta\phi$ . The effect on the left hand side of (12) is:

$$\Delta\theta wD_1\hat{Q}_{t+\Delta t} + \Delta\phi wD_2\hat{Q}_{t+\Delta t} + Z(\hat{\theta}, \hat{\phi})\Delta Q_{t+\Delta t}$$

while the effect on the right hand side of (12) is:

$$(-\Delta\theta(1-w)D_1 - \Delta\phi(1-w)D_2)Q_t$$

We can equate these and solve for  $\Delta Q_{t+\Delta t}$  as:

$$\Delta Q_{t+\Delta t} = -Z^{-1}(\Delta\theta D_1 + \Delta\phi D_2) \left( w\hat{Q}_{t+\Delta t} + (1-w)Q_t \right) \quad (13)$$

Defining

$$\begin{aligned} R_1 &= Z^{-1}D_1 \left( w\hat{Q}_{t+\Delta t} + (1-w)Q_t \right) \\ R_2 &= Z^{-1}D_2 \left( w\hat{Q}_{t+\Delta t} + (1-w)Q_t \right) \end{aligned}$$

and setting  $\Delta Q_{t+\Delta t} = P(t + \Delta t) - \mathbb{1}^T \hat{Q}_{t+\Delta t}$ , we obtain the condition:

$$P(t + \Delta t) - \mathbb{1}^T \hat{Q}_{t+\Delta t} = -\Delta\theta \mathbb{1}^T R_1 - \Delta\phi \mathbb{1}^T R_2 \quad (14)$$

We also need the volatility condition to hold for time  $t + \Delta t$ . We can use the quotient rule to determine the effect of perturbing  $\hat{\theta}(t + \Delta t)$  and  $\hat{\phi}(t + \Delta t)$  on the level of volatility:

$$\Delta V = \frac{-\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T (\Delta Q_{t+\Delta t}^u - \Delta Q_{t+\Delta t}^d)}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T \Delta Q_{t+\Delta t} \right] \quad (15)$$

Using (13) applied to  $Q^u$  and  $Q^d$ , we can rewrite (15) as:

$$\begin{aligned} \Delta V &= \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T Z^{-1} (\Delta\theta D_1 + \Delta\phi D_2) \left( w\hat{Q}_{t+\Delta t}^u + (1-w)Q_t^u - w\hat{Q}_{t+\Delta t}^d - (1-w)Q_t^d \right)}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} \right. \\ &\quad \left. - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T Z^{-1} (\Delta\theta D_1 + \Delta\phi D_2) (w\hat{Q}_{t+\Delta t} + (1-w)Q_t) \right] \\ &= \Delta\theta \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T R_3}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T R_1 \right] \\ &\quad + \Delta\phi \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T R_4}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T R_2 \right] \end{aligned} \quad (16)$$

where  $R_1$  and  $R_2$  are as given above and:

$$R_3 = Z^{-1} D_1 (w\hat{Q}_{t+\Delta t}^u + (1-w)Q_t^u - w\hat{Q}_{t+\Delta t}^d - (1-w)Q_t^d)$$

$$R_4 = Z^{-1} D_2 (w\hat{Q}_{t+\Delta t}^u + (1-w)Q_t^u - w\hat{Q}_{t+\Delta t}^d - (1-w)Q_t^d)$$

Equating (16) to  $V(t + \Delta t) - \hat{V}(t + \Delta t)$ , where  $\hat{V}(t + \Delta t)$  is given by

$$\hat{V}(t + \Delta t) = \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{\mathbb{1}^T \hat{Q}_{t+\Delta t}}$$

we obtain a second condition for  $\Delta\theta$  and  $\Delta\phi$ . We can summarise the two conditions as:

$$\begin{bmatrix} \kappa_1 & \kappa_2 \\ \kappa_3 & \kappa_4 \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta\phi \end{bmatrix} = \begin{bmatrix} P(t + \Delta t) - \mathbb{1}^T Q(t + \Delta t) \\ V(t + \Delta t) - \hat{V}(t + \Delta t) \end{bmatrix} \quad (17)$$

where

$$\begin{aligned} \kappa_1 &= -\mathbb{1}^T R_1 \\ \kappa_2 &= -\mathbb{1}^T R_2 \\ \kappa_3 &= \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T R_3}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T R_1 \right] \\ \kappa_4 &= \frac{\sigma(r_0, 0)}{2\Delta r(t + \Delta t)} \left[ \frac{\mathbb{1}^T R_4}{\mathbb{1}^T \hat{Q}_{t+\Delta t}} - \frac{\mathbb{1}^T (\hat{Q}_{t+\Delta t}^u - \hat{Q}_{t+\Delta t}^d)}{(\mathbb{1}^T \hat{Q}_{t+\Delta t})^2} \mathbb{1}^T R_2 \right] \end{aligned}$$

Stepping forward in time thus consists of alternating between updating  $(\hat{Q}_{t+\Delta t}, \hat{Q}_{t+\Delta t}^u, \hat{Q}_{t+\Delta t}^d)$  using (12) and solving (17) to update  $\hat{\theta}(t + \Delta t)$  and  $\hat{\phi}(t + \Delta t)$ . Again, using  $\theta(t)$  and  $\phi(t)$  as starting conditions for  $\hat{\theta}(t + \Delta t)$  and  $\hat{\phi}(t + \Delta t)$ , this process only takes one or two iterations to converge to a reasonable level of precision.

As before, we note that this algorithm is general across all finite difference techniques, and that once  $\theta(t)$  and  $\phi(t)$  are known, the backward equation can be used to price American securities.

## 4 Numerical examples

We consider two test problems to demonstrate our new technique. In both cases, we make use of the Crank-Nicholson forward equation algorithm, described in section 1.2, and compare the algorithm's prices to those produced by the trinomial approach described in section 1.1 along with the backward

Maturity	Yield
0.5	0.03430
1.0	0.03824
1.5	0.04183
2.0	0.04512
2.5	0.04812
3.0	0.05086

Exhibit 1: Yield curve for examples in section 4.1. All rates expressed with continuous compounding.

equation style approaches described in Uhrig and Walter [1996] and Vetzal [1998], wherein actual zero coupon bond prices (as opposed to Arrow-Debreu prices) are generated at each time, and parameters are calibrated to match these to the yield curve.

## 4.1 Calibration to yield curve

For this example, we consider the standard Hull-White model given by (4) and implement the procedures described in section 2. For the forward-equation approach, we follow the method outlined in this paper. The backward-equation technique is implemented according to Vetzal [1998], using the calibration technique outlined in section 2.1. The yield curve is given in exhibit 1, and is assumed to be piecewise linear for points between those given. We consider the pricing of a two year call option, written on a three year zero coupon bond with face value \$1, where the option has strike price \$0.943. We assume a mean reversion rate of  $\lambda = 0.1$  and a volatility of  $\sigma = 0.01$ . A closed form solution exists for options on zero coupon bonds in this model (see Hull and White [1990]). The correct price for this security is 0.0028.

In exhibit 2 we present relative pricing error, as the time step is decreased.<sup>4</sup> Since the Crank-

---

<sup>4</sup>For all three methods, we price the option (once calibration is complete) using backward equations and backward induction, so the times reported here are representative of American option pricing times. For the forward equation approach and the trinomial tree, a European security could be priced faster using  $Q$  evaluated at the security's maturity.

Timestep (Tri)	Timestep (BE/FE)	Error (Tri)	Error (FE)	Error (BE)	Time (Tri)	Time (FE)	Time (BE)
0.5000	0.5000	0.0363	0.0155	0.0140	0.0017	0.0029	0.0025
0.2500	0.5000	0.0292	0.0155	0.0140	0.0065	0.0027	0.0025
0.1000	0.2500	0.0044	-0.0027	-0.0030	0.0269	0.0094	0.0090
0.0500	0.2000	-0.0071	0.0025	0.0022	0.0915	0.0169	0.0165
0.0250	0.1429	0.0030	-0.0028	-0.0029	0.3635	0.0421	0.0415
0.0100	0.1000	0.0009	-0.0006	-0.0007	2.2428	0.1166	0.1205
0.0050	0.0667	0.0004	0.0002	0.0001	8.9697	0.3781	0.3766
0.0025	0.0500	0.0002	0.0000	-0.0000	35.9485	0.8813	0.8825

Exhibit 2: Performance of algorithms for exercise in section 4.1. Tri refers to the Hull-White trinomial tree, while FE and BE refer to the Forward Equation and Backward Equation Crank-Nicholson algorithms respectively. For the trinomial tree, we assume that  $\Delta r = \sigma\sqrt{3}\Delta t$ , with boundaries set at  $\pm \frac{0.184}{\lambda\Delta t}\Delta r$  as suggested in Hull and White [1994]. For the Crank-Nicholson method,  $\Delta r = \sigma\sqrt{2}\Delta t$ , with boundaries set at  $\pm \frac{5\sigma}{\sqrt{2}\lambda}$ , as suggested by Cairns [2004]. Time taken is in seconds running MATLAB on an Intel Core 2 Duo E8400.

Nicholson method (in either forward equation or backward equation form) is second order, we use a timestep chosen to be approximately the square-root of the time step used by the Trinomial tree. Given the second order accuracy in time, this should give comparable accuracy to the trinomial tree.

Exhibit 2 gives a clear picture of the advantages of using the Crank-Nicholson algorithm. For large time steps (low precision) the advantage of needing to neither solve the implicit equations nor iterate to find  $\alpha$  renders the trinomial method more efficient, however, as more accuracy is required, the second order convergence of the Crank-Nicholson method more than outweighs its more involved calculations.<sup>5</sup> Given that the backward equation can be implemented in a forward induction fashion here, using only two equation solutions for each iteration on  $\hat{\alpha}(t)$ , the performance of the forward equation approach and the backward equation approach are almost identical.

---

<sup>5</sup>It should be noted that while the Crank-Nicholson method does require more work than the Trinomial tree, the effort required to calculate one step is still  $O(N)$  where  $N$  is the number of spatial steps in the lattice.

Maturity	Yield	Volatility
1	0.0525	0.006825
2	0.055	0.0066
3	0.0575	0.006325
4	0.06	0.006
5	0.0625	0.005625
6	0.065	0.0052

Exhibit 3: Yield curve and volatility curve for examples in section 4.2. All rates expressed with continuous compounding. Volatility is absolute. The short rate ( $r_0$ ) is 0.05, while the short rate volatility ( $\sigma$ ) is chosen so that  $\sigma r_0^\beta = 0.14r_0$ .

## 4.2 Calibration to yield curve and volatility

Our second numerical example follows the methodology of section 3, using a yield curve and volatility curve as given in exhibit 3. The short rate is assumed to follow (7). We consider pricing a 4 year call option written on a 5 year zero coupon bond, with strike price 0.933. We consider four levels for  $\beta$ : 0, 0.5, 1.0 and 1.5. The first case ( $\beta = 0$ ) corresponds to the standard Hull-White extended Vasicek model.  $\beta = 0.5$  gives the extended Cox-Ingersoll-Ross model. Setting  $\beta = 1.0$  obtains a geometric model for interest rates. Lastly,  $\beta = 1.5$  is consistent with the estimation work of Chan, Karolyi, Longstaff, and Sanders [1992], who conclude this most accurately explains US treasury rates.

In this case, the backward equations will not commute, and so the backward equation approach will require full pricing of the time  $t$  bond to evaluate the error implied by a particular choice of  $\phi(t)$  or  $\theta(t)$ . Given this, our Newton-style updating of  $\phi(t)$  and  $\theta(t)$  *cannot* be adapted to the backward induction approach. We therefore make use of MATLAB’s “lsqnonlin” command, which allows us to calibrate  $\phi(t)$  and  $\theta(t)$  while applying constraints to prevent the optimiser wandering into regions where the Crank-Nicholson algorithm is unstable (i.e. avoiding very high values of  $\theta(t)$  or  $\phi(t)$ ).

Timestep (Tri)	Timestep (BE/FE)	Error (Tri)	Error (FE)	Error (BE)	Time (Tri)	Time (FE)	Time (BE)
0.5000	0.5000	0.0649	-0.0150	-0.0170	0.0012	0.0156	0.0874
0.2500	0.5000	0.0646	-0.0150	-0.0170	0.0029	0.0156	0.0874
0.1000	0.2500	0.0270	-0.0038	-0.0042	0.0138	0.0655	0.5352
0.0500	0.2000	0.0068	-0.0028	-0.0031	0.0499	0.1274	1.1707
0.0250	0.1429	0.0074	-0.0012	-0.0013	0.1930	0.3279	3.8393
0.0100	0.1000	0.0030	-0.0006	-0.0006	1.3631	0.9116	10.5540
0.0050	0.0667	0.0010	-0.0006	-0.0007	8.1016	3.4059	50.4267
0.0025	0.0500	0.0004	-0.0003	-0.0003	60.8680	8.5314	168.0676

Exhibit 4: Performance of algorithms for exercise in section 4.2.1. Tri refers to the Hull-White trinomial tree, while FE and BE refer to the Forward Equation and Backward Equation Crank-Nicholson algorithms respectively. For the trinomial tree, we assume that  $\Delta r = \sigma\sqrt{3\Delta t}$ , with boundaries set at  $r_0 \pm \frac{0.184}{0.05\Delta t}\Delta r$  as suggested in Hull and White [1994], and consistent with a level of mean reversion of 0.05. For the Crank-Nicholson method, we use  $\Delta r = \sqrt{2}\sigma\Delta t$ , with boundaries set at  $r_0 \pm \frac{5\sigma}{\sqrt{0.1}}$ , consistent with Cairns' analysis, also with a level of mean reversion of 0.05. Time taken is in seconds running MATLAB on an Intel Core 2 Duo E8400.

#### 4.2.1 Constant volatility ( $\beta = 0$ )

For this case, a closed form solution is available for bond options (see Hull and White [1990]). The analytical price for this security is 0.0019. Again, we consider using a larger time step for the Crank-Nicholson methods, since they have a higher order convergence rate. Results are given in exhibit 4.

Here some additional numerical costs are borne by the Crank-Nicholson approaches. For the forward equation method, the volatility calibrating algorithm requires 7 matrix inversions for each iteration on  $\phi$  and  $\theta$  (once each for evaluating  $Q$ ,  $Q^u$  and  $Q^d$  along with one inversion for each of  $R_1, \dots, R_4$ ) as compared to two inversions for the yield curve calibrating algorithm (once for  $Q$  and once for  $R$ ). The backward equation method, in contrast, requires  $t/\Delta t$  inversions for each iteration on  $\phi$  and  $\theta$  (one for each step required to evaluate the value of the time  $t$  bond).

The difference between orders of convergence and orders of computational time required becomes

very apparent in exhibit 4. All three methods reduce errors as  $\Delta t$  declines. However, both the Trinomial method and the backward equation approach require rapidly increasing amounts of time to achieve these improvements relative to the forward equation approach. In the case of the Trinomial tree, this follows because the method is only first order accurate in time. For the backward equation method, although it is second order accurate in time, the effort required to calibrate  $\theta$  and  $\phi$  is second order. The results of this are striking: to achieve accuracy of approximately 0.003 would take 1.36 seconds ( $\Delta t = 0.01$ ) with the tree, 1.17 seconds ( $\Delta t = 0.2$ ) for the backward equation and only 0.13 seconds ( $\Delta t = 0.2$ ) with the forward equation approach.

#### 4.2.2 Non-constant volatility ( $\beta > 0$ )

For  $\beta > 0$ , no closed form solutions are available. For the case  $\beta = 0.5$ , closed form solutions are available for bond prices (see Hull and White [1990]) and for option prices, provided the ratio  $\theta(t)/\sigma(t)^2$  is a constant integer (see Maghsoodi [1996]). However, in general, numerical solutions must be used to obtain option prices.

For the trinomial tree approach, we transform variables so that volatility becomes constant, as discussed in Hull and White [1993] and Hull and White [1991]. Setting  $r^* = \frac{r_0^\beta r^{1-\beta}}{1-\beta}$  for  $\beta \neq 1$  and  $r^* = r_0 \log(r)$  for  $\beta = 1$  converts (7) to:

$$dr^* = (\theta(t) - \phi(t)) \left(\frac{r_0}{r}\right)^\beta dt + \sigma r_0^\beta dW$$

This allows the regular Hull-White tree to be constructed without negative probabilities. We perform

a similar transform for the Crank-Nicholson methods.<sup>6</sup> Exhibit 5 presents results for the option pricing with  $\beta \neq 0$ . In order to obtain a “true” option price, we performed Richardson extrapolation on the trinomial solutions with stepsizes 0.005 and 0.0025.

Interestingly, the backward equation has low errors relative to the other two techniques for comparable levels of  $\Delta t$ . Nevertheless, the extreme time cost required under this method generally outweighs its advantages in accuracy for a given level of  $\Delta t$ . Examining our target of 0.3% accuracy, we find that for  $\beta = 0.5$ , the forward approach achieves its target in 0.30 seconds, while the backward equation takes 3.5 times as long (1.05 seconds) and the trinomial tree takes approximately 15 times as long (4.45 seconds). For  $\beta = 1.0$ , the backward equation performs best at this level of accuracy, taking 0.53 seconds, compared to 3.61 and 7.82 seconds for the forward equation and trinomial tree respectively. For the case of  $\beta = 1.5$ , 35 time steps for the backward equation achieves the target, where 75 steps are required by the forward equation and 1000 steps for the trinomial tree. Time-wise, however, the forward equation still dominates the field, with a time taken of 3.8 seconds, compared to 4.4 seconds for the backward equation and 5.5 seconds for the tree.

We note that in all cases, the forward equation and trinomial tree methods have monotonic convergence to the true value, so Richardson extrapolation could be applied to increase the convergence rate. In contrast, the backward equation method often has oscillatory convergence, which is not

---

<sup>6</sup>Under this transform, the forward equation is given by:

$$\begin{aligned} & \frac{\partial Q}{\partial t} + \left( \theta(t) - \phi(t)r - \frac{1}{2}\beta\sigma^2r^{2\beta-1} \right) \left( \frac{r_0}{r} \right)^\beta \frac{\partial Q}{\partial r^*} \\ & + \left( (\beta - 1)\phi(t) - \frac{\theta(t)\beta}{r} + \frac{1}{2}\beta(1 - \beta)\sigma^2r^{2\beta-2} \right) Q - \frac{\sigma^2}{2}r_0^{2\beta} \frac{\partial^2 Q}{\partial r^{*2}} = -rQ \end{aligned}$$

and the backward equation is given by:

$$\frac{\partial Q}{\partial t} + \left( \theta(t) - \phi(t)r - \frac{1}{2}\beta\sigma^2r^{2\beta-1} \right) \left( \frac{r_0}{r} \right)^\beta \frac{\partial Q}{\partial r^*} + \frac{1}{2}\sigma^2r_0^{2\beta} \frac{\partial^2 Q}{\partial r^{*2}} = rQ$$

$\beta = 0.5$  (true price 0.0021)

Timestep (Tri)	Timestep (BE/FE)	Error (Tri)	Error (FE)	Error (BE)	Time (Tri)	Time (FE)	Time (BE)
0.5000	0.5000	0.1410	0.0236	-0.0243	0.0011	0.0141	0.0855
0.2500	0.5000	0.0927	0.0236	-0.0243	0.0030	0.0140	0.0852
0.1000	0.2500	0.0131	0.0080	-0.0036	0.0122	0.0613	0.4890
0.0500	0.2000	0.0174	0.0067	-0.0005	0.0391	0.1104	1.0499
0.0250	0.1429	0.0038	0.0029	-0.0008	0.1345	0.2998	3.3015
0.0100	0.1000	0.0036	0.0020	0.0002	0.7845	0.8064	9.4879
0.0050	0.0667	0.0014	0.0005	-0.0003	4.4531	3.1154	46.3460
0.0025	0.0500	0.0007	0.0004	-0.0000	26.1898	8.2620	146.2942

$\beta = 1.0$  (true price 0.0024)

Timestep (Tri)	Timestep (BE/FE)	Error (Tri)	Error (FE)	Error (BE)	Time (Tri)	Time (FE)	Time (BE)
0.5000	0.5000	0.2042	0.6596	-0.0051	0.0011	0.0220	0.0917
0.2500	0.5000	0.1000	0.6596	-0.0051	0.0028	0.0218	0.0912
0.1000	0.2500	0.0381	0.0192	0.0023	0.0130	0.0646	0.5284
0.0500	0.2000	0.0173	0.0112	-0.0001	0.0468	0.1276	1.1508
0.0250	0.1429	0.0104	0.0041	-0.0016	0.1803	0.3601	3.2369
0.0100	0.1000	0.0038	0.0034	0.0006	1.2999	0.9014	10.5274
0.0050	0.0667	0.0021	0.0016	0.0004	7.8196	3.6096	49.9042
0.0025	0.0500	0.0011	0.0008	0.0002	59.5069	9.2330	169.4497

$\beta = 1.5$  (true price 0.0028)

Timestep (Tri)	Timestep (BE/FE)	Error (Tri)	Error (FE)	Error (BE)	Time (Tri)	Time (FE)	Time (BE)
0.5000	0.5000	0.2493	0.2618	0.0285	0.0013	0.0151	0.1000
0.2500	0.5000	0.0864	0.2618	0.0285	0.0036	0.0150	0.0993
0.1000	0.2500	0.0486	0.0220	0.0077	0.0156	0.0673	0.5029
0.0500	0.2000	0.0164	0.0133	0.0049	0.0514	0.1118	1.0261
0.0250	0.1429	0.0080	0.0075	0.0029	0.1789	0.2779	4.3809
0.0100	0.1000	0.0034	0.0039	0.0017	1.0379	0.8518	14.0293
0.0050	0.0667	0.0029	0.0013	0.0003	5.4418	3.8377	46.2266
0.0025	0.0500	0.0015	0.0013	0.0008	29.8245	8.3548	209.5588

Exhibit 5: Performance of algorithms for exercise in section 4.2.2. Tri refers to the Hull-White trinomial tree, while FE and BE refer to the Forward Equation and Backward Equation Crank-Nicholson algorithms respectively. For the trinomial tree, we assume that  $\Delta r^* = \sigma\sqrt{3\Delta t}$ , with boundaries set at  $r_0^* \pm \frac{0.184}{0.05\Delta t}\Delta r^*$  as suggested in Hull and White [1994] for mean reversion 0.05. For the Crank-Nicholson method, we use  $\Delta r^* = \sqrt{2}\sigma r_0^\beta \Delta t$ , with boundaries set at  $r_0^* \pm \frac{5\sigma r_0^\beta}{\sqrt{0.1}}$ , consistent with Cairns' analysis, also with a level of mean reversion of 0.05. In both cases, boundary conditions are moved inward to ensure that  $r$  is real and positive for all points. Time taken is in seconds running MATLAB on an Intel Core 2 Duo E8400.

suitable to extrapolation techniques.

## 5 Conclusion

We have presented a methodology for extending the trinomial tree approach to calibrating short rate models to more general classes of finite difference techniques. Our methodology is applicable across a wide range of different models, covering most of the widely used no-arbitrage short rate models. Our method, in contrast to extant approaches has the advantage of being second order accurate in time steps, while only requiring a first order computation cost.

Being able to use higher order methods that are frequently used in equity options pricing results in a far more efficient pricing of securities. Given that many fixed income securities have very long lives (often up to 30 years) computational speed is extremely important. For researchers in this area, the ability to evaluate bond prices quickly is also enormously useful, when dealing with large fixed income data sets, or where the algorithm must be implemented many times, such as when other model parameters must be estimated using time-series data.

## References

- Black, Fischer, Emanuel Derman, and William Toy (1990), A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options, *Financial Analysts Journal* 46, 33–39.
- Black, Fischer and Piotr Karasinski (1991), Bond and Option Pricing when Short Rates are Lognormal, *Financial Analysts Journal* 47, 52–59.

- Cairns, Andrew J. G. (2004), *Interest Rate Models: An Introduction* (Princeton University Press, Princeton, New Jersey).
- Chan, K. C., G. Andrew Karolyi, Francis A. Longstaff, and Anthony B. Sanders (1992), An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance* 47, 1209–1227.
- Crank, J. and P. Nicholson (1947), A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type, *Mathematical Proceedings of the Cambridge Philosophical Society* 43, 50–67.
- Ho, Thomas S.Y. and Sang Bin Lee (1986), Term Structure Movements and the Pricing of Interest Rate Claims, *Journal of Finance* 41, 1011–1029.
- Hull, John and Alan White (1990), Pricing Interest Rate Derivative Securities, *Review of Financial Studies* 3, 573–592.
- (1991), Valuing Derivative Securities Using the Explicit Finite Difference Method, *Journal of Financial and Quantitative Analysis* 25, 87–100.
- (1993), One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities, *Journal of Financial and Quantitative Analysis* 28, 235–254.
- (1994), Numerical procedures for implementing term structure models I: Single-factor models, *Journal of Derivatives* 2, 7–16.
- (2001), The general Hull-White model and supercalibration, *Financial Analysts Journal* 57, 34–43.

- Jamshidian, Farshid (1991), Forward induction and construction of yield curve diffusion models, *Journal of Fixed Income* 1, 62–74.
- Maghsoodi, Yoosef (1996), Solution of the Extended-CIR term structure and bond option valuation, *Mathematical Finance* 6, 89–109.
- Tavella, Domingo and Curt Randall (2000), *Pricing Financial Instruments: The Finite Difference Approach* (John Wiley and Sons, New York).
- Uhrig, Marliese and Ulrich Walter (1996), A new numerical approach for fitting the initial yield curve, *Journal of Fixed Income* 5, 82–90.
- Vetzal, Kenneth R. (1998), An improved finite difference approach to fitting the initial term structure, *Journal of Fixed Income* 7, 62–81.