Available online at www.sciencedirect.com
SCIENCE DIRECT

ELSEVIER

Journal of Economic Dynamics & Control 27 (2003) 1801–1831

JOURNAL OF Economic Dynamics & Control

www.elsevier.com/locate/econbase

Two-factor convertible bonds valuation using the method of characteristics/finite elements ☆

Giovanni Barone-Adesi, Ana Bermudez, John Hatgioannides *

Centre for Mathematical Trading and Finance, City University Business School, Frobisher Crescent, Barbican Centre, London EC2Y 8HB, UK

Abstract

In this paper we solve a two-factor convertible bonds model that fits the observed term structure, calibrates the volatility parameters to market data and allows for correlation between the state variables. We propose the method of characteristics together with finite elements for time and space discretization. An empirical investigation into the pricing of National Grid Group's convertible issue produced prediction errors of less than 5% for 215 successive trading days.

JEL classification: C63; G12; G13

Keywords: Convertible bonds; Method of characteristics; Finite elements; Two-colored options; Correlated Brownian motions

1. Introduction

Convertible bonds are sophisticated financial instruments playing a major role in the financing of companies. Typically, they are corporate debt securities or structured products that offer investors the right to forgo future coupon and/or principal payments in exchange to a specified number of shares of common stock. This hybrid feature of convertible bonds provides investors with the downside protection of ordinary bonds and the upside return of equities and fund managers with asset allocation strategies that take advantage of both fixed-income and equity markets.

Since the important work by Ingersoll (1977a,b) and Brennan and Schwartz (1977) the contingent claims approach (Black and Scholes, 1973; Merton, 1973) to pricing

☆ An earlier version of the paper titled "A General Characteristics/Finite Element Approach for PDE Derivatives Pricing Models. Valuation of Convertible Bonds" was presented at the First SIAM-EMS Conference held at Berlin, September 2–6, 2001.

* Corresponding author. Tel.: +44-20-70408973; fax: +44-20-70408885.
E-mail address: j.hatgioannides@city.ac.uk (J. Hatgioannides).

PII: S0165-1889(02)00083-0

convertible bonds is the definitive choice.[1] Convertible bonds are viewed as derivatives of the underlying equity[2] and, ideally, interest rates since they have long lifespans, meaning that the commonly used assumption of a flat term structure is not valid. As such, the theoretical equilibrium price of a convertible bond is defined as the value that offers no arbitrage opportunity to either the holder or the issuer. Usual provisions such as the possibility of early conversion, callability by the issuer and putability by the holder, make the issuer to follow a call policy (referred to as optimal call) that minimizes the value of a convertible bond and the investor to follow conversion (referred to as optimal conversion) and redemption (referred to as optimal redemption) strategies that maximize the value of the convertible bond at each point in time.

Because of the complexity of convertibles, the resulting pricing equation can be solved only numerically. More often than not however, a number of the challenging features of convertible bonds are assumed away in the literature either due to inherent limitations of the adopted numerical scheme or, simply, in order to reduce the difficulty of implementing and using the model. Thus, the focus of our work is on the practical and commercially usable application of a contingent claim pricing model that can be solved with numerical methods and which incorporates most of the innovative characteristics of convertible products.

In this paper, we extend the previous literature by presenting a characteristics/finite element (CFE) numerical approach combined with a duality method to deal with the early exercise/free boundary problem for solving a two-factor convertible bond pricing model that fits the observed term structure of interest rates and allows for correlation between the state variables. For the first time in the convertibles' literature, we use a variant of Hull and White's (1990) (HW) framework[3] for the

[1] Traditional methods such as "break-even period" analysis, "discount cashflow" analysis and "synthetics" have serious shortcomings as discussed by Cheung and Nelken (1994).

[2] The early work by Ingersoll (1977a,b) and Brennan and Schwartz (1977, 1980) assumed that the firm's value is the (or one of) underlying stochastic variable(s) that the convertible bond depends upon. Although theoretically very attractive, the existence of senior debt, preferred equity and multiple classes of common equity in a typical firm's capital structure makes the valuation of convertibles in such a context intractable in practice. Furthermore, non-synchronous trading and the availability of credible data on non-publicly traded issues pause serious additional problems. King (1986) and Carayannopoulos (1996) discuss alternative simplifications, quite unsatisfactory and impractical in our view, with a view to keep the firm's value as the underlying state variable.

[3] One might correctly argue that the Heath et al. (1992) (HJM) interest rate framework is more general than the HW's. However, given the potential complexity of the calibration and—especially, for American-options'—evaluation procedure within the HJM framework, the latter's comparative advantage over our adopted HW's can be profitably split between one and multi-factor (for the interest rate process alone) implementations. Rebonato (1998, Chapters 13 and 17) shows that the benefits of the HJM approach for one factor interest rate models, are indistinguishable from the HW approach. He carries on by demonstrating that this picture changes radically in moving to multi-factor interest rate approaches, where the HJM approach has a very strong appeal, especially for those users who feel that the options they have to price and risk manage require explicit accountability of the imperfect correlation among interest rates. We believe that the imperfect local correlation among interest rates is of secondary importance to the price of a convertible bond with American-style exercise features, and in any case, it would have required a three-factor model (one stochastic process for equity, two correlated interest rate processes) which would have induced further, perhaps unnecessary for the problem at hand, complexities.

dynamics of the stochastic interest rate process. The latter, (i) incorporates deterministically mean-reverting features, (ii) allows for perfect matching of an arbitrary input yield curve via the introduction of time-dependent parameters, and (iii) permits an exact matching of an arbitrary term structure of volatilities (at least as seen from the present time). To that end, model calibration to simple volatility-dependent instruments such as caps and floors could be carried out in a very efficient way. Coupled with the fact that the other state variable in our convertibles model is the stock price as opposed to the overall firm's value, we may use implied volatilities from stock options to produce a convertible bond pricing framework which is compatible with the current market data for both equity and interest rates.

Previous numerical work has focused on finite difference schemes (see for example, Brennan and Schwartz, 1980; McConnel and Schwartz, 1986; Tsiveriotis and Fernandes, 1998; Nyborg, 1996; Wilmott, 1998; Tavella and Randall, 2000) or lattice methodologies (see for example Cheung and Nelken, 1994, 1995, 1996; Carayannopoulos, 1996; Ho and Pteffer, 1996).[4] However, given the specifications of the financial valuation problem at hand, there are clear advantages of our approach:

First, in the valuation of convertible bonds, a partial differential inequality has to be solved. Conversion, call and put provisions impose inequality constraints in the numerical solution that have to be taken into account in order to avoid arbitrage opportunities. It is a so-called free boundary problem. Given that it is almost always impossible to find a useful explicit solution to any given free boundary problem, the primary aim is to construct efficient and robust numerical methods for its computation. Two possible approaches have been suggested in the relevant literature: (i) linear complementary problems, which are usually linked with finite differences, and (ii) variational inequalities, which involve finite elements. The latter has been followed in our work for three reasons: (i) variational inequalities are the mathematical tool of functional analysis that best suits the rigorous formulation of early exercise problems; (ii) they provide an excellent framework to deal with issues such as existence and uniqueness of the solution; (iii) they are appropriate to analyze the error incurred in the numerical methods (numerical analysis).

The most common method of handling the early exercise condition is simply to advance the discrete solution over a timestep ignoring the restriction and then to apply the constraint explicitly (see for example Clewlow and Strickland, 1998). Although very easy to implement, it has the disadvantage that the solution is in an inconsistent state at the beginning of each timestep, or in other words, a discrete form of the linear complementary problem or the variational inequality is not satisfied. Moreover, Forsyth and Vetzal (2001) compared the efficiency and the accuracy of an implicit penalty method for valuing American-style options with the commonly used technique

[4] It is interesting to notice that discretized versions of continuous-time valuation models can themselves be interpreted as discrete probabilistic models. In fact, Brennan and Schwartz (1977) were the first to show that the explicit finite differences method is equivalent to a binomial lattice approach and the implicit finite differences method corresponds to a multinomial lattice where, in the limit, the underlying variable can move from its (current) value to infinite possible values at next timestep. A proof of this equivalence in a more general setting can be found in Lapeyre et al. (2003).

of handling the American constraint explicitly in the lattice methodologies and they found that the partial differential equation (PDE) method is asymptotically superior to the binomial lattice method, even if the solution is confined at a single point.

In previous numerical work using FD, the treatment of the early exercise in the three guises that it appears in convertibles' valuation (call, put and conversion) is most of the times explicit, with all consequent inaccuracy problems. In lattice methodologies the treatment of American features is always explicit and subject to inherent numerical difficulties (see, Leisen and Reimer, 1996; Baradaran-Seyed, 2000).

In order to deal with free boundary problems we propose an iterative algorithm in which the solution of the variational inequality is approximated by a sequence of solutions of variational equalities. Our algorithm enjoys great generality in that it can accommodate any type of early exercise provisions that may be a function of time and/or state variable(s). More important from a financial perspective is the fact that our algorithm allows to keep tracking the free boundary surfaces for every time step. Therefore, we can not only solve for the convertibles' value at any time until expiration but we can determine ex-ante, for which levels of the underlying asset and the interest rate the embedded conversion, call and put options, will become in-the-money.

Second, contract specifications for convertibles' are very complex and vary a lot across issues; consequently, the FE method provides greater flexibility and has some clear advantages over FD regarding computational practicalities:[5] (i) FE is very suitable for modular programming; (ii) a solution for the entire domain is computed with FE instead of isolated nodes as with FD codes; (iii) FE provide accurate Greeks (risk management parameters) as a by-product; (iv) FE can easily deal with irregular domains, whereas in a FD setting the placing of gridpoints is difficult;[6] (v) FE provide more flexibility in terms of incorporating final conditions (the payoff function of the derivative) and handling boundary conditions. Very often Neumann conditions are easier to obtain than Dirichlet conditions when estimating the behavior of the option as the underlying asset goes to infinity. However, boundary conditions involving derivatives (Neuman conditions) are difficult to handle with FD, whereas FE can incorporate them easily; (vi) FE can easily deal with high curvature. In most FE codes this is achieved by adaptive remeshing, a technique well developed in theory and in practise; (vii) The FE method provides greater flexibility over that of FD in that it allows for unstructured meshing, and therefore better precision via local refinement. As Zvan et al. (1998a) have shown, unstructured meshing can be applied to a wide variety of financial models. The idea is that an accurate solution of the pricing PDE requires in many occasions a fine mesh spacing in certain regions of the domain. Some studies have indicated for example, the need for small mesh spacing near barriers (Figlewski and Gao, 1997;

[5] The distinction between FE and FD is relevant at the theoretical level, i.e. when dealing with issues in numerical analysis. Once the discrete scheme is written and one is left with algebraic transformations of values at the grid points, the distinction vanishes. On structured meshes, FD and FE + numerical integration can be shown equivalent. The contrast should be seen more as variational methods versus FD rather than FE versus FD.

[6] Irregular domains may arise when knock-out barriers are imposed on a multiple-asset option. Also when it is necessary to solve numerically on some parts of the domain only.

Zvan et al., 2000a). Pooley et al. (2000) proved that the finite element method with standard unstructured meshing techniques can lead to significant efficiency gains over structured meshes with a comparable number of vertices. Pironneau and Hecht (2000) present and test the modified metric Voronoi of mesh adaptation for a problem with a free boundary that arises in finance for the pricing of American options, leading to satisfactory results. [7]

Finally, the convertibles' valuation PDE becomes convection dominated (in the sense that convection is big relatively to diffusion) in many regions of the domain. Convection dominance is further reinforced by the theoretically necessary choice of a mean reverting process for the interest rate. It is widely known that in such situations second order centred time-discretization schemes may lead to spurious oscillations. Previous work with explicit finite difference schemes did not make explicit account for the convection dominance. In the lattice framework this is equivalent to saying that the local drift is so large that branching into the usual binomial or trinomial tree will result into negative probabilities. Hull and White (1993) have solved this with their alternative branching technique. In a PDE framework one has to resort to first order upwind time-differencing or to the most recent Eulerian (including flux limiters) and characteristics techniques as described in Ewing and Wang (2001). As they point out, the method of characteristics symmetries and stabilizes the transport PDE, greatly reducing temporal errors. Therefore it allows for large timesteps (an essential feature when dealing with long-dated financial instruments as convertibles) without loss of accuracy. [8]

We do not address, at least formally, the issue of credit risk of the issuer. [9] Credit risk is typically incorporated in a convertibles' model by adding a constant option-adjusted spread or effective credit spread to the riskless interest rate as, for example, in Ho and Pteffer (1996), and use this throughout the normal discounting procedures. This approach is in the spirit of the work by Duffie and Singleton (1999). However, we believe that this approach unnecessarily penalizes the credit risk-free equity upside of the convertible bond. As Tsiveriotis and Fernandes (1998) point out, the value of a convertible bond has components of different default risks: the equity upside has zero default risk since the issuer can always deliver its own stock. On the other hand, coupon, principal payments and any put provisions depend on the issuer's timely access to the required cash amounts—which, crucially are not known in advance—and thus

[7] They use a characteristics/FE method for the space discretization and the Brennan and Schwartz algorithm to deal with American features.

[8] The main drawback of the characteristics method is that is just of order one as opposed to the second order central scheme for the first derivative. However, a high-order characteristic/finite element method has been proposed by Boukir et al. (1997) and could be applied in our case.

[9] The assumption that the value of the convertible depends upon the value of the issuer's common stock precludes the possibility of bankruptcy. Structural valuation approaches, as reviewed by Nyborg (1996), account for credit risk but use the total value of the firm as the stochastic variable and, despite their theoretical appeal, involve many unobservable parameters (notably, the volatility of the firm's value instead of the underlying equity) which make them impractical to use. As we discuss in footnote 2, many simplifications that have been proposed do not tackle the practical issues adequately.

introduce credit risk.[10] Although we can easily incorporate both approaches for credit risk in our empirical model, we feel that it is beyond the scope of the present paper to investigate further the credit risk issue.

Finally, since the method of characteristics/finite elements for solving two-factor arbitrage-free convertible bond models is general, flexible, and uniform enough to be used for pricing a wide array of exotic options (such as Asians, two-colored options, etc.), it may lead to the development of fast and efficient commercial software that obviates the need for a separate numerical technique, and hence, software, to be used—as it is predominantly the case—for each class of exotic option pricing models.

The remainder of this paper is organized as follows. In Section 2 we present our two-factor valuation model for convertible bonds. In Section 3 we solve numerically the theoretical model and we provide theoretical convertible bond prices for different contract specifications. In Section 4 we validate our numerical approach using empirical data. Section 5 concludes the paper.

2. A two-factor convertible bond model

2.1. The theoretical model

Let $V(r,S,t;T)$ be the price of a convertible bond with maturity date $T>t$, which is a measurable function of the underlying stock price S, the spot interest rate r, and time t. The dynamics for equity and term structure are given by the following diffusion processes:

$$\mathrm{d}S = [\mu S - D(S,t)]\,\mathrm{d}t + \sigma S\,\mathrm{d}Z_S, \tag{1}$$

$$\mathrm{d}r = u(r,t)\,\mathrm{d}t + w(r,t)\,\mathrm{d}Z_r, \tag{2}$$

$$\mathrm{E}(\mathrm{d}Z_r\,\mathrm{d}Z_S) = \rho(r,S,t)\,\mathrm{d}t \quad \text{with } -1 \leqslant \rho(r,S,t) \leqslant +1,$$

where μ and σ are the expected rate of return and volatility of the underlying stock,[11] $D(S,t)$ is the dividend yield, and u and w are the expected rate of return and volatility of the spot interest rate which may be time-dependent. This latter feature of the interest rate process distinguishes our model to Brennan and Schwartz (1980), or Longstaff and Schwartz (1995) by ensuring that the bond valuation can be made consistent with the

[10] Tsiveriotis and Fernandes (1998) carry on to define a hypothetical derivative security, “the cash-only part of the convertible bond-COCB” which follows the same dynamics as the convertible’s value. The resulting valuation equation for the COCB should explicitly involve the issuer’s credit spread. On the other hand, the part of the value of the convertible bond related to payments in equity should be discounted using the risk-free rate. Tsiveriotis and Fernandes use a single factor for the pricing of convertible issues—a flat term structure of interest rates is assumed—and they use an explicit finite difference scheme in the numerical solution of their model. Naturally, our general characteristics/finite elements numerical methodology can incorporate their approach as well.

[11] As we mentioned in the introductory section, the key advantage of using equity instead of the overall firm value as one of the underlying factors is that stock price volatility can be inferred from observed option prices.

market time value of money. The two Wiener processes dZ_S and dZ_r are both drawn from normal distributions with zero mean, variance dt, and correlation coefficient ρ.

Following the no-arbitrage arguments by Brennan and Schwartz (1980), it can be shown (see Kwok (1998) or Wilmott (1998)) that the fair value of the convertible bond satisfies the following PDE (in order to keep the notation light, we suppress functional dependencies):

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho S \sigma w \frac{\partial^2 V}{\partial S \partial r} + \frac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + (rS - D(S,t))\frac{\partial V}{\partial S} + (u - \lambda_r w)\frac{\partial V}{\partial r} - rV = 0, \tag{3}$$

where $\lambda_r(r,t)$ is the market price of interest rate risk (see Vasicek, 1977) and appears in the valuation equation because the state variable r is not a traded asset itself.

2.2. The conversion, call and put conditions

A rational investor seeks to maximize the value of the convertible bond at any point in time. Following McConnel and Schwartz (1986), the value of a convertible bond must be greater or equal than its conversion value:

$$V(r,S,t) \geqslant nS, \tag{4}$$

where n is the number of shares of the issuer's common stock into which the convertible can be converted (also known as the conversion ratio).

The optimal conversion condition implies that at each point in time t and each level of the interest rate r there is a particular value of $S = S_f(r,t)$ which marks the boundary between the holding region and the conversion region. We assume that this value is unique and we refer to it as optimal exercise price. This is what is known in the literature as a free boundary problem, similar to the valuation of American-style vanilla options, which gives rise to the following partial differential inequality:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho S \sigma w \frac{\partial^2 V}{\partial S \partial r} + \frac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + (rS - D(S,t))\frac{\partial V}{\partial S} + (u - \lambda_r w)\frac{\partial V}{\partial r} - rV \leqslant 0. \tag{5}$$

When it is optimal to hold the convertible bond, the equality in (3) is valid and the strict inequality in (4) must be satisfied. Otherwise, it is optimal to convert the bond and only the inequality in (5) holds and the equality in (4) is satisfied.[12] The

[12] In the special case where there are no coupons paid on the bond and no dividend paid on the underlying stock, the conversion is not optimal till expiry and the convertible bond can be value explicitly as a combination of cash and a European call option. An increase in the dividend yield makes early exercise more likely, whereas an increase in the coupon payment makes conversion less probable. If the underlying stock pays dividends, before expiry there is a large range of asset values for which the solution of the governing valuation equation (3) is less than the conversion value nS.

free-boundary problem also arises from extra provisions in the convertible bond's indenture agreement. A call feature, which gives the issuing company the right to buy back the convertible issue at any time (or during specified periods, known as intermittent calls) for a specified cash amount (which can be time-varying as well), say M_C, places an upper bound to the convertible's no-arbitrage price.

$$V(r,S,t) \leqslant M_C. \tag{6}$$

In practice however, the call policy followed by managers to induce conversion is not consistent with the theoretical work of Ingersoll (1977a) and Brennan and Schwartz (1977, 1980). Ingersoll (1977b) and Constantinides and Grundy (1987) provide evidence that firms delay calling convertible bonds till long after the market price has exceeded the call price. Jalan and Barone-Adesi (1995) demonstrate that the inequal tax treatment of debt and equity and the need to tap the financial markets, justifies firms to delay calling in order to induce conversion. This allows for a formal linkage between the ex-ante need to issue callable convertible bonds, as a way to increase the residual equity value of the firm, and the observed reluctance to call ex-post. Hence, we will modify the above call condition by writing:

$$V(r,S,t) \leqslant kM_C, \tag{6$'$}$$

where k is a conveniently chosen factor bigger then one.[13]

Similarly, a put feature which gives the right to the holder of the convertible to sell it back to the issuer for a cash amount, say M_P (which can be time dependent), at any time (or again, during intermittent periods) places a lower bound to the convertibles' no-arbitrage price:

$$V(r,S,t) \geqslant M_P. \tag{7}$$

Clearly, convertible bonds with call features worth less than convertibles without. On the contrary, put features increase the value of the convertible to the holder.

Unilateral conditions such as (4), (6), (7) suggest that at each time there are in general two stock prices where downside and upside constraints start becoming binding. These limiting stock prices are unknown and are part of the problem's desired solution; in other words, they are free boundaries beyond which the governing equation (3) does not apply. When the value of the bond is strictly between the up and down bounds, the equality in (3) holds. If the upper bound is reached, the equal sign is replaced by "greater than" and if the lower bound is reached the "less than" sign becomes into place. More precisely, the valuation problem to be solved consists of finding two functions V and P such that:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \rho S\sigma w \frac{\partial^2 V}{\partial S \partial r} + \frac{1}{2} w^2 \frac{\partial^2 V}{\partial r^2} + (rS - D(S,t))\frac{\partial V}{\partial S}$$
$$+ (u - \lambda_r w)\frac{\partial V}{\partial r} - rV - P = 0 \tag{8}$$

[13] In view of usual corporate policy to call back convertibles when its price exceeds by 30% the set call price, in the empirical implementation of our model we will choose $k = 1.3$.

and

$$V(r,S,t) \geqslant nS,$$

$$V(r,S,t) \leqslant M_C,$$

$$V(r,S,t) \geqslant M_P$$

together with final and boundary conditions. P is the Lagrange multiplier which adds or subtracts value in order to ensure that the constraints are being met.

2.3. The term structure model

The most important criticism of the two-factor models by Brennan and Schwartz (1980) and Longstaff and Schwartz (1995) is that they fail to ensure that the convertible bond valuation is consistent with the time value of money observed in the market. To overcome this shortcoming, Ho and Pteffer (1996) proposed a two-dimensional binomial lattice which takes as inputs both the observed Treasury and stock prices. They constructed their quadro-tree (the terminology is due to Cheung and Nelken (1994) who suggested a similar approach) so that when the stock movement is ignored, the two-dimensional lattice is identical to the one-factor, arbitrage-free, term structure model described by Ho and Lee (1986) —hereafter HL—and Black et al. (1990) —hereafter BDT. Besides computational difficulties that arise from the explosive increase in the number of node points in each discrete time step, convergence problems and the choice of a meaningful time step—since convertibles have long life spans, there is a very important drawback of the quadro-tree methodology from a financial perspective: the handling of mean reversion of the interest rate process.

There are two distinct ways of imparting to the spot interest rate process the mean reverting feature which is needed in order to bring about a realistic description of the dynamics of the observed term structure: the first way is to impose a decaying behavior to the diffusive component of the process—this is the approach taken by HL and BDT in their algorithmically constructed lattices; the second is to assign an explicitly mean-reverting component to the deterministic part of the spot interest rate process—this is the approach taken by Hull and White (1990) who extend previous work by Vasicek (1977) and Cox et al. (1985). As Rebonato (1998) points out, it is always possible to choose the parameters of the volatility-decaying process and of the deterministically mean-reverting model in such a way that, as seen from the present time, both distributions will appear identical. The same is no longer true, however, if one considers the distributions obtainable, using the same parameters, from a later time. The volatility-decaying process will produce a new distribution (as seen from the later time) with much lower variance per unit time than it was obtained initially. If the future time step is considerably apart from the present point, in order to obtain a stationary distribution—a distribution whose variance does not grow as time goes to infinity—the forward rate process for the short interest rate would have to be almost deterministic. Clearly this can have serious implications for pricing long-dated American-type options, as they appear in three guises in convertible structures (i.e., convert, call and put).

In order to overcome this important shortcoming of the HL–BDT models evident in quadro-tree approaches, for the first time in the convertibles' literature, we use the Hull and White (1990) framework in our empirical parametrization of the interest rate process (2) which (i) incorporates deterministically mean reverting features for the spot interest rate process, (ii) allows for perfect matching of an arbitrary input yield curve via an introduction of time-dependent parameters, and (iii) permits for an exact conditional calibration to an arbitrary term structure of volatilities.

By setting the risk-neutral interest rate drift $(u(r,t) - \lambda(r,t)w(r,t))$ in (2) equal to $(n(t) - \gamma r)$ we obtain the Hull and White model:

$$\mathrm{d}r = (n(t) - \gamma r)\,\mathrm{d}t + w\,\mathrm{d}Z_r, \tag{9}$$

where w determines the overall volatility of the short rate process and γ determines the relative volatility of long and short rates.

Both γ and w can be inferred from market prices of actively traded interest rate options. Suppose we have a set of M interest rate options, the market price of which we denote by $market_i (i = 1, \dots, M)$. Also assume that there is an interest rate option valuation model that admits closed form solution under the Hull and White specification. Let us write $model_i(\gamma, w)$ for the theoretical option values.

One way to calibrate is to solve the following minimization problem:

$$\underset{\gamma,\sigma}{minimize} \sqrt{\sum_{i=1}^{M} \left(\frac{model_i(\gamma, w) - market_i}{M} \right)^2}.$$

Then, once γ and w have been estimated, we choose $n = n^*(t)$ at a reference time t^* so that theoretical-model-prices and market prices of an array of input discount bonds coincide.

Under the risk neutral process (9), the value of zero-coupon bonds is

$$Z(r,t;T) = \mathrm{e}^{A(t;T) - rB(t;T)} \tag{10}$$

where

$$\begin{aligned} A(t;T) = &- \int_t^T n^*(s)B(s;T)\,\mathrm{d}s \\ &+ \frac{w^2}{2\gamma^2}\left(T - t + \frac{2}{\gamma}\,\mathrm{e}^{-\gamma(T-t)} - \frac{1}{2\gamma}\,\mathrm{e}^{-2\gamma(T-t)} - \frac{3}{2\gamma} \right), \end{aligned} \tag{11}$$

$$B(t;T) = \frac{1}{\gamma}(1 - \mathrm{e}^{-\gamma(T-t)}). \tag{12}$$

In order to fit the yield curve at a reference time t^*, $n^*(t)$ has to satisfy:

$$\begin{aligned} A(t^*;T) &= - \int_t^T n^*(s)B(s;T)\,\mathrm{d}s + \frac{w^2}{2\gamma^2}\left(T - t^* + \frac{2}{\gamma}\mathrm{e}^{-2\gamma(T-t^*)} - \frac{3}{2\gamma} \right) \\ &= \log(Z_M(t^*;T)) + r^*B(t^*;T) \end{aligned} \tag{13}$$

for $Z_M(t^*;T)$ the market price of discount bond expiring at T as of time t^*.

Expression (12) is an integral equation which can be solved by differentiating twice [14] with respect to time $t > t^*$:

$$n^*(t) = -\frac{\partial^2}{\partial t^2}\log(Z_M(t^*;t)) - \gamma\frac{\partial}{\partial t}\log(Z_M(t^*;t)) + \frac{w^2}{2\gamma}(1 - \mathrm{e}^{-2\gamma(t-t^*)}). \qquad (14)$$

3. The numerical solution of the two-factor convertible bond model

3.1. Numerical valuation of convertible bonds

The numerical solution using the CFE method of the valuation PDE for convertible bonds can be considered as a special case of the more general two-colored option pricing problem discussed in Appendix A. More precisely, the governing valuation equation for convertibles in (8) is a special case of equation (A.2) in Appendix A for the choices:

$$x_1 = r,$$

$$x_2 = S,$$

$$A_{11} = \frac{1}{2}w^2,\ A_{12} = A_{21} = \frac{1}{2}\rho\sigma S w,\ A_{22} = \frac{1}{2}\sigma^2 S^2,$$

$$B_1 = u - \lambda_r w,\ B_2 = rS - D(S,t).$$

Moreover, unilateral conditions such as the conversion provision (4), the call provision (6) and the put provision (7) fit into the general form of conditions (A.3)–(A.6) in Appendix A for:

$$R_1(r,S,t) = \max\{nS, M_P\}, \qquad (15a)$$

$$R_2(r,S,t) = \max\{nS, M_C\}. \qquad (15b)$$

Indeed,

$$\max\{nS, M_P\} \leqslant V \leqslant \max\{nS, M_C\}, \qquad (16a)$$

$$\max\{nS, M_P\} < V < \max\{nS, M_C\} \Rightarrow P = 0, \qquad (16b)$$

$$V = \max\{nS, M_P\} \Rightarrow P \leqslant 0, \qquad (16c)$$

$$V = \max\{nS, M_C\} \Rightarrow P \geqslant 0. \qquad (16d)$$

[14] Alternatively, Laplace transform methods can be used.

3.2. A note regarding the unilateral condition

Notice that the restriction

$$nS \leqslant V \leqslant M_C$$

implies

$$S \leqslant \frac{M_C}{n}.$$

If

$$S = \frac{M_C}{n},$$

then

$$nS = V = M_C,$$

and according to the discussion in Section 2, the holder of the convertible will either convert into shares or, upon call, will give the bond back to the issuer to get M_C in cash. In any case, we do not have a convertible product any more. Therefore, if the bond is callable we need to solve just for:

$$S \in \left[0, \frac{M_C}{n}\right]. \tag{17}$$

Thus, if the bond is callable then the spatial domain is the rectangle:

$$\Omega = (0, \infty) \times \left(0, \frac{M_C}{n}\right).$$

The Dirichlet boundary conditions on $S = M_C/n$ would be

$$V(r, S, t) = M_C = nS \quad \text{for } S = \frac{M_C}{n}. \tag{18}$$

However, since M_C may depend on time (by definition or when for example we consider accrued interest), to work with $0 \leqslant S \leqslant M_C/n$ would obey us to change the domain in each time step. In order to avoid that, we extend the solution by nS for $M_C/n \leqslant S \leqslant \infty$. This may be achieved by setting

$$R_2(r, S, t) = \max\{nS, M_C\}.$$

The solution could also be extended by M_C for $M_C/n \leqslant S \leqslant \infty$. The problem is that, in such a case, the lower unilateral restriction, $V(r, S, t) \geqslant nS$, would not be satisfied. Notice that for $nS > M_C$ also $nS > M_P$ and therefore

$$R_1(r, S, t) = nS.$$

This implies that

$$nS \leqslant V(r, S, t) \leqslant nS \Rightarrow V(r, S, t) = nS.$$

If there is no call we set $M_C = \infty$. The domain becomes

$$\Omega = (0, \infty) \times (0, \infty)$$

and

$$R_2(r,S,t) = \max\{nS, M_C\} = \infty,$$

i.e., there is no upper restriction.

If there is no put we set $M_P = 0$. In that way

$$R_1(r,S,t) = \max\{nS, M_P\} = nS,$$

i.e., there is no lower restriction.

3.3. Final, boundary and jump conditions

1. Coupons are paid discretely (typically every quarter or half-year); no-arbitrage arguments lead to the jump condition:

$$V(r,S,t_c^-) = V(r,S,t_c^+) + K(r,t_c), \tag{19}$$

where $K(r,t_c)$ is the amount of discrete coupon paid on date t_c. Such discrete cashflows may be incorporated in the governing valuation equation (8) by adding the Dirac delta function term $-K\delta(t-t_c)$.

2. The final condition for the convertible bond is

$$V(r,S,T) = \max(nS, F), \tag{20}$$

where F is the par value of the bond. If we take into account the embedded options and the possibility of coupons payments, it becomes:

$$V(r,S,T) = \min\{\max\{nS, M_P, F + K(T)\}, M_C\}. \tag{21}$$

Although in (21) we have taken into account call and put provisions, convertible contracts in the market do not allow the holder to put back the bond at expiration. Furthermore, upon call at expiry the issuer pays to the holder not the agreed call price but the redemption value plus the coupon; the same holds if the holder chooses to redeem the bond at its final date to get the principal. Redemption value and face value are not necessarily equal. Therefore, in the numerical implementation we will use:

$$V(r,S,T) = \max\{nS,\ \textit{Redemption Value}\ + K(T)\}. \tag{22}$$

3. At an exceeding high share price, it is almost certain that the bond will be converted. Hence the following boundary condition is considered for $S \to \infty$

$$V(r,S,t) = nS \quad \text{as } S \to \infty. \tag{23}$$

4. At an infinite interest rate, the straight bond component tends to zero and we are left just with the call, the put and the conversion feature. Therefore we should have

$$V(r,S,t) = \min\{\max\{nS, M_P\}, M_C\} = \max\{nS, M_P\} \quad \text{for } r \to \infty.$$

However this definition is not consistent with the extension $V = nS$ for $S > M_C/n$ (It would be appropriate if we extend instead by $V = M_C$). Therefore we have to define

$$V(r,S,t) = \min\{\max\{nS, M_P\}, \max\{nS, M_C\}\} \quad \text{for } r \to \infty. \tag{24}$$

5. At zero share price, the convertible behaves like an ordinary bond:

$$V(r,0,t;T)=W(r,t;T), \tag{25}$$

where $W(r,t;T)$ is the value of the corresponding bond without the convertibility feature, and therefore, in the general case must be found as the solution of a PDE with the instantaneous interest rate as the only spatial variable, and subject to appropriate auxiliary conditions. As Pironneau and Hecht (2000) point out, the above Dirichlet condition for $S=0$ is implicitly defined in the PDE (3). In fact, by setting $S=0$ in Eq. (3) the one-factor valuation for the ordinary bond is obtained. When solving numerically we have considered the natural Neumann condition:

$$\frac{\partial V}{\partial n_A}=0 \quad \text{on } S=0, \tag{26}$$

where [15]

$$\frac{\partial V}{\partial n_A}=\sum_{i,j=1}^{2} a_{ij}\,\frac{\partial V}{\partial x_j}\,n_i.$$

6. It appears quite tricky to specify the boundary condition for very small an interest rate. The boundary condition at zero interest rates it is not clearly specified in any of the works published. Zvan et al. (1998a,b, 2001) proposes a PDE on this boundary, but besides this complicates the numerical solution, no financial justification is given. Wilmott (1998) states that this condition depends on the IR model specification and suggests assuming a finite partial derivative, i.e.,

$$\lim_{r\to 0^+}\frac{\partial V}{\partial r}(r,S,t)<\infty. \tag{27}$$

However this information is not enough when coming to the implementation. Moreover the compatibility between boundary conditions and inequality constraints, due to optimal call, put and conversion (as defined by Brennan and Schwartz (1977) and Ingersoll (1977a)) is not straightforward and yet unspecified in previous work.

We have decided to use the natural Neumann condition,

$$\frac{\partial V}{\partial n_A}=0 \quad \text{on } r=0, \tag{28}$$

where

$$\frac{\partial V}{\partial n_A}=\sum_{i,j=1}^{2} a_{ij}\,\frac{\partial V}{\partial x_j}\,n_i.$$

which in particular implies a finite partial derivative with respect to the interest rate [16] in (27).

[15] The notation for this expression is defined in the appendix.

[16] The Neumann condition provides a good numerical solution which does not differ significantly from the one obtained when a Dirichlet condition is used instead. We may point out here that we have tried both, linear interpolation between the values for zero and infinite asset price and also the intrinsic value of the conversion, call and put embedded options.

Table 1

Maturity	Yield
1 m	0.07
6 m	0.07447
1 yr	0.07016
2 yr	0.06631
5 yr	0.06224
7 yr	0.06121
10 yr	0.06037
30 yr	0.05990

3.4. The system of characteristics

The method of characteristics for time discretization is described in Appendix A for the more general two-colored option valuation problem. It requires the solution of the system of characteristics (A.27a,b) subject to conditions (A.28a,b). For the Hull and White interest rate model (9), if we assume a constant correlation coefficient between the underlying equity and the interest rate process, i.e., $\rho(r,S,T)=\rho$, and a constant dividend yield D_0, the system of characteristics becomes:

$$\dot{\phi}_1(\tau)=\frac{1}{2}\rho\sigma w-n(t)+\gamma\phi_1(\tau), \tag{29a}$$

$$\dot{\phi}_2(\tau)=(\sigma^2-\phi_1(\tau)+D_0)\phi_2(\tau). \tag{29b}$$

The solution of the above system is: [17]

$$\phi_1(\tau_n)=-\delta+\mathrm{e}^{-\gamma\Delta t}[r+\delta]+\mathrm{e}^{\gamma\tau_n}\int_{\tau_n}^{\tau_{n+1}}\mathrm{e}^{-\gamma\tau}n(\tau)\,\mathrm{d}\tau, \tag{30a}$$

$$\phi_2(\tau_n)=S\exp[-(\sigma^2+D_0)\Delta\tau]\times\exp\left[\int_{\tau_n}^{\tau_{n+1}}\phi_1\,\mathrm{d}\tau\right], \tag{30b}$$

where

$$\delta=\frac{1}{2\gamma}\rho\sigma w. \tag{30c}$$

In order to compute (30a–c), we approximate the integrals numerically.

3.5. Contract specifications and theoretical convertible bond prices

We consider here the pricing of a set of theoretical convertible bonds with different contract specifications using our numerical approach. The Hull and White interest rate model (9) is fitted to the term structure as in Table 1: [18] with

$$\gamma=0.1,\ w=0.02.$$

We assume a constant correlation between the spot interest rate and the underlying stock $\rho=0.1$. The volatility of the underlying stock is 15% and it's continuous dividend

[17] Details of these calculations are available upon request.

[18] This theoretical term structure as well as input value are taken from Epstein et al. (2000).

Table 2

Contract characteristics	Value
Zero-coupon CB on non-dividend paying stock	1.0243
Zero-coupon CB on dividend paying stock	1.0000
Coupon-bearing CB on dividend paying stock	1.0869
Coupon-bearing, callable CB on dividend paying stock	1.0498
Coupon-bearing, callable, puttable CB on dividend paying stock	1.0810
Coupon-bearing, puttable CB on dividend paying stock	1.1296

yield is 4%. We value a convertible bond with face value of 1 currency unit, 3.5 years to maturity, which can be converted into 1 unit of the asset and pays a semi-annual coupon of 3%. The bond can be called back at any time for 1.15 and it is continuously puttable for 0.95. The results as shown in Table 2 were obtained for asset value $S=1$ and spot rate $r=7\%$, using 100 steps with spatial domain $[0,2]\times[0,4]$ for non-callable bonds and $[0,2]\times[0,1.15]$ for callable ones, and a regular mesh that takes 40 points not equally spaced on each axis.

The convertible bond value increases monotonically as the issuer's stock price increases and when the spot rate decreases. The value of the convertible declines for dividend paying stocks; this occurs because a higher dividend yield implies a lower expected rate of stock price appreciation and because the value of dividends is not impounded in the bond's price since the convertibles' investor does not receive dividend payments. Adding coupons to the bond increases, as expected, the value of the contract, and makes the probability of conversion lower. The callable feature is valuable to the issuer, hence the convertible decreases in value, whereas when the redemption option is added the contract's value increases; the two effects are however non-symmetric.

4. Empirical results

In order to validate our two-factor convertible bond price model we carry out a comparison of our numerical solution against market quotes for the National Grid $4-\frac{1}{4}\%$ convertible issue (rated Aa3 by Moody's, A+ by S&P) maturing on 17/02/2008.

More specifically, we compare daily market quotes for the convertible's clean price with the projected price given by our model on the 21st of August 2000 for 215 days, i.e., from 21st of August 2000 to 15th of June 2001. As of the starting time in our sample, that is, the 21st of August 2000, the expiration of the convertible expressed in years is

$$T=7.49589.$$

The bond has face value:

$$F=£1,000.00,$$

and can be redeemed at expiration for £1,209.31.

The National Grid's issue can be converted at any time at

$$n=239.8082.$$

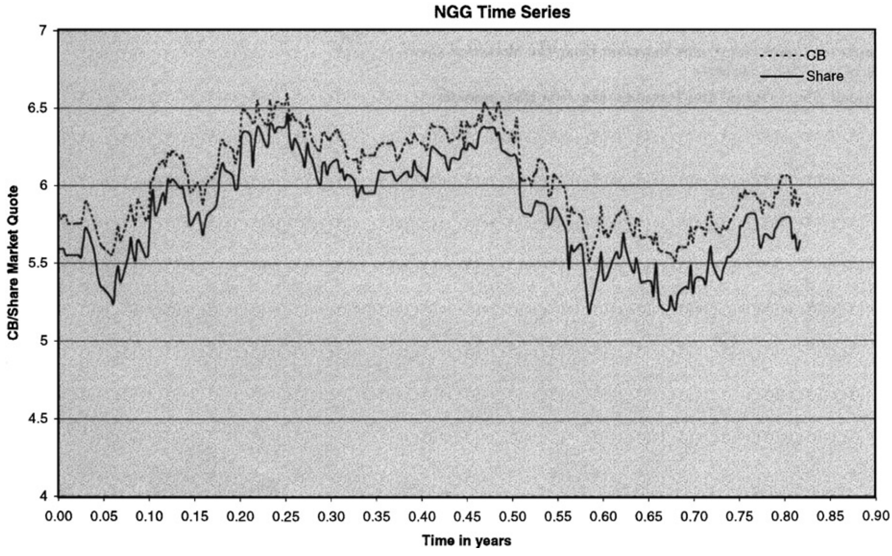

Fig. 1. National Grid share and clean convertible bond market prices from 21st of August 2000 (reference day) to 15th of June 2001 (end of sample). Convertible prices are expressed in a per share basis. Daily data is used throughout.

Table 3

From date	To date	Call price
17-Feb-03	17-Aug-03	108.975
17-Aug-03	17-Feb-04	110.022
17-Feb-04	17-Aug-04	111.100
17-Aug-04	17-Feb-05	112.209
17-Feb-05	17-Aug-05	113.350
17-Aug-05	17-Feb-06	114.525
17-Feb-06	17-Aug-06	115.734
17-Aug-06	17-Feb-07	116.978
17-Feb-07	17-Aug-07	118.258
17-Aug-07	17-Feb-08	119.575

Fig. 1 plots the Convertible bond–National Grid share prices from 21st August 2000 to 15th of June 2001.

Furthermore, the bond is continuously callable at a variety of rates as shown in Table 3.

We have used as a proxy for the instantaneous interest rate the UK spot rate (see Duffee (1996) for an interesting discussion of alternative interest rate series). The historical correlation between the share price of the National Grid Group and the UK spot rate was calculated using daily data for the last five years:[19]

$$\rho = 0.1317.$$

[19] Both time series were obtained from Datastream.

Table 4

T_i	σ_i	R_i (%)
1	10.50	0.064959
2	13.40	0.065984
3	15.40	0.066527
4	16.20	0.066567
5	17.00	0.066677
6	17.10	0.066546
7	17.20	0.066416
8	17.27	0.066222
9	17.33	0.065890
10	17.40	0.065289

As input for the underlying stock's volatility, we have used at-the-money implied volatility ($\sigma_{\rm NG}$) for the vanilla put option on National Grid Group as of the 21st of August 2000:

$$\sigma_{\rm NG} = 35.05\%.$$

The share has an annual dividend yield:

$$D_{\rm NG} = 2.51\%.$$

The Hull and White interest rate model in (9) has been fitted and calibrated to market data as of the 21st of August 2000. Values for the overall interest rate volatility parameter w and the relative (long/short) volatility parameter γ have been chosen using actively traded caps, with tenor of 0.25 years, and with maturities running from 1 to 10 years. Liquid cap data with expiration (T_i), rate (R_i) and at-the-money volatility (σ_i) for the 21st of August 2000 are as shown in Table 4.

The above data set reveals one important advantage of imposing mean-reversion directly in the deterministic part of the interest rate as opposed to the volatility structure. As we discussed in Section 2.2, algorithmically constructed lattices in the spirit of Black et al. (1990) require a decreasing volatility structure for mean reversion of the interest rate to take place. Clearly, this pattern is not evident in the caps data above so the quadro-tree approach of Ho and Pteffer (1996) or Cheung and Nelken (1994) for market-consistent pricing of convertible bonds fails to impose mean reversion in the interest rate process.

We have chosen as inputs of market interest rates, the zero spot curve [20] with expirations ranging from zero to ten years, equally spaced by $\tau = 0.25$ (compatible with the tenor of the caps). Fig. 2 depicts the zero curve.

[20] The zero spot interest rate curve for the 21st of August 2000 was taken from Datastream.

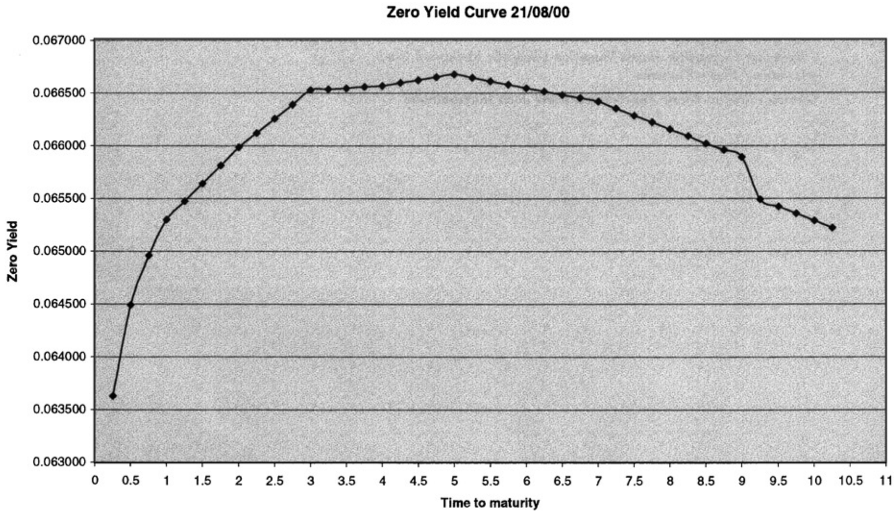

Fig. 2. Zeron spot yield curves for 21st of August 2000 (reference day).

These rates have been used to approximate via cubic splines the logarithm of the market zero bond price of arbitrary maturity t as of reference time t^* (i.e., the 21st of August 2000), $Z_M(t^*,t)$. Once the function $\log(Z_M(t^*,t))$ has been built, $n(t^*,t)$—see expression (14)—can be evaluated and the model is guaranteed to fit observed market bond prices. Note that we do not add a constant credit spread to the riskless term structure as it is usually done in the literature (for example, Ho and Pteffer, 1996). As we discussed in the introductory section, adding a constant option-adjusted spread or effective credit spread to the riskless interest rate penalizes the credit risk-free equity upside potential of the convertible bond. How to account optimally for the credit risk of the issuer is an interesting avenue of further research.

After calibration, the following values were obtained for the interest rate volatility parameters:

$$\gamma = 0.00628,$$

$$w = 0.01025.$$

We are using 2736 daily time steps (since the convertibles' expiration is the 17th of February 2008) in our numerical solution.[21] This provides a clear advantage of our numerical methodology compared to quadro-tree approaches which, because of the inherent explosion of the number of nodes at each time step, can only accommodate a much smaller number of time steps, thus reducing the accuracy of the calculations. We have chosen a spatial domain of $[0,2] \times [0,20]$.

[21] We have considered all inputs in a per share basis, i.e., we have normalised the conversion ratio to unity, and we have divided all other inputs (face value, redemption value, coupons, call price) by the given conversion ratio. On a per share basis, the historical share and bond prices, as well as all other inputs, fall in the range $[0,10]$.

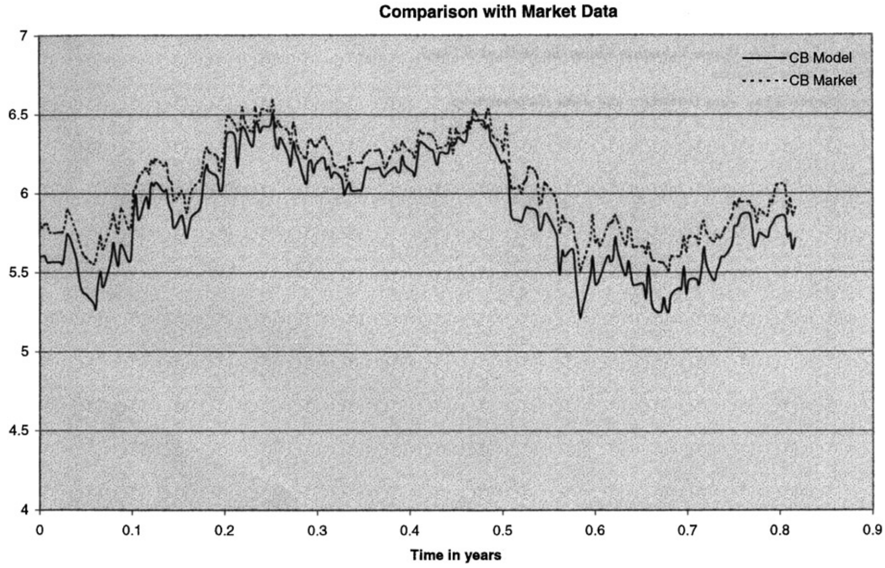

Fig. 3. Market quotes of the NGG convertible bond and model predicted prices from 21st of August 2000 to 15th of June 2001. Both are clean prices and expressed in a per share basis. Daily intervals for 214 trading days.

As we have seen above, the National Grid Group convertible is not callable before the 23rd of February 2003 and afterwards the call price varies with time-to-maturity. Therefore, we have made the up unilateral constraint time dependent.

Fig. 3 plots our numerical valuation results as of the 21st of August 2000 against actual markets quotes for 215 successive trading days.

As it can be seen in the graph, our two-factor model systematically underestimates the market. Two reasons can explain this deviation; first, it is well known that issuers of convertible bonds do not actually follow what we define as rational call policy. Instead, they wait until the share price is well above the call price in order to exercise their right. We have used a value of 30% to account for this delayed call practice by issuers, which of course, is an open matter. Second, we did not take into account in our valuations the accrued interest (*AccIR*) which must be paid by the issuer upon call and by the holder upon put. In that case, the unilateral constraints in expression (16) should read as:

$$\max\{nS, M_P + AccIR\} \leqslant V \leqslant \max\{M_C + AccrIR, nS\},$$

where

$$AccIR(t) = K(t_C^{i+1})\,\frac{t - t_C^i}{t_C^{i+1} - t_C^i},$$

and t_C^i, t_C^{i+1} are successive coupon payments such that $t \in [t_C^i, t_C^{i+1}]$. Omission of the accrued interest clearly underestimates the convertible bond's value.

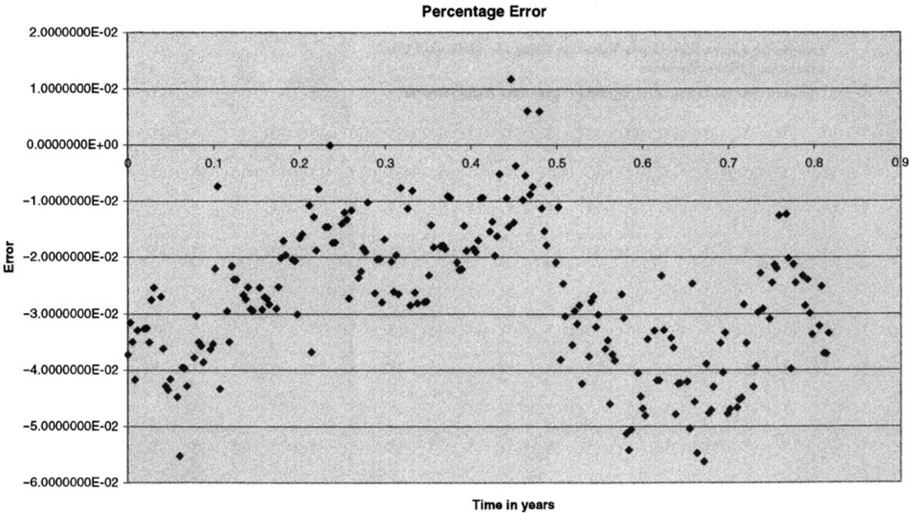

Fig. 4. Percentage predicted errors in the NGG convertible bond from 21st of August 2000 to 15th of June 2001. Daily intervals for 214 trading days.

Overall, our valuation results appear to be very promising. As it can be seen in Fig. 4, almost all of model predictions fall within 5% of market values.

This is a considerable improvement in the accuracy of valuation results compared to the 10% average overpricing and 12.90% overpricing that King (1986) and Carayannopoulos (1996) report, respectively.[22]

5. Conclusion

In this paper we extend the previous literature on the valuation of convertible bonds by solving a two-factor model that fits the observed yield curve, imposes mean reversion in the interest rate process directly in the drift function, calibrates both interest rate and underlying equity volatilities to market observables and allows for correlation between the state variables.

We have applied the method of characteristics together with finite elements for time and space discretization of the convertibles' valuation PDE. There are clear advantages of our numerical scheme compared with the traditionally used finite differences and

[22] We do not claim that we have performed a concise empirical investigation as King (1986) did in the context of a single-factor model and Carayannopoulos (1996) reported in the context of a two-factor convertibles' bond model. However, there are clear advantages of our framework, both theoretical and numerical, over theirs, so it is a matter of further investigation if our very promising results for National Grid extent to other convertible issuers as well. One particular point in support of our methodology is that we provided 215 consecutive days forecasts for the convertibles' price which is far longer and more finely spaced compared to both King's (1986) and Carayannopoulos's (1996) predictions. This is related to the ability of our numerical scheme to accommodate a large number of steps.

lattice methodologies in terms of its (i) flexibility in incorporating final conditions (the payoff function of the contingent claim), boundary conditions (at zero, infinity, or at a barrier) and jump conditions arising from discrete intermediate payoffs of the state variables, (ii) generality to pricing a wide array of exotic options and (iii) accuracy, especially when two-dimensional valuation problems are posited.

We enhance the numerical efficiency of our method by developing an iterative algorithm that approximates the solution of the variational inequality (present in securities with early-exercise features) by a sequence of solutions of variational equalities. Since our algorithm allows keeping track of the free-boundary surfaces for every discrete time step, it provides not just the solution for the price of a convertible bond at any time but also determines ex-ante for which levels of the underlying asset and the short-term interest rate the embedded conversion, call and put option will become in-the-money.

Empirical investigation into the pricing of National Grid Group's convertible issue produced prediction errors of less than 5% for 215 successive trading days, a substantial improvement compared to the 10–12% biases reported in the empirical studies of King (1986) and Carayannopoulos (1996).

A useful direction for future research is to investigate the effect of credit risk on convertible prices and test our methodology for a wide array of convertible issues. Although our approach is general enough to accommodate the usual methodology of adding a credit risk spread to the discounting procedure, we believe that this penalises the upside-equity potential of convertibles which is credit risk-free.

6. For further reading

The following references may also be of interest to the reader: Bliss and Ronn, 1998; Butler, 1995; Butler and Waldvogel, 1996; Crabbe and Nikuolis, 1997; Glowinski and Marroco, 1975; Harrison and Kreps, 1979; Wilmott et al., 1993; Zvan et al., 1997, 2000b.

Acknowledgements

The authors acknowledge insightful comments from Berc Rustem (the editor), three anonymous JEDC referees, and from participants at the First SIAM-EMS Conference. Of course, all errors are the authors' responsibility.

Appendix A. The method of characteristics/finite elements for the numerical solution of two-colored contingent claims

Let the value of a two-colored contingent claim, V, be a function of time, t, and two stochastic variables x_1, x_2 whose evolution is given by the system of stochastic differential equations (SDEs):

$$\mathrm{d}x_i = \mu_j(x_1, x_2, t)\,\mathrm{d}t + \sigma_j(x_1, x_2, t)\,\mathrm{d}Z_j, \quad j = 1, 2,$$

where Z_1, Z_2 are two Wiener processes with correlation coefficient ρ.

Following the standard dynamic no-arbitrage arguments by Black and Scholes (1973), Merton (1973) the following partial differential equation (PDE) for the value of V can de derived:

$$\frac{\partial V}{\partial t} + \sum_{i,j=1}^{2} A_{ij}\frac{\partial^2 V}{\partial x_i\,\partial x_j} + \sum_{j=1}^{2} B_j\frac{\partial V}{\partial x_j} + A_0 V + F = 0 \quad \text{in } \Omega \times (0,T), \tag{A.1}$$

where A_{ij}, B_i, A_0 and F are given measurable functions of x_1, x_2, t. Typically, x_1, x_2 represent quantities such as the underlying asset's value or interest rates that are non-negative, therefore the computational domain Ω can be restricted to be $[0,\infty) \times [0,\infty)$.

In the case that V has to satisfy some unilateral constraints, such as American-early exercise in the form of conversion, call and put provisions for convertible bonds, partial differential inequalities have to be considered. More precisely, the valuation problem to be solved consists of finding two functions V and P such that:

$$\frac{\partial V}{\partial t} + \sum_{i,j=1}^{2} A_{ij}\frac{\partial^2 V}{\partial x_i\,\partial x_j} + \sum_{j=1}^{2} B_j\frac{\partial V}{\partial x_j} + A_0 V + F = P \quad \text{in } \Omega \times (0,T), \tag{A.2}$$

and

$$R_1 \leqslant V \leqslant R_2, \tag{A.3}$$

$$R_1 < V < R_2 \Rightarrow P = 0, \tag{A.4}$$

$$V = R_1 \Rightarrow P \leqslant 0, \tag{A.5}$$

$$V = R_2 \Rightarrow P \geqslant 0 \tag{A.6}$$

together with final and boundary conditions which depend upon the specific derivative product and where $R_1(x_1, x_2, t)$, $R_2(x_1, x_2, t)$ are given functions. P is the Lagrange multiplier which adds or substracts value in order to ensure that the constraints are being met.[23]

In order to introduce the so-called weak formulation of the problem, equation (A.2) must be written in divergence form:

$$\frac{\partial V}{\partial t} - \sum_{i,j=1}^{i,j=2} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial V}{\partial x_j}\right) + \sum_{j=1}^{j=2} b_j\frac{\partial V}{\partial x_j} + a_0 V - F + P = 0, \tag{A.2$'$}$$

where a_{ij} is a symmetric matrix of functions ($a_{ij} = a_{ji}$).

The new coefficients a_{ij}, b_i, a_0 are defined as:

$$a_{11} = A_{11},\ a_{22} = A_{22},\ a_{12} = a_{21} = \frac{1}{2}(A_{12} + A_{21}),$$

$$b_1 = \sum_{i=1}^{i=2}\frac{\partial a_{i1}}{\partial x_i} - B_1 = \frac{\partial A_{11}}{\partial x_1} + \frac{1}{2}\frac{\partial (A_{12} + A_{21})}{\partial x_2} - B_1,$$

[23] Notice that in the region where $P = 0$, the identity in (A.1) holds.

$$b_2 = \sum_{i=1}^{i=2} \frac{\partial a_{i2}}{\partial x_i} - B_2 = \frac{\partial A_{22}}{\partial x_2} + \frac{1}{2}\,\frac{\partial (A_{12}+A_{21})}{\partial x_1} - B_2,$$

$$a_0 = -A_0.$$

We proceed to write a weak formulation of the valuation problem in order to carry out discretization in space using the finite elements method. For this purpose we first introduce boundary conditions. Let us call Γ the boundary of the computational domain Ω, which is assumed to be made up of two parts Γ_R and Γ_D. Let us denote $\vec{n}$ a unit outward normal vector to Γ. Then the following additional equations are considered:

$$\frac{\partial V}{\partial n_A} + aV = G \quad \text{on } \Gamma_R, \tag{A.7}$$

$$V = H \quad \text{on } \Gamma_D, \tag{A.8}$$

where

$$\frac{\partial V}{\partial n_A} = \sum_{i,j=1}^{2} a_{ij} \frac{\partial V}{\partial x_j} n_i.$$

Let $a(t;\cdot,\cdot)$ be the family of bilinear forms:

$$a(t;V,W) = \sum_{i,j=1}^{2} \int_\Omega a_{ij}(x_1,x_2,t)\,\frac{\partial V}{\partial x_j}\,\frac{\partial U}{\partial x_i}\,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_\Omega a_0(x_1,x_2,t) VU\,\mathrm{d}x_1\,\mathrm{d}x_2 + \int a(x_1,x_2,t) VU\,\mathrm{d}\Gamma, \tag{A.9}$$

and let $L(t;\cdot)$ be the family of linear forms

$$L(t;W) = \int_\Omega F(x_1,x_2,t) U\,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_{\Gamma_R} G(x_1,x_2,t) U\,\mathrm{d}\Gamma. \tag{A.10}$$

Also, in order to apply the method of characteristics, we introduce an artificial velocity field $\vec{b} = (b_1, b_2)$ such that we can express the material or total derivative of V with respect to (inverse) time t and the velocity field $\vec{b}$ as

$$\dot{V} = \frac{\partial V}{\partial t} + \vec{b} \cdot grad V. \tag{A.11}$$

The following two equivalent weak formulation forms are considered:

1. Primal formulation in which the Lagrange multiplier P can be eliminated leading to a variational inequality of the first kind:

Find $V(t) \in K(t)$ such that

$$\int_\Omega \dot{V}(t)(W - V(t))\,\mathrm{d}x_1\,\mathrm{d}x_2 + a(t;V(t), W - V(t)) \geqslant L(t, W - V(t))$$

$$\text{for all } W \in K(t); \tag{A.12}$$

2. Mixed formulation which involves the two unknowns V and P:
Find $V(t) \in X$ and $P(t)$ satisfying conditions (2.3)–(2.6), (2.10) such that:

$$\int_\Omega \dot{V}(t) U \, dx_1 \, dx_2 + a(t; V(t), U) + \int_\Omega PU \, dx_1 \, dx_2 = L(t, U) \text{ for all } U \in X_0 \tag{A.13}$$

where

$$X = \{U(x_1, x_2) \in C^0(\bar{\Omega}) \colon U \text{ is continuously piecewise differentiable in } \Omega\},$$

$$X_0 = \{U \in X \colon U_{/\Gamma_D} = 0\},$$

and $K(t)$ is the family of convex sets of functions defined, for each t in $[0, T]$, by:

$$K(t) = \{W(x_1, x_2) \colon R_1(x_1, x_2, t) \leqslant W(x_1, x_2) \leqslant R_2(x_1, x_2, t),$$

$$W(x_1, x_2) = H(x_1, x_2, t) \text{ on } \Gamma_D\}.$$

Theory for the existence of solution for evolutionary variational inequalities can be found in reference books such as Duvaut and Lions (1972), Glowinski et al. (1973), Bensoussan and Lions (1978). Most existence theorems have been proved under the assumption of coerciveness of the bilinear form (A.9). However, it turns out that PDEs arising in finance are usually degenerated because some of their coefficients vanish as any of the independent variables goes to zero. Even if variational theory does not strictly apply in this case, existence and uniqueness of solution for degenerated variational equations can be still examined using the recent technique of viscosity solutions proposed by Crandall et al. (1992). Particular applications to financial PDEs can be found in Barles et al. (1995).

A.1. Solving the free-boundary problem: a Lagrange multiplier method

Taking the mixed formulation of the valuation problem in (A.13) as the starting point, we will show that the solution of a variational inequality can be approximated by a sequence of solutions of variational equalities through a specific iterative algorithm. This algorithm is a particular application of the one introduced by Bermúdez and Moreno (1981) in a more general abstract framework. It has not been applied in finance before, but has been used extensively in other fields such as computational fluid mechanics.

Recall that inequalities (A.3)–(A.6) establish a relation between P and V which can be written in a more compact way by introducing the following family (indicated by x_1, x_2, t) of maximal monotone multivalued functions of Y:

$$G(x_1, x_2, t)(Y) = \left\{ \begin{array}{l} \emptyset \text{ if } Y < R_1(x_1, x_2, t) \\ (-\infty, 0] \text{ if } Y = R_1(x_1, x_2, t) \\ 0 \text{ if } R_1(x_1, x_2, t) < Y < R_2(x_1, x_2, t) \\ [0, \infty) \text{ if } Y = R_2(x_1, x_2, t) \\ \emptyset \text{ if } Y > R_2(x_1, x_2, t). \end{array} \right\}. \tag{A.14}$$

Then it is straightforward to show that inequalities (A.4)–(A.6) are equivalent to the relation:

$$P(x_1,x_2,t) \in G(x_1,x_2,t)(V(x_1,x_2,t)). \tag{A.15}$$

The problem with the above relation is that it is difficult to implement since $G(x_1,x_2,t)$ is multivalued. However we have the following result which characterizes the elements belonging to the graph of the multivalued operator.

Lemma. *The following two expressions are equivalent*

$$\text{(I) } U \in G(x_1,x_2,t)(W), \tag{A.16}$$

$$\text{(II) } U = G_\lambda(x_1,x_2,t)(W + \lambda U) \textit{ for all } \lambda > 0, \tag{A.17}$$

where $G_\lambda(x_1,x_2,t)$ *is the Yosida approximation of* $G(x_1,x_2,t)$ *defined by*:

$$G_\lambda(x_1,x_2,t)(Y) = \left\{ \begin{array}{l} \frac{1}{\lambda}(Y - R_1(x_1,x_2,t)) \textit{ if } Y \leqslant R_1(x_1,x_2,t) \\ 0 \textit{ if } R_1(x_1,x_2,t) \leqslant Y \leqslant R_2(x_1,x_2,t) \\ \frac{1}{\lambda}(Y - R_2(x_1,x_2,t)) \textit{ if } Y \geqslant R_2(x_1,x_2,t) \end{array} \right\}.$$

Proof. See Bermúdez and Moreno (1981). □

The advantage of statement (II) above compared to (I) is that it is an equality. However, as a counterweight, in statement II, U appears also in the right-hand-side of the expression. In view of the above Lemma, relations (A.3)–(A.6) are equivalent to the following equality:

$$P(x_1,x_2,t) = G_\lambda(x_1,x_2,t)(V(x_1,x_2,t) + \lambda P(x_1,x_2,t)), \tag{A.18}$$

where λ is an arbitrarily chosen positive real number.

We are in a position now to introduce the following iterative algorithm:

(1) At the beginning P_0 is given arbitrarily.

(2) At iteration m an approximation of the Lagrange multiplier is known and we proceed as follows: First, we work out a new approximation of V, V_{m+1}, by solving the variational equality:

$$\begin{aligned} &\int_\Omega \dot{V}_{m+1} U \,\mathrm{d}x_1\,\mathrm{d}x_2 + \sum_{i,j=1}^{2} \int_\Omega a_{ij} \frac{\partial V_{m+1}}{\partial x_j} \frac{\partial U}{\partial x_i} \,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_\Omega a_0 V_{m+1} U \,\mathrm{d}x_1\,\mathrm{d}x_2 \\ &\quad + \int_\Omega P_m U \,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_{\Gamma_R} \alpha V_{m+1} U \\ &\qquad = \int_\Omega F U \,\mathrm{d}x_1\,\mathrm{d}x_2 + \int_{\Gamma_R} G U \,\mathrm{d}\Gamma, \end{aligned} \tag{A.19}$$

together with the initial condition

$$V_{m+1}(x_1,x_2,0)=V^0(x_1,x_2). \tag{A.20}$$

Then, we update the Lagrange multiplier P by using equality (A.19). More precisely P_{m+1} is defined as:

$$P_{m+1}(x_1,x_2,t)=G_\lambda(x_1,x_2,t)(V_{m+1}(x_1,x_2,t)+\lambda P_m(x_1,x_2,t)), \tag{A.21}$$

where for convergence, λ has to be greater than some positive value which depends on the coefficients a_{ij},a_0 (see Bermúdez and Moreno, 1981).

We proceed next to solve the general valuation problem numerically. To that end, a discretization must be done, i.e., the problem has to be replaced by a new one with a finite number of degrees of freedom or unknowns. This can be done either for the primal formulation (A.12) or, equivalently, over the mixed formulation (A.13). First, we introduce a semi-discretization in time, by replacing the total derivative of V with respect to time by a two-point formula which involves the value of V at the previous time step evaluated at the position where the material point was one time step ago. Second, we shall carry out space discretization using a finite element method. The combination of both discretization processes is called the characteristics/finite element method or the Lagrange–Galerkin method. In the context of continuum mechanics it has been introduced in the eighties by Benqué et al. (1983), Pironneau (1982), Douglas and Russel (1982). A recent application in finance has been developed by Vázquez (1998) to solve the one-factor valuation problem of vanilla American options. Third, we will apply the above iterative numerical algorithm to the full discretized valuation problem.

A.2. Time discretization: method of characteristics

We have introduced earlier a velocity field $\vec{b}=(b_1,b_2)$ and we identified the total derivative of V with respect to this velocity field, namely

$$\dot{V}(x,t)=\frac{\partial V}{\partial t}(x,t)+\sum_{j=1}^{2}b_j(x,t)\frac{\partial V}{\partial x_j}(x,t). \tag{A.22}$$

The total derivative with respect to time of a scalar field V is defined by:

$$\dot{V}(x,t)=\frac{\partial}{\partial\tau}V(\phi(x,t;\tau))|_{\tau=t}, \tag{A.23}$$

where

$$\tau\longrightarrow\phi(x,t;\tau)$$

represents the trajectory described by the material point that occupies position x at time t.

The trajectories associated to the velocity field $\vec{b} = (b_1, b_2)$ can be found as the solution to the ordinary differential equation (ODE):

$$\dot{\phi}(\tau) = \vec{b}(\phi_1(\tau), \phi_2(\tau), \tau). \tag{A.24}$$

If $\vec{b} = (b_1, b_2)$ is, let us say, a Lipschitz function with respect to $x = (x_1, x_2)$ and continuous with respect to t, the ODE in (A.24) has a unique solution for each final condition

$$\phi(t) = x. \tag{A.25}$$

Let us consider a partition of the time interval $[0, T]$, $0 < t_1 < \cdots < t_N = T$. Then, definition (A.23) suggests the following first-order in time backward approximation of $\dot{V}$ at time t_{n+1}:

$$\dot{V}(x, t_{n+1}) = \frac{V(x, t_{n+1}) - V(\phi(x, t_{n+1}; t_n), t_n)}{t_{n+1} - t_n}. \tag{A.26}$$

Notice that at time step $(n+1)$ and in order to compute $V(x, t_{n+1})$ we have to solve the system of ordinary differential equations (one for each point x in the computational domain Ω):

$$\frac{\partial \phi_1}{\partial \tau}(\tau) = b_1(\phi_1(\tau), \phi_2(\tau), \tau), \tag{A.27a}$$

$$\frac{\partial \phi_2}{\partial \tau}(\tau) = b_2(\phi_1(\tau), \phi_2(\tau), \tau), \tag{A.27b}$$

with the final conditions

$$\phi_1(t_{n+1}) = x_1, \tag{A.28a}$$

$$\phi_2(t_{n+1}) = x_2, \tag{A.28b}$$

on the time interval $[t_n, t_{n+1}]$, backwards in time. In fact, we are interested in the solution just at time t_n, because V must be evaluated just at this point (see expression (A.26)). Details of this numerical method can be found in Bercovier et al. (1982) and Bermúdez and Durany (1987).

A.3. Space discretization: finite elements

Finite elements methods are obtained by restricting the test function involved in the variational formulation to be in a finite dimensional space. This space is usually made up of globally continuous functions that are polynomials in each element of a polygonal mesh of the domain Ω. In the present work, we consider the finite element space consisting of continuous piecewise linear functions on a triangular mesh of the domain Ω. Let us denote by τ_h a family of triangulations of the domain Ω, where the parameter h tends to zero and represents the size of the mesh. Linked to the triangulation τ_h, we define a family of finite-dimensional spaces of functions, namely

$$X_h = \{W_h \in C(\bar{\Omega}) \colon W_{h/K} \in \wp_1, \forall K \in \tau_h\}, \tag{A.29}$$

where, as usual, $C(\Omega)$ denotes the space of continuous functions defined in Ω and $\wp_1$ represents the space of polynomials of degree less or equal than one in two variables. Also, as in the continuous problem, we define:

$$X_{0,h} = \{W_h \in X_h\colon\ W_h(Q) = 0,\ \forall Q \text{ vertex on } \Gamma_D\} \text{ and} \tag{A.30}$$

$$K_h(t_n) = \{W_h \in X_h\colon\ R_1(Q,t_n) \leqslant W_h(Q) \leqslant R_2(Q,t_n) \text{ and } W_h(Q) = H(Q),\ \forall Q \text{ vertex of } \tau_h\}. \tag{A.31}$$

Having chosen the space X_h we can define a discrete counterpart of the problem (A.12) and (A.13). For a more detailed description of this method, see Zienkiewicz (1983).

The last step is to apply the algorithm we have introduced earlier to the full discretized problem.

References

Baradaran-Seyed, F., 2000. Pricing of convertible securities. M.Sc. Dissertation, City University Business School, London.

Barles, G., Daher, C., Romano, M., 1995. Convergence of numerical schemes for problems arising in finance theory. Mathematical Models and Methods in Applied Science 5, 125–143.

Benqué, J.P., Esposito, P., Labadie, G., 1983. New decomposition finite element methods for the Stokes problem and the Navier–Stokes equations. In: Proceedings of Numerical Methods in Laminar and Turbulent Flow Conference, Seattle, WA. Pineridge, Swansea, pp. 553–563.

Bensoussan, A., Lions, J.-L., 1978. Applications des Inéquations Variationneles en Contrôle Stochastique. Dunod, Paris.

Bercovier, M., Pironneau, O., Sastri, V., 1982. Finite elements and characteristics for some parabolic–hyperbolic problems. Applied Mathematical Modelling 7, 89–96.

Bermúdez, A., Durany, J., 1987. La methode des caracteristiques pour les problemes de convection-diffusion stationnaires. Mathematical Modelling and Numerical Analysis 21 (1), 7–26.

Bermúdez, A., Moreno, C., 1981. Duality methods for solving variational inequalities. Computers and Mathematics with Applications 7, 43–58.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Black, F., Derman, E., Toy, B., 1990. A one factor model of interest rates and its application to Treasury bond options. Financial Analyst Journal 46, 33–39.

Bliss, R.R., Ronn, E.I., 1998. Callable US Treasury bonds: Optimal calls, anomalies, and implied volatilities. Journal of Business 71, 211–252.

Boukir, K., Maday, Y., Metivet, B., Razanfindrakoto, E., 1997. A high order characteristics/finite element method for the incompressible Navier–Stokes equations. International Journal for Numerical Methods in Fluids 25, 1421–1454.

Brennan, M.J., Schwartz, E.S., 1977. Convertible bonds: valuation and optimal strategies for call and conversion. Journal of Finance 32, 1699–1715.

Brennan, M.J., Schwartz, E.S., 1980. Analysing convertible bonds. Journal of Financial and Quantitative Analysis 15 (4), 907–929.

Butler, H.-J., 1995. Evaluation of callable bonds: finite difference methods, stability and accuracy. Economic Journal 105, 374–384.

Butler, H.-J, Waldvogel, J., 1996. Pricing callable bonds by means of Green's function. Mathematical Finance 6, 53–88.

Carayannopoulos, P., 1996. Valuing convertible bonds under the assumption of stochastic interest rates: an empirical investigation. Quarterly Journal of Business and Economics 35 (3) 17–31.

Cheung, W., Nelken, I., 1994. Costing the converts. Risk 7 (7), 47–49.

Cheung, W., Nelken, I., 1995. Costing the converts. In: Over the Rainbow, Vol. 46. Risk Publications, pp. 313–317.

Cheung, W., Nelken, I., 1996. Fervour of the convert. AsiaRISK.

Clewlow, L., Strickland, C., 1998. Implementing Derivative Models. Wiley, New York.

Constantinides, G.M., Grundy, B.D., 1987. Call and conversion of convertible corporate bonds: theory and evidence. Working paper, University of Chicago.

Cox, J., Ingersoll, J.E., Ross, S., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–467.

Crabbe, L.E., Nikuolis, P., 1997. The puttable bond market: structure, historical experience, and strategies. Journal of Fixed Income 7, 47–60.

Crandall, M.G., Ishii, H., Lions, P.-L., 1992. User's guide to viscosity solutions of second order partial differential equations. Bulletin American Mathematical Society 27 (1), 1–67.

Douglas, J., Russel, T., 1982. Numerical methods for convection dominated diffusion problems based on combining methods of characteristics with finite element methods or finite differences. SIAM Journal on Numerical Analysis 19 (5), 871.

Duffee, G.R., 1996. Idiosyncratic variation of Treasury Bill Yield. Journal of Finance 51, 527–551.

Duffie, D., Singleton, K.J., 1999. Modeling term structure of defaultable bonds. Review of Financial Studies 12, 687–720.

Duvaut, G., Lions, J.-L., 1972. Les inéquations en mécanique et en physique. Travaux et Recherches Mathématiques 21, 382.

Epstein, D., Haber, R., Wilmott, P., 2000. Pricing and hedging convertible bonds under non-probabilistic interest rates. Journal of Derivatives 7 (4), 31–41.

Ewing, R.E., Wang, H., 2001. A summary of numerical methods for time-dependent advention-dominated partial differential equations. Journal of Computational and Applied Mathematics 128, 423–445.

Figlewski, S., Gao, B., 1997. The adaptive mesh model: a new approach to efficient option pricing, Working paper, Stern School of Business, New York University.

Forsyth, P.A., Vetzal, K., 2001. Quadratic convergence for valuing American options using a penalty method. Discussion paper, University of Waterloo.

Glowinski, R., Marroco, A., 1975. Sur l'approximation par éléments finis d'ordre un, et la résolution par pénalisation-dualité, d'une classe de problèmes de Dirichlet non-lineaires. R.A.I.R.O Modélisation Mathématique et Analyse Numérique R-2, 41–76.

Glowinski, R., Lions, J.-L., Trémolières, R., 1973. Analyse Numérique des Inéquations Variationnelles. Dunod, Paris.

Harrison, J.M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.

Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. Econometrica 60, 77–105.

Ho, T.S.Y., Lee, S.-B., 1986. Term structure movements and pricing interest rate contingent claims. Journal of Finance 42, 1129–1142.

Ho, T.S.Y., Pteffer, D.M., 1996. Convertible bonds: model, value, attribution and analytics. Financial Analyst Journal 52 (5), 35–44.

Hull, J.C., White, A., 1990. Pricing interest rate derivative securities. Review of Financial Studies 3, 573–592.

Hull, J.C., White, A., 1993. Efficient procedures for valuing European and American path dependent options. Journal of Derivatives 1, 21–31.

Ingersoll Jr., J.E., 1977a. A contingent claim valuation of convertible securities. Journal of Financial Economics 4, 289–322.

Ingersoll, J., 1977b. An examination of corporate call policies on convertible securities. Journal of Finance 32, 463–478.

Jalan, P., Barone-Adesi, G., 1995. Equity financing and corporate convertible bond policy. Journal of Banking and Finance 19, 187–206.

King, R., 1986. Convertible bond valuation: an empirical test. Journal of Financial Research 9, 53–69.

Kwok, Y.-K., 1998. Mathematical Models of Financial Derivatives. Springer-Finance, Berlin.

Lapeyre, B., Sulem, A., Talay, D., 2003. Understanding Numerical Analysis for Option Pricing. Cambridge University Press, Cambridge, forthcoming.

Leisen, D., Reimer, M., 1996. Binomial models for option valuation examining and improving convergence. Applied Mathematical Finance 3, 319–346.

Longstaff, F.A., Schwartz, E.S., 1995. Valuing risky debt: a new approach. Journal of Finance 50 (3), 789–819.

McConnel, J.J., Schwartz, E.S., 1986. LYON taming. The Journal of Finance XLI (3), 561–577.

Merton, R.C., 1973. Theory of rational option pricing. Bell Journal of Economics and Management Science 4, 141–183.

Nyborg, K.G., 1996. The use and pricing of convertible bonds. Applied Mathematical Finance 3, 167–190.

Pironneau, O., 1982. On the transport-diffusion algorithm and its application to the Navier–Stokes equations. Journal of Numerical Mathematics 38, 309–332.

Pironneau, O., Hecht, F., 2000. Mesh adaption for the Navier–Stokes equations. East–West Journal of Numerical Mathematics 8 (1), 25–35.

Pooley, D.M., Forsyth, P.A., Vetzal, K.R., Simpson, R.B., 2000. Unstructured meshing for two asset barrier options. Applied Mathematical Finance 7, 33–60.

Rebonato, R., 1998. Interest Rate Option Models. Wiley, New York.

Tavella, D., Randall, C., 2000. Pricing Financial Instruments: The Finite Difference Method. Wiley, New York.

Tsiveriotis, K., Fernandes, C., 1998. Valuing convertible bonds with credit risk. The Journal of Fixed Income 8 (2), 95–102.

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

Vázquez, C., 1998. An upwind numerical approach for an American and European option pricing model. Applied Mathematics and Computation 97, 273–286.

Wilmott, P., 1998. Derivatives: The Theory and Practice of Financial Engineering. Wiley, New York.

Wilmott, P., Dewynne, J., Howison, S., 1993. Option Pricing: Mathematical Models and Computation. Oxford Financial Press, Oxford.

Zienkiewicz, O.C., 1983. Finite Elements and Apro. Wiley, New York.

Zvan, R., Forsyth, P., Vetzal, K., 1997. A finite element approach to the pricing of discrete lookbacks with stochastic volatility. Discussion paper, University of Waterloo.

Zvan, R., Forsyth, P.A., Vetzal, K.R., 1998a. A general finite element approach for PDE option pricing model. Proceedings of Quantitative Finance'98, New York.

Zvan, R., Forsyth, P., Vetzal, K., 1998b. Robust numerical methods for PDE models of Asian options. Journal of Computational Finance 1, 39–78.

Zvan, R., Forsyth, P.A., Vetzal, K.R., 2000a. PDE methods for pricing barrier options. Journal of Economic Dynamics and Control 24, 1563–1590.

Zvan, R., Forsyth, P., Vetzal, K., 2000b. Convergence of lattice and PDE methods for pricing Asian options. Discussion paper, University of Waterloo.

Zvan, R., Forsyth, P.A., Vetzal, K.R., 2001. A finite volume approach for contingent claims valuation. IMA Journal of Numerical Analysis 21, 703–721.