An Application of Finite Elements to Option Pricing

MICHAEL J. TOMAS III*
KISHORE K. YALAMANCHILI

This study applied the finite element method (FEM) to pricing options. The FEM estimates the function that satisfies a governing differential equation through the assembly of piecewise continuous functions over the domain of the problem. Two common representations, a variational functional representation, and a weighted residual representation are used in the application of the method. The FEM is a versatile alternative to other popular lattice methods used in option pricing. Advantages include the abilities to directly estimate the Greeks of the option and allow nonuniform mesh construction. As an illustration of the advantages that the FEM offers, the method was used to price European put options and discrete barrier knock-out put options. © 2001 John Wiley & Sons, Inc. Jrl Fut Mark 21:19–42, 2001

INTRODUCTION

Even though option pricing problems continue to increase in complexity, OTC markets continue to create innovative option structures that preclude analytic solutions. In a study of several popular numerical techniques, Geske and Shastri (1985) concluded that no one method is best;

The authors thank Mark Holder, Hari Krishnan, Robert Webb, and two anonymous referees for their comments.
*Correspondence author, Babson College, Finance Division, Babson Park, Massachusetts 02457-0310; e-mail: mtomas@babson.edu

Received March, 1999; Accepted April, 2000

■ *Michael J. Tomas III is an Assistant Professor at Babson College in Babson Park, Massachusetts.*

■ *Kishore K. Yalamanchili is a Vice President at State Street Research and Management in Boston, Massachusetts.*

The Journal of Futures Markets, Vol. 21, No. 1, 19–42 (2001)

each has advantages and disadvantages in regard to certain applications. Practitioners generally favor lattice-based methods, as they are intuitive and fairly efficient. Finite difference methods and binomial models are generally easy to specify for most option problems with one or two state variables. In most cases, the results from these methods are reasonably accurate for most applications.

Certain exotic options prove difficult for standard lattice-based methods to handle, however. For instance, path dependency or extreme nonlinearity of the pricing function may be difficult to handle with standard lattice methods. This article presents a method, popular in engineering applications, that has advantages in pricing some of these options. This method, the finite element method (FEM), has been used by aerospace and civil engineers to handle structural problems that tax the more traditional finite difference method.

The FEM, like modern option pricing, is a relatively recent field of study. The term *finite element method* was first used in a article by Ray Clough (1960) to describe work done at Boeing several years earlier. In short, the FEM estimates the function that satisfies the underlying differential equation through the assembly of piecewise continuous functions or elements over the domain of the problem.

This method, when properly applied, has advantages over other discretization methods (e.g., the method of finite differences). Some advantages of the FEM (vs. the standard finite difference method) include its flexibility to effectively accommodate irregular domains (in particular, nonuniform meshes), the ability to directly estimate derivative values (delta and gamma), and the ability of higher order elements to better estimate nonlinear pricing functions. The purpose of this article is to demonstrate the approach of the FEM to option pricing and to illustrate some of the advantages described previously. In addition to applying the method to the well-known Black–Scholes partial differential equation (Black & Scholes, 1973), a pedagogical benchmark, as a second illustration we price discrete barrier options (as per Ahn, Figlewski, & Gao, 1999), highlighting the cost savings that can be achieved with a nonuniform mesh application of the FEM.

FOUNDATIONS AND RELATIONS TO OTHER WORK

History of Finite Element Analysis

Just as the field of option pricing is relatively new, so too is the numerical technique of finite element analysis. The exact start of the FEM is diffi-

cult to pinpoint, as several facets of the development of the method were happening simultaneously in different places. Many theories and techniques from mathematics, continuum mechanics, and structural analysis went into the development of the FEM. Courant (1943) used the idea of piecewise polynomial interpolations over triangular subregions to study the Saint-Venant torsion problem. However, the work was somewhat impractical, as actual application on a reasonable scale had to wait until the advent of digital computers. There is some consensus that the FEM as applied to engineering problems began in the 1950s. Turner, Clough, Martin, and Topp (1956) developed the technique to evaluate the stiffness of a delta airplane wing while at Boeing.

The first use of the term *finite element method* can be traced to an article by Clough (1960), which summarized the work done at Boeing. In 1963, the FEM was recognized as having sound mathematical foundations; the FEM can be regarded as the solution of a variational problem through the minimization of a functional (Cook, Malkus, & Plesha, 1989). During the mid-1960s, the FEM was extended to problems of heat transfer and fluid flow (Allaire, 1985).

The FEM as it applies to option pricing was briefly discussed by Wilmott, Dewynne, and Howison (1993), who outlined the method with simple linear functions (elements) and showed that the solution with such functions would be the same as that of the standard finite difference method. Their work did not reveal some of the benefits that the FEM has over the conventional finite difference method, however (it was not the focus of the study). Tomas (1996) applied the FEM, by means of the variational functional approach, to price European and American puts and European and American barrier options (continuous barrier). However, his work failed to show the advantages that the FEM has near irregular boundaries (nonuniform mesh construction), and it was limited to the variational form (which, as discussed later, cannot be applied directly to parabolic partial differential equations of the form seen in most option pricing applications).

In this article, more complex elements are presented that allow for the direct calculation of piecewise continuous delta and gamma functions. In addition, this study applies the FEM with two approaches: (a) the variational functional approach as discussed by Wilmot et al. and (b) the more general weighted residual approach (in its Galerkin form). Also, we illustrate the ease with which the FEM can handle nonuniform mesh construction to reduce the number of computations needed for a given level of accuracy.

Other Numerical Techniques in Option Pricing

Several techniques that have become popular for the pricing of option problems include (a) the binomial method, (b) the method of finite differences, (c) simulation, and (d) numerical integration.

Cox, Ross, and Rubinstein (1979) developed the binomial model as an early alternative to the closed-form solutions presented by Black and Scholes (1973). The binomial model is an elegant and conceptually appealing model for valuing options. Since Cox et al.'s article was published, a number of extensions, modifications, and applications have been put forth. For example, Boyle (1988) extended the binomial model to price options where there are two underlying state variables. Hull and White (1988) presented a generalized multivariate multinomial extension to Cox et al.'s binomial model for pricing American options. Boyle, Evnine, and Gibbs (1989) also presented an extension to the binomial model to value multivariate contingent claims. More recently, Thompson (1995) extended the binomial framework to value claims with multiple exercise decisions over time. More recently, Figlewski and Gao (in press) and Ahn, Figlewski, and Gao (1999) illustrated a new method for pricing options, the adaptive mesh model. Their method greatly reduces nonlinearity error by grafting one or more small sections of fine mesh onto a tree with coarser time and price steps.

The option pricing models developed by Black and Scholes (1973) and Merton (1973) were solutions to partial differential equations with boundary conditions. Many option pricing models developed since then can also be expressed as partial differential equations with boundary conditions. Garman (1976) and Cox, Ingersoll, and Ross (1985) showed that under the assumption that the underlying asset follows a generalized Wiener process, the option pricing problem will be the solution to a resulting parabolic partial differential equation. One technique, the method of finite differences, has long been popular in mathematics, physics, and engineering for solving such equations.

Brennan and Schwartz (1978) provided a discussion of both the explicit and implicit finite difference methods as they relate to one-dimensional option pricing problems. Geske and Shastri (1985) provided a comparison of several lattice approaches, including the explicit and implicit finite difference methods, concluding that the explicit finite difference method with log transformation is the most efficient method for valuing large numbers of options. Hull and White (1990) modified the explicit finite difference method to guarantee convergence when valuing derivative securities.

Another numerical method for pricing options and an alternative to the discretization methods discussed previously is that of Monte Carlo simulation. Boyle (1977) described the Monte Carlo method and its application to option pricing. Recent extensions to the basic Monte Carlo method include works such as Barraquand and Martineau's (1995) article, in which a partitioning method was combined with the simulation method to price American options.

Many options problems have solutions that can be expressed in integral form, the integration of which may be difficult or tedious. Numerical integration has been used in several option pricing studies. For instance, Parkinson (1977) used numerical integration to value American put options. Kim (1990) derived a valuation formula for American options written on futures contracts. After the optimal exercise boundaries were established, the solution to his problem was achieved with numerical integration. Chen and Scott (1992) priced interest rate options based on a two-factor Cox–Ingersoll–Ross model (Cox et al., 1985). Their initial solution was expressed as multivariate integrals. They showed how to modify the problem so that univariate numerical integration could be used to achieve a solution.

All these methods are used in problems where estimating the solution is often too difficult to arrive analytically. The FEM is closest to the method of finite differences, as we discuss next.

AN OVERVIEW OF THE FEM

The FEM is similar in spirit to the method of finite differences. Both are discretization methods that attempt to estimate values for an unknown function that satisfies a differential equation or a partial differential equation. Both attempt to estimate values of the function at discrete points (*nodes*) over the domain (or subset of the domain) of the problem. However, beyond this the methods differ considerably.

In the FEM, the partial differential equation (*strong form*) itself is not estimated, as with the finite difference method. Like the finite difference method, the set of differential equations in terms of unknown variables is replaced with a related but approximate set of algebraic equations where the unknown variables are evaluated at nodal points. With the FEM, an equivalent integral formulation of the differential equation, or *weak form* (known as a *variational functional* or *variational formulation*), is often used in the estimation. Finding the unknown function that minimizes the integral expression is equivalent to satisfying the partial differential equation. Another popular approach, known as the *method of*

weighted residuals, is applied in situations where it is difficult or impossible to find a variational functional form. In this study, both methods were employed to price options on the basis of the Black–Scholes (Black & Scholes, 1973) differential equation.

In the FEM, the domain of the problem is divided into subdomains or finite elements. Each finite element relates values in the domain of the element to nodal values on the boundary and interior of the element. Commonly, the relationship between nodal values and values in the element domain are expressed as polynomial interpolation functions or shape functions.

The FEM is done in such a way as to ensure (at least) the approximating function is continuous where the elements meet (*nodal boundary points*). Element shape functions may be specified to allow continuity of the derivatives for the function as well. The ability of the FEM to easily and accurately estimate values for the derivatives of the functions is an advantage that the FEM has over other numerical techniques, such as the finite differences method. Differencing solution values from the binomial and finite difference method can lead to errors in the estimates for delta and gamma (Pelsser & Vorst, 1994).

The FEM results in a system of equations, the solution of which yields a function whose values approximate those of the exact solution. If the set of differential equations is linear, the resulting system of equations will be linear and solvable with matrix techniques. The solution will also be stable and convergent (if the problem is properly specified).

FINITE ELEMENT SUBDIVISION AND APPLICATION

In the FEM, the solution function u to the set of differential equations is replaced with an approximate function of the form:

$$u \approx \sum_{j=1}^{N} u_j N_j \tag{1}$$

where u_j is the approximate value of u (or possibly derivative) at each element node and N_j is the interpolation function (usually a piecewise polynomial function). A difference between the FEM and other numerical methods, such as the method of finite differences or classical variational methods, is that the approximation for u is established on subdomains or finite elements.

Element Formulation

Over each element, the function of interest, u, is related to nodal points, u_i (for a finite number i), through interpolation functions, also known as *shape functions*. The number and location of nodal points per element will depend on the geometry of the element (dimensions, shape, and properties of the problem to be retained) and the type of interpolation. Generally, the interpolation used is a polynomial expansion based on the coordinate system (known variables). The number of terms used in the polynomial expansion will determine the number of nodal unknowns needed. The shapes, dimensions, and properties of the problem that need to be retained will help determine the location of the nodes for the element.

As mentioned before, the function that satisfies differential equations is estimated with a set of equivalent algebraic equations. The two most common approaches used in the FEM are (a) the variational functional approach and (b) the weighted residual approach. Both are presented in turn in the application section.

FINITE ELEMENT SHAPE FUNCTIONS

Although the FEM involves the establishment of a system of equations in the solution process, the heart of the method is the discretization process. The FEM discretization makes assumptions about how the unknown function behaves on a regional (element) basis. It is not surprising, then, that much of the development of the FEM has been in the derivation of different families of finite element shape functions.

Three general families of polynomial elements exist. The first two families, *Lagrange elements* and *serendipity elements,* are similar in that for all orders of shape functions, only the continuity of the function at the nodal boundaries is enforced. Slope values are modeled at least as a constant but in increasing order as the shape functions become more complex. The difference in the Lagrange and serendipity family of elements is most easily seen in higher dimension problems (two dimensions and greater), where the number of nodes used in the elements differ. In higher dimensional problems, the Lagrange family of elements has more interior nodes on the domain of the element.

The last family is *Hermite elements.* In the Hermite family of elements, the continuity of the function and the derivative(s) of the function with respect to the coordinate(s) (the dependent variables) are modeled. The difference between Hermite elements and the Lagrange–serendipity

family is that for the latter, each node represents the unknown function, whereas with the Hermite family, each node represents the unknown function and one or more derivative values. In the case of option pricing, this feature allows the FEM to model the price, delta, and gamma functions simultaneously.

As an illustration, a simple linear Lagrange–serendipity element, a cubic Lagrange–serendipity element, and a fifth-order Hermite element were used in this study. The shape functions are presented in the Appendix A.

A natural question to ask is whether a large number of simple elements is better than a small number of complex elements. Usually, for the same mesh, the higher the order of the elements used is, the better the approximation will be. However, total economy is an important consideration. Although more complex elements may reduce the number of equations to be solved (e.g., two quadratic elements may be more efficient than five linear elements), they also require greater computational time in the initial formulation of the system of equations. Therefore, the tradeoff can be very problem dependent.

APPLYING THE FEM

This section presents the application of the FEM to the standard option pricing problem with one underlying asset. This section illustrates the FEM in the pricing of European put options under the Black–Scholes (Black & Scholes, 1973) framework. Later in the section, we apply the FEM to price discrete barrier options and illustrate the computational savings that can be achieved with a nonuniform mesh.

Variational Functional Approach for the Black–Scholes Equation

In this method, the differential equation, or strong form, is represented by a scalar quantity or functional (Π) weak form, which is defined in an integral form:

$$\Pi(u) = \int_v F(u, \ldots)dV$$

A strong form enforces conditions at all points, whereas a weak form enforces conditions in an average or integral sense. Direct substitution of finite elements to the differential equation will not usually result in a

properly defined system of equations (the system will not have a unique solution). Application to the weak form does allow for a system of equations with a unique solution, however. The integrand F contains terms in u and its derivatives, but the derivatives will usually be of a lower order than the differential equation itself.

Minimization of the functional with respect to the unknown function is equivalent to satisfying the differential equation. In variational calculus terms, the solution sought is a function u such that for small changes in u (δu), the value of Π remains the same (is stationary). Thus, the solution in the continuum problem results in a variation of zero. The difficulty with this method lies in finding a functional associated with a continuum problem.

The original Black–Scholes (Black & Scholes, 1973) equation does not have a variational form because of the first-order derivative term in the state variable. However, with a transformation of variables, the Black–Scholes differential equation can be reformulated as a one-dimensional heat equation with the thermal diffusivity constant equal to 1. Fortunately, a functional form for the heat equation does exist and is used in the FEM formulation as discussed later.

Recall that the Black–Scholes (Black & Scholes, 1973) differential equation for a put option is of the form

$$\frac{\sigma^2}{2} S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP + \frac{\partial P}{\partial t} = 0 \tag{2}$$

with the boundary conditions

$$P(0, t) = E$$
$$P(S, T) = \max(E - S, 0)$$

P is the value of the put option, S is the value of the stock, E is the exercise price, r is the risk-free interest rate, σ is the volatility of the stock price, and t is the time to maturity. The domain of the problem includes $S \geq 0$ and all $T \geq 0$.

With the transformation outlined in Wilmott et al. (1993), the equation can be rewritten as

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial \tau} = 0 \tag{3}$$

with the following boundary at maturity:

$$u(x,\ 0)\ =\ \max(e^{1/2(k_1-1)x}\ -\ e^{1/2(k_1+1)x},\ 0) \tag{4}$$

The transformed equation now has the domain $-\infty \leq x \leq \infty$ and $\tau \geq 0$.[1]

When applying the FEM, we use the variational functional form of the Laplace differential equation, which has the following formulation (see Allaire, 1985, for methods of derivation):

$$\Pi\ =\ \int\left[\frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\ +\ (u)\left(\frac{\partial u}{\partial \tau}\right)\right]dx \tag{5}$$

Finite Element Formulation for the Variational Approach

The FEM seeks a solution to the previously mentioned variational problem by constructing an approximate (discrete) problem for which a piecewise continuous function can be estimated. We begin by dividing the continuous domain of the real problem into finite elements. Each element will be a small piece of continuum connected to another element only at its nodes. As stated before, the function u will be represented by nodal values u_i, from which values on the interior of the element can be interpolated. We can express the interpolating relationship for any element as

$$u\ =\ [\mathbf{N}]^T\ [\mathbf{q}] \tag{6}$$

where [**N**] is the vector of shape functions (to be discussed shortly) and [**q**] is the vector of nodal option values (and possibly derivative values). The number and types of terms in [**N**] and [**q**] will depend on the family of elements and the polynomial order of interpolation chosen. Therefore, as part of this process, the choice of element type, number of elements, and dimensions of the element are necessary.

Appendix A lists the shape functions with respect to a natural coordinate ξ for each of the elements listed previously. Every element has

[1]Additionally, it will be useful for to know the following relationships when solving for Hermite elements:

$$\frac{\partial P}{\partial S}\ =\ (e^{(a-1)x+\beta\tau})\left(\alpha u\ +\ \frac{\partial u}{\partial x}\right)$$

$$\frac{\partial^2 P}{\partial S^2}\ =\ \frac{1}{E}\ (e^{(a-2)x+\beta\tau})\left(\frac{\partial^2 u}{\partial x^2}\ +\ (2\alpha\ -\ 1)\ \frac{\partial u}{\partial x}\ +\ (\alpha^2\ -\ \alpha)u\right)$$

where

$$\alpha\ =\ \frac{-(k_1\ -\ 1)}{2},\ \beta\ =\ \frac{-(k_1\ +\ 1)^2}{4},\ \text{and}\ k_1\ =\ \frac{2r}{\sigma^2}$$

endpoints that lie at $\xi = \pm 1$, regardless of its physical length. To rescale the element from ξ to x, a simple linear relationship is established as follows:

$$\xi = \mu x + \beta$$
$$\frac{\partial \xi}{\partial x} = \mu$$

The values for μ and β can be obtained with the ending nodal values for the element (Appendix B contains pseudocode for a cubic element in natural coordinates). Using this information, we have the following:

$$\mu = \frac{2}{b - a}$$
$$\beta = \frac{-(a + b)}{b - a}$$

where a and b are the values for x at the ends of the element. The value of the derivative of ξ with respect to x is $\partial\xi/\partial x = \mu$ (this is also known as the *inverse of the Jacobian*). This relationship is used shortly.

The integrand in Equation 5 contains terms in μ and $\partial\mu/\partial x$ (the term $\partial\mu/\partial\tau$ is ignored for now). Using the shape functions, we can establish a relationship between u and its nodal values contained in $[\mathbf{q}]$. We can similarly define a relationship for the slope term as

$$\frac{\partial u}{\partial x} = \left[\frac{\partial N}{\partial x}\right]^T [\mathbf{q}] \tag{7}$$

Recall that we defined the shape functions with respect to the standardized coordinate ξ. Using the rescaling relationship between ξ and x combined with

$$\frac{\partial N_i}{\partial x} = \frac{\partial N_i}{\partial \xi}\frac{\partial \xi}{\partial x}$$

allows us to rewrite Equation 7 as

$$\frac{\partial u}{\partial x} = \left[\frac{\partial N}{\partial x}\right]^T [\mathbf{q}] = \left[\frac{\partial N}{\partial \xi}\frac{\partial \xi}{\partial x}\right]^T [\mathbf{q}] = \mu\left[\frac{\partial N}{\partial \xi}\right]^T [\mathbf{q}] \tag{8}$$

We are now ready to substitute the finite element formulation into the variational functional form of the problem. For any particular element, the functional becomes

$$\prod = \frac{1}{2}[\mathbf{q}]^T[\mathbf{g}][\mathbf{q}] + [\mathbf{q}]^T[\mathbf{h}][\dot{\mathbf{q}}] \tag{9}$$

where

$$[\mathbf{g}] = \int_a^b \left[\frac{\partial N}{\partial x}\right]\left[\frac{\partial N}{\partial x}\right]^T dx = \int_{-1}^1 \left[\frac{\partial N}{\partial \xi}\right]\left[\frac{\partial N}{\partial \xi}\right]^T \mu^2 \frac{d\xi}{\mu}$$

$$[\mathbf{h}] = \int_a^b [\mathbf{N}][\mathbf{N}]^T dx = \int_{-1}^1 [\mathbf{N}][\mathbf{N}]^T \frac{d\xi}{\mu}$$

and $[\dot{\mathbf{q}}]$ represent the derivative of u with respect to τ at various nodal values.

In the variational formulation, we are interested in values of Π such that for small changes in u, the value of Π does not change; this can be expressed as

$$\frac{\partial \Pi}{\partial u_i} = 0$$

Taking the derivative of Equation 9 with respect to all the μ_i and rearranging yields, we obtain

$$\left\{\int_{-1}\left[\frac{\partial N}{\partial \xi}\right]\left[\frac{\partial N}{\partial \xi}\right]^T \mu^2 \frac{d\xi}{\mu}\right\}[\mathbf{q}] + \left\{\int_{-1}^1 [\mathbf{N}][\mathbf{N}]^T \frac{d\xi}{\mu}\right\}[\dot{\mathbf{q}}] = 0 \tag{10}$$

To handle the time derivative, a finite difference scheme in the time dimension is employed. Discretizing $[\dot{\mathbf{q}}]$, we have

$$\left\{\int_{-1}^1 \left[\frac{\partial N}{\partial \xi}\right]\left[\frac{\partial N}{\partial \xi}\right]^T \mu^2 \frac{d\xi}{\mu}\right\}[\mathbf{q}] + \left\{\int_{-1}^1 [\mathbf{N}][\mathbf{N}]^T \frac{d\xi}{\mu}\right\}[\mathbf{q} - \mathbf{q}^*]\frac{1}{\Delta\tau} = 0$$

Rearranging results in

$$\left\{\int_{-1}^1 \left(\left[\frac{\partial N}{\partial \xi}\right]\left[\frac{\partial N}{\partial \xi}\right]^T \mu^2 = \frac{1}{\Delta\tau}[\mathbf{N}][\mathbf{N}]^T\right)\frac{d\xi}{\mu}\right\}[\mathbf{q}]$$

$$= \left\{\int_{-1}^1 \frac{1}{\Delta\tau}[\mathbf{N}][\mathbf{N}]^T \frac{d\xi}{\mu}\right\}[\mathbf{q}^*] \tag{11}$$

At the initiation of the finite element solution, $\mathbf{q}^*$ will be the values for the option at maturity. Afterward, they will be the solved values of $\mathbf{q}$ in succession. Equation 11 is of the form

$$[k]_{\text{element}}\ [\mathbf{q}]_{\text{element}} = [\mathbf{Q}]_{\text{element}} \tag{12}$$

The matrix $[k]_{\text{element}}$ is a coefficient matrix whose values are dependent on the coordinate system and the form of the functional; $[\mathbf{Q}]_{\text{element}}$ is the vector of nodal right-hand side values (which are known). This provides a system of linear equations over the domain of one element.

Galerkin Approach to the Black–Scholes Equation

In this approach, a residual based on the differential equation is formed. For a differential equation of the form

$$Lu - f = 0$$

where L is a differential operator, u the dependent variable, and f a function of the independent variable or a constant, a residual can be defined as

$$R(a, x) = L\tilde{u}(a, x) - f$$

where $\tilde{u}$ is a discretized representation of u. The values of a must be determined to make the residual as close to zero as possible. Several different methods of weighted residuals exist (Cook et al., 1989). One that is very popular and is used in this article is Galerkin's method. In this method, weight functions, $W_i(x)$, are selected to set the residual to zero, resulting in equations of the form

$$\int_x W(x)R(a, x)dx = 0$$

If the residuals are small, $\tilde{u}$ is considered a good solution. If $\tilde{u}$ is an exact solution to u, the residual will be zero at all points.

The advantage of residual methods is that they are applicable to all differential equations. The resulting coefficient matrix will not be symmetric, however, as with the variational approach.

To apply Galerkin's method, we multiply the original Black–Scholes (Black & Scholes, 1973) differential equation by weight functions and minimize the residual of this equation as

$$\int W\left(\frac{\sigma^2}{2} S^2 \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP + \frac{\partial P}{\partial t}\right)dS = 0 \tag{13}$$

where the integration is over the domain of interest.

It is standard practice to begin with integration by parts to reduce the order of differentiation in the integrand, which allows the use of linear shape functions. If we divide through by S^2, the first term in parenthesis in the integrand can be written as

$$\frac{\sigma^2}{2}\int W\frac{\partial^2 P}{\partial S^2}\,dS = \frac{\sigma^2}{2}\,W\frac{\partial P}{\partial S}\Big]_{\text{boundary}} - \frac{\sigma^2}{2}\int\frac{\partial W}{\partial S}\frac{\partial P}{\partial S}\,dS$$

With the proper choice of weight functions, the first term on the right hand side of the equation becomes zero. Other terms in the integrand do not require integration by parts, as they are of low enough order. This allows the full expression to be written as

$$\int -\frac{\sigma^2}{2}\frac{\partial W}{\partial S}\frac{\partial P}{\partial S} + r\frac{1}{S}W\frac{\partial P}{\partial S} - rW\frac{1}{S^2}P + W\frac{1}{S^2}\frac{\partial P}{\partial t}\,dS = 0 \quad (14)$$

Finite Element Formulation for the Galerkin Approach

As with the variational functional form, we can substitute the finite element formulation into Equation 14. As before, the time derivative is handled via an implicit difference scheme. This results in the following equation:

$$\left\{\int_{-1}^{1}\left(-\frac{\sigma^2}{2}\left[\frac{\partial N}{\partial\xi}\right]\left[\frac{\partial N}{\partial\xi}\right]^T\mu^2 + \frac{r}{S}[\mathbf{N}]\left[\frac{\partial N}{\partial\xi}\right]^T\mu - \left(r - \frac{1}{\Delta t}\right)\right.\right.$$
$$\left.\left.\frac{1}{S^2}[\mathbf{N}][\mathbf{N}]^T\right)\frac{d\xi}{\mu}\right\}[\mathbf{q}] = \left\{\frac{1}{\Delta t}\int_{-1}^{1}\frac{1}{S^2}[\mathbf{N}][\mathbf{N}]^T\frac{d\xi}{\mu}\right\}[\mathbf{q}^*] \quad (15)$$

Again, at the initiation of the finite element solution, $\mathbf{q}^*$ will be the values for the option at maturity. Afterward, they will be the solved values of $\mathbf{q}$ in succession. Equation 15 is also of the same form as Equation 12.

Assembly and Solution

In either case, variational formulation or residual formulation, a system of equations for each element is established. The desired solution covers the entire domain, however. So the next step is to form a global linear system of equations whose solution will be the desired function (and possibly derivatives) at the nodes over the entire domain. For both cases, the global linear system is established by the assembly of the element

linear systems. The global coefficient matrix is formed from the element right-hand side vectors. The assembly is based on the idea that the solution (and possibly its derivatives) is continuous at the interelement boundaries. A local–global numbering system is employed in the computer routine to keep track of element nodal position. Continuity is assured by overlap (addition of terms) as the element solutions are placed into the global system (Appendix B contains pseudocode for local and global coefficient matrix construction). The result is a system of equations of the same form as previously shown but on the domain of the entire problem:

$$[k]_{\text{global}}\ [\mathbf{q}]_{\text{global}} = [\mathbf{Q}]_{\text{global}} \tag{16}$$

The resulting system (Equation 16) is solved for $[\mathbf{q}]_{\text{global}}$ at each time increment; all other values are known.

Under the variational form, values on the vector [**q**] must be transformed back to the (S,t) coordinate system. The value for $P(S,t)$ can be easily obtained by multiplying through the values $\mu(x,\tau)$ by $Ee^{\alpha x+\beta\tau}$ (from the transformation).[2] As the Galerkin method is applied directly to the Black–Scholes (Black & Scholes, 1973) differential equation, the vector [**q**] will contain the solution values directly.

EUROPEAN OPTION RESULTS

Table I shows error estimates for the FEM solutions for a European put option without a dividend with the various elements. In addition, error estimates for put option delta values and put option gamma values for the Hermite elements are presented, as these elements allow for direct calculation of these derivatives. Two different error estimates, root mean squared errors (RMSEs) and mean absolute errors, were calculated for each element and method type against the closed-form Black–Scholes (Black & Scholes, 1973) solution for a set of 27 European put options. The 27 combinations result from strike prices of \$35, \$40, and \$45; maturities of 1, 4, and 7 months; and volatilities of 0.2, 0.3, and 0.4. The error results for the FEM calculations are presented for three different element–time-step runs: (a) 50 elements and 50 time steps, (b) 100 elements and 100 time steps, and (c) 200 elements and 200 time steps. The mesh for each element type ranged from \$10 to \$70. The boundary value

[2]For the Hermite family of elements, the solution vector will contain function values and derivatives values (stacked alternately). To handle derivative terms for the Hermite family of elements, the equations in Footnote 1 must also be used.

TABLE I

European Put Error Estimates

	Linear Variation	*Linear Galerkin*	*Cubic Variational*	*Cubic Galerkin*	*Hermite Variational*	*Hermite Galerkin*
50 elements/50 time steps						
RMSE price	0.01943	0.00778	0.00585	0.00563	0.00491	0.01141
MAE price	0.013312	0.006343	0.004600	0.004316	0.003946	0.008674
RMSE delta	NA	NA	NA	NA	0.00100	0.00127
MAE delta	NA	NA	NA	NA	0.000834	0.001040
RMSE gamma	NA	NA	NA	NA	0.00032	0.00127
MAE gamma	NA	NA	NA	NA	0.000261	0.000643
100 elements/100 time steps						
RMSE price	0.00345	0.00324	0.00244	0.00301	0.00278	0.00407
MAE price	0.002440	0.002445	0.001892	0.002269	0.002164	0.003134
RMSE delta	NA	NA	NA	NA	0.00048	0.00051
MAE delta	NA	NA	NA	NA	0.000380	0.000403
RMSE gamma	NA	NA	NA	NA	0.00024	0.00041
MAE gamma	NA	NA	NA	NA	0.000169	0.000241
200 elements/200 time steps						
RMSE price	0.00138	0.00162	0.00150	0.00096	0.00143	0.00180
MAE price	0.001000	0.001192	0.000987	0.000794	0.001020	0.001372
RMSE delta	NA	NA	NA	NA	0.00027	0.00029
MAE delta	NA	NA	NA	NA	0.000205	0.000215
RMSE gamma	NA	NA	NA	NA	0.00009	0.00029
MAE gamma	NA	NA	NA	NA	0.000075	0.000195

Note: The table compares the performance of variational and Galerkin approaches to the finite element method with three elements: (1) linear Lagrange, (2) cubic Lagrange, and (3) quintic Hermite. Root mean squared errors (RMSEs) and mean absolute errors (MAEs) were calculated for each element and method type against the closed form Black–Scholes (Black & Scholes, 1973) solution for a set of 27 European put options. In addition, because as the Hermite element also solved directly for the delta and gamma of the option, error rates for these are given. All calculations were based on an initial stock price of $40 and a riskless interest rate of 5.0% simple interest. The 27 combinations resulted from strike prices of $35, $40, and $45; maturities of 1, 4, and 7 months; and volatilities of 0.2, 0.3, and 0.4.

at $10 was kept at the present value of the strike price minus the stock price. This value should be approximately reasonable for a deep-in-the-money put option (the intrinsic value is too high for a European option). The boundary valued at $70 was kept at zero. For the variational form, the state variable was partitioned evenly in the transformed variable (x). For the Galerkin method, the state variable, S, was partitioned evenly. When the estimates did not fall on the node, values in the table came from interpolations via the shape functions.

As seen in Table I, the FEM provides accurate option pricing values. One of the advantages of the Hermite elements is the direct calculation of delta and gamma values (Lagrange–serendipity elements and finite difference solutions can be used to estimate delta and gamma values, *ex*

post, from the solution values for the option prices). The tables show that these values are also accurate.

DISCRETE BARRIER OPTIONS

As an illustration of some of the advantages the FEM offers over the more widely used binomial and finite difference methods, we applied the technique to the pricing of discrete barrier options. Barrier options have become one of the more popular exotic options in practice, especially in foreign currency markets. In general, there are two basic types of barrier options, *out options* and *in options*. In the case of out options, when the barrier is hit, the option expires worthless (or pays some fixed rebate). For in options, the barrier must be struck before the option can be exercised with value.

Barrier options have many variations; in this section, we focus on those variations involving the frequency at which the barrier is checked. For European barrier options with a continuously monitored barrier, closed-form solutions exist. As the monitoring of the barrier becomes noncontinuous, it becomes necessary (in most cases) to use numerical techniques. Discrete monitoring reduces the probability of hitting the barrier, thereby increasing the value of out options and reducing the value of in options. Standard binomial and trinomial lattice models can be adapted to price these options, but as Ahn et al. (1999) pointed out, convergence may be problematic. They applied the adaptive mesh method to price these types of barrier options and achieved good convergence results.

The FEM can also be applied to price-discrete barrier options. For this example, we focus on up and out barrier put options. Two examples were chosen to illustrate the methods accuracy: (a) a discrete barrier option with a single knock-out date halfway through the life of the option and (b) a discrete barrier with three knock-out dates evenly spaced through the life of the option.

A simple modification can be made to price discrete barrier options with the FEM. As with other lattice-based or tree-based methods, a simple substitution routine can be added to check values on the lattice past the barrier. In this instance, at the discrete time step for which the barrier is checked, if the node corresponds to a stock price equal to or greater than the proscribed barrier, a zero (or rebate) is substituted for the option value.

As with the European put options, we ran 27 combinations for a variety of element–time steps to estimate RMSEs. The additional param-

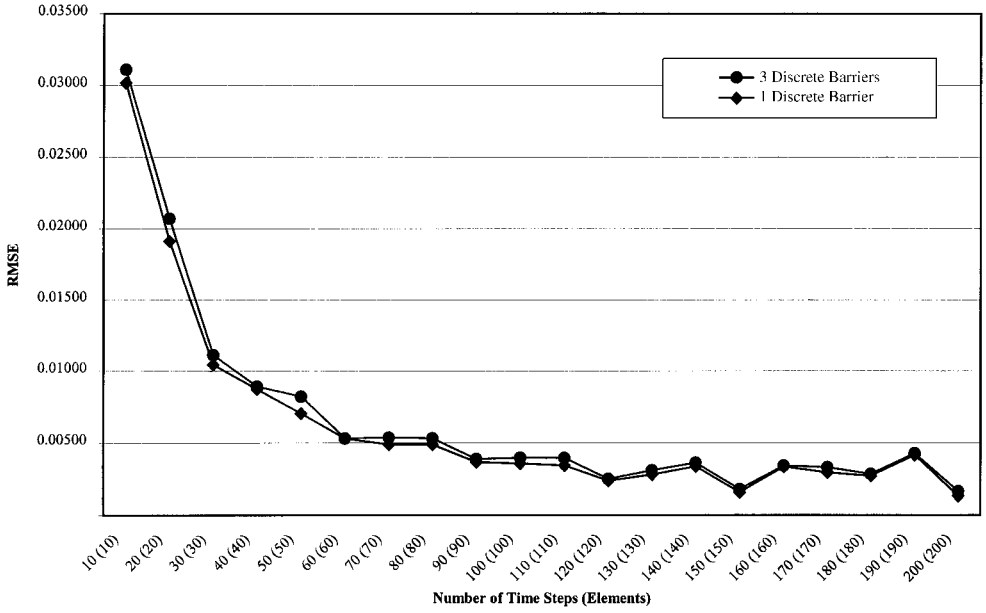


FIGURE 1
Root mean squared error (RMSE) for finite element method discrete barrier solutions (uniform mesh).

eter added to each of the 27 runs was a knock-out barrier of $50. As a benchmark, we used a modified binomial model. Because the relative position of the barrier to the nodal values on the tree can influence the outcome, we ran the binomial model seven times and averaged the solution values to reduce the error.[3] The earlier results suggest the FEM is accurate for a variety of elements; for ease we chose to focus on the cubic Galerkin element for this section, although any of the elements and either method (variational or weighted residual) presented before would work.[4] Figure 1 presents a chart of the RMSE for element–time-step combinations from 10 to 200 for both the one-discrete-barrier and three-discrete-barrier problems. Error rates fell very quickly to under .005 for 90 elements and time steps and beyond.

One of the advantages the FEM offers over standard finite difference and binomial (trinomial) routines is the ability to handle a nonuniform mesh. This can greatly reduce the number of computations needed to achieve desired levels of accuracy.

[3]The model is run with 397, 398, 399, 400, 401, 402, and 403 jumps.
[4]The cubic element is midway between the linear and Hermite elements in terms of complexity, offering the simplicity of the Lagrange family with the flexibility of a higher order polynomial, and was chosen as representative of an average element.

We revisited the one-discrete-barrier and three-discrete-barrier cases, this time reducing the number of elements in the region of the domain away from the desired solution value. As in the previous examples, the domain ran from \$10 to \$70, with desired option solutions for the 27 combinations at the stock price of \$40 with a \$50 barrier. Rather than generating a uniform mesh across the whole domain, the domain from \$10 to \$28 and from \$52 to \$70 was handled by two elements each, whereas the domain between \$28 and \$52 was refined as in the earlier examples.

We again generate RMSEs for an increasing number of time steps and elements. This time, however, to achieve the same nodal spacing for the refined internal region of the domain as in the uniform mesh examples, we required an almost 60% reduction in the number of elements. For example, 84 elements achieved the same refined mesh spacing around the desired solution node as 200 elements in the uniform mesh (80 elements in refined region and 4 in the coarse region). As seen in Figure 2 and Table II, the nonuniform mesh spacing achieved very similar results to the uniform mesh at a greatly reduced computational cost. Although the solution values in the coarse mesh are not as accurate, the flexibility of the cubic element is sufficient to allow the solution value at the desired node (\$40 stock price) to be accurate.

The only change required in programming the nonuniform mesh FEM is in the initial mesh-generation process. That is, the spacing in the underlying asset values at the nodes will be greater in the coarse regions (in the pseudocode, the values would be assigned at a different spacing in the vector **global asset values** [**m**]). The code used to solve for the FEM solutions (examples of which are provided in the pseudocode) requires no modification.

CONCLUSION

This study applied the FEM to the pricing of two options. As a pedagogical illustration, the method is applied to the familiar Black–Scholes (Black & Scholes, 1973) differential equation to price European put options. In addition, using a simple modification, we applied the FEM to the pricing of discrete barrier options. We presented the advantages of the FEM, including nonuniform mesh construction and direct derivative valuation. The FEM is a versatile alternative to other popular lattice methods used in option pricing.

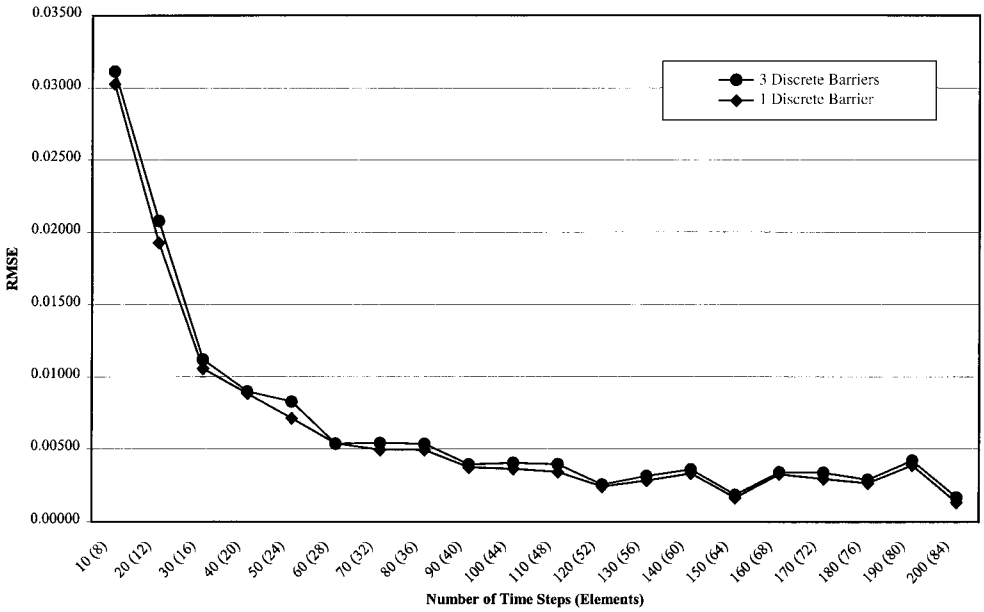


FIGURE 2
Root mean squared error (RMSE) for finite element method discrete barrier solutions (nonuniform mesh).

TABLE II

Discrete Barrier Error Estimates

Time-Step/Element Uniform Mesh (Element Nonuniform Mesh)	*One-Discrete-Barrier Uniform Mesh*	*One-Discrete-Barrier Nonuniform Mesh*	*Three-Discrete-Barrier Uniform Mesh*	*Three-Discrete-Barrier Nonuniform Mesh*
10/10 (8)	0.03017	0.03028	0.03110	0.03116
20/20 (12)	0.01913	0.01925	0.02072	0.02078
50/50 (24)	0.00703	0.00710	0.00822	0.00826
100/100 (44)	0.00356	0.00362	0.00396	0.00403
150/150 (64)	0.00157	0.00162	0.00179	0.00183
200/200 (84)	0.00133	0.00131	0.00165	0.00167

Note: The table compares the performance of uniform and nonuniform applications of the finite element method for the cubic Lagrange element with the Galerkin method. Root mean squared errors (RMSEs) against the benchmark binomial solution for a set of 27 European put options are given. All calculations were based on an initial stock price of $40, a barrier of $50, and a riskless interest rate of 5.0% simple interest. The 27 combinations result from strike prices of $35, $40, and $45; maturities of 1, 4, and 7 months; and volatilities of 0.2, 0.3, and 0.4.

APPENDIX A: SHAPE FUNCTIONS

Linear Lagrange–Serendipity Shape Functions

$$N_1 = \frac{1}{2}(1 - \xi)$$

$$N_2 = \frac{1}{2}(1 + \xi)$$

Cubic Lagrange/Serendipity Shape Functions

$$N_1 = \frac{1}{16}(1 - \xi)(9\xi^2 - 1)$$

$$N_2 = \frac{9}{16}(1 - \xi^2)(1 - 3\xi)$$

$$N_3 = \frac{9}{16}(1 - \xi^2)(1 + 3\xi)$$

$$N_4 = \frac{1}{16}(1 + \xi)(9\xi^2 - 1)$$

Quintic Hermite Shape Functions

$$N_1 = \left(\frac{1}{2} - \frac{15}{16}\xi + \frac{5}{8}\xi^3 - \frac{3}{16}\xi^5\right)$$

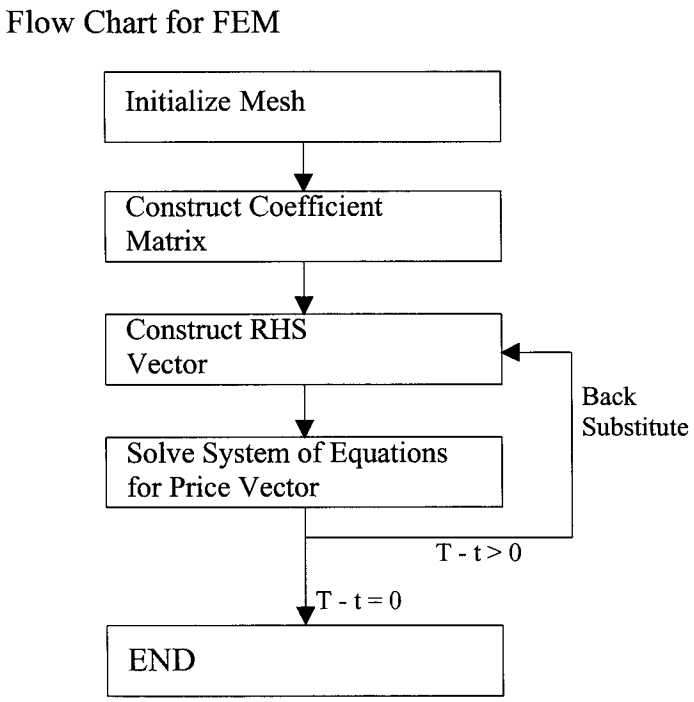
$$N_2 = \left(\frac{5}{16} - \frac{7}{16}\xi - \frac{3}{8}\xi^2 + \frac{5}{8}\xi^3 + \frac{1}{16}\xi^4 - \frac{3}{16}\xi^5\right)$$

$$N_3 = \left(\frac{1}{16} - \frac{1}{16}\xi - \frac{1}{8}\xi^2 + \frac{1}{8}\xi^3 + \frac{1}{16}\xi^4 - \frac{1}{16}\xi^5\right)$$

$$N_4 = \left(\frac{1}{2} + \frac{15}{16}\xi - \frac{5}{8}\xi^3 + \frac{3}{16}\xi^5\right)$$

$$N_5 = \left(-\frac{5}{16} - \frac{7}{16}\xi + \frac{3}{8}\xi^2 + \frac{5}{8}\xi^3 - \frac{1}{16}\xi^4 - \frac{3}{16}\xi^5\right)$$

$$N_6 = \left(\frac{1}{16} + \frac{1}{16}\xi - \frac{1}{8}\xi^2 - \frac{1}{8}\xi^3 + \frac{1}{16}\xi^4 + \frac{1}{16}\xi^5\right)$$



The system of equations is based on Equation 16. At the initial time step the RHS vector contains the maturity values of the option.

Subroutine Global_Coeff_Matrix_Construction (global_coeff_matrix)

```
for (k = 1 to number_of_elements)
  for (m = 1 to 4)
  global_number(m) = 3.0*(k − 1) + m
  element_asset_values(m) = global_asset_values(global_number(m))
  next m
  Call Elemental_Coeff_Matrix_Construction (element_asset_values, element_coeff)
  for (i = 1 to 4)
    for (j = 1 to 4)
    global_coeff_matrix (global_number(i), global_number(j)) =
      global_coeff_matrix(global_number(i), global_number(j)) +
      element_coeff(i, j)
    next j
  next I
next k
```

Subroutine Elemental_Coeff_Matrix_Construction (element_asset_value, element_coeff)

```
nat_coord = quadrature_point
Call Cubic_element_in_natural_coordinate (element_asset_value, nat_coord, . . .)
for (i = 1 to 4)
  for (j = 1 to 4)
    element_coeff = element_coeff +
      (−volatility^2/2.0) *(shape_funct_derivative(i)*(shape_funct_derivative(j)*mu^2) +
      (riskfreerate/interpolated_asset_value)*(shape_funct(i)*shape_funct(j)*mu) +
      (−riskfreerate/interpolated_asset_value^2)*(shape_funct(i)*shape_funct(j))
    element_coeff = element_coeff* gaussian_quadrature_weights/mu
  next j
next i
```

Subroutine Cubic_element_in_natural_coordinate (element_asset_value, nat_coord, shape_funct, shape_funct_derivative, mu, beta, interpolated_asset_value)

shape_funct(1) = (1.0/16.0)*(1.0 − nat_coord)*(9.0*nat_coord^2 − 1.0)
shape_funct(2) = (9.0/16.0)*(1.0 − nat_coord^2)*(1.0 − 3.0*nat_coord)
shape_funct(3) = (9.0/16.0)*(1.0 − nat_coord^2)*(1.0 + 3.0*nat_coord)
shape_funct(4) = (1.0/16.0)*(1.0 + nat_coord)*(9.0*nat_coord^2 − 1.0)
shape_funct_derivative(1) = (−27.0*nat_coord^2/16.0) + (18.0*nat_coord/16.0) + (1.0/16.0)
shape_funct_derivative(1) = (81.0*nat_coord^2/16.0) − (18.0*nat_coord/16.0) − (27.0/16.0)
shape_funct_derivative(1) = (−81.0*nat_coord^2/16.0) − (18.0*nat_coord/16.0) + (27.0/16.0)
shape_funct_derivative(1) = (27.0*nat_coord^2/16.0) + (18.0*nat_coord/16.0) − (1.0/16.0)

mu = 2.0/(element_asset_value(4) − element_asset_value(1))
beta = −(element_asset_value(1) + element_asset_value(4))/(element_asset_value(4) − element_asset_value(1))

interpolated_asset_value = shape_funct(1)*element_asset_value(1) + shape_funct(2)*element_asset_value(2) + shape_funct(3)*element_asset_value + shape_funct(4)*element_asset_value(4)

BIBLIOGRAPHY

Ahn, D.-H., Figlewski, S., & Gao, B. (1999, Summer). Pricing discrete barrier options with an adaptive mesh model. Journal of Derivatives, 6 (4), 33–43.

Allaire, P. E. (1985). Basics of the finite element method, solid mechanics, heat and fluid mechanics. Dubuque, IA: Brown.

Barraquand, J., & Martineau, D. (1995). Numerical valuation of high dimensional multivariate American securities. Journal of Financial and Quantitative Analysis, 30, 383–405.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81, 637–659.

Boyle, P. (1977). Options: A Monte Carlo approach. Journal of Financial Economics, 4, 323–338.

Boyle, P. (1988). A lattice framework for option pricing with two state variables. Journal of Financial and Quantitative Analysis, 23, 1–12.

Boyle, P., Evnine, J., & Gibbs, S. (1989). Numerical evaluation of multivariate contingent claims. Review of Financial Studies, 2, 241–250.

Brennan, M. J., & Schwartz, E. S. (1978). Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis. Journal of Financial and Quantitative Analysis, 13, 461–474.

Chen, R.-R., & Scott, L. (1992). Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. Review of Financial Studies, 5, 613–636.

Clough, R. W. (1960). The finite element method in plane stress analysis. In ASCE 2nd Conference on Electronic Computation.

Cook, R. D., Malkus, D. S., & Plesha, M. E. (1989). Concepts and applications of finite element analysis (3rd ed.). New York: Wiley.

Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibration. Bulletin of the American Mathematical Society, 49, 1–43.

Cox, J., Ingersoll, J., & Ross, S. (1985). An intertemporal general equilibrium model of asset prices. Econometrica, 53, 363–384.

Cox, J., Ross, S., & Rubinstein, M. (1979). Option pricing: A simplified approach. Journal of Financial Economics, 7, 229–263.

Figlewski, S., & Gao, B. (in press). The adaptive mesh model: A new approach to efficient option pricing. Journal of Financial Economics, 53 (3), 313–351.

Garman, M. (1976). A general theory of asset valuation under diffusion state processes (Working Paper No. 5). Berkeley: University of California–Berkeley.

Geske, R., & Shastri, K. (1985). Valuation by approximation: A comparison of alternative option valuation techniques. Journal of Financial and Quantitative Analysis, 20, 45–74.

Hull, J., & White, A. (1988). The use of control variate technique in option pricing. Journal of Financial and Quantitative Analysis, 23, 237–251.

Hull, J., & White, A. (1990). Valuing derivative securities using the explicit finite difference method. Journal of Financial and Quantitative Analysis, 25, 87–99.

Kim, I. J. (1990). The analytic valuation of American options. Review of Financial Studies, 3, 547–572.

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

Parkinson, M. (1977). Option pricing: The American put. Journal of Business, 50, 21–36.

Pelsser, A., & Vorst, T. (1994, Spring). The binomial model and the Greeks. Journal of Derivatives, 1 (3), 45–49.

Thompson, A. C. (1995). Valuation of path-dependent contingent claims with multiple exercise decisions over time: The case of take-or-pay. Journal of Financial and Quantitative Analysis, 30, 271–292.

Tomas, M. J., III. (1996). Finite element analysis and option pricing. Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.

Turner, M. J., Clough, R. W., Martin, H. C., & Topp, J. L. (1956). Stiffness and deflection analysis of complex structures. Journal of Aeronautical Science, 23, 805–823.

Wilmott, P., Dewynne, J., & Howison, S. (1993). Option pricing: Mathematical models and computation. Oxford Financial Press.