Mathematical Finance, Vol. 10, No. 1 (January 2000), 53–75

ON THE RATE OF CONVERGENCE OF DISCRETE-TIME CONTINGENT CLAIMS

STEVE HESTON

Goldman Sachs & Co., New York

GUOFU ZHOU

Washington University, St. Louis

This paper characterizes the rate of convergence of discrete-time multinomial option prices. We show that the rate of convergence depends on the smoothness of option payoff functions, and is much lower than commonly believed because option payoff functions are often of all-or-nothing type and are not continuously differentiable. To improve the accuracy, we propose two simple methods, an adjustment of the discrete-time solution prior to maturity and smoothing of the payoff function, which yield solutions that converge to their continuous-time limit at the maximum possible rate enjoyed by smooth payoff functions. We also propose an intuitive approach that systematically derives multinomial models by matching the moments of a normal distribution. A highly accurate trinomial model also is provided for interest rate derivatives. Numerical examples are carried out to show that the proposed methods yield fast and accurate results.

KEY WORDS: option price, interest rate derivatives, binomial, trinomial, multinomial, smoothness

1. INTRODUCTION

Since the seminal work of Cox, Ross, and Rubinstein (1979), binomial models and various discrete time generalizations have received wide attention from both finance researchers and practitioners. These discrete models provide an easy way to understand how uncertainties are resolved in a continuous-time asset pricing model, and how contingent claims can be hedged or spanned by available assets. In fact, many interesting insights on a continuous-time model, which might not be available otherwise, can be understood by taking the limits of certain discrete-time models. Other than as an excellent pedagogical tool, the discrete-time models also play an important role in practice for valuing most contingent claims for which simple closed-form solutions of continuous-time models are usually not available. Because of the importance and usefulness of discrete-time models, there has been extensive research in extending them, examples of which include Hull and White (1988), Boyle, Evnine, and Gibbs (1989), Madan, Milne, and Shefrin (1989), and Rubinstein (1994). Broadie and Detem-

We are grateful to Mark Broadie, Phil Dybvig, Raymond Kan, Junping Wang, an associate editor, and the editor for helpful discussions or comments. We are especially grateful to two anonymous referees for providing many insightful and detailed suggestions that substantially improved the paper.

Manuscript received January 1998; final revision received March 1999.

Address correspondence to Prof. Guofu Zhou, John M. Olin School of Business, Washington University, St. Louis, MO 63130; e-mail: zhou@zhoufin.wustl.edu.

ple (1996, 1998) provide some of the latest developments on American options and discrete-time approximations in general.

Although theoretical proofs of convergence of discrete-time models to their continuous-time analogues are given by He (1990) and Amin and Khanna (1994), among many others, the *rate* of convergence, which appears to be one of the central properties of a discrete-time model, has received little attention. The rate or order of convergence measures the (asymptotic) speed/accuracy trade-off of a numerical method. Almost any method can give fast inaccurate results, and, given enough computational time, many methods can give arbitrarily accurate results. Given competing discrete-time models, the order can help to rank which of the models is a preferred one. Moreover, there are at least three important reasons in practice to know the correct rate of convergence. First, discrete-time models with higher rates of convergence are required to compute many contingent claim prices on a real-time basis. Second, extrapolation is a useful technique for increasing the accuracy of a discrete-time solution in practice, but it is no longer useful if the rate of convergence is unknown or if it is not a fixed constant. Third, insights on the rate of convergence can potentially lead to lower hedging costs for derivatives because highly accurate solutions are obtained by using the minimum number of periods or transactions.

This paper characterizes the rate of convergence of discrete-time models to their continuous-time limit. First, we provide a theoretical proof that the widely used binomial model cannot converge to its limit as fast as one might believe from available theorems of the finite difference method (see Ames 1992). As the binomial method is a special case of the standard finite difference method and the latter usually converges at the $1/n$ rate, it seems obvious that the n-period binomial solution should converge to its limit, the Black–Scholes formula value, at a rate of $1/n$. Furthermore, asymptotic expansions of the error can be written as a constant function of $1/n$ plus higher order terms, and the standard Richardson extrapolation can be used to obtain solutions with higher convergence rate. But this is true and can be shown only under the condition that the payoff function of the contingent claim is smooth.

In finance, however, the payoff functions are often of all-or-nothing type, and hence are not continuously differentiable. This implies that the usual theorems in a standard numerical analysis book (such as Ames 1992) are not applicable to interesting option pricing problems in finance. We give an example that shows the rate of convergence of the binomial model cannot be faster than $1/\sqrt{n}$ at the nodes near expiration. However, this does not claim that the binomial solution at current node cannot converge at the $1/n$ rate. It simply states that it cannot do so uniformly on the binomial tree. We also provide a theorem that shows the binomial model can achieve at least the $1/\sqrt{n}$ rate. Hence, it is theoretically possible, and indeed we show in numerical examples, that the convergence rate of the binomial model fluctuates between $1/\sqrt{n}$ and $1/n$ across the nodes of the binomial tree. Although at the current node the solution may still have the $1/n$ rate of convergence, the nonsmoothness of the payoff functions can have an impact sufficient to cause the well-known oscillatory pattern of the binomial prices at the current node, making invalid the standard Richardson extrapolation.

Intuitively, as the payoff function is not continuously differentiable, the binomial model carries at its start a larger error. This error may or may not offset in the backward deduction process. As a result, it sometimes has an overall error no better than $1/\sqrt{n}$ at some nodes of the binomial tree which may or may not carry to the current node. To overcome the problem caused by the nonsmoothness of the payoff functions, we propose both a smoothing approach and an adjustment approach such that the resulting solution achieves its maximum possible convergence rate, $1/n$, across all the nodes of the binomial tree.

Section 2 provides both theoretical proofs and numerical examples to illustrate the ideas and their potential practical applications. In Section 3, we extend the analysis to discrete-time multinomial approximations to continuous-time models. We show that various high-order multinomial methods, such as the trinomial and pentanomial models, can be motivated and obtained by matching higher order moments of the discrete-time model to those of a continuous-time one. Like the binomial case, it is important to emphasize that the multinomial models achieve their higher rate of convergence only if the payoff functions are smooth enough. For example, contrary to popular belief, the trinomial model cannot always have a faster rate of convergence than the binomial model in option pricing due to the nonsmoothness of the payoff function. In Section 4, we show how the proposed approaches can be applied to American options to obtain fast and accurate results. Additionally, we propose a simple algorithm to locate the optimal exercise boundary for an American option. This algorithm has a rate of convergence of $1/\sqrt{n}$, which does not appear to be recognized in the literature. In Section 5, we provide a new and highly accurate trinomial model for interest rate derivatives. Conclusions are offered in the final section.

2. THE BINOMIAL MODEL

Although our approach is generally applicable, for pedagogical reasons we will focus the discussion on valuing a standard call price when the stock price, S, follows a geometric Brownian motion. In this simplified setting, the price of a contingent claim with payoff $g(S)$ at expiration time T, by Black and Scholes (1973), must satisfy the familiar partial differential equation

$$\frac{1}{2}\sigma^2 S^2 C_{SS} + rSC_S - rC + C_t = 0, \tag{2.1}$$

with (*terminal*) condition $C(S,T) = g(S)$. Making a transformation of variables, $x = [\log S - (r - \frac{1}{2}\sigma^2)t]$ and $\tau = T - t$, then $U(x,\tau) = e^{r(T-t)}C(S,t)$ satisfies the standard diffusion (heat) equation

$$U_\tau = \frac{1}{2}\sigma^2 U_{xx}, \tag{2.2}$$

with initial condition $U(x,0) = g(\exp[x + (r - \frac{1}{2}\sigma^2)T])$. In particular, if $g(S) = \max(S - K, 0)$, the payoff of a standard call option on the stock with strike price K, then the (terminal) condition becomes $C(S,T) = \max(S - K, 0)$ or $U(x,0) = \max(\exp(x + (r - \frac{1}{2}\sigma^2)T) - K, 0)$.

To analyze the rate of convergence of the binomial model, it is easier to link it to the finite difference method than otherwise. The classical explicit finite difference method computes the value of $U(x,\tau)$ by replacing the derivatives in equation (2.2) by finite difference approximations,

$$\frac{U(x,\tau+\Delta_\tau) - U(x,\tau)}{\Delta_\tau} = \frac{1}{2}\sigma^2 \frac{U(x+\Delta_x,\tau) - 2U(x,\tau) + U(x-\Delta_x,\tau)}{\Delta_x^2}$$

where Δ_τ and Δ_x are small increments of the time and x variable, respectively. Starting from the initial values (or terminal values of the payoff function), the solution can be iteratively computed from

$$U(x,\tau+\Delta_\tau) = qU(x+\Delta_x,\tau) + (1 \quad 2q)U(x,\tau) + qU(x-\Delta_x,\tau),$$

where $q = \sigma^2\Delta_\tau/2\Delta_x^2$. When $\Delta_x = \sigma\sqrt{\Delta_\tau}$ or $q = \quad /2$, equation (2.4) gives rise to the familiar binomial solution,[1]

$$\hat{C}(S,t-\Delta) = e^{-r\Delta_t}\left[\hat{C}(uS,t) + \hat{C}(dS,t)\right]/2,$$

where $\Delta = \Delta_t = \Delta_\tau$, $u = e^{(r-\sigma^2/2)\Delta_t + \sigma\sqrt{\Delta_t}}$, and $d = e^{(r-\sigma^2/2)\Delta_t - \sigma\sqrt{\Delta_t}}$.

In an n-period binomial model, the time to maturity (from today t to maturity date T) is divided into n subintervals, $t_0 = t < t_1 < \cdots < t_n = T$. The length of all of the subintervals is $\Delta_\tau = t_{i+1} - t_i = (T-t)/n$. The call price at each node of the standard binomial tree, $S_{ij} = Se^{i\sigma\sqrt{\Delta_\tau} + j(r-\sigma^2/2)\Delta_\tau}$, is computed recursively from the discounted expected payoff relationship, equation (2.5). The binomial prices will be different from the continuous-time solution of equation (2.2). This is because the differentials in equation (2.2) are approximated by their finite difference analogues in the binomial model. The approximation error is the so-called *discretization error* or *truncation error*, which is well known to be of order $O(\Delta_\tau + \Delta_x^2)$ (assume the smoothness of the continuous-time solution). For the n-period binomial model, this translates into an order of $1/n$. In other words, apart from an error of magnitude of $1/n$, the continuous-time solution also satisfies equation (2.3).

Because the truncation error is of order $1/n$, the *local error* for the solutions in equation (2.5), which is obtained after multiplying equation (2.3) by $\Delta_\tau = T/n$, is of the order $1/n^2$. As there are n steps in the binomial recursions, the local errors will be accumulated n times. Hence, one may expect that the difference between the binomial prices and the continuous-time solution should be of the order $n \times 1/n^2 = 1/n$, the same order as the truncation error. Indeed, if the payoff function has continuous derivatives up to the second order, we have the following proposition.

PROPOSITION 2.1. *Let $\hat{C}_n$ and C be the binomial and continuous-time prices of a European option with terminal payoff $g(S)$. If $g(S)$ is continuously differentiable up to the second order, then*

$$\hat{C}_n = C + O(1/n)$$

Proof. For the n-period binomial model, $\Delta_x^2 = \sigma\Delta_\tau = O(1/n)$. The smoothness assumption on the payoff function ensures that C has bounded derivatives up to the fourth order (with respect to S) inside its domain (e.g., see Friedman 1964). Hence, in terms of this continuous-time solution, equation (2.5) can be written as

$$C(S,t-\Delta \quad = e^{-r\Delta_t}[C(uS,t) + C(dS,t)]/2 + O(1/n^2),$$

[1] Another version is Cox et al.'s (1979) binomial model, $\hat{C}(S,t-\Delta_t) = [p\hat{C}(uS,t) + (1-p)\hat{C}(dS,t)]/(1+r^*)$, where $u = e^{\sigma\sqrt{\Delta_t}} - 1$, $d = e^{-\sigma\sqrt{\Delta_t}} - 1$, $r^* = e^{r\sqrt{\Delta_t}} - 1$, and $p = (r-d)/(u-d)$. This is the explicit finite difference method applied to the log transform of equation (2.1). Although the results hold for both versions, we provide here only for the first one, which appears to have simpler notations.

where Δ_τ and Δ_x are small increments of the time and x variable, respectively. Starting from the initial values (or terminal values of the payoff function), the solution can be iteratively computed from

$$U(x,\tau+\Delta_\tau) = qU(x+\Delta_x,\tau) + (1-2q)U(x,\tau) + qU(x-\Delta_x,\tau),$$

where $q=\sigma^2\Delta_\tau/2\Delta_x^2$. When $\Delta_x = \sigma\sqrt{\Delta_\tau}$ or $q=1/2$, equation (2.4) gives rise to the familiar binomial solution,[1]

$$\hat{C}(S,t-\Delta) = e^{-r\Delta_t}\left[\hat{C}(uS,t)+\hat{C}(dS,t)\right]/2, \tag{2.5}$$

where $\Delta=\Delta_t=\Delta_\tau$, $u=e^{(r-\sigma^2/2)\Delta_t+\sigma\sqrt{\Delta_t}}$, and $d=e^{(r-\sigma^2/2)\Delta_t-\sigma\sqrt{\Delta_t}}$.

In an n-period binomial model, the time to maturity (from today t to maturity date T) is divided into n subintervals, $t_0=t<t_1<\cdots<t_n=T$. The length of all of the subintervals is $\Delta_\tau = t_{i+1}-t_i=(T-t)/n$. The call price at each node of the standard binomial tree, $S_{ij}=Se^{i\sigma\sqrt{\Delta_\tau}+j(r-\sigma^2/2)\Delta_\tau}$, is computed recursively from the discounted expected payoff relationship, equation (2.5). The binomial prices will be different from the continuous-time solution of equation (2.2). This is because the differentials in equation (2.2) are approximated by their finite difference analogues in the binomial model. The approximation error is the so-called *discretization error* or *truncation error*, which is well known to be of order $O(\Delta_\tau+\Delta_x^2)$ (assume the smoothness of the continuous-time solution). For the n-period binomial model, this translates into an order of $1/n$. In other words, apart from an error of magnitude of $1/n$, the continuous-time solution also satisfies equation (2.3).

Because the truncation error is of order $1/n$, the *local error* for the solutions in equation (2.5), which is obtained after multiplying equation (2.3) by $\Delta_\tau = T/n$, is of the order $1/n^2$. As there are n steps in the binomial recursions, the local errors will be accumulated n times. Hence, one may expect that the difference between the binomial prices and the continuous-time solution should be of the order $n\times 1/n^2 = 1/n$, the same order as the truncation error. Indeed, if the payoff function has continuous derivatives up to the second order, we have the following proposition.

PROPOSITION 2.1. *Let $\hat{C}_n$ and C be the binomial and continuous-time prices of a European option with terminal payoff $g(S)$. If $g(S)$ is continuously differentiable up to the second order, then*

$$\hat{C}_n = C + O(1/n)$$

Proof. For the n-period binomial model, $\Delta_x^2=\sigma\Delta_\tau = O(1/n)$. The smoothness assumption on the payoff function ensures that C has bounded derivatives up to the fourth order (with respect to S) inside its domain (e.g., see Friedman 1964). Hence, in terms of this continuous-time solution, equation (2.5) can be written as

$$C(S,t-\Delta_t) = e^{-r\Delta_t}[C(uS,t)+C(dS,t)]/2 + O(1/n^2)$$

[1] Another version is Cox et al.'s (1979) binomial model, $\hat{C}(S,t-\Delta_t) = [p\hat{C}(uS,t)+(1-p)\hat{C}(dS,t)]/(1+r^*)$, where $u=e^{\sigma\sqrt{\Delta_t}}-1$, $d=e^{-\sigma\sqrt{\Delta_t}}-1$, $r^*=e^{r\sqrt{\Delta_t}}-1$, and $p=(r-d)/(u-d)$. This is the explicit finite difference method applied to the log transform of equation (2.1). Although the results hold for both versions, we provide here only for the first one, which appears to have simpler notations.

where $O(1/n^2)$, the local error that results from multiplying the discretization error by $1/n$, is bounded by c_0/n^2 for some constant c_0 independent of the price nodes. Let $S_{ji} = u^i d^{n-j} S$, $V(j,i) = \hat{C}(S_{ji}, t_i) - C(S_{ji}, t_i)$ be the error at each node, and V_i be the maximum absolute error at each time t_i. Then, from equations (2.5) and (2.7), we have

$$|V \quad i)| \leq |V(j + \quad i+1)|/2 + |V(j \; i+1)|/2 + c_0/n^2$$

$$\leq V_i \quad + c_0/n \qquad\qquad n \quad 1)$$

implying that $V_i \leq V_{i+1} + c_0/n^2$. Since $V_n = 0$, it follows that $V(j,i)| \leq V_i \leq V \quad + c_0/n^2 \leq V_{i+2} + 2c_0/n^2 \leq V_n + (n \; - i)c_0/n^2 \leq c_0/n = O(1/n)$. □

Although the above proof is well known in the numerical analysis literature, it is instructive to provide it here to show how errors are transmitted over time. Proposition 2.1 suggests that, for smooth payoff functions, the rate of convergence is as fast as $1/n$. Moreover, it is seen from the proof that the convergence is *uniform* in the sense that the solution has the same rate of convergence at all nodes of the binomial tree. Furthermore, it is easy to show that the error also admits an asymptotic expansion in terms of a constant function of $1/n$ and higher order terms so that Richardson extrapolation applies. Though shown above only for a geometric Brownian motion, the proposition can be extended to allow a fairly general stochastic process for the underlying asset. From the literature on finite difference methods, it seems that similar theoretical results also hold for exotic options, such as the barrier options and other path-dependent options. However, the conclusions may have to be dependent on the specific type of options, and their rigorous proofs remain interesting problems to be explored.

Proposition 2.1 provides the rate of convergence only for option prices with smooth payoff functions. However, payoff functions that are of practical interest in finance are often of the all-or-nothing type and are not continuously differentiable. In this case, Proposition 2.1 is no longer applicable. Indeed, for nonsmooth payoff functions, the errors at a time close to the maturity can be as large as the order of $1/\sqrt{n}$. To illustrate, consider a special case of the binomial model for the standard call option. Let $x = [\log K - (r - \frac{1}{2}\sigma^2)(T - \Delta_\tau)]$ (or $S = K$) be a node point at time $T - \Delta_\tau$. Then, at this node, the binomial price $\hat{C}(K, T - \Delta_\tau)$ is, by equation (2.5),

$$\hat{C}(K,T \quad \Delta_\tau) \quad K(e^{\Delta} \quad {}^{\Delta_\tau/2} \qquad /2$$

$$\frac{1}{2}\sigma^2\Big)(T \quad t)K/\sqrt{n} + O(1/n)$$

On the other hand, the exact value at the same node is, by the Black–Scholes formula,

$$C(K,T \quad \Delta \qquad K\sigma/(\sqrt{2\pi}\sqrt{n} \qquad O(1/n$$

A comparison of (2.9) with (2.9′) shows that $\hat{C}(0, T - \Delta_\tau)$ converges to $C(0, T - \Delta_\tau)$ at the rate $1/\sqrt{n}$.

The above example shows theoretically that the solution cannot converge faster than the $1/\sqrt{n}$ rate, at least at some nodes of the binomial tree. The question is whether or not this rate is achievable. In what follows, we show that, for the standard call option, we have the following proposition.

PROPOSITION 2.2. *Let $\hat{C}_n$ and C be the binomial and continuous-time call prices of a standard European call option with the terminal payoff* $\max(S - K, 0)$. *Then,*

$$\hat{C}_n = C + O(1/\sqrt{n} \tag{2.10}$$

Proof. Following Cox et al. (1979), the n-period binomial call price can be written as

$$\hat{C}_n = SP([a], n, p_1) - Ke^{-r\tau}P([a], n, p_2), \tag{2.11}$$

where $P(\quad)$ is the right-tailed binomial distribution, $[a]$ is the integer part of

$$\frac{n}{2} + \frac{1}{2.}\frac{\ln(K/S) - (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}/\sqrt{n}} \qquad p_2 = \frac{\ }{2}, \qquad p_1 = \frac{1}{2}e^{\sigma\sqrt{\tau}/\sqrt{n} - \sigma^2\tau/2n}$$

Now, by using the normal distribution to approximate the binomial one, we get (e.g., see Johnson, Kotz, and Kemp 1992, p. 114)

$$P\left(np + \sqrt{np(1-p)}\, z, n, p\right) = \quad N(z) + O(1/\sqrt{n} \tag{2.12}$$

where $N(z)$ is the standard normal distribution function. Hence, it follows that

$$P([a], n, p_2 \qquad - N(b) + O(1/\sqrt{n} \tag{2.13}$$

where $b = [\ln(K/S) - (r - \sigma^2/2)\tau]/\sigma\sqrt{\tau}$. Now notice that $a = np_1 + \sqrt{np_1(1 - p_1)}$ $(b - \sigma\sqrt{\tau} + O(1/\sqrt{n}))$ and $N(z + O(1/\sqrt{n})) = N(z) + O(1/\sqrt{n})$; we have

$$P([a], n, p_1) = \; -N\left(b - \sigma\sqrt{\tau} + O\left(\frac{1}{\sqrt{n}}\right)\right) + O\left(\frac{1}{\sqrt{n}}\right), \tag{2.14}$$

which implies

$$\hat{C}_n = S(1 - N(b - \sigma\sqrt{\tau})) - Ke^{-r\tau}(1 - N(b)) + O\left(\frac{1}{\sqrt{n}}\right) \tag{2.15}$$

The first two terms on the right-hand side are equivalent to the continuous-time call price given by the Black–Scholes formula. □

Proposition 2.2 states that the rate of convergence is at least as fast as $1/\sqrt{n}$ for the standard European call option. Though the results are presented only for the call, it is clear that the same conclusion holds for the standard European put option as well. From the numerical analysis literature for general initial functions (e.g., Brenner, Thomée, and Wahlbin 1975), Proposition 2.2 actually holds for all payoff functions that are piecewise smooth with only a finite number of continuous jumps. However, it is not clear from a theoretical standpoint whether the same conclusion should hold for the barrier options and other path-dependent options. Nevertheless, numerical solutions in practice tend to suggest that Proposition 2.2 is widely applicable, although the proof may be quite complex. On the other hand, the $1/\sqrt{n}$ rate is the best possible rate of uniform convergence for those piecewise smooth functions because a certain order of uniform convergence must imply a certain degree of the smoothness (see Brenner et al. 1975). In particular, $1/\sqrt{n}$ is the best possible uniform convergence rate for the standard European call option. However, this does not necessarily imply that the binomial solution at the current node cannot converge faster than the $1/\sqrt{n}$ rate.

Indeed, Broadie and Detemple (1996, 1998) and Leisen and Reimer (1996)[2] seem to provide convincing evidence that the binomial solution for the standard call option at the current node can converge faster than $1/\sqrt{n}$ to achieve the $1/n$ rate. However, it is not clear that this holds for general nonsmooth payoff functions. Even for the standard call option, the error as shown earlier is of order $1/\sqrt{n}$ near expiration. By the proof of Proposition 2.1, the error can potentially be transmitted over time, causing the accuracy of the solution at early times to become less accurate. The $1/n$ convergence rate as suggested by Broadie and Detemple and Leisen and Reimer is striking in that it says that the transmitted errors only cause the commonly observed oscillatory behavior of the binomial price but do not affect the order of convergence at the current node.

Thus far, we have shown that the accuracy or rate of convergence of the binomial method depends crucially on the smoothness of the payoff function. Smooth payoff functions generally enjoy a much higher rate of convergence than nonsmooth payoff functions. Unfortunately, most payoff functions of practical interest are not smooth. The open question is whether it is possible to achieve the maximum possible rate, $1/n$, for nonsmooth payoff functions such that there are no oscillatory behaviors. We propose two approaches to this end.

Notice that the greater error for nonsmooth payoff functions is caused by less accurate prices at the time prior to the end of the tree. Clearly, if we can have accurate values there, the proof of Proposition 2.1 implies that solutions after that time should be very accurate. This suggests using our first method, which replaces the binomial prices prior to the end of the tree by the Black–Scholes values, and computing the rest of binomial prices as usual. Intuitively, this is equivalent to replacing the discretization over the last time step with an infinitely fine grid. As a result, the errors introduced by the first-order discontinuities of the payoff function are reduced.[3] Of course, it is unnecessary to do so for the standard European call price, because one already has the continuous-time solution and does not need it to obtain a discrete-time approximation. However, the procedure is useful for American options and exotic options for which the Black–Scholes is an excellent approximation prior to maturity. If the option price has bounded continuous derivatives up to the fourth order, the Black–Scholes adjustment clearly works and the proof is similar to that of Proposition 2.1. Unfortunately, the bounded derivative condition is not easily shown for various contingent claims. Violation of this condition makes it very difficult for us to give a rigorous proof for its validity, but our extensive numerical experiments uniformly convince us that the simple adjustment approach works well in practice. Independently, Broadie and Detemple (1996) also show the effectiveness of similar adjustments. But, as shown in Section 3, the adjustments may not work for high-order discrete-time approximations, such as the trinomial model.

Our second approach is to smooth the payoff function. As our earlier example for the $1/\sqrt{n}$ rate convergence shows, the inaccuracy occurs at singular points of the payoff function. Intuitively, if we can smooth the payoff function at these points, the

[2] We are grateful to an associate editor and Mark Broadie for bringing our attention to these papers.

[3] This also suggests that for a call or a put, the replacement may need only be done at the at-the-money node. However, our calculations are based on replacement at all the nodes one step away from maturity.

binomial recursion might be more accurate. Indeed, let $g(x)$ be the payoff function, and $G(x)$ is the smoothed one,

$$G(x) = \frac{1}{2\Delta_x} \int_{-\Delta_x}^{\Delta_x} g(x-y)\, dy \tag{2.16}$$

This is a rectangular smoothing of $g(x)$. The smoothed function, $G(x)$, can be easily computed analytically for most payoff functions used in practice. Applying the binomial model to $G(x)$ instead of $g(x)$ yields a rather surprising and interesting result: The associated binomial prices converge now at the $1/n$ rate to its continuous-time limit and this convergence is uniform across the nodes of the binomial tree. A rather lengthy and intricate proof of this is provided by Thomée and Wahlbin (1974), who in turn build their work on many earlier ones. In comparison with the Black–Scholes adjustment, the smoothing approach is more general and applicable to all European options on assets whose prices are complex diffusion processes.

To illustrate how the approaches work in actual computations, we provide in Table 2.1 a simple numerical example that computes the value of a European call option on a stock whose price is 100. The strike is 100, the time to maturity is one year, the volatility is 40%, and the continuously compounded annual interest is 6%. There are three panels in the table. The first two columns of the first panel are the number of periods (n) of the binomial model and the exact Black–Scholes price.[4] The next three columns are the error of the the binomial price, the ratio of the errors, and the error of the extrapolated solution. The other two panels are the corresponding results obtained by using the Black–Scholes adjustment and the smoothing procedure, respectively.

The error of the usual binomial model does not necessarily go down as n increases. For example, the error, as measured by the difference between the binomial price and the Black–Scholes (BS) price, is 0.0044 when $n = 640$, worse than an error of -0.0019 when $n = 40$. In contrast, the error decreases as n increases with either the BS adjustment or the smoothing method. The BS adjustment is strikingly accurate. It achieves penny accuracy when $n = 40$, corresponding roughly to weekly intervals of the binomial price tree. This same accuracy will take about daily intervals for the standard binomial method to achieve.

For the example under consideration, the smoothing procedure has roughly the same magnitude of errors as the binomial method. This is a little surprising since theoretically it should converge at a much faster rate. There are two likely reasons. First, the constant in the theoretical rate may be large for this particular example. Second, the rate of convergence is not fixed for the standard binomial model. Intuitively, when the stock price node lies exactly at the singular point (this happens for some particular n's and strike levels), the singularity may not matter at all, and hence it will converge to the solution at a higher rate, say $1/n$, which is the convergence rate of the smoothing procedure. When the node is away from the singular point, the convergence is slowed down, an observation consistent with our earlier theoretical example. In other words, the binomial solution oscillates irregularly around its limit, a well-known pattern observed by practitioners (e.g., Derman et al. 1995). Hence, as n varies, the binomial

[4] A single Black–Scholes price accurate up to the reported digit can be computed by using any symbolic software, such as Mathematica. But it is impractical to do so if a great number of such values are needed as required by the first approach. A simple solution is to use Strecok's (1968) algorithm, which is easily incorporated into any program, to compute Black–Scholes prices to the desired accuracy.

TABLE 2.1
Binomial Model

Consider the valuation of a European call option on a stock whose price is 100. The strike is 100, the time to maturity is one year, the volatility is 40%, and the continuously compounded annual interest is 6%. The first two columns of the first panel are the number of periods (n) of the binomial model and the Black–Scholes (BS) price. The next three columns are the error of the binomial price, the ratio of the errors, and the error of the extrapolated solution. The other two panels are the corresponding results obtained by using the Black–Scholes adjustment and the smoothing procedure, respectively.

n	Exact	Binomial	Error ratio	Extrapolated
10	18.47260446	−0.18552506		
20	18.47260446	−0.05399578	3.43591784	0.07753350
40	18.47260446	−0.00191274	28.22959146	0.05017031
80	18.47260446	0.01410255	−0.13563052	0.03011785
160	18.47260446	0.01496727	0.94222596	0.01583199
320	18.47260446	0.01034461	1.44686708	0.00572194
640	18.47260446	0.00446058	2.31911889	−0.00142345
1280	18.47260446	−0.00100635	−4.43243516	−0.00647328

n	Exact	BS Adjustment	Error ratio	Extrapolated
10	18.47260446	0.08592189		
20	18.47260446	0.04385530	1.95921356	0.00178870
40	18.47260446	0.02210462	1.98398798	0.00035394
80	18.47260446	0.01102834	2.00434776	−0.00004795
160	18.47260446	0.00545794	2.02060394	−0.00011246
320	18.47260446	0.00269819	2.02281302	−0.00006155
640	18.47260446	0.00136043	1.98333827	0.00002267
1280	18.47260446	0.00070582	1.92745730	0.00005120

n	Exact	Smoothing	Error ratio	Extrapolated
10	18.47260446			
20	18.47260446		1.97585053	0.00606740
40	18.47260446		1.98736102	0.00159783
80	18.47260446		1.99350098	0.00041214
160	18.47260446		1.99662566	0.00010718
320	18.47260446		1.99812100	0.00002987
640	18.47260446		1.99933536	0.00000528
1280	18.47260446		2.00082367	−0.00000327

solution may sometimes have roughly the same magnitude of errors as the smoothing procedure.

However, a nice feature of the smoothing procedure is that it has steadily declining errors, and its true advantage lies in the accuracy of its extrapolated solution. Because it has an order of convergence $1/n$, twice the solution minus the solution with half the number of periods, $2C_{2n} - C_n$, should converge to the continuous-time solution at an even faster rate. Indeed, as the last column in Table 2.1 shows, the errors of the extrapolated solutions based on the 40-, 20-, and 10-period solutions, are 0.0015 and 0.0061, within one penny of the exact solution. With n as small as 40, the smoothing procedure is much more easily implemented, say in a spreadsheet program, than other procedures. Strikingly, the table shows that the extrapolated solutions converge even faster than those based on the Black–Scholes adjustment.

Information for assessing further the rate of convergence of the three numerical methods is available from the ratios of the errors. For the standard binomial solution, the ratios fluctuate greatly. For example, the ratio is 28.2396 when $n = 40$, but only 2.3191 when $n = 640$. Then, it becomes -4.4324 when $n = 1280$. In contrast, the ratios from both the Black–Scholes adjustment and the smoothing procedure are quite stable. If the rate of convergence is $1/n$ and if the usual Richardson expansion holds, then the ratios should be close to 2. Indeed, both the Black–Scholes adjustment and the smoothing procedure work so well that the ratios are almost identical to 2 for all the n's.

In summary, because of the high rate of convergence of both the Black–Scholes adjustment and the smoothing procedure, it is not surprising that their extrapolated solutions work exceptionally well. They are more accurate than the original solutions, and the accuracy increases steadily as n increases. However, this is not the case for the standard binomial method. Its extrapolated solutions fluctuate without a clear pattern, and the accuracy is often worse than the original solutions. As noted earlier, the binomial model does not have a uniform $1/n$ rate of convergence. Its solution oscillates irregularly between a rate of $1/n$ and $1/\sqrt{n}$ at the nodes of the bionomial tree. This seems the fundamental reason why extrapolation fails in the standard binomial model.

Unlike some papers in the literature, we compare the effectiveness of the methods by analyzing their order of convergence, rather than solely their errors. We do this because it is possible for a relatively inaccurate method to produce a value with a relatively smaller error by coincidence for a particular lattice size. In contrast, the order of convergence indicates the asymptotic speed of computation as increased accuracy is demanded. For a method with linear convergence, the accuracy shrinks proportionally to the time step. But the lattice computations increase with the square of the number of time steps (in a single dimension). Consequently, it takes four times the computational work to double the accuracy. This problem is even worse in higher dimensions, and makes methods with low order of convergence extremely slow for highly accurate results. Therefore, the order of convergence provides an asymptotically valid way to compare the relative speed and accuracy of numerical methods.

3. MULTINOMIAL MODELS

For a given step size $\Delta = \Delta_\tau$, we define a multinomial approximation in terms of a "probability vector" $p = (p_1, p_2, \dots, p_k)$ and a multinomial distribution vector $h = (h_1, h_2, \dots, h_k)$,

$$C(S, t - \Delta) = e^{-r\Delta} \sum_{i=1}^{k} p_i C(S e^{(r - \sigma^2/2)\Delta + h_i \sigma \Delta^{1/2}}, t). \tag{3.1}$$

A particular case of equation (3.1) is $k = 2$, $p_1 = p_2 = 1/2$, and $h_1 = -1$ and $h_2 = 1$. In this case, one can verify that equation (3.1) is exactly the binomial model as described in equation (2.5).

The p vector resembles the risk-neutral probabilities in standard discrete-time models, but our setup here is more general. Any real vector p, some components of which may be negative, can be used to define the approximation scheme as long as $\sum_{i=1}^{k} p_i = 1$. The h vector represents the spacing of the multinomial nodes. Both p and h are fixed constants chosen below to obtain potentially higher order accurate discrete-time approximations to the continuous-time solution.

For illustration, our discussion will be focused on a one-dimensional case where the stock price follows the lognormal distribution. To analyze the properties of the multinomial approximation, it is convenient to make a change of variables,

$$U(z,t) = e^{-rt} C(S,t$$

where $z = (\log(S) - (r - \sigma^2/2)t)/\sigma$. In probability terms, since S is lognormal, z is the standardized Wiener process. The transformation simplifies both the price process and the associated option pricing equation. In terms of the transformed option price, the multinomial approximation has a simpler expression:

$$U(z,t) = \sum_{i=1}^{k} p_i U\big(z \quad h_i \Delta^{1} \quad t + \Delta\big)$$

The following proposition characterizes its accuracy.

PROPOSITION 3.1. *If the multinomial model* (p, h) *matches the first* q *moments of a normal distribution, that is,*

$$\sum_{i=1}^{k} p_i h_i^j = 0, \quad \text{for all odd } j \leq q$$

and

$$\sum_{i=1}^{k} p_i h_i^j = \frac{}{2^{j/2}(j/2)!} \quad \text{for all even } j \leq q,$$

and if the terminal payoff function is $2q$ *times continuously differentiable, then the multinomial approximation* (3.3) *has a local error of* $O(\Delta^{(q+1)/2})$, *and the associated discrete-time solution,* $\hat{C}$, *converges to the continuous-time solution,* C, *at a rate of* $O(\Delta^{(q+1)/2-1})$; *that is,*

$$\hat{C} = C + O(\Delta^{(q+1)/2} \tag{3.6}$$

Proof. Let U be the transformed continuous-time solution with the transformation given by equation (3.2). As it is the case for the proof of Proposition 2.1, the smoothness assumption on the payoff function guarantees that U is $2q + 2$ times differentiable and the derivatives are bounded in the time region of interest. To demonstrate that the local error is of order $\Delta^{(q+1)/2}$, it suffices to show

$$U(z, t - \Delta) = \sum_{i=1}^{k} p_i U\big(z + h_i \Delta^{1/2}, t\big) + O(\Delta^{(q+1)/2})$$

First, expanding the left-hand side into a Taylor series about (z, t), we get

$$U(z,t-\Delta) = U(z,t) + \sum_{j=1}^{[q/2]} \Delta^j \frac{\partial^j}{\partial t^j} U(z,t)/j! + O(\Delta^{(q+1)/2})$$

Now, let Π denote the first term in the right-hand side of equation (3.7). We want to show that Π deviates from the right-hand side of equation (3.8) with an error at most $O(\Delta^{(q+1)/2})$. Expanding Π into a Taylor series about (z, t), we have

$$\Pi = U(z,t) + \sum_{i=1}^{k} \sum_{j=1}^{q} p_i h_i^j \Delta^{j/2} \frac{\partial^j}{\partial z^j} U(z,t)/j! + O(\Delta^{(q+1)/2})$$

Substituting the first q moment assumptions shows

$$\Pi = U(z,t) + \sum_{j=1}^{[q/2]} \Delta^j \frac{1}{2^j} \frac{\partial^{2j}}{\partial z^{2j}} U(z,t)/j! + O(\Delta^{(q+1)/2} \tag{3.10}$$

Now, since U satisfies $U_{zz}(z,t)/2 + U_t(z,t) = 0$, differentiating this equation once with respect to t or twice with respect to z, we have $U_{tt}(z,t) = -\frac{1}{2}U_{zzt}(z,t)$ and $U_{zzt}(z,t) = -\frac{1}{2}U_{zzzz}(z,t)$. Hence, $U_{tt} = -\frac{1}{2} \times (-\frac{1}{2})U_{zzzz} = \frac{1}{4}U_{zzzz}$. More generally, differentiating j times with respect to t and $2j$ times with respect to z, we get

$$\frac{\partial^j}{\partial t^j} U(z,t) = \frac{1}{2^j} \frac{\partial^{2j}}{\partial z^{2j}} U(z,t) \tag{3.11}$$

Hence, substituting equation (3.11) into equation (3.10) yields the desired result on the local error. To complete the proof, it remains to show the rate of convergence of the discrete-time solution. This is easily done following the proof of Proposition 2.1. □

Table 3.1 shows the number of moments matched by various multinomial approximations. The first row of the table shows an asymmetric trinomial procedure suggested in a multidimensional context by He (1990) and Amin (1991). The next two rows show the binomial and trinomial procedures. It is interesting that the usual binomial model not only matches the first two moments of a normal distribution, but a third one as well. The last row shows a higher order pentanomial procedure that appears new in the literature.[5]

The rate of convergence of the multinomial models, like the binomial one, depends crucially on the smoothness of the payoff function. For payoff functions that are smooth enough, high-order convergence is guaranteed by Proposition 3.1.[6] However, for those payoff functions that are not continuously differentiable, like the standard call payoff function, the rate of convergence is the smaller of the approximate rate at the boundary and the rate of the multinomial scheme's truncation error. For example, for smooth payoff functions, the rate of convergence of the trinomial model is $\Delta^2 = 1/n^2$. But it still has only a (uniform) rate of $1/\sqrt{n}$ for the standard call option on the

[5] The pentanomial procedure uses step sizes of $\pm 4.94842\sqrt{\Delta}$, etc., because it is impossible to match the desired moments of a normal distribution using step sizes of $\pm\sqrt{\Delta}$, $\pm 2\sqrt{\Delta}$, and $\pm 3\sqrt{\Delta}$.

[6] Proposition 3.1 implies the result of Proposition 2.1, but has a stronger assumption on the smoothness.

TABLE 3.1

Order of Multinomial Models

The table provides the order of the local error and the rate of convergence of multinomial models when the payoff functions are smooth enough. The p vector and h vector are given in the second and third columns. The step size is Δ, which is proportional to $1/n$ in an n-period multinomial model.

Method	p	h	q	Local error	Convergence
Asymmetric trinomial	$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$	$(-\frac{\sqrt{2}}{2},\sqrt{2},-\frac{\sqrt{2}}{2})$	2	$O(\Delta^{3/2})$	$O(\Delta^{1/2})$
Binomial	$(\frac{1}{2},\frac{1}{2})$	$(-1,1)$	3	$O(\Delta^2)$	$O(\Delta)$
Trinomial	$(\frac{1}{6},\frac{2}{3},\frac{1}{6})$	$(-\sqrt{3},0,\sqrt{3})$	5	$O(\Delta^3)$	$O(\Delta^2)$
Pentanomial	$(p_1,p_2,p_3,p_2,p_1)^*$	$(-a_1,-a_2,0,a_2,a_1)^\dagger$	7	$O(\Delta^4)$	$O(\Delta^3)$

$^*p_1 = 0.00261961$, $p_2 = 0.181415$, $p_3 = 0.636647$.
$^\dagger a_1 = 4.94842$, $a_2 = 1.64947$.

binomial tree. With the smoothing method, the rate goes up to $1/n$. As shown by Thomée and Wahlbin (1974), if a high-order smoother is used, the accuracy can be improved up to the rate of the truncation error, $1/n^2$.

To illustrate, consider the European call option valued earlier by using binomial methods. The first panel of Table 3.2 reports the numerical results by using the standard trinomial method (the trinomial model in Table 3.1). Like Table 2.1, the five columns of the first panel of Table 3.2 provide the number of periods, the exact Black–Scholes price, the error of the trinomial price, the ratio of the errors, and the error of the extrapolated solution. Since the $1/n^2$ rate of convergence is considered for the trinomial model, the extrapolated solution is computed as $(4C_{2n} - C_n)/3$, rather than $2C_{2n} - C_n$ as the case in the binomial model. Table 3.2 shows that the ratio varies from 3.1644 to 6.4835, and to -0.4443 as n increases. There is no indication of fast convergence at all. For example, when $n = 1280$, the trinomial price still has an error of -0.0024, almost of the same magnitude as the binomial error (see Table 2.1). This should not happen with smooth payoff functions, but occurs here due to the non-smoothness of the call payoff function. Furthermore, the extrapolated solutions show little improvement in reducing the error.

Although not reported in the table, the earlier BS adjustment and the smoothing method do not appear here to achieve the potential $1/n^2$ rate of convergence of the trinomial model.[7] This is because, with these two procedures, either the solution or the payoff function is still not smooth enough. Intuitively, the closer the time of the BS adjustment to today, the more accurate the solution should be. In the special case where the adjustment time is today, we get the exact BS price. Hence, the BS adjustment with a step size larger than previously used may cause the trinomial model to have $1/n^2$ rate of convergence. For simplicity, we apply the BS adjustment at time $[T\sqrt{n/10}]/n$, where $[T\sqrt{n/10}]$ is the integer part of $T\sqrt{n/10}$. The time to maturity

[7] For example, the errors from the BS adjustment are not much different from the binomial case, and the ratio varies from -0.0067 to -30.3617, and to 0.0830 as n increases from 320 to 640, and to 1280.

TABLE 3.2
Trinomial Model

Consider the valuation of a European call option on a stock whose price is 100. The strike is 100, the time to maturity is one year, the volatility is 40%, and the continuously compounded annual interest is 6%. In the table, the first two columns of the first panel are the number of periods (n) of the trinomial model and the Black–Scholes price. The next three columns are the error of the trinomial price, the ratio of the error, and the error of the extrapolated solution. The second panel reports the corresponding results obtained by using a large-step Black–Scholes (LBS) adjustment.

n	Exact	Trinomial	Error ratio	Extrapolated
10	18.47260446	−0.19310730		
20	18.47260446	−0.06102509	3.16439174	−0.01699769
40	18.47260446	−0.00941243	6.48345738	0.00779179
80	18.47260446	0.00720457	−1.30645215	0.01274357
160	18.47260446	0.00922698	0.78081594	0.00990112
320	18.47260446	0.00584260	1.57925997	0.00471447
640	18.47260446	0.00105423	5.54207021	−0.00054190
1280	18.47260446	−0.00237290	−0.44427702	−0.00351528

n	Exact	LBS adjustment	Error ratio	Extrapolated
10	18.47260446	−0.00438731		
20	18.47260446	−0.00107186	4.09319744	0.00003330
40	18.47260446	−0.00021631	4.95512945	0.00006887
80	18.47260446	−0.00005450	3.96907669	−0.00000056
160	18.47260446	−0.00001351	4.03454030	0.00000016
320	18.47260446	−0.00000340	3.97881976	−0.00000002
640	18.47260446	−0.00000085	3.99579080	0.00000000
1280	18.47260446	−0.00000021	3.98838098	0.00000000

from $[T\sqrt{n/10}]/n$ to T is only a fraction of T, and approaches zero as n increases without bound. In other words, for large n, the time of the BS adjustment is still close to maturity. Numerical results from this large-step BS adjustment (LBS adjustment) are reported in the second panel of Table 3.2. Interestingly, with the LBS adjustment, the solutions are highly accurate. For example, the binomial solution without adjustment has an error -0.0024 for $n = 1280$, but the error of the LBS adjustment is only -0.00000021. The rate of convergence of the LBS adjustment is clearly $1/n^2$ as the ratios are close to 4. The high accuracy is also confirmed by the extrapolated solutions, which are accurate up to the fifth digit with n as small as $n = 40$.

4. AMERICAN OPTIONS

Thus far our discussion has been focused on European options. In fact, both the BS adjustment and the smoothing procedure are also applicable to American options. Consider the valuation of an American put option price with the same parameters as in the European call option example. The results are reported in Table 4.1. As before, the

TABLE 4.1
American Option

Consider the valuation of an American put option on a stock whose price is 100. The strike is 100, the time to maturity is one year, the volatility is 40%, and the continuously compounded annual interest is 6%. In the table, the first column of the first panel is the number of periods (n) in the binomial model. The next four columns are the usual binomial put price, its difference over time, the ratio of the differences, and the extrapolated solution. The other two panels are the corresponding results obtained by using the Black–Scholes (BS) adjustment and the smoothing procedure, respectively.

n	Binomial	Difference	Difference ratio	Extrapolated
10	13.27497441			
20	13.31784106	0.04286664		13.36070770
40	13.32555369	0.00771263	5.55797769	13.33326632
80	13.32176043	−0.00379326	−2.03324744	13.31796717
160	13.31410339	−0.00765704	0.49539499	13.30644636
320	13.30661139	−0.00749200	1.02202811	13.29911939
640	13.30038206	−0.00622933	1.20269808	13.29415273
1280	13.29565580	−0.00472626	1.31802574	13.29092954
n	BS adjustment	Difference	Difference ratio	Extrapolated
10	13.37783463			
20	13.34536217	−0.03247246		13.31288971
40	13.32367643	−0.02168574	1.49741103	13.30199070
80	13.31117601	−0.01250042	1.73480059	13.29867559
160	13.30399931	−0.00717670	1.74180560	13.29682261
320	13.30007858	−0.00392073	1.83045198	13.29615786
640	13.29803658	−0.00204200	1.92003938	13.29599458
1280	13.29694381	−0.00109277	1.86865232	13.29585104
n	Smoothing	Difference	Difference ratio	Extrapolated
10				
20		−0.11261697		13.31850681
40		−0.06195137	1.81782869	13.30722105
80		−0.03430790	1.80574632	13.30055661
160		−0.01859832	1.84467703	13.29766787
320		−0.00968624	1.92007598	13.29689370
640		−0.00560373	1.72853621	13.29537249
1280		−0.00264434	2.11913684	13.29568753

first column is the number of periods (n). Since there is no exact solution for the American put, the next four columns of the first panel report the standard binomial put price, their differences, and the ratios of the differences. The other two panels are the corresponding results obtained by using the Black–Scholes adjustment and the smoothing procedure, respectively.

Since the exact error cannot be computed as there are no available analytical formulas for the American put price, the differences of the numerical solutions are relied upon to indicate the speed of convergence. As n increases, the differences of the binomial solutions oscillate somewhat, but are generally within penny accuracy. This is also true for both the BS adjustment and the smoothing procedure. The differences demonstrate that the numerical solutions converge fairly well to the continuous limit. But this does not mean that all of the three procedures have the same rate of convergence. Information on the rate of convergence is contained in the ratio of the differences. As is known, if any of the procedures has a convergence rate of $1/n$, the corresponding ratios should be close to 2. Indeed, the ratios from either the BS adjustment or the smoothing procedure are fairly close to 2. In contrast, the ratios from the binomial model oscillate irregularly between -2.0332 and 5.5580.

The numerical experiments suggest that both the BS adjustment and the smoothing procedure appear to have a convergence rate of $1/n$ even for American options. In contrast, the rate of convergence of the standard binomial model, like its applications to European options, oscillates irregularly. This is further confirmed by examining the extrapolated solutions, which are given in the last column where it can be seen that the differences of the extrapolated solutions based on either the BS adjustment or the smoothing procedure shrink steadily. In contrast, the differences of the extrapolated binomial solutions oscillate, making it difficult to determine their accuracy.

In comparison with other numerical approaches, our methods have known rates of convergence that are useful for extrapolation, while the rates of convergence of other approaches are mostly unknown (e.g., see Huang, Subrahmanyam, and Yu 1996). Moreover, our methods are very simple to implement and easy to understand, and are applicable to a general stock price process. In contrast, many existing approaches are fairly complex and suitable only for specific problems.

Although there is extensive literature on valuing American options, little is available on the computation of the critical price at which an option should be optimally exercised. In what follows we apply both the BS adjustment and the smoothing procedure to obtain the critical price. In contrast with the $1/n$ rate of convergence for the option price, it is striking that the critical price converges at a rate of only $1/\sqrt{n}$. We will illustrate this by computing the critical price of the previous put option example.

Theoretically, it is well known that the critical price, S^*, of an American put is the root of

$$f(S^*) \equiv P(S^*) - (K - S^*) = 0, \tag{4.1}$$

where $P(S^*)$ is the American put option price when the stock price equals S^*. As $f(S) > 0$ when $S > S^*$ and $f(S) < 0$ when $S < S^*$, we can use the standard bisection method to find the root of the nonlinear function $f(S)$. Since the exercise price is bounded between the strike K and $2rK/(\sigma^2 + 2r)$ (Ingersoll 1987, p. 375), we only need to search in the range $[2rK/(\sigma^2 + 2r), K]$. With the American put option price, $P(S)$, computed by either the BS adjustment or the smoothing procedure, the bisection method can be easily implemented. Hence, the root of (4.1) provides us a numerical approximation of the exact critical price.

The numerical results are provided in Table 4.2. There are two convincing "pieces of" evidence supporting the $1/\sqrt{n}$ rate. First, the price ratios are approaching $\sqrt{2} = 1.41$ as n increases. For example, when $n = 160$, the ratios are 1.431 and 1.363, already close to $\sqrt{2}$. Second, the extrapolated solution, $(\sqrt{2}\,S_{2n} - S_n)/(\sqrt{2} - 1)$, converges steadily.

TABLE 4.2
Critical Price

Consider the valuation of the critical or exercise price for an American put option on a stock. The strike price is 100, the time to maturity is one year, the volatility is 40%, and the continuously compounded annual interest is 6%. In the table, the first column of the first panel is the number of periods (n) in the binomial model with the Black–Scholes (BS) adjustment prior to maturity. The next four columns are the critical price, the difference of the computed critical prices, the ratio of the differences, and the extrapolated solution. The second panel provides the same results by using the smoothing procedure.

	BS Adjustment	Difference	Difference ratio	Extrapolated
10	64.13948376			
20	63.01636124	1.12312253		60.30490360
40	62.25847626	0.75788497	1.48191687	60.42878009
80	61.74383577	0.51464049	1.47264932	60.50138372
160	61.38429845	0.35953732	1.43139659	60.51629858
320	61.12979222	0.25450623	1.41268573	60.51535983
640	60.94894032	0.18085191	1.40726317	60.51232519
1280	60.82004496	0.12889536	1.40309091	60.50886403
n	Smoothing	Difference	Difference ratio	Extrapolated
10				
20		0.99015123		60.46008717
40		0.65944077	1.50150138	60.59905207
80		0.48748350	1.35274481	60.52671015
160		0.33700952	1.44649773	60.55297697
320		0.24730681	1.36271831	60.52223165
640		0.17506599	1.41264911	60.52157044
1280		0.12642396	1.38475322	60.51257872

Indeed, the differences are smaller than a few pennies when $n > 160$, and are less than one penny when $n > 320$. It is difficult to show theoretically why the rate of convergence of the critical price is $1/\sqrt{n}$. But a simple error analysis reveals that the accuracy of the critical prices depends not only on the option price approximations, but also on approximations to the first-order derivative. This may help clarify why there is a difference in numerical accuracy between an American option price and the associated critical price.

5. INTEREST RATE CON NGENT CLAIMS

The analysis on the rate of convergence also applies to interest rate contingent claims and other derivatives. Instead of the Black–Scholes formula, an appropriate European option pricing formula can be used for the adjustment prior to maturity. Alternatively, especially in the absence of a closed-form solution to the European option, the smoothing technique can be used to improve the rate of convergence of existing binomial and trinomial models for interest rate derivatives.

Although both the adjustment and smoothing methods are straightforward to apply to interest rate options, a major problem in valuing interest rate contingent claims is to compute the price of a zero coupon bond. As the payoff function of such a bond is smooth, neither the adjustment nor the smoothing method is needed in using existing binomial and trinomial models to compute the bond price.

However, those existing binomial and trinomial models are applications of explicit finite difference methods to the valuation equation of the interest rate derivatives (e.g., see Hull and White 1990), and their accuracy is limited to Δ or $1/n$. In what follows, we propose a new trinomial model that has $1/n^2$, a much faster rate of convergence. Then we provide a numerical example to compare its accuracy with standard binomial and trinomial models.

Suppose that the instantaneous interest rate, $r(z,t)$, is a function of a Wiener process z (in the risk-neutral probabilities) at time t. It is well known (Cox, Ingersoll, and Ross 1985) that any interest rate contingent claim will be a function of z and t, and its value, $V(z,t)$, must satisfy a valuation equation. For a wide class of specifications, examples of which include Vasicek (1977), Dothan (1978), Cox et al. (1985), Longstaff (1989) (corrections in Beaglehole and Tenney 1992), and Black and Karasinski (1991), this valuation equation can be transformed to the following partial differential equation

$$V_t(z,t) = -\frac{1}{2}V_{zz}(z,t) + r(z,t)V(z,t),$$

with appropriate boundary conditions. To obtain the highly accurate trinomial method, we differentiate this equation with respect to t to obtain

$$V_{tt}(z,t) = -\frac{1}{2}V_{zzt}(z,t) + [r(z,t)V(z,t)]_t$$

Then, differentiating equation (5.1) twice with respect to z and substituting the result, $V_{zzt} = V_{tzz} = -V_{zzzz}/2 + (rV)_{zz}$, into equation (5.2), we have a useful expression for V_{tt}:

$$V_{tt} = \frac{1}{4}V_{zzzz}(z,t) - \frac{1}{2}[r(z,t)V(z,t)]_{zz} + [r(z,t)V(z,t)]_t$$

Now, by using a second-order Taylor series expansion, we can relate the value of V at time $t - \Delta$ to those at t as follows:

$$V(z,t-\Delta) = V(z,t) - \Delta V_t(z,t) + \frac{\Delta^2}{2}V_{tt}(z,t) + O(\Delta^3)$$

The idea is to replace the derivatives of equation (5.4) by discrete approximations and then solve for $V(z,t-\Delta)$. To do this, we denote by $\psi(V)$ the left-hand side of the following important finite difference approximation,

$$\frac{V(z-\sqrt{3\Delta},t) - 2V(z,t) + V(z+\sqrt{3\Delta},t)}{3\Delta}$$
$$= V_{zz}(z,t) + \frac{\Delta}{4}V_{zzzz}(z,t) + O(\Delta^2)$$

Making use of equations (5.1) and (5.3), and the finite difference operator ψ, we obtain from equation (5.4) a complete discrete approximation to the valuation function,

$$(5.6)\quad V(z,t-\Delta)$$
$$= V(z,t) + \frac{\Delta^2}{2}\psi(V)$$
$$\frac{\Delta^2}{4}\left[\psi(rV + \frac{r(z,t)V(z,t) - r(z,t-\Delta)V(z,t-\Delta)}{\Delta}\right]$$

Solving equation (5.6) for $V(z,t-\Delta)$, we obtain a trinomial model for the contingent claim,

$$(5.7)\quad V(z,t-\Delta) = p_1 V(z-\sqrt{3\Delta},t) + p_2 V(z,t) + p_3 V(z+\sqrt{3\Delta},t) + O(\Delta^3),$$

ere

$$p_1 = \frac{}{6}\,\frac{-r(z-\sqrt{3\Delta},t)\Delta/2}{1+r(z,t-\Delta)\Delta/2}$$
$$p_2 = \frac{2}{3}\,\frac{1-r(z,t)\Delta/2}{1+r(z,t-\Delta)\Delta/2},$$
$$p_3 = \frac{}{6}\,\frac{1-r(z+\sqrt{3\Delta},t)\Delta/2}{1+r(z,t-\Delta)\Delta/2}$$

This trinomial model discounts the future values with weights p_1, p_2, and p_3. The p's are approximately the "risk neutral probabilities" as they are positive and their sum, $p_1 + p_2 + p_3$, are approximately equal to $(1 - r\Delta)$. If the payoff function is smooth, such as the payoff function of a zero-coupon bond, the trinomial model should converge to the continuous-time solution at a rate $1/n^2$.

In contrast, existing binomial or trinomial methods for interest contingent claims have a rate no better than $1/n$. This is because, applying the standard explicit finite difference methods with $\Delta_x = \sqrt{\Delta}$ to equation (5.1), one has a trinomial model similar to equation (2.4),

$$V(z,t-\Delta) = \frac{1}{2}V(z-\sqrt{\Delta},t) + \frac{1}{2}V(z+\sqrt{\Delta},t) - r(z,t)\Delta V(z,t) + O(\Delta^2),$$

but the local error is $O(\Delta^2)$ rather than $O(\Delta^3)$. It is observed that replacing $r(z,t)\Delta V(z,t)$ by $r(z,t-\Delta)\Delta V(z,t-\Delta)$ will not change the order of the local error. As a result, one obtains a binomial model with the same theoretical accuracy,

$$V(z,t-\Delta) = \frac{\frac{1}{2}V(z-\sqrt{\Delta},t) + \frac{1}{2}V(z+\sqrt{\Delta},t)}{1+r(z,t-\Delta)\Delta} + O(\Delta^2)$$

This is Heston's (1995) binomial model. However, the trinomial model (5.8) is not the one often used. Motivated from analogues to the trinomial model for stock options, Hull and White (1990) suggest the use of spacing $\Delta_x = \sqrt{3\Delta}$ in applications to term structure models. With this spacing, it is easy to show that the trinomial model (5.8) becomes,

$$V(z,t-\Delta) = \frac{1}{6}V(z-\sqrt{\Delta},t)\left[\frac{2}{3} - r(z,t)\right]\Delta V(z,t) \tag{5.10}$$

$$6 \quad z+\sqrt{\Delta},t) + O(\Delta^2$$

This is the trinomial model used by Hull and White (1990) and others.

To illustrate the rate of convergence, we consider the valuation of a zero coupon bond in the Cox et al. (1985) square-root model, where the instantaneous interest rate, $r(z,t)$ follows the diffusion process:

(5.
$$dr = \kappa(\theta - r)dt + \sigma\sqrt{r}\,dz,$$

where κ, θ, and σ are parameters of the model, and dz is the standard Gaussian–Wiener process. Let $\kappa = 0.15$, $\theta = 0.054$, and $\sigma = 0.18$. These are not unreasonable values for the parameters. Suppose the current rate is $r_0 = 6\%$. We compute the price of a zero-coupon bond with face value 1 and maturities of 5 and 30 years.

The results are provided in Table 5.1. The first column is $n = 1/\Delta$, the number of periods in discrete-time models. There are numerical prices. The first one is computed by using the standard binomial model (5.8). The second is from the trinomial model (5.10), and the third is from the new trinomial model (5.7). The second column of the table is the analytical price given by Cox et al. (1985), and the rest of the columns are the errors and ratios of the three discrete-time models. Consider the 5-year bond first; the results are in the first panel of the table. Theoretically, the binomial and the trinomial models should converge at the same rate, but the numerical results show that the binomial model performs better despite the fact that it uses less information than the trinomial model. Because the payoff function is smooth and the bond price is infinitely differentiable, the ratios behave so nicely that they are almost exactly equal to 2. For the new trinomial model, it is remarkable that the accuracy is up to the fourth digit with n as small as 20. In contrast, the same accuracy has to take $n = 1280$ to achieve for the binomial model and the trinomial model. Again because of the smoothness of the payoff function, the ratios are almost 4, confirming a theoretical $1/n^2$ rate of convergence. For the 30-year bond, the same observations hold as seen from the results in the second panel of the table. However, as the maturity lengthens, a larger n is required to obtain the same accuracy. It is easy to verify that the ratios are almost 2 and 4 as n increases beyond $n = 1280$.

6. CONCLUSIONS

This paper analyzes the rate of convergence of discrete-time models to their continuous-time analogues. We find that the rate of convergence depends on both the local error of the discrete-time models and the smoothness of the payoff functions of a specific derivative. As most option payoff functions are often of all-or-nothing type,

TABLE 5.1
Bond Price Valuation

The table provides the price of a zero-coupon bond with face value 1 and maturities of 5 and 30 years. The first column is the number of periods of discrete-time models. The second column is the exact price given by the Cox et al. (1985) formula in which the instantaneous interest rate follows the square-root process,

$$dr = \kappa(\theta - r)\,dt + \sigma\sqrt{r}\,dz,$$

where the parameters are $\kappa = 0.150$, $\theta = 0.054$, $\sigma = 0.180$, and $r_0 = 6\%$. The rest of the columns are the error and the ratio of the errors of three discrete-time models: the standard binomial model, the standard trinomial model, and the new trinomial model.

		Binomial		Trinomial		New Trinomial	
	Exact	Error	Error ratio	Error	Error ratio	Error	Error ratio
5-year bond							
10	0.76340934	−0.00655702		−0.00037357		−0.00037357	
20	0.76340934	−0.00322681	2.03204151	0.00711609	2.01899450	−0.00009375	3.98452582
40	0.76340934	−0.00160006	2.01668273	0.00354038	2.00997877	−0.00002346	3.99614239
80	0.76340934	−0.00079665	2.00849207	0.00176570	2.00508514	−0.00000587	3.99875825
160	0.76340934	−0.00039747	2.00428460	0.00088172	2.00256331	−0.00000147	3.99926055
320	0.76340934	−0.00019852	2.00214819	0.00044058	2.00128572	−0.00000037	4.00039047
640	0.76340934	−0.00009921	2.00107961	0.00022022	2.00064465	−0.00000009	3.99796903
1280	0.76340934	−0.00004959	2.00053715	0.00011009	2.00032225	−0.00000002	3.99975543
30-year bond							
10	0.29544480	−0.18926386		−0.11465518		−0.11465518	
20	0.29544480	−0.14181983	1.33453735	0.22198694	1.46836545	−0.06596376	1.73815402
40	0.29544480	−0.09147413	1.55038177	0.13644313	1.62695570	−0.03100200	2.12772582
80	0.29544480	−0.05075034	1.80243389	0.07512791	1.81614452	−0.01179293	2.62886367
160	0.29544480	−0.02565978	1.97781642	0.03823067	1.96512135	−0.00366776	3.21529586
320	0.29544480	−0.01263399	2.03101196	0.01889883	2.02291194	−0.00099492	3.68648273
640	0.29544480	−0.00623147	2.02744905	0.00933472	2.02457430	−0.00025455	3.90852260
1280	0.29544480	−0.00309056	2.01629201	0.00463243	2.01507871	−0.00006399	3.97788257

their numerical approximations may not have the rate of convergence commonly expected. In particular, we show that the rate of convergence of the standard binomial model will not converge to the Black–Scholes formula value at the $1/n$ rate at all nodes of the binomial tree. Solutions to the nonsmoothness of the payoff functions are proposed. In particular, a Black–Scholes adjustment and a smoothing procedure are provided to make the binomial model converge uniformly at the $1/n$ rate so that the standard Richardson extrapolation can be used to obtain solutions with higher order of accuracy.

The proposed procedures are useful not only in binomial models and multinomial models, but also useful to any other discrete-time models that face the nonsmoothness problem of the payoff functions. In particular, our studies point out the need to analyze the impact of the nonsmoothness problem on various Monte Carlo schemes and the theoretical rate of convergence of many simulation procedures which are of interest both in theory and practice. Additionally, the nonsmoothness boundary conditions also appear to affect the rate of convergence of numerical solutions to stochastic differential equations which are widely used to value derivatives. These are interesting topics for future research.

REFERENCES

AMES, W. A. (1992): *Numerical Methods for Partial Differential Equations*, 3rd ed. New York: Academic Press.

AMIN, K. (1991): On the Computation of Continuous-Time Option Prices Using Discrete-Time Models, *J. Financial Quant. Anal.* 26, 477–496.

AMIN, K., and A. KHANNA (1994): Convergence of American Option Values from Discrete-to Continuous-Time Financial Models, *Math. Finance* 4, 289–304.

BEAGLEHOLE, D., and M. TENNEY (1992): Corrections and Additions to "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *J. Financial Econ.* 32, 345–353.

BLACK, F., and P. KARASINSKI (1991): Bond and Option Pricing when Short Rates Are Lognormal, *Financial Anal. J.* (July-August), 52–59.

BOYLE, P. P., J. EVNINE, and S. GIBBS (1989): Valuation of Options on Several Underlying Assets, *Rev. Financial Stud.* 2, 241–250.

BRENNER, P., V. THOMÉE, and L. WAHLBIN (1975): *Besov Spaces and Applications to Difference Methods for Initial Value Problems*. New York: Springer-Verlag.

BROADIE, M., and J. DETEMPLE (1996): American Option Valuation: New Bounds, Approximations, and a Comparison of Existing Methods, *Rev. Financial Stud.* 9, 1211–1250.

BROADIE, M., and J. DETEMPLE (1998): Recent Advances in Numerical Methods for Pricing Derivative Securities; in *Numerical Methods in Finance*, L. C. G. Rogers and D. Talay, eds. Cambridge: Cambridge University Press.

COX, J. C., J. E. INGERSOLL, and S. ROSS (1985): A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385–407.

COX, J. C., S. A. ROSS, and M. RUBINSTEIN (1979): Option Pricing: A Simple Approach, *J. Financial Econ.* 7, 229–263.

DERMAN, E., I. KANI, D. ERGENER, and I. BARDHAN (1995): Enhanced Numerical Methods for Options with Barriers, *Financial Anal. J.* (Nov-Dec), 65–74.

DOTHAN, M. (1978): On the Term Structure of Interest Rates, *J. Financial Econ.* 7, 229–264.

FRIEDMAN, A. (1964): *Partial Differential Equation of Parabolic Type*. Englewood Cliffs, NJ: Prentice-Hall.

HE, H. (1990): Convergence from Discrete to Continuous Time Contingent Claims Prices, *Rev. Financial Stud.* 3, 523–546.

HESTON, S. (1995): Discrete Versions of Continuous-Time Interest Rate Models, *J. Fixed Incomes*, 86–88.

HUANG, J., M. SUBRAHMANYAM, and G. YU (1996): Pricing and Hedging American Options: A Recursive Integration Method, *Rev. Financial Stud.* 9, 277–300.

HULL, J., and A. WHITE (1988): The Use of Control Variate Technique in Option Pricing, *J. Financial Quant. Anal.* 23, 237–251.

HULL, J., and A. WHITE (1990): Valuing Derivative Securities Using the Explicit Finite Difference Method, *J. Financial Quant. Anal.* 25, 87–99.

INGERSOLL, J. E. (1987): *Theory of Financial Decision Making*. Totowa, NJ: Rowman and Littlefield.

JOHNSON, N. L., S. KOTZ, and A. W. KEMP (1992): *Univariate Discrete Distributions*, 2nd ed. New York: Wiley.

LEISEN, D., and M. REIMER (1996): Binomial Models for Option Valuation—Examining and Improving Convergence, *Appl. Math. Finance* 3, 319–346.

LONGSTAFF, F. A. (1989): A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates, *J. Financial Econ.* 23, 195–224.

MADAN, D. B., F. MILNE, and H. SHEFRIN (1989): The Multinomial Option Pricing Model and its Brownian and Poisson Limits, *Rev. Financial Stud.* 2, 251–265.

RUBINSTEIN, M. (1994): Implied Binomial Trees, *J. Finance* 49, 771–818.

STRECOK, A. J. (1968): On the Calculation of the Inverse of the Error Function, *Math. Computation* 22, 144–158.

THOMÉE, V., and L. WAHLBIN (1974): Convergence Rates of Parabolic Difference Schemes for Non-smooth Data, *Math. Computation* 28, 1–13.

VASICEK, O. (1977): An Equilibrium Characterization of the Term Structure, *J. Financial Econ.* 5, 177–188.