

Multivariate Binomial Approximations for Asset Prices with Nonstationary Variance and Covariance Characteristics

Teng-Suan Ho
Richard C. Stapleton
Lancaster University

Marti G. Subrahmanyam
New York University

In this article, we suggest an efficient method of approximating a general, multivariate log-normal distribution by a multivariate binomial process. There are two important features of such multivariate distributions. First, the state variables may have volatilities that change over time. Second, the two or more relevant state variables involved may covary with each other in a specified manner, with a time-varying covariance structure. We discuss the asymptotic properties of the resulting processes and show how the methodology can be used to value a complex, multiple exerciseable option whose payoff depends on the prices of two assets.

In practice, many problems in the valuation of derivative assets are solved by using binomial approximations to continuous distributions. In this article, we

Earlier versions of this article have been presented at the European Finance Association Conference, Lisbon, Portugal; Aarhus University, Denmark; Melbourne University, Australia; University of Michigan, and Virginia Polytechnic Institute, USA. We are grateful to participants at these presentations and to John Chang, Chi-fu Huang (the editor), Apoorva Koticha, Ser-Huang Poon, Steven Tsay, and an anonymous referee for comments on previous drafts. Address correspondence and reprint requests to T. S. Ho, Department of Accounting and Finance, The Management School, Lancaster University, Lancaster LA1 4YX, England, U.K.

suggest an efficient method of approximating a general, multivariate lognormal distribution by a multivariate binomial process. There are two features of such multivariate lognormal distributions that are of interest. First, the state variables may have volatilities that change over different time intervals (i.e., exhibit a term structure of volatility), either because of changing volatility or mean reversion. Second, there may be two or more relevant state variables involved (e.g., a commodity price and a foreign exchange rate) which may covary with each other in a specified manner.

The binomial approach to the valuation and hedging of options has become increasingly important with the creation of new exotic derivative products. Many options with pathdependent payoffs (e.g., American options and Asian or average-rate options) may be valued using the binomial methodology.¹ Also, in the analysis of interest rate derivative products, it is often useful to model the construction and evolution of the term structure of interest rates using a binomial process. The present paper provides a general methodology for the construction of binomial approximations of multivariate lognormal distributions, both across state variables and across time.

In order to illustrate the type of problems that can be solved using the methods presented in this paper, consider the following example of an option on oil.² The contract is an American-style option which allows the holder to buy oil at a specified price denominated in Japanese yen. The payoff on the option in yen depends on two variables: the U.S. dollar-denominated oil price and the yen/dollar foreign exchange rate. At time 0, suppose that the volatility of the oil price is high, but is expected to decline over the life of the option. Also, suppose that there is significant mean reversion in the price of oil. In contrast, in this example, the yen/dollar exchange rate is assumed to have constant volatility and no mean reversion. Further, suppose that the correlation between the price of oil and the yen/dollar exchange rate is high, but is expected to fall.

We wish to value this oil option by approximating the true joint distribution of the variables with a bivariate binomial process. The method developed in this article is designed to solve this general problem of changing volatilities, mean reversion, and multiple state

¹ In the cast of many pathdependent options, the problem, however, is that the number of paths explodes with the number of stages in the binomial tree. Hence, such options can only be valued for quite small n -size binomial trees.

² Other examples of multivariate options include the "delivery option" that is embedded in many bond futures contracts, a "cocktail" option to receive the principal payment on a bond in one of many currencies at specified exchange rates, and a long-term American-style currency option under stochastic interest rates. For other examples, see Boyle, Evnine, and Gibbs (1989) and Stulz (1982).

variables, whereas existing models in the literature only deal with one or two aspects of the problem at a time.³ For example, Nelson and Ramaswamy (1990) show how to approximate various univariate distributions with binomial processes. Amin (1991) derives an alternative approach to the problem of a changing volatility function. However, neither article deals with changing volatility in a multivariate context. Amin (1991) and Boyle (1988) provide models of the multivariate case, but they deal with either one asset in the context of changing volatility or multiple assets with a constant correlation structure. Amin, for example, assumes a constant covariance matrix when dealing with the multivariate problem explicitly. The incremental contribution of our article is that it deals with the general multivariate problem, where the assets may have volatilities changing over time but differently across assets, a changing covariance matrix, and differential mean reversion.

In the literature, one method suggested to solve both the univariate and multivariate problem involves a change in the conditional probability computed on a node-by-node basis. In this article, we explicitly recognize that the two problems are really two aspects of the same general issue. In other words, we need, in general, to choose the probabilities so that both the time series (term structure of volatility) and cross-sectional (correlation structure) characteristics of the underlying asset prices are satisfied. We show that an application of the linear multiple regression property of joint normally distributed variables leads to a formula for the appropriate conditional probabilities. Application of this linear relationship yields a multivariate binomial distribution which is simple to compute and apply.

In Section 1, we introduce the problem in a formal manner and define our notation. In Section 2, we review the literature on the topic of binomial approximation of lognormally distributed variables and relate our work to the previous research. Section 3 describes the method in the context of a single variable that has both volatility changes and mean reversion. Section 4 derives and demonstrates a similar methodology for the general multivariate case described above. Our conclusions are summarized in Section 5.

1. Statement of the Problem and Notation

We assume that the prices of each of the underlying assets, $X_1, X_2, \dots, X_J, \dots, X_J$ follows a lognormal diffusion process:

$$d \ln X_j = \mu_j(X_j, t)dt + \sigma_j(t)dZ_j, \quad j = 1, 2, \dots, J, \quad (1)$$

³ Note that a term structure of volatility can result either from mean reversion, hanging conditional volatility, or both.

where μ_j and σ_j are the instantaneous drift and volatility of $\ln X_j$, and dZ_j is a standard Brownian motion. The instantaneous correlation between the Brownian motions dZ_j and dZ_k is $\rho_{j,k}(t)$. The instantaneous drift in Equation (1) is a function of X and t , which allows for mean reversion that may change over time. We assume that $\mu(X_j, t)$ is linear in X and the instantaneous variances and covariances are non-stochastic functions of time. Hence, the asset prices are lognormally distributed at any time t . There is a finite number (m) of future dates in the time interval $[0, T]$ at which we are interested in the asset prices. The dates are numbered $t_1, t_2, \dots, t_i, \dots, t_m$, where $t_m = T$, and these are determined by the requirements of the option valuation problem that is being solved.⁴ We are interested in the joint distribution of the prices of the assets on these dates.

We denote the unconditional mean (at time 0) of the logarithmic asset return at time i as $\mu_{i,j}$. The conditional volatility over the period $i-1$ to i is denoted $\sigma_{i-1,i,j}$, and the unconditional volatility is $\sigma_{0,i,j}$. Also, the covariances across assets at time t_i are, respectively,

$$\begin{aligned}\text{cov}_0(\ln X_{i,j}, \ln X_{i,k}) &= \rho_{0,i,j,k} \sigma_{0,i,j} \sigma_{0,i,k} \equiv \sigma_{0,i,j,k}, \\ \text{cov}_{i-1}(\ln X_{i,j}, \ln X_{i,k}) &= \rho_{i-1,i,j,k} \sigma_{i-1,i,j} \sigma_{i-1,i,k} \equiv \sigma_{i-1,i,j,k},\end{aligned}$$

where the correlation coefficients $\rho_{0,i,j,k}$ and $\rho_{i-1,i,j,k}$ are specified exogenously.⁵

The conditional and unconditional volatilities depend, in general, on the functional form of $\mu_j(X_j, t)$ and $\sigma_j(t)$. For example, if μ_j is a constant, $\ln X_j$ follows a random walk with

$$\sigma_{0,i,j} = \sqrt{\frac{1}{t_i} \int_0^{t_i} \sigma_j^2(t) dt}$$

and

$$\sigma_{i-1,i,j} = \sqrt{\frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \sigma_j^2(t) dt}.$$

In general, however, the volatilities $\sigma_{0,i,j}$ and $\sigma_{i-1,i,j}$ are complex functions of the instantaneous volatilities $\sigma_j(t)$. In this article, instead of starting with a particular specified process for X_j , we assume, more

⁴ For example, if, as in Breen (1991), we are approximating an American option using the Geske and Johnson (1984) technique, using a trivariate distribution, we would be interested in three dates: t_1, t_2, T .

⁵ Note that, given our assumptions, the volatilities and covariances are state independent.

generally that the relevant means, volatilities, and conditional volatilities are given exogenously.⁶

We wish to approximate the process in Equation (1) with a sequence of binomial distributions such that the mean, variance, and covariance characteristics converge to these given values. The sequence of binomial distributions yields vectors of outcomes of $\mathbf{X}_{t,j}$, where $\mathbf{X}_{t,j}$ is the price of asset j at time t , $t = 1, 2, \dots, m$. Conditional probabilities have to be so chosen that the volatility specifications and correlations are satisfied for each time period.

Since we are concerned only with a finite set of dates (m), the data input required consists of m vectors of exogenously given means $(\bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_t, \dots, \bar{\mu}_m)$, where $\bar{\mu}_t = (\mu_{t,1}, \dots, \mu_{t,J})$, and m unconditional covariance matrices $(\Sigma_{0,1}, \Sigma_{0,2}, \dots, \Sigma_{0,m})$, where⁷

$$\Sigma_{0,t} = \begin{bmatrix} \sigma_{0,t,1}^2 & \cdots & \sigma_{0,t,1,k} & \cdots & \sigma_{0,t,1,J} \\ & & & & \\ & & \sigma_{0,t,k}^2 & \cdots & \sigma_{0,t,k,J} \\ & & & & \\ & & & & \sigma_{0,t,J}^2 \end{bmatrix}, \quad (2)$$

and $m - 1$ conditional covariance matrices where

$$\Sigma_{t-1,t} = \begin{bmatrix} \sigma_{t-1,t,1}^2 & \cdots & \sigma_{t-1,t,1,k} & \cdots & \sigma_{t-1,t,1,J} \\ & & & & \\ & & \sigma_{t-1,t,k}^2 & \cdots & \sigma_{t-1,t,k,J} \\ & & & & \\ & & & & \sigma_{t-1,t,J}^2 \end{bmatrix}. \quad (3)$$

The exogenously given drift and volatility terms are satisfied asymptotically by choosing the number of binomial stages between 0 and t_1 , t_1 and t_2 , etc. There are n_1 binomial stages between 0 and t_1 , t_2 stages between t_1 and t_2 , and so on. As we discuss below, the ability to choose the “denseness” of the binomial tree between any two dates represents a significant difference between our model and the previous binomial approximations available in the literature.

⁶ There are, however, natural restrictions that have to be imposed on the variance-covariance matrices. In particular, the relevant variance-covariance matrices have to be positive semidefinite.

⁷ Note that it is not restrictive in any sense to confine our attention to the series of date $t_1, t_2, \dots, t_1, \dots, t_m$ because the time intervals between the dates can be made as small as desired.

In our method, we first construct a binomial approximation for each of the state variables on which the option payoff depends. We then model the given covariance structure by adjusting the conditional probabilities of $X_{i,k}$ given the outcomes of $X_{i,j}$.⁸ Since our methodology is closely related to, but significantly different from, the previous work cited above, we now discuss the contribution of these papers.

2. Relationship to the Literature on Binomial Approximations and Option Pricing

The method we use to approximate the multivariate process is closely related to previous contributions by Amin (1990, 1991), Amin and Bodurtha (1995), Boyle (1988), He (1990), and Nelson and Ramaswamy (1990) [henceforth, NR].⁹ The basic idea in NR is to approximate a given univariate process for the price for the underlying asset by a "simple" binomial process. In the NR context, "simple" means that the number of nodes of the binomial process increases linearly with time. In order to ensure that the process has the desired variance characteristics, while remaining simple, NR suggest an adjustment in the conditional probabilities of the binomial process over time.

In this article, we also construct simple binomial processes. However, in contrast to NR, we allow the number of binomial stages n_i between any two points t_{i-1} and t_i to be greater than 1. This means that, in the univariate case, the NR method can be regarded as a special case of ours, where $t_i - t_{i-1}$ is the same for all i and also $n_i = 1$ for all i . In our approach, even in the multivariate case, we can, if we wish, accommodate a change in volatility by allowing n_i to vary. This follows from a suggestion of Amin (1991), but with an important difference. We do not need to change the time intervals (t_{i-1}, t_i) in our more general setting. This means that our method is readily extendible to the multivariate case.

It is well known that a multivariate process across two or more assets can also be modeled by changing the conditional probabilities associated with the nodes. Boyle (1988) shows how to construct a trinomial distribution for each variable and then combine them in a lattice so that the variables have the required variance and covariance characteristics. Amin (1991) models the covariance characteristics of the assets assuming a constant variance-covariance matrix (i.e., one where the asset prices have constant volatilities over time). He then

⁸ An alternative method when the correlation between the variables is high is to orthogonalize the variables and then construct the binomial tree.

⁹ A method similar to that in Nelson and Ramaswamy (1990) was suggested and implemented in Stapleton and Subrahmanyam (1988).

compares the efficiency of Boyle's (1988) method against a range of alternatives, including two where the variables are joint binomially distributed. He finds that, from the standpoint of computational efficiency, the latter approaches are marginally inferior. However, if the factors generating the asset prices ("shocks" in his Example 2) are chosen so as to be independent, the binomial model has the advantage that it does not suffer from the problem of "negative probabilities" that arises in the Boyle approach. Amin (1990) and He (1990) use a trinomial distribution to model a bivariate distribution of two factors. They do it in such a way that the factors are uncorrelated, but not independent. He (1990) shows that the distribution converges to the multivariate Brownian motion assumed and that option prices also converge. These models do not incorporate changing conditional variances or mean reversion.

In our model, we choose the multivariate binomial method as suggested by Amin's work. The major difference between our model and the work by Amin and Boyle is that unlike their approaches, the n_i in our case is not constrained to be one through time. We are then able to model simultaneously the effects of both changing volatility and mean reversion in the case of two or more underlying assets.

In summary, we extend the existing literature in two directions. First, in the univariate case, we are able to extend the NR and Amin models by allowing the number of binomial stages to exceed one and to vary within the time intervals (t_{i-1}, t_i) . This means that we can control the "denseness" of the binomial tree between the multiple exercise dates of the option. Our technique can then be used, for instance, to extend Breen's (1991) accelerated binomial option pricing model to the case where volatility changes and/or the asset price mean reverts. Secondly, we are able to extend Amin's work by modeling a general multivariate process, where the individual assets have different rates of change of variance and mean reversion.

3. A Method for Constructing a Univariate Binomial Process with Specified Variances

The problem, in the univariate case, is to approximate with a binomial process the true process for X_i , given the means μ_i , conditional volatilities $\sigma_{i-1,t}$, and the unconditional volatilities $\sigma_{0,t}$.¹⁰ The conditional volatilities of the approximated binomial process will be de-

¹⁰ We drop the subscript j in this section. For instance, instead of $\sigma_{0,t,j}$ and $\sigma_{i-1,t,j}$ we simply write $\sigma_{0,t}$ and $\sigma_{i-1,t}$, respectively. Also, instead of $X_{i,j}$ we simply write X_i .

noted $\hat{\sigma}_{i-1,i}(n_i)$ since they will be a functions of n_i , the number of binomial stages. We require that

$$\lim_{n_i \rightarrow \infty} \hat{\sigma}_{i-1,i}(n_i) = \sigma_{i-1,i}, \quad \forall i. \quad (4)$$

The unconditional volatility of the approximated process over the period $(0, t_i)$ is similarly denoted $\hat{\sigma}_{0,i}(n_1, n_2, \dots, n_i)$, since it is, in general, a function of the number of binomial stages over each of the subperiods $t_1, t_2, \dots, t_i$. Here, we require

$$\lim_{n_i \rightarrow \infty} \hat{\sigma}_{0,i}(n_1, n_2, \dots, n_i) = \sigma_{0,i}, \quad \forall i, l, \quad l = 1, \dots, i. \quad (5)$$

In addition, of course, we constrain the mean of the approximated process to be equal to μ_i for each i . In other words,

$$\lim_{n_i \rightarrow \infty} \hat{\mu}_i = \mu_i, \quad \forall i. \quad (6)$$

Our method involves the construction of m separate binomial distributions, where the time periods are denoted $t_1, t_2, \dots, t_i, \dots, t_m$. The set of these distributions forms a discrete stochastic process for X_i :

$$(X_1, X_2, \dots, X_m),$$

where X_i is only defined at the times t_i . An example, where $m = 2$ is shown in Figure 1. X_1 is an $N_1 + 1$ vector where at node r

$$X_{1,r} = X_0 u_1^{N_1-r} d_1^r,$$

where $N_1 = n_1$. X_2 is an $N_2 + 1$ dimensional vector with

$$X_{2,r} = X_0 u_2^{N_2-r} d_2^r$$

at node r , where $N_2 = n_1 + n_2$. We have to choose the up and down movements u_1, u_2, d_1 , and d_2 and the probabilities so that Equations (4), (5) and (6) are satisfied. We denote

$$x_i = \ln(X_i/X_0)$$

and the conditional probabilities of x_2 given a node r at t_1 as

$$q(x_2 | x_1 = x_{1,r})$$

or $q(x_2)$ for short.

In general, we need to choose values for the u_i, d_i , and $q(x_2)$ so that the convergence Equations (4), (5), and (6) are satisfied. We first establish a lemma which guarantees that the conditional volatility and the unconditional mean converge to their values in Equations (4) and (6).

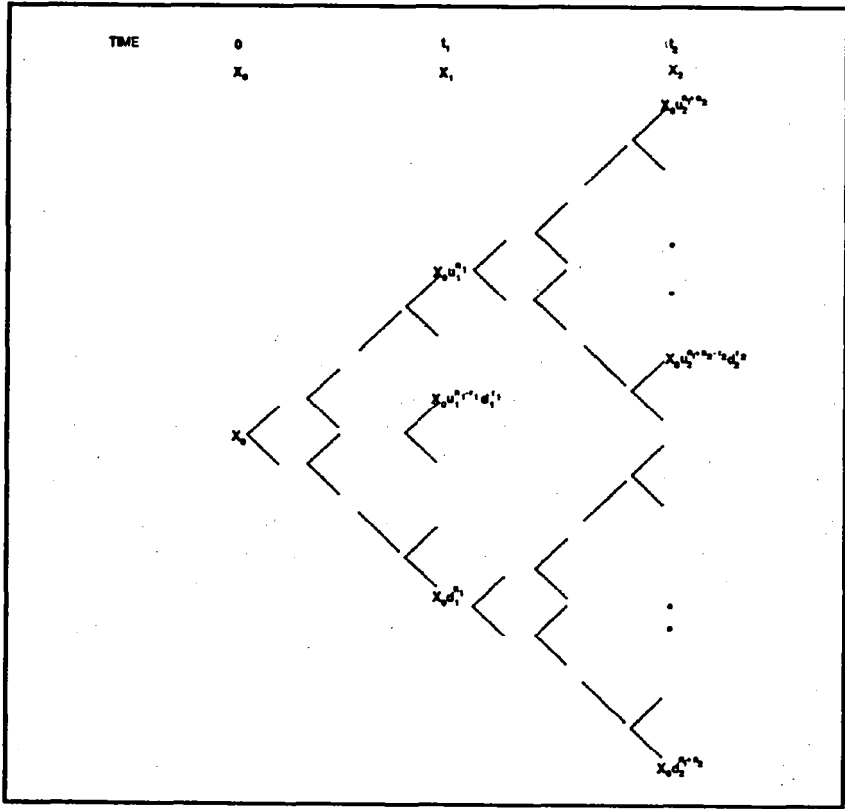


Figure 1

A discrete process for X_1, X_2

There are $n_1 + 1$ nodes at t_1 numbered $r = 0, 1, \dots, n_1$. There are $n_1 + n_2 + 1$ nodes at t_2 numbered $s = 0, 1, \dots, n_1 + n_2$. X_0 is the starting price, X_1 is the price at time t_1 , X_2 is the price at time t_2 . u_1, d_1, u_2 , and d_2 are the proportionate up and down movements. Although X_1 and X_2 are generated by binomial processes, this method gives two vectors ($n_1 + 1$ dimensional for X_1 and $n_2 + n_1 + 1$ dimensional for X_2). Intermediate values at time $t \in (0, t_1)$, $t \in (t_1, t_2)$ are not defined.

Lemma 1. Suppose that the up and down movements u_i and d_i are chosen so that

$$d_i = \frac{2(E(X_i)/X_0)^{\frac{1}{N_i}}}{1 + \exp(2\sigma_{i-1,i}\sqrt{(t_i - t_{i-1})/n_i})}, \quad i = 1, 2, \dots, m, \quad (7)$$

$$u_i = 2(E(X_i)/X_0)^{\frac{1}{N_i}} - d_i, \quad i = 1, 2, \dots, m, \quad (8)$$

where $N_i = \sum_{l=1}^i n_l$, then if for all i , the conditional probability $q(x_i) \rightarrow 0.5$ as $n_i \rightarrow \infty$, for $l = 1, 2, \dots, i$, then the unconditional mean and the conditional volatility of the approximated process

approach respectively their true values:

$$\lim_{\substack{n_i \rightarrow \infty \\ i=1,2,\dots,t}} \frac{\hat{E}(X_i)}{X_0} \rightarrow \frac{E(X_i)}{X_0}, \quad \lim_{n_i \rightarrow \infty} \hat{\sigma}_{t-1,t} \rightarrow \sigma_{t-1,t}.$$

Proof. See Appendix A. ■

The up and down movements u_i and d_i are analogous to those in Cox, Ross, and Rubinstein (1979), in that they are chosen to match the true mean and conditional volatility. Of course, in our case, since the conditional volatilities are allowed to change over time, the u_i s and d_i s also change correspondingly.

Given Lemma 1, we can guarantee that the mean and conditional volatility of the approximated process converge to their required values by choosing u_i and d_i using the given conditional volatilities $\sigma_{t-1,t}$.¹¹ The remaining problem is to choose the conditional probabilities $q(x_i)$ in such a manner that the *unconditional* volatility converges to the true value as in Equation (5). First note that, by Lemma 1, we are free to choose $q(x_i)$ without affecting the convergence of the conditional volatility, as long as $q(x_i) \rightarrow 0.5$ as $n_i \rightarrow \infty$. Thus we can choose the probabilities so that both the unconditional and conditional volatilities converge.

since $x_i = \ln(X_i/X_0)$ is normally distributed, it follows that the regression

$$x_i = a_i + b_i x_{i-1} + \varepsilon_i, \quad E_{t-1}(\varepsilon_i) = 0,$$

is linear with

$$b_i = \sqrt{[t_i \sigma_{0,t}^2 - (t_i - t_{i-1}) \sigma_{t-1,t}^2] / t_{i-1} \sigma_{0,t-1}^2},$$

and¹²

$$a_i = E(x_i) - b_i E(x_{i-1}).$$

We now choose the conditional probabilities $q(x_i)$ so that

$$E_{t-1}(x_i) = a_i + b_i x_{i-1} \tag{9}$$

¹¹ Note that the convergence of the approximated mean to the true mean implies that the approximated logarithmic mean $\hat{\mu}_i$ also converges to the true logarithmic μ_i .

¹² Also, since X_i is lognormally distributed, $E(x_i) = \ln(\frac{E(X_i)}{X_0}) - \frac{1}{2}\sigma_{0,t}^2$.

also holds for the approximated variables x_i and x_{i-1} .¹³ In Appendix B we show how to choose the conditional probability $q(x_i)$ so as to guarantee that Equation (9) holds. We can now establish:

Theorem 1. Suppose that the X_i are joint lognormally distributed. If the X_i are approximated with binomial distributions with $N_i = N_{i-1} + n_i$ stages and u_i and d_i given by Equations (7) and (8), and if the conditional probability of an up movement at node r at time $i - 1$ is

$$q(x_i | x_{i-1} = x_{i-1,r}) = \frac{a_i + b_i x_{i-1,r} - (N_{i-1} - r) \ln u_i - r \ln d_i}{n_i (\ln u_i - \ln d_i)} - \frac{\ln d_i}{\ln u_i - \ln d_i}, \quad \forall i, r \quad (10)$$

then $\hat{\mu}_i \rightarrow \mu_i$ and $\hat{\sigma}_{0,i} \rightarrow \sigma_{0,i}$ and $\hat{\sigma}_{i-1,i} \rightarrow \sigma_{i-1,i}$ as $n_i \rightarrow \infty$, $\forall i$.

Proof. First note that, as $n_i \rightarrow \infty$

$$q(x_i) \rightarrow \frac{-\ln d_i}{\ln u_i - \ln d_i} \rightarrow 0.5.$$

Hence, the conditions of Lemma 1 are fulfilled and $\hat{\sigma}_{i-1,i}$ converges to $\sigma_{i-1,i}$.

We now prove the convergence of $\hat{\sigma}_{0,i} \rightarrow \sigma_{0,i}$ by induction. First, we choose for $i = 1$, $q(x_1) = 0.5$. Hence from Lemma 1, the theorem is true for $i = 1$. We need now to show that $\hat{\sigma}_{0,2} \rightarrow \sigma_{0,2}$ as n_1 and $n_2 \rightarrow \infty$. Then, it follows by induction that $\hat{\sigma}_{0,i} \rightarrow \sigma_{0,i}$ for all i . Lemma 1 guarantees that $\hat{\sigma}_{1,2} \rightarrow \sigma_{1,2}$ if $q(x_2)$ is chosen by Equation (10). In Appendix B we show that, with this choice of $q(x_2)$, then¹⁴

$$\hat{E}_i(x_2) = a_2 + b_2 x_1.$$

The volatility $\hat{\sigma}_{0,2}$ is therefore

$$\hat{\sigma}_{0,2} = \sqrt{[b_2^2 \hat{\sigma}_{0,1}^2 t_1 + (t_2 - t_1) \hat{\sigma}_{1,2}^2] / t_2}.$$

Now, since $\hat{\sigma}_{0,1} = \sigma_{0,1}$ and $\hat{\sigma}_{1,2} \rightarrow \sigma_{1,2}$ as $n_2 \rightarrow \infty$, then

$$\hat{\sigma}_{0,2} \rightarrow \sqrt{[b_2^2 \sigma_{0,1}^2 t_1 + (t_2 - t_1) \sigma_{1,2}^2] / t_2}.$$

¹³ Note that Equation (9) has to hold for the approximated process, while it is true by assumption for the original variables.

¹⁴ Note that the proof in Appendix B refers to the trivariate case with coefficients α , β , and γ .

and substituting

$$b_2^2 = [t_2 \sigma_{0,2}^2 - (t_2 - t_1) \sigma_{1,2}^2] / t_1 \sigma_{0,1}^2,$$

we find that $\hat{\sigma}_{0,2} \rightarrow \sigma_{0,2}$. It follows that if $n_i \rightarrow \infty$, $\hat{\sigma}_{0,i} \rightarrow \sigma_{0,i}$, $\forall i$. ■

3.1 Properties of the approximation for small n

In practical applications we are forced to approximate the stochastic process of X_i using small values of n_i . The smaller the n , the more rapid will be the computation of option prices. When n_i is small, the conditional probability at a node suggested by the formula in Equation (10) could be outside the natural bounds for probability (i.e., q could exceed 1 or be less than 0).¹⁵

In Table 1, we show the effect of the restriction $0 \leq q(x_i | x_{i-1} = x_{i-1,r}) \leq 1$, $\forall i, r$. We simulate the model for three cases. In the first case, the process for X_t has the *mean reversion* property. In the second case, the process for X_t exhibits *declining conditional volatility*. Finally, the third case combines the mean reversion property of case 1 and the changing volatility of case 2.

The simulations all show, as we would expect, that the accuracy of the approximation improves as n increases. The first panel indicates that, for a constant variance process, mean reversion is easily captured by the method, even for relatively small n . This is also true for cases of declining volatility in panel 2. The method is accurate for quite small n size and converges rapidly for reasonable changes in volatility. When the volatility declines more rapidly from 10 percent in the first period to 7 percent in the second, then the accuracy begins to decline somewhat. This is due to the reliance of the method on the conditional volatility for the calculation of X_2 . In extreme cases, however, a change in the size of n_i can be made which can improve the accuracy.¹⁶

Theorem 1 allows us to approximate a process with given mean, variance, and covariance characteristics over the periods $(0, t_i)$ and (t_{i-1}, t_i) , where $i = 1, 2, \dots, m$. We can therefore construct a process using all the dates $t_1, t_2, \dots, t_m$. Successive application of Theorem 1 guarantees that the volatilities converge to their given values over each

¹⁵ Note that this is always a "small n " problem, since $q(x_2)$ in Equation (10) limits to 0.5 as $n_i \rightarrow \infty$. This can be seen from Equation (10) where the second term goes rapidly to 0.5 as n_i increases and the first term can be kept within the range -0.5 to 0.5 by choosing a large enough size of n_i .

¹⁶ Simulations of the case where X_t is fitted over three periods, t_1, t_2, t_3 , yield similar results. For example, assuming constant conditional volatility of 10 percent and mean reversion such that $\sigma_{0,2} = 0.09$, $\sigma_{0,3} = 0.08$, we found that for $n = 10$, $\hat{\sigma}_{0,3} = 0.07974(26)$; $n = 30$, $\hat{\sigma}_{0,3} = 0.07991(9)$; $n = 50$, $\hat{\sigma}_{0,3} = 0.07995(5)$. Similar results were obtained for the case of changing volatility and where the process has both mean reversion and volatility change.

Table 1

Convergence of binomial approximation in a two-period example with (1) mean reversion, (2) changing volatilities, and (3) mean reversion and changing volatilities

			Volatility of approximated process $\hat{\sigma}_{0,2}^1$		
$\sigma_{0,1}$	$\sigma_{1,2}$	$\sigma_{0,2}$	$n = 10$	$n = 30$	$n = 50^2$
(1) Mean reversion ³					
0.10	0.10	0.10	0.10000	0.10000	0.10000
0.10	0.10	0.09	0.08987(13) ⁴	0.08996(4)	0.08997(3)
0.10	0.10	0.08	0.07930(70)	0.07977(23)	0.07986(14)
(2) Changing volatilities ⁵					
0.10	0.09	0.09513	0.09510(3)	0.09512(1)	0.09513(0)
0.10	0.08	0.09055	0.09044(11)	0.09051(4)	0.09053(2)
0.10	0.07	0.08631	0.08605(26)	0.08622(9)	0.08626(5)
(3) Mean reversion and changing volatilities ⁶					
0.10	0.09	0.08513	0.08510(3)	0.08512(1)	0.08513(0)
0.10	0.08	0.08055	0.08055(0)	0.08055(0)	0.08055(0)
0.10	0.07	0.07631	0.07626(5)	0.07629(2)	0.07630(1)

¹ In each case the periods are $t_1 = 1$ year, $t_2 = 2$ years.

² In all cases n is the same for year 1 and year 2 (i.e., $n_1 = n$, $n_2 = n$).

³ All volatilities are annualized. The volatility over the first period $\sigma_{0,1}$ is exact, by construction.

⁴ In parentheses we show $(\sigma_{0,2} - \hat{\sigma}_{0,2}) \times 100,000$. This is the approximation error. This is also the size of the error $(\sigma_{1,2} - \hat{\sigma}_{1,2})$.

⁵ In this case we assume that X_t follows random walk with changing variance; hence, $\sigma_{0,2}^2 = (\sigma_{0,1}^2 + \sigma_{1,2}^2)/2$.

⁶ In this case we assume that X_t follows a mean reverting process with changing conditional variance. In these examples, we choose $\sigma_{0,2} = \sqrt{(\sigma_{0,1}^2 + \sigma_{1,2}^2)/2} - 0.01$.

time period. The resulting binomial process can be used to evaluate multiple exercisable options with a given number of exercise dates as in Breen (1991).

4. The Multivariate Case

There are two ways in which an option valuation problem may involve a multivariate probability distribution. The first is when the payoff on the option depends on the outcome of two or more random variables. This is the case for the yen oil option discussed in the introduction. The payoff there depends on the dollar oil price and the yen/dollar exchange rate. The second motivation stems from the consideration of the value of an *American-style* option where the intermediate value of the option depends on the price of the underlying asset and other variables, such as the rate of interest. In this case we need to model the covariance of the asset price and the rate of interest, as well as the time series properties of the two variables. In this section we consider the general problem of approximating the multivariate distribution of

two or more variables each of which has the properties (of mean reversion and changing volatility) discussed in Section 3.¹⁷

Initially, we consider the case of a single time period $(0, T)$. This is relevant for *European-style* options where the payoff depends on two or more variables. We wish to approximate the joint distribution of the variables $(X, Y, Z, \dots)$ with a multivariate binomial distribution. Hence, we choose the conditional probabilities (of Y given X , and of Z given Y and X) given the covariance between the variables.¹⁸ We confine ourselves to the case where there are three relevant random variables (X, Y, Z) . The first step involves approximating the first random variable X_T with a binomial distribution using the method outlined in Section 3. We then build a vector of Y_T values using Equations (7) and (8), and using the conditional volatility of Y_T given X_T . We then choose the conditional probability of Y_T given the outcome X_T . This conditional probability is denoted $q(y_T | x_T = x_{T,r})$, where r is the node of the X binomial tree, and $x_{T,r}$ is the value of x_T at the r th node. In Appendix B we show that the appropriate value of $q(y_T)$ is¹⁹

$$q(y_T | x_T = x_{T,r}) = \frac{\alpha_y + \beta_y x_{T,r} - n_y \ln d_y}{n_y (\ln u_y - \ln d_y)}, \quad (11)$$

where α_y and β_y are the coefficients from the simple regression of Y_T on X_T , and n_x and n_y are the number of binomial stages in the x and y approximations, respectively.

In the next stage, we build a vector of Z_T using the conditional volatility of Z_T given X_T and Y_T . We then choose the probability of Z given an X and a Y node. This conditional probability is shown in Appendix B [Equation (B3)] to be

$$q(z_T | x_T = x_{T,r}, y_T = y_{T,s}) = \frac{\alpha_z + \beta_z x_{T,r} + \gamma_z y_{T,s} - n_z \ln d_z}{n_z (\ln u_z - \ln d_z)}, \quad (12)$$

where α_z , β_z , and γ_z are the coefficients from the multiple regression of Z_T on X_T and Y_T . If the conditional probabilities are chosen by Equations (11) and (12), it follows by an extension of Theorem 1 that

¹⁷ Boyle (1988) develops a technique for the multivariate case based on a trinomial distribution. Our method differs from Boyle's in that we use a simpler binomial distribution. Our method allows us to solve the general problem of constructing a multivariate process with arbitrary cross-sectional variances and covariances, and conditional variances and covariances over time.

¹⁸ An alternative method suggested by Amin (1991) models the covariance using factor loadings on two independent random variables. A problem with this approach is that the number of Y and Z nodes increases rapidly with the n size making the computation of complex option prices difficult.

¹⁹ This follows from Equation (B3) with $z = y$ and $\gamma_z = 0$.

the volatilities of Y and Z converge, as n_y and n_z increase, to their values σ_y and σ_z .

We now consider the accuracy of the covariance and correlation between the approximated X and Y variables. We have:

Theorem 2. *Suppose the up and down movements of two correlated random variables, X and Y , are chosen by Equations (7) and (8) and that the conditional probabilities are given by Equation (11). Then the approximated value of the conditional covariance is equal to its true value. Also, when there are three or more correlated variables, for example X , Y , and Z where X is the first fitted variable, the covariance of the pairs (X, Y) and (X, Z) are exact and that of (Y, Z) limits to its true value.*

Proof. See Appendix A. ■

Note that, in the case of just two variables X and Y , this cross-sectional property is an *exact* one and holds for all values of n and not just for the limiting values. The approximation in the case of three or more variables is due to the fact that Y has only an approximately correct variance which limits to its true value as n increases. Also, the covariance property holds in the multiperiod case for all the conditional covariances. However, since our method approximates the variances of the variables, the resulting correlation is only asymptotically correct.

4.1 The multiperiod multivariate case

We now turn to the general problem of a multiperiod, multivariate process. In this case we may be concerned with say two variables (the dollar price of oil and the yen/dollar exchange rate) at a series of dates $(t_1, t_2, \dots)$. Our general problem is to construct a multiperiod, multivariate process over periods $[(0, t_1), (t_1, t_2), \dots]$. For simplicity, we illustrate the general methodology with an example of a two-period $[(0, t_1), (t_1, t_2)]$, bivariate (X, Y) case. This case is sufficient, for example, if we wish to value a compound option whose payoff depends on two variables.

Suppose that (X_1, X_2) and (Y_1, Y_2) are multivariate lognormally distributed with volatilities $(\sigma_{0,1,x}, \sigma_{0,2,x}, \sigma_{1,2,x})$ and $(\sigma_{0,1,y}, \sigma_{0,2,y}, \sigma_{1,2,y})$. Also, assume that the correlation between x_1 and y_1 is $\rho_{0,1,x,y}$. The conditional correlation between x_2 and y_2 is denoted $\rho_{1,2,x,y}$. Note that, for the joint lognormal distribution, the conditional correlation is nonstochastic. However, in this method there is no need to restrict the conditional correlation to be the same at each point in time.

The steps in the computation for the general case are

1. Compute the up and down movements for X_1 and Y_1 , and X_2 and Y_2 , independently using the methodology of Section 3. Specifically, the vector of Y_1 is constructed using the conditional volatility of Y_1

given X_1 , X_2 requires the conditional volatility $\sigma_{1,2,x}$, and Y_2 requires the conditional volatility of Y_2 given both Y_1 and X_2 .

2. Compute the conditional probability of an up movement in Y_1 given a value of X_1 using Equation (11). We denote these probabilities as $q(y_1 | x_1 = x_{1,r})$ since they are of Y_1 given that X_1 is at a node r at time t_i .

3. Compute the conditional probability of an up movement in X_2 given a value of X_1 using the methodology of Section 3 and Equation (10) in particular. These probabilities are denoted $q(x_2 | x_1 = x_{1,r})$.

4. Compute the conditional probability of Y_2 given both Y_1 and X_2 . We denote this probability as $q(y_2 | y_1 = y_{1,r}, x_2 = x_{2,s})$. This probability must satisfy both the time-series (Y_2 on Y_1) and cross-sectional (Y_2 on X_2) properties of Y_2 . In other words, it must satisfy the volatilities $(\sigma_{0,1,y}, \sigma_{0,2,y}, \sigma_{1,2,y})$ and the conditional correlation $\rho_{1,2,x,y}$. In Appendix B, we show that in this case

$$q(y_2) = \frac{a_2 + b_2 y_{1,r} + c_2 x_{2,s} - [(N_1 - r) \ln u_{2,y} + r \ln d_{2,y}] - n_2 \ln d_{2,y}}{n_2 \ln u_{2,y} - \ln d_{2,y}}, \quad (13)$$

where the probability $q(y_2) = q(y_2 | y_1 = y_{1,r}, x_2 = x_{2,s})$, and where a_2 , b_2 , and c_2 are the multiple regression coefficients from the regression of y_2 on y_1 and x_2 .

Again, an extension of Theorem 1 can be used to show that when $q(y_2)$ and $q(x_2)$ are chosen in this manner, both the variances and the correlations of the multivariate process converge to their given values. However, the probabilities in Equations (10), (11), and (13) could be greater than one or less than zero when the n size is finite. As we see now, this is particularly likely when the correlation between the variables is high.

It is important now to consider the limitations on the accuracy of the multivariate method when the natural limits are placed on the conditional probabilities in Equations (10), (11), and (13). In Table 2, we show the results of simulations with and without a nonnegativity constraint placed on the conditional probabilities. For simplicity, we show just the case of a single time period $(0, t_i)$ with $t_i = 1$ year. Even when the conditional probability is constrained, the simulations show that the accuracy of the approximation is not adversely affected by the nonnegativity constraint for correlations under 0.8. For high values of ρ (e.g., $\rho = 0.9$) some inaccuracy is apparent, but this reduces for large n values.²⁰

²⁰ In fact the inaccuracy can be eliminated by appropriate choke of n size for the dependent variable.

Table 2

Accuracy of the binomial approximation in a multivariate case with and without nonnegativity constraint on probabilities

Correlation ² $\rho_{0,1,x,y}$	Volatility of approximated process, $\hat{\sigma}_{0,1,y}$ ¹			
	Binomial stages, $n_{1,x} = n_{1,y} = 10$		Binomial stages, $n_{1,x} = n_{1,y} = 20$	
	Estimated $\hat{\sigma}_{0,1,y}$	Approximation error ³	Estimated $\hat{\sigma}_{0,1,y}$	Approximation error
0.0	0.1000	0	0.1000	0
0.1	0.0999	1	0.1000	0
0.2	0.0998	2	0.0999	1
0.3	0.0995	5	0.0998	2
0.4	0.0992	8	0.0986	4
0.5	0.0987	13	0.0984	6
0.6	0.0982	18	0.0991	9
0.7	0.0975	25	0.0988	12
0.8	0.0960	40	0.0984	16
0.8 ⁴	(0.0967)	(33)	(0.0984)	(16)
0.9	0.0878	122	0.0961	39
0.9 ⁴	(0.0959)	(41)	(0.0979)	(21)

¹ In all cases, each of the volatilities (conditional and unconditional) for X and Y is 10 percent.

² $t_1 = 1$ year.

³ The approximation error is shown $(\sigma_{0,1,y} - \hat{\sigma}_{0,1,y}) \times 100,000$.

⁴ The number in parentheses below the constrained estimated volatility $\hat{\sigma}_{0,1,y}$ is the corresponding unconstrained volatility estimate.

5. Applications of the Methodology to Option Pricing

In applying the methodology to the pricing of options we need to approximate the *risk neutral* distribution of the asset prices. For the problems considered here, this distribution is joint lognormal where the means of the variables are their respective forward prices and the volatilities are the same as those of the actual distribution. It follows from Theorem 1 that if the u_i and d_i are chosen as in Equations (7) and (8) with the additional requirement that the mean equals the forward price, then the distribution converges to the required risk neutral distribution. In this section, we price multiple exercisable options on two variables using these approximating distributions. However, before applying the method we need to establish that the option payoffs and prices also converge as the n size increases.

In the case of a finite n size, we propose the option payoffs and values only as approximations to their limiting values. The method makes no use of the (unknown) true mean of the underlying asset and does not assume a complete market in which no-arbitrage option values are determined. We show that the option values we com-

pute converge to the prices in the complete market, continuous-time economy.²¹

We will first discuss and prove convergence in the case of a European-style call option on a single asset: the Black-Scholes case. We then extend the proof to the case of American-style options. Our procedure in the univariate case is as follows. If the option has a maturity t , the asset price on which the option is written has a lognormally distributed price X_t with volatility σ . We construct a binomially distributed variable $X_t(n)$ with binomial probability $q = 0.5$, an expected value $E[X_t(n)] = F_{0,t}$; the asset forward price, and u and d chosen by the formula in Equations (7) and (8). It follows from Lemma 1 that $X_t(n)$ converges in distribution to the risk neutral density of X_t ; that is, to a distribution with a mean $F_{0,t}$ and volatility σ . We now establish

Theorem 3. Define $X_t(n)$ using Equations (7) and (8) with $E[X_t(n)] = F_{0,t}$.

(a) $X_t(n)$ converges in distribution to X_t , where X_t is lognormal with mean $F_{0,t}$ and logarithmic variance $\sigma^2 t$.

(b) For a European-style call option with payoff function $g(X_t) = \max[X_t - K, 0]$, $g[X_t(n)] \rightarrow g(X_t)$.

(c) $E[g(X_t(n))] \rightarrow E[g(X_t)] = C_0/B_{0,t}$, is the Black-Scholes value of the call option.

Proof:

(a) Follows directly from Lemma 1.

(b) Follows since $g(\cdot)$ is a continuous function; and Billingsley (1986), Theorem 25.7, Corollary 1.

(c) Follows from the fact that the sequence $g[X_t(n)]$ is uniformly integral, given the lognormal distribution; and Billingsley (1986), Theorem 25.12. ■

Parts (b) and (c) of Theorem 3 follow quite closely the analysis of Duffie (1988, pp. 244-248). An important implication of the theorem is that hedge ratios and other risk parameters, such as the deltas, thetas, vegas, and gammas (i.e., the derivatives of C_0), also converge as the

²¹ The prices which our finite n -size economy values converge to are thus the no-arbitrage prices. Our method is in the spirit of Rajasingham (1990) who shows that it is not necessary for the finite economies to be complete when the economies are used as approximations. Our task, therefore, differs somewhat from that of Cox, Ross, and Rubinstein (1979) and Duffie (1988), who assume knowledge of the asset's true mean and show that the no-arbitrage, finite economy prices converge to the continuous economy limit. It is interesting to note that Amin (1990) and He (1990) both suggest an alternative three-state approach to the valuation of options on multivariate processes which allows a complete market, no-arbitrage valuation in the finite economy. In cut bivariate binomial case, the finite economy is incomplete, however, our approximated values still be shown to converge to the continuous-time limiting values.

n size increases. This allows the approximation of prices and hedge ratios with the binomial method.

5.1 American-style options

We now extend the discussion of convergence to the case of American-style options. Consider first a call option that is exerciseable twice; at time t_1 and at time t_2 . Let the exercise price at time t_1 be K_1 , and at time t_2 be K_2 . The payoff at time t_2 , if the option is not exercised at t_1 , is

$$g(X_2) = \max[X_2 - K_2, 0].$$

The value of the option, if alive at time t_1 , is given by the Black-Scholes function

$$C_1 = C_1(X_1).$$

Just prior to the exercise decision at time t_1 , the value of the option can be written

$$g(X_1) = \max[C_1, X_1 - K_1].$$

The value at time 0 of the claim is

$$C_0 = B_{0,1} E[g(X_1)],$$

where the expectation is again taken with respect to the risk neutral density.²² We now approximate C_0 using a binomial distribution for X_1 and X_2 .

To approximate C_0 , we first construct $X_1(n)$ and $X_2(n)$ using Equations (7) and (8) with $E(X_1) = F_{0,1}$ and $E(X_2) = F_{1,2}$. We then compute $g(X_2(n))$ and $C_1(n)$ in the same way as in the case of the European-style option above. It then follows:

Corollary 1. *The payoff and value of a twice exerciseable option converge to their continuous limits as n increases.*

Proof: From Theorem 3, it follows that $C_1(n)$ converges to the Black-Scholes value C_1 . We know also the payoff (if exercised) at t_1 , $X_1(n) \rightarrow X_1$. Hence $g(X_1(n)) \rightarrow g(X_1)$. Finally, since $g(X_1)$ is again uniformly integral it follows that the sequence $E[g(X_1(n))]$ converges to $E[g(X_1)]$, and $C_0(n)$ converges to C_0 . ■

The above extension of Theorem 3 shows that our approximation converges for a twice exerciseable option. By induction it must follow that an m times exerciseable option can be approximated in a similar way choosing the u and d proportions according to Equations (7)

²² This follows from Geske and Johnson (1984).

and (8). Hence, in the limit, as m increases the values converge to the continuously exercisable American option price.²³

The proof in Theorem 3 and Corollary 1 extend to the multivariate case, as shown for example in He (1990). Here, what is required again is that the option payoff is uniformly integral where the underlying state space is multivariate.

5.2 An example: the valuation of a twice exercisable, quality option

Consider the following quality option contract. The option is to acquire either asset X or asset Y . The holder has the additional option to exercise at time t_2 or time t_1 . If he exercises at t_2 , he can buy either asset X at a price $K_{2,x}$ or asset Y at a price $K_{2,y}$. Alternatively, he can exercise his option at time t_1 and buy asset X for $K_{1,x}$ or asset Y for $K_{1,y}$. This is a quality option with an *American* feature.²⁴

The payoffs on the option are as follows:

At time t_2 ,

$$\begin{aligned} g(X_2, Y_2) &= \max[X_2 - K_{2,x}, Y_2 - K_{2,y}, 0], \\ &\quad \text{if option not exercised at } t_1, \\ &= 0, \quad \text{if option exercised at } t_1. \end{aligned}$$

At time t_1 the value of the option is

$$g(X_1, Y_1) = \max[X_1 - K_{1,x}, Y_1 - K_{1,y}, C_1],$$

where C_1 is the value of the option at time t_1 , if it is not exercised at time t_1 . In order to value this option, we make the following assumptions: (1) interest rates in the economy are nonstochastic, (2) $t_1 = 1$ and $t_2 = 2$, (3) the relevant prices of the assets X_1 , X_2 , Y_1 , and Y_2 are joint lognormally distributed, and (4) a preference-free valuation relationship exists for the valuation of the option.

The fourth assumption implies that the option can be valued using a *risk neutral* density function where the means of the stochastic variables are their respective forward prices. This implies that the

²³ Strictly, this argument requires that the strike price of the option be a continuous function of time. Otherwise, it is possible, for example, that the option payoff could be positive if exercised at an irrational date and zero otherwise. An alternative rigorous proof of convergence for American options is provided by Amin and Khanna (1994).

²⁴ An example of a quality option is the option available in most bond futures contracts. When we use the term "American" feature we mean that it can be exercised (at least once) before the final maturity.

Table 3
A twice exercisable option: sensitivity analysis

	Number of binomial stages ¹ , <i>n</i>				
Correlation Coefficient ² , <i>ρ</i>	4	8	12	16	30
0.0	0.0881	0.0884	0.0884	0.0885	0.0885
0.2	0.0845	0.0849	0.0850	0.0850	0.0849
0.4	0.0802	0.0808	0.0810	0.0810	0.0810
0.6	0.0750	0.0759	0.0761	0.0763	0.0764
0.8	0.0677	0.0694	0.0700	0.0702	0.0706
Time ³	<i>t</i> ₁ = 1 year		<i>t</i> ₂ = 2 years		
Forward prices ⁴	<i>F</i> _{0,1,<i>x</i>} = <i>F</i> _{0,1,<i>y</i>} = 1		<i>F</i> _{0,2,<i>x</i>} = <i>F</i> _{0,2,<i>y</i>} = 1		
Zero bond prices ⁵	<i>B</i> _{0,1} = 0.9		<i>B</i> _{0,2} = 0.81		
Volatility ⁶ :					
Conditional (<i>x</i>)	<i>σ</i> _{0,1,<i>x</i>} = 0.1		<i>σ</i> _{1,2,<i>x</i>} = 0.1		
Conditional (<i>y</i>)	<i>σ</i> _{0,1,<i>y</i>} = 0.08		<i>σ</i> _{1,2,<i>y</i>} = 0.07		
Unconditional ⁷ (<i>x</i>)			<i>σ</i> _{0,2,<i>x</i>} = 0.09		
Unconditional ⁸ (<i>y</i>)			<i>σ</i> _{0,2,<i>y</i>} = 0.07517		
Strike prices	<i>K</i> _{1,<i>x</i>} = <i>K</i> _{1,<i>y</i>} = 0.96		<i>K</i> _{2,<i>x</i>} = <i>K</i> _{2,<i>y</i>} = 1		

- ¹ The n size applies to n_1 for X and Y , and to n_2 for both X and Y .
- ² This is the value of ρ both for time t_1 and for time t_1 to t_2 .
- ³ The following input data for the twice exercisable quality option on two assets is assumed for all option price calculations.
- ⁴ The fact that forward prices are the same for delivery at years 1 and 2 implies that the assets pay dividends.
- ⁵ These prices imply a constant rate of interest of 11.11 percent per year.
- ⁶ These volatilities are all quoted on an annualized basis.
- ⁷ This lower unconditional volatility implies mean reversion in the X_t process.
- ⁸ $\sigma_{0,2,y} = \sqrt{(\sigma_{0,1,y}^2 + \sigma_{1,2,y}^2)/2}$. Hence, y follows a random walk with changing variance.

means of the variables are

$$\begin{aligned} E(X_1) &= F_{0,1,x}, & E(Y_1) &= F_{0,1,y}, \\ E(X_2) &= F_{0,2,x}, & E(Y_2) &= F_{0,2,y}, \end{aligned}$$

where $F_{0,t}$ refers to the forward price of the asset for delivery at time t . We approximate the required distributions with a binomial distribution which has n stages from 0 to t_1 and from t_1 to t_2 .

In Table 3, we show the input data for the option valuation. Note that the unconditional volatility for asset X_t is less than its (constant) conditional volatility, indicating mean reversion. Also the conditional volatility of Y_t is nonconstant. For convenience we choose the forward prices of X_t and Y_t to be constant, equal to each other, and equal to 1. The strike prices are also 1 at time t_2 . At time t_1 , the strike prices are assumed to be lower (0.96) in order to allow the possibility of early exercise, and to make the valuation problem more interesting.

In Table 3, we illustrate the convergence of the option price as the number of binomial stages (both n_1 and n_2 for X and Y) increases. Note that reasonably accurate answers are obtainable with $n = 8$ even when the correlation between the variables is high. This is due to the fact that with two variables and two periods, there are $(n + 1)^4$ states. Table 3 also shows that this *quality option* declines in value as the assumed correlation between the underlying assets increases: with $n = 30$, the value is 8.85 percent in the uncorrelated case and only 7.06 percent when $\rho = 0.8$.

This example illustrates the rapid convergence, of the method. Even for relatively small values of n and highly correlated variables, the errors of the approximated process are relatively small. The resulting option prices are accurate and can be computed rapidly given the small values of n .

6. Conclusions

In this article, we have shown that binomial distributions with changing probability parameters can be constructed to approximate the joint distribution of the asset price at various (exercise) dates. The method can be extended, using state-dependent probabilities, to approximate the covariance characteristics in the multivariate case, where there is more than one state variable determining the option payoff. Tests of the accuracy of the multivariate binomial approximation for variables with changing variance and covariance characteristics show it to be quite accurate even with a small number of binomial stages. The accuracy also extends to cases where the correlation between the variables is high. The method offers a fast and efficient computational method for multiple exercisable options prices which can be extended to American-style option prices using the Geske and Johnson (1984) approximation technique. This would facilitate a generalization of the work of Breen (1991). Alternatively, it can be used to value a compound option whose payoff depends on two or more variables.

The principal feature of our methodology is its computational efficiency compared with alternative option pricing methods using numerical integrations. This advantage is clear-cut for American options where exercise is possible on two or more dates; our approach is less computationally intensive. This is because in our method, the number of nodes increases only linearly with the number of exercise dates. For instance, numerical integration along the lines of Geske and Johnson (1984) would involve n^3 instead of $(3n + 1)$ nodes in our binomial method in the thrice-exercisable case. For options involving two or more state variables, and/or more exercise dates, the comparative efficiency of our methodology is even more significant. For example, if

Amin's (1991) orthogonal factor method is used with three state variables, $(n + 1)^3$ nodes are generated for the third variable, whereas each variable has just $n + 1$ nodes using our method. This leads to a dramatic reduction in the number of required integrations. For some options our method can make the difference between feasible and nonfeasible computation. These computational advantages become even more important when hedge ratios and other risk management sensitivity parameters are required in addition to the option values.

Appendix A: Proof of Lemma 1 and Theorem 2

Proof of Lemma 1. (a) For a given N_i , the mean of the approximated process is

$$\frac{\hat{E}(X_t)}{X_0} = \prod_{i=1}^t [q(x_i)u_i + \{1 - q(x_i)\}d_i]^{n_i}.$$

Since, by the condition of the theorem, $q(x_i) \rightarrow 0.5$ as $n_i \rightarrow \infty$, $i = 1, 2, \dots, t$,

$$\lim_{\substack{n_i \rightarrow \infty \\ i=1,2,\dots,t}} \frac{\hat{E}(X_t)}{X_0} = [0.5(u_i + d_i)]^{N_i}.$$

From Equation (8), however,

$$u_i + d_i = 2 \left[\frac{E(X_i)}{X_0} \right]^{\frac{1}{n_i}}.$$

Hence,

$$\lim_{\substack{n_i \rightarrow \infty \\ i=1,2,\dots,t}} \frac{\hat{E}(X_t)}{X_0} = \frac{E(X_t)}{X_0}.$$

(b) The conditional variance can be written in the limit, since $q(x_i) \rightarrow 0.5$, when $n_i \rightarrow \infty$,²⁵ as

$$\begin{aligned} \lim_{n_i \rightarrow \infty} [\hat{\sigma}_{t-1,t}(n_i)]^2 &= \frac{1}{t_i - t_{i-1}} n_i (0.5)^2 (\ln u_i - \ln d_i)^2 \\ &= \frac{0.25 n_i}{t_i - t_{i-1}} \left[4 \sigma_{t-1,t}^2 \frac{(t_i - t_{i-1})}{n_i} \right], \end{aligned}$$

that is, $\lim_{n_i \rightarrow \infty} [\hat{\sigma}_{t-1,t}(n_i)]^2 = \sigma_{t-1,t}^2$. ■

²⁵ See Cox and Rubinstein (1985), p. 201.

Proof of Theorem 2. In the case of just two variables (X, Y) it is sufficient to establish the theorem for the special case of $n_x = 1$. For $n_x = 1$, the means of the two variables X and Y , where Y is chosen to be conditional on X , are given by

$$\begin{aligned}\mu_x &= (\ln u_x + \ln d_x)/2, \\ \mu_y &= q \ln u_y + (1 - q) \ln d_y, \\ q &= (q_0 + q_1)/2,\end{aligned}$$

where $q_r = q(y | x = x_r)$ as defined by Equation (11).

The conditional covariance between X and Y is given by

$$\hat{\sigma}_{xy} = \left(\frac{1}{2}\right)^2 (q_0 - q_1)(\ln u_x - \ln d_x)(\ln u_y - \ln d_y).$$

But from Equation (11), q_0 and q_1 are given by

$$q_r = \frac{\alpha + \beta[r \ln d_x + (1 - r) \ln u_x] - \ln d_y}{\ln u_y - \ln d_y}, \quad r = 0, 1$$

so that the approximated covariance $\hat{\sigma}_{x,y}$ can be written as

$$\hat{\sigma}_{xy} = \left(\frac{1}{2}\right)^2 \beta (\ln u_x - \ln d_x)^2 = \beta \sigma_x^2,$$

that is, $\hat{\sigma}_{xy} = \sigma_{xy}$. ■

A similar argument can be used for the case of three or more variables. For instance, consider the case of three variables X, Y , and Z , where X is the first fitted variable. First, the proof for the case of two variables applies directly to the covariances between X and Y , and X and Z . Next, since in the limit, the binomial probability of Y converges to 0.5, then the covariance of Z with Y converges to its true value.

Appendix B: The Binomial Probability in the Case of a Multivariate Lognormal Stochastic Process

B.1 The general problem

The method used in the article can be applied to construct a binomial tree as a discrete time approximation for any multivariate lognormal distribution. We first consider the general problem of approximating variables with a given covariance structure. We then apply this general method to the problem in the text. To see the specific details of the

method consider the case of a trivariate lognormal distribution of three variables X , Y , and Z with the following variance covariance matrix (between the logarithms of the variable):²⁶

$$\Omega = \begin{pmatrix} \sigma_x^2 & \sigma_{x,y} & \sigma_{x,z} \\ \sigma_{y,x} & \sigma_y^2 & \sigma_{y,z} \\ \sigma_{z,x} & \sigma_{z,y} & \sigma_z^2 \end{pmatrix}$$

For notational convenience we use lowercase letters to denote natural logarithms (i.e., $x = \ln X$, $y = \ln Y$, $z = \ln Z$).

Since x , y , and z are normally distributed, the multiple regression

$$z = \alpha_z + \beta_z x + \gamma_z y + \varepsilon,$$

where

$$\beta_z = \frac{\sigma_{z,x}\sigma_y^2 - \sigma_{y,z}\sigma_{x,y}}{\sigma_x^2\sigma_y^2 - \sigma_{x,y}^2},$$

$$\gamma_z = \frac{\sigma_{y,z}\sigma_x^2 - \sigma_{x,y}\sigma_{z,x}}{\sigma_x^2\sigma_y^2 - \sigma_{x,y}^2},$$

and

$$\alpha_z = E(z) - \beta_z E(x) - \gamma_z E(y),$$

is linear, and the conditional expectation of z is

$$E(z | x, y) = \alpha_z + \beta_z x + \gamma_z y. \quad (\text{B1})$$

First, we construct separate binomial trees for the variable x , y , z , using the method described in the text. We then choose the conditional probability of an up movement in z so that Equation (B1) is satisfied at each node. Given that z is a binomial process this implies that

$$\alpha_z + \beta_z x_r + \gamma_z y_s = n_z \{q(z) \ln u_z + [1 - q(z)] \ln d_z\}, \quad (\text{B2})$$

where $q(z) = q(z | x = x_r, y = y_s)$ is the probability of an up movement in z given that x is at node r and y is at node s of their respective binomial distributions. In Equation (B2) n_z is the number of stages in the binomial process of z , and u_z and d_z are the up and down movements of z . Solving Equation (B2) we find

$$q(z) = \frac{\alpha_z + \beta_z x_r + \gamma_z y_s - n_z \ln d_z}{n_z (\ln u_z - \ln d_z)}. \quad (\text{B3})$$

²⁶ The method is readily generalized to the n -variable case. However, the notation is complex.

B.2 The the series problem with two time periods

In the text we consider the problem of constructing a process for X_t with given volatility characteristics at two specified points in time t_{i-1} and t_i . In this case we construct binomial trees of $z = x_i$ and $x = x_{i-1}$. The resulting values of X_{i-1} and X_i are illustrated in Figure 1 in the text. The difference in this case is that x_i is a time-series (i.e. a cumulative) variable. In this case, the conditional expectation of x_i given node r at time t_{i-1} is

$$a_i + b_i x_{i-1} = n_i \{q(x_i) \ln u_i + [1 - q(x_i)] \ln d_i\} + (N_{i-1} - r) \ln u_i + r \ln d_i. \quad (B4)$$

In Equation (B4) the first term represents the drift from time t_{i-1} to time t_i . The second term is the expected value of x_i , given that the variable is in state r at time t_{i-1} , if the drift from time t_{i-1} to t_i is zero.²⁷

Solving Equation (B4) for $q(x_i)$ we have in this case

$$q(x_i) = \frac{\alpha_i + \beta_i x_{i-1} - [(N_{i-1} - r) \ln u_i + r \ln d_i] - n_i \ln d_i}{n_i (\ln u_i - \ln d_i)},$$

where $q(x_i) = q(x_i | x_{i-1} = x_{i-1}, r)$ is the probability of an up movement in x_i given that x_{i-1} is at node r .

B.3 The multivariate time-series problem

In the text we consider the example where there are two relevant variables (X_t, Y_t) and two time periods $(t-1, t)$. In this case we suggest first constructing the binomial process for (X_{t-1}, X_t) and the relevant conditional probabilities $q(x_i)$ using the methods described above. The remaining problem is to compute the conditional probability of Y_2 given both Y_1 and X_2 . This must reflect both the time-series properties of the Y_t process and the conditional correlation of the two variables $(\rho_{1,2,x,y})$.

This is a multivariate problem involving three variables (Y_t, Y_{t-1}, X_t) . Hence we can use the general case in Section 1 of Appendix B with $z = Y_t$, $x = Y_{t-1}$, and $y = X_t$. Again, we recognize that y_t is a time-series variable. In this case, therefore, the conditional expectation of

²⁷ Note that, in the special case where the volatility is constant over time $u_x = u_y$ and Equation (B4) becomes

$$a_i + b_i x_{i-1} = n_i \{q(x_i) \ln u_i + [1 - q(x_i)] \ln d_i\} + x_{i-1}.$$

In this case the conditional expectation of x_i is the drift plus x_{i-1} .

y_i , given node r of y_{i-1} and node s of x_i , is

$$a_i + b_i y_{i-1} + c_i x_i = n_i \{ q(y_i) \ln u_{i,y} + [1 - q(y_i)] \ln d_{i,y} \} + (N_{i-1} - r) \ln u_{i,y} + r \ln d_{i,y}. \quad (B5)$$

Solving Equation (B5) for $q(y_i)$ we have

$$q(y_i) = \frac{a_i + b_i y_{i-1} + c_i x_i - [(N_{i-1} - r) \ln u_{i,y} + r \ln d_{i,y}] - n_i \ln d_{i,y}}{n_i (\ln u_{i,y} - \ln d_{i,y})},$$

where $q(y_i) = q(y_i | y_{i-1} = y_{i-1,r}, x_i = x_{i,s})$ is the conditional probability of an up movement in y_i given that y_{i-1} is at node r and x_i is at node s .

References

- Amin, K. I., 1990, "A Simplified Discrete Time Approach for the Pricing of Derivative Securities with Stochastic Interest Rates," working paper, University of Michigan Ann Arbor.
- Amin, K. I., 1991, "On the Computation of Continuous Time Option Prices Using Discrete Approximations," *Journal of Financial and Quantitative Analysis*, 26, 477-495.
- Amin, K. I., and J. N. Bodurtha, Jr., 1995, "Discrete Time Valuation of American Options with Stochastic Interest Rates," *Review of Financial Studies*, 8, 193-234.
- Amin, K. I., and A. Khanna, 1994, "Convergence of American Option Values from Discrete to Continuous-time Financial Models," *Mathematical Finance*, 4, 289-304.
- Billingsley, P., 1986, *Probability and Measure*, 2nd ed., John Wiley and Sons, New York.
- Breen, R., 1991, "The Accelerated Binomial Option Pricing Model," *Journal of Financial and Quantitative Analysis*, 26, 153-164.
- Boyle, P. P., 1988, "A Lattice Framework for Options for Pricing when there are Two State Variables," *Journal of Financial and Quantitative Analysis*, 23, 1-12.
- Boyle, P. P., J. Evnine, and S. Gibbs, 1989, "Numerical Evaluation of Multivariate Contingent Claims," *Review of Financial Studies*, 2, 241-250.
- Bunch, D. S., and H. Johnson, 1992, "A Simple and Numerically Efficient Valuation Method for American Puts Using a Modified Geske-Johnson Approach," *Journal of Finance*, 47, 809-816.
- Cox, J. C., and M. Rubinstein, 1985, *Options Markets*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Cox, J. C., S. A. Ross, and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229-263.
- Duffie, D., 1988, *Security Markets*, Academic Press, San Diego, Calif.
- Geske, R., 1979, "The Valuation of Compound Options," *Journal of Financial Economics*, 7, 63-81.
- Geske, R., and H. Johnson, 1984, "The American Put Valued Analytically," *Journal of Finance*, 39, 1511-1542.
- He, H., 1990, "Convergence from Discrete to Continuous-Time Contingent Claims Prices," *Review of Financial Studies*, 3, 523-546.

Nelson, D. B., and K. Ramaswamy, 1990, "Simple Binomial Processes as Diffusion Approximations in Financial Models," *Review of Financial Studies*, 3, 393-430.

Rajasingham, A., 1990, "Convergence of Discrete Time Event Trees to Continuous Time Economies and the Pricing of Contingent Claims," Technical Report, International Monetary Fund, Washington, D.C.

Stapleton, R. C., and M. G. Subrahmanyam, 1988, "The Valuation of Options on Portfolios," *Proceedings of the 15th Annual Conference of the European Finance Association*, Madrid, Spain.

Stulz, R. M., 1982, "Options on the Minimum or the Maximum of Two Risky Assets: Analysis and Applications," *Journal of Financial Economics*, 10, 161-185.