INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313. 761-4700 800. 521-0600

Order Number 9236335

Three essays on contingent claims pricing

Li, Anlong, Ph.D.

Case Western Reserve University, 1992

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

THREE ESSAYS ON CONTINGENT CLAIMS PRICING

by

ANLONG LI

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Thesis Advisor: Peter Ritchken

Department of Operations Research

Case Western Reserve University

May 1992

CASE WESTERN RESERVE UNIVERSITY

GRADUATE STUDIES

We hereby approve the thesis of

ANLONG LI

candidate for the Ph. D.

degree.*

Signed: P. Riecken
(Chairman)

Hamilton Gannon

Edw. Guay

17 m [Signature]

Date 3/20/92

*We also certify that written approval has
been obtained for any proprietary material
contained therein.

I grant to Case Western Reserve University the right to use this work, irrespective of any copyright, for the University's own purposes without cost to the University or to its students, agents and employees. I further agree that the University may reproduce and provide single copies of the work, in any format other than in or from microforms, to the public for the cost of reproduction.

Anthony E.

THREE ESSAYS ON CONTINGENT CLAIMS PRICING

Abstract

by

ANLONG LI

This dissertation consists of three research topics in contemporary financial option pricing theories and their applications. The common theme of these topics involves the pricing of financial claims whose values become path-dependent when using the usual lattice approximating schemes.

The first essay explores the potential of transformation and other schemes in constructing a sequence of simple binomial processes that weakly converges to the desired diffusion limit. Convergence results are established for the valuation of both European and American contingent claims when the underlying asset prices are approximated by simple binomial processes. It is also demonstrated how to construct reflecting or absorbing binomial processes to approximate diffusions with boundaries. Numerical examples demonstrate that the proposed simple approximations not only converge, but also give more accurate results than existing methods such as Nelson and Ramaswamy (1990), especially for longer

maturities.

Our purpose in essay 2 is two-fold. First, we extend some of the simple lattice-approximation methods for one-dimensional diffusions to higher dimensions and develop special lattices to approximate perfectly correlated diffusions. We then examine current modeling issues of the term structure of interest rates, and demonstrate how to apply the approximation techniques developed here to handle path-dependence and multi-sources of uncertainty in these models.

The last essay analyzes the investment decisions of insured banks under fixed-rate deposit insurance. The model takes into account the charter value and allows banks to dynamically revise their asset portfolios. Trade-offs exist between preserving the charter and exploiting deposit insurance. The optimal bank portfolio problem is solved analytically for a constant charter value. In any audit period, banks maximize their risk exposure before some critical time and act cautiously thereafter. The corresponding deposit insurance is shown to be a put option that matures at this critical time rather than at the audit date.

TABLE OF CONTENTS

Abstract	ii
List of Figures	vii
List of Tables	vii

ESSAY 1

BINOMIAL APPROXIMATION IN FINANCIAL MODELS: COMPUTATIONAL SIMPLICITY AND CONVERGENCE	1
1. Introduction	1
2. Weak Convergence and Diffusion Approximation	4
3. Binomial Approximation	8
3.1. Complexity of the Binomial Lattice	8
3.2. Adjusted Binomial Lattices	11
3.3. Binomial Lattices Generated by Transformations	12
4. Singular Diffusions	17
4.1. Reflecting Boundary	18
4.2. Absorbing Boundary	21
5. Contingent Claim Approximation	22
5.1. European Options	22
5.2. Discount Bonds	25
5.3. American Options	26

6. Numerical Examples	30
6.1. Bond Pricing	31
6.2. Stock Options	38
7. Concluding Remarks	41
Appendices to Essay 1	43
A. An Invariant Result	43
B. An Adjusted Binomial Lattice	45
C. Pseudo Path-Independent Model	46
D. Proof of Equations (3.13) and (3.16)	51
E. Controlling Step Size	52
E1. Reflecting Boundary	52
E2. Absorbing/Reflecting Boundary	55

ESSAY 2

APPROXIMATION OF MULTIDIMENSIONAL DIFFUSIONS AND THE TERM STRUCTURE OF INTEREST RATES	57
1. Introduction	57
2. Simple Approximation of Multi-Dimensional Diffusions	59
2.1. The Transformation Methods	60
2.2. The Two-Dimensional Case	63
2.2.1. A Quadrinomial Lattice	63
2.2.2. A Trinomial Lattice	65
2.2.3. The Degenerate Case: $\rho = 1$	67
2.3. Extension to n Perfectly Correlated Diffusions	69
3. Theory of the Term Structure of Interest Rates	71
3.1. Notation and Expectations Hypotheses	71

3.2.	Equilibrium Models of Interest Rates	73
3.3.	Arbitrage-Free Models of Interest Rates	75
3.3.1.	The Risk-Neutralized Forward Process	75
3.3.2.	The Forward Risk-Adjusted Measure	79
4.	Approximating the Term Structure	83
4.1.	Approximating the One-Factor Model	85
4.2.	A Multi-Factor Term Structure Model	88
5.	Concluding Remarks	92
Appendices to Essay 2		93
A.	Proof of Theorem 1	93
B.	Proof of Theorem 2	94
C.	Proof of Theorem 3	97
 <i>ESSAY 3</i>		
OPTIMAL BANK PORTFOLIO CHOICE		
UNDER FIXED-RATE DEPOSIT INSURANCE		100
1.	Introduction	100
2.	The Static Model - No Portfolio Revision	103
3.	Continuous Portfolio Revision	107
4.	Trinomial Approximation	113
5.	A Second Look at the Charter Value	116
6.	Concluding Remarks	120
Appendix to Essay 3		122
 References		130

List of Figures

2.1	Quadrinomial Lattice for 2-Dimensional Diffusion	63
2.2	Trinomial Lattice for 2-Dimensional Diffusion	66
2.3	Trinomial Lattice for Degenerate 2- D Diffusion	68
2.4	$(n+1)$ -Nomial Lattice for Degenerate n - D Diffusion	70
3.1	Optimal Portfolio Policies	111
3.2	Alternative Payoff Functions	116
3.3	Some Specific Payoff Functions	118
3.4	Optimal Policies Under the Payoffs in Figure 3.3	119

List of Tables

1.1.	Discount Bond Prices	36
1.2.	Discount Bond Prices (Long Maturities)	37
1.3.	Call and Put Option Prices	40

**Binomial Approximation in Financial Models:
Computational Simplicity and Convergence**

1. Introduction

Binomial models were first introduced by Sharpe (1978) and Cox, Ross, and Rubinstein (1979) to price options on assets with lognormal prices. This approach is attractive for valuing both American contingent claims and options with alternative asset price processes for which a closed-form option pricing formula, such as that of Black and Schole (1973), is not available. Cox and Rubinstein (1985) extend their model to approximate general diffusion processes. However, the resulting lattice is complicated by the fact that the number of states grows exponentially from one period to the next. A simple way to avoid such complexity is to transform the process into one that can be easily approximated by computationally simple binomial lattices whose nodes grow linearly in number from period to period. This approach has been used by Nelson and Ramaswamy (1990) in binomial models, and by Hull and White (1990a) in the explicit finite-difference method. Amin (1991) suggests transforming the time scale to overcome the computational complexity caused by time-dependent volatilities.

In this essay, we investigate simple binomial approximations from several perspectives. First we identify the class of diffusions that can be simply approximated using the popular binomial models of Cox and Rubinstein (1985, chapter 7) with no transformation. This results in a much larger set of diffusions that can be used as the transformed processes; thus, Nelson and Ramaswamy's (1990) method is generalized. We then explore the possibility of achieving computational simplicity by directly adjusting the Cox and Rubinstein (1985) binomial model. It turns out that the adjusted binomial lattice is a second truncation of the transformation method, further confirming our belief that transformation is, in principle, essential for achieving computational simplicity. However, when the transformation is analytically intractable the adjusted binomial model can serve as an approximation.

We also propose a different approach to resolve the singularity problem associated with the boundary of a diffusion. Such diffusions are approximated here by reflecting or absorbing binomial processes. Although Nelson and Ramaswamy (1990) have developed a multiple-jump scheme for such cases, unfortunately, numerical examples show that their approximations become coarse as the maturity lengthens. Theoretically, both approaches guarantee convergence; however, the method developed here does not become coarse for longer maturities.

Actually, the time increment can be chosen to make the binomial chain purely reflecting or absorbing. Thus, the binomial process will reach an approximating boundary in a given number of steps. The process is either reflected or it stays at the approximating boundary, depending upon the nature of the boundary. This is particularly attractive when applied to the implicit finite-difference method, because it prevents the process from getting too close to the ultimate boundary, and the calculated transition probability will stay within the interval $[0,1]$.

If asset prices can be approximated by binomial processes, then the corresponding options on such assets can be approximated using the same lattice. For European options with a continuous payoff function, the continuous mapping theorem of weak convergence guarantees that the option price sequence obtained from the binomial lattice will converge to its continuous-time counterpart, as long as the binomial processes weakly converge to the diffusion limit. For American options, one has to show that the sequence of optimal exercise strategies obtained from the binomial approximation converges to the optimal exercise strategy in the diffusion limit. This is an issue that has not been thoroughly studied. Assuming the optimal strategies are the same for both the approximating binomial processes and the diffusion limit, one can use the intuitive argument that, before the early exercise, the limit of the option price sequence satisfies the

partial differential equation for the option price in continuous time. However, the optimal strategies are not known beforehand, and it remains to be shown whether the optimal strategies on the approximation lattice converge to the optimal strategy in continuous time.

The essay is organized as follows: Section 2 reviews the basics of diffusion approximation. Section 3 discusses how transformation methods can be used to achieve computationally simple binomial approximations. Section 4 focuses on approximating diffusions with boundaries. Section 5 deals with convergence in approximating both European and American contingent claims. Section 6 provides numerical examples, and section 7 concludes the essay. Proofs can be found in the appendices.

2. Weak Convergence and Diffusion Approximation

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbb{R}^d$ be the d -dimensional Euclidean vector space. Let $W(t)$ be an $\mathbb{R}^d$ -valued Wiener process defined on $(\Omega, \mathcal{F}, P)$. Fix a finite time interval $[0, T]$. Then $\mathcal{F}_t = \mathcal{F}(W(s), 0 \leq s \leq t) \subset \mathcal{F}$ for all $0 \leq t \leq T$. An $\mathbb{R}^d$ -valued diffusion process $Y(t)$ can be defined by the following stochastic differential equation (SDE):

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t))dW(t), \quad Y(0) = Y_0, \quad (2.1)$$

where $\mu(t, Y(t))$ and $\sigma(t, Y(t))$ are the instantaneous mean and

standard deviation of $Y(t)$, respectively. Let $C^d[0,T]$ be the space of $\mathbb{R}^d$ -valued continuous functions on $[0,T]$. Then $W(t)$ and $Y(t)$ have sample paths in $C^d[0,T]$.

The processes used to approximate $Y(t)$ do not necessarily have continuous paths. Let $D^d[0,T]$ be the space of $\mathbb{R}^d$ -valued functions on $[0,T]$, which are right continuous and have left-hand limits, and let $Y^{(n)}(t)$ be a sequence of processes with sample paths in $D^d[0,T]$. We use " $\Rightarrow$ " to denote weak convergence. Then $Y^{(n)} \Rightarrow Y$ if $PY^{(n)} \Rightarrow PY$, where $PY^{(n)}$ and PY are the measures induced by $Y^{(n)}$ and Y , respectively. On the other hand, if $Y^{(n)}(t) \Rightarrow Y(t)$, every finite distribution of $Y^{(n)}(t)$ will converge to that of $Y(t)$. For a more detailed discussion of weak convergence, see Billingsley (1968).

In financial models, diffusions are usually approximated by binomial or multinomial processes in discrete-time. Such processes are characterized by the following definitions.

Definition 1. (Multinomial tree) Let $J_1, \dots, J_m$ be functions from $\mathbb{R}^d$ to $\mathbb{R}^d$, and let $0 = t_0 < t_1 < \dots < t_n = T$ and $Y_0 \in \mathbb{R}^d$. An m -ary tree is constructed as follows. At time t_0 , the starting node (the "root") is labeled Y_0 . At time $t_k < T$, each node Y at the beginning of period k branches into m nodes (the "sons"), labeled $J_1(Y, t_k), \dots, J_m(Y, t_k)$. We call such an arrangement an n -period, m -nomial tree (lattice), and refer to $(J_1, \dots, J_m)$ as the generator of this tree.

Definition 2. (Multinomial process) Let $\{Y_k, 0 \leq k \leq n\}$ be a discrete Markov chain in $\mathbb{R}^d$ with transition function $\nu(y_k, t_k, \Gamma) = P\{Y_{k+1} = \Gamma | Y_k = y_k\}$. If Γ takes values only from $J_1(Y_k, t_k), \dots, J_m(Y_k, t_k)$, where $J_1, \dots, J_m$ are given functions from $\mathbb{R}^d$ to $\mathbb{R}^d$, then Y_k is called an m -nomial Markov chain on the m -nomial tree generated by $(J_1, \dots, J_m)$. The process $Y(t) = Y_{[nt]}$ is called an m -nomial Markov process, where $[nt]$ is the largest integer less than or equal to nt .

In graph theory terminology, levels are used to measure the distance from a node to the root in a tree. Here we use periods instead. We also draw the tree on two-dimensional Cartesian coordinates, with time on the x -axis and state on the y -axis. In this case, we picture the tree growing from left to right.

In each period k (or time t_k), there are m^k nodes. The number of nodes grows exponentially from one period to the next, but this number can be dramatically reduced if we combine those nodes that have the same labels (values). Graphically, we no longer have a tree once this combination is performed; thus, the more general term "lattice" is used. Such a lattice is considered computationally simple if, after combination, the number of nodes grows linearly from period to period.

In these definitions, if $J_i(Y_k, t_k)$ and $\nu(Y_k, t_k, \Gamma)$ are not dependent on the time index t_k , then the m -nomial tree, the Markov

chain $\{Y_k\}$, and the Markov process $Y(t)$ are nonhomogeneous.

For diffusion approximation, we present in lemma 1 a modified version of corollary 7.4.2 of Ethier and Kurtz (1986, pp. 355-356). This is the theoretical basis for the discrete approximations used in recent finance literature (see Nelson [1990] and He [1991]).

Lemma 1. Suppose the SDE (2.1) has an a.e. unique solution for any given Y_0 . Let $Y_k^{(n)}$, $0 \leq k \leq n$, be an m -nomial Markov chain with lattice generator $(J_1, \dots, J_m)$ and transition function $v(y, t, \Gamma)$. Set

$$\mu_n(y, t) = \frac{1}{n} \sum_{i=1}^m [J_i(y, t) - y] v(y, t, J_i(y, t)), \text{ and} \quad (2.2)$$

$$\sigma_n^2(y, t) = \frac{1}{n} \sum_{i=1}^m [J_i(y, t) - y] [J_i(y, t) - y]^T v(y, t, J_i(y, t)). \quad (2.3)$$

Suppose for every $r > 0$,

$$\sup_{\|y\| \leq r} \|J_i(y, t) - y\| \rightarrow 0 \quad \forall i \quad (2.4)$$

$$\sup_{\|y\| \leq r} \|\mu_n(y, t) - \mu(y, t)\| \rightarrow 0 \quad (2.5)$$

$$\sup_{\|y\| \leq r} \|\sigma_n^2(y, t) - \sigma^2(y, t)\| \rightarrow 0. \quad (2.6)$$

Define $Y^{(n)}(t) = Y^{(n)}([nt])$. Then the sequence $\{Y^{(n)}(t)\}$ converges in distribution to the solution of equation (2.1).

In applying lemma 1, one first has to check whether the underlying diffusion equation (2.1) has an a.e. unique solution (or whether the corresponding martingale problem is well posed). Because most diffusions in financial models have this property, condition (2.4) is trivially satisfied in most cases. Conditions (2.5) and (2.6), which are often referred to as the consistency conditions, state that the first two local moments of the discrete process converge to that of the diffusion.

The next lemma is a direct consequence of weak convergence and is useful in proving convergence in option approximation.

Lemma 2. Let g be a real-valued, bounded, and continuous function on $D^d[0, T]$. Then $Y^{(n)} \Rightarrow Y$ implies $g(Y^{(n)}) \rightarrow g(Y)$.

3. Binomial Approximation

3.1. Complexity of the Binomial Lattice

In the rest of this essay, we consider only one-dimensional diffusions. Divide the time interval $[0, T]$ into n subintervals $[t_i, t_{i+1}]$ of equal length $h = T/n$, where $t_i = ih$, $i = 0, 1, \dots, n$. From definition 1, a binomial tree is generated by two functions, J^+ and J^- . Graphically, the building block at node y looks like

$$\begin{array}{c}
 y \text{ --- } \left[\begin{array}{l} \text{--- } y_h^+ = J^+(y, t, \sqrt{h}) \\ \text{--- } y_h^- = J^-(y, t, \sqrt{h}) \end{array} \right.
 \end{array} \tag{3.1}$$

For convenience, we call J^+ and J^- the up and down jumps, respectively. If the up-jump probability is $q(y,t)$, then the down-jump probability is $1 - q(y,t)$. After J^+ and J^- are constructed, $q(y,t)$ is calculated to satisfy consistency conditions (2.5) and (2.6).

The building block (3.1), together with the calculated transition probability $q(y,t)$, generates a binomial model. If the number of nodes grows linearly from period to period, then the model is computationally simple. In particular, the model is path-independent if, starting at any node, the binomial chain reaches the same state by following different paths, as long as these paths have the same number of up and down jumps. The model is considered stable if $q(y,t)$ falls between 0 and 1.

Assume J^+ and J^- are twice differentiable with respect to $\sqrt{h}$ for any given y and t . Then conditions (2.5) - (2.6) require that

$$J^\pm(y,t,\sqrt{h}) = y \pm \sigma(y,t)\sqrt{h} + O^\pm(h). \quad (3.2)$$

Omitting the term $O^\pm(h)$, we have the state-symmetric binomial model of Cox and Rubinstein (1985, chapter 7),

$$y_h^\pm = y \pm \sigma(y,t)\sqrt{h}, \quad (3.3)$$

with up-jump probability

$$q(y,t) = \frac{1}{2} \left[1 + \frac{\mu(y,t)\sqrt{h}}{\sigma(y,t)} \right]. \quad (3.4)$$

An alternative is the probability-symmetric model, where both jumps have probability of 0.5

$$y_h^\pm = y \pm \sigma(y, t)\sqrt{h} + \mu(y, t)h. \quad (3.5)$$

Both models have advantages. The state-symmetric model does not incorporate the drift term $\mu(y, t)$ in the jumps. This coincides with the notion that option price does not depend on the expected stock return. However, state-symmetric models may not be stable, whereas probability-symmetric models always are.¹

For Brownian motion and the geometric Wiener process, both models are computationally simple, with only $n + 1$ nodes in period n . In fact, the state-symmetric model is path-independent for homogeneous diffusions if and only if $\sigma(y)$ is linear in y .²

¹ Trigeorgis (1991) has developed a binomial model for Brownian motion that is both state-symmetric and stable.

² To see this, note that

$$\begin{aligned} y_h^{+-} - y_h^{-+} &= [2\sigma(y, t) - \sigma(y^+, t+h) - \sigma(y^-, t+h)]\sqrt{h} \\ &= -[\sigma_{yy}'(y, t)\sigma^2(y, t) + \sigma_t'(y, t)]\sqrt{h}^3 + O(h^2). \end{aligned}$$

For homogeneous diffusions, $\sigma(y, t) = \sigma(y)$. Thus, $y_h^{+-} - y_h^{-+} = 0$ implies $\sigma_{yy}''(y) = 0$, i.e., $\sigma(y)$ is linear. On the other hand, if $\sigma(y) = a + by$, then

$$\begin{aligned} y^{+-} - y^{-+} &= y^+ - \sigma(y^+)\sqrt{h} - [y^- + \sigma(y^-)\sqrt{h}] \\ &= (y^+ - y^-) - [2a + b(y^+ + y^-)]\sqrt{h} \\ &= 2(a + by)\sqrt{h} - 2(a + by)\sqrt{h} \\ &= 0. \end{aligned}$$

However, this is not the case in general. For the state-symmetric model presented here, a three-period binomial lattice looks like

where

$$\begin{cases} y_h^{+-} = y^+ - \sigma(y^+, t+h)\sqrt{h} \\ y_h^{-+} = y^- + \sigma(y^-, t+h)\sqrt{h}. \end{cases} \quad (3.7)$$

Generally, y_h^{+-} and y_h^{-+} are not equal, so we have to use different nodes to represent them. The number of nodes in the lattice grows geometrically from period to period, exceeding one million in as few as 20 periods. As a result, we have a computationally complex lattice. The next two subsections discuss ways to resolve this problem.

3.2. Adjusted Binomial Lattices

To reduce the complexity of the non-recombining lattice for general diffusions, we make the following adjustment:

$$y_h^\pm(y, t) = y \pm \sigma(y, t)\sqrt{h} + \lambda(y, t)h, \quad (3.8a)$$

where

$$\lambda(y, t) = \sigma(y, t) \left[\frac{1}{2} \sigma'_y(y, t) + \int \frac{\sigma'_t(y, t)}{\sigma^2(y, t)} dy + C \right], \quad (3.8b)$$

and C is any constant. The binomial lattice in equation (3.8a) is computationally simple not only for linear volatility functions, but also for the square root volatility function $\sigma(y, t) = \sigma\sqrt{y}$.

It can be shown that equation (3.8b) is a necessary condition for path-independence. (See appendix B for details.) Generally, we need more adjustment terms in the lattice to close the gap between y_h^{+-} and y_h^{-+} . In the Constant Elasticity of Variance (CEV) model with $\gamma = 1 - k^{-1}$ (k is any positive integer), we need to include terms up to the order of $(\sqrt{h})^k$ in the adjustment to accomplish this. However, the difference $y_h^{+-} - y_h^{-+}$ in the adjusted lattice (3.8a) is usually as small as $o(h^2)$, and we can force the nodes to reconnect to obtain computational simplicity.³ (See appendix C for further discussion.)

3.3. Binomial Lattices Generated by Transformations

In section 3.1, we showed that the Cox and Rubinstein binomial model (3.1) is computationally simple if and only if the volatility function of diffusion (2.1) is linear. For diffusions

³ For example, we can take the average of y_h^{+-} and y_h^{-+} or simply pick either one of them.

with general volatility functions, computational simplicity can be achieved through transformation. To do this, first identify a function f such that the transformed process $X(t) = f^{-1}(X(t), t)$ has a linear volatility. Then construct a sequence of simple binomial processes $X^{(n)}(t)$ that weakly converges to $X(t)$. If f is continuous, $Y^{(n)}(t) = f(X^{(n)}(t), t) \Rightarrow Y(t)$.

We first consider the case in which $X(t)$ has unity volatility.⁴ To identify the transformation $f: X(t) \mapsto Y(t)$, let $g = f^{-1}$. Applying Itô's formula, we have

$$dX(t) = [\mu(Y, t) \frac{\partial g}{\partial Y} + \frac{1}{2} \sigma^2(Y, t) \frac{\partial^2 g}{\partial Y^2} + \frac{\partial g}{\partial t}] dt + \sigma(Y, t) \frac{\partial g}{\partial Y} dW(t). \quad (3.9)$$

Choose g such that $\sigma(Y, t) \frac{\partial g}{\partial Y} = 1$, which gives

$$x = g(y, t) = \int_0^y \frac{dz}{\sigma(z, t)}. \quad (3.10)$$

For convenience, we set the lower limit of the integral to zero. This gives $g(0, t) = 0$. As long as $\sigma(y, t) > 0$, g will be strictly increasing in y . Thus, the transformation f , which is the inverse of g , exists and is strictly increasing in x . Since $X(t)$ can be approximated by the simple model $x_h^\pm = x \pm \sqrt{h}$, the corresponding binomial model for $Y(t)$ is

⁴ This is the case examined in Nelson and Ramaswamy (1990) and Hull and White (1990a).

$$y_h^\pm = f(x \pm \sqrt{h}, t + h), \quad (3.11)$$

with up-jump probability

$$q_h(y, t) = \begin{cases} \frac{\mu(y, t)h + y - y_h^-}{y_h^+ - y_h^-} & \text{if } y^+ \neq y^- \\ 1 & \text{if } y^+ = y^-. \end{cases} \quad (3.12)$$

Using Taylor's expansion for equation (3.11), we have

$$y_h^\pm = y \pm \sigma(y, t)\sqrt{h} + \left[\frac{1}{2}\sigma_y'(y, t)\sigma(y, t) + \int_0^y \frac{\sigma_t'(y, t)}{\sigma^2(y, t)} dy \right] h + o(h). \quad (3.13)$$

Clearly, the adjusted lattice (3.8) is a truncation of equation (3.13). The number of possible states of the binomial lattice (3.11) is at most $2n + 1$ for n partitions of the time interval $[0, T]$. Let $\mathcal{Y}^{(n)}$ be the state space for the binomial Markov chain $Y_k^{(n)}$; then

$$\mathcal{Y}^{(n)} = \{f(f^{-1}(y_0, 0) + k\sqrt{h}, mh) \mid -m \leq k \leq m, 0 \leq m \leq n\}. \quad (3.14)$$

Assumption 1. The diffusion equation (2.1) has an a.e. unique solution $Y(t)$ on $[0, T]$ for any given y_0 .

Assumption 2. $\mu(y, t)$ is continuous, $\sigma(y, t)$ is nonnegative and twice differentiable, and the integral in equation (3.10) exists.

Assumption 3. For every $r > 0$, there exists an $h^* > 0$ such that for all $0 < h < h^*$, $0 \leq q_h(y, t) \leq 1$ for all $y \in \mathcal{Y}^{(n)}$.

The purpose of assumption 1 is obvious. Assumption 2 validates the use of the transformation. Assumption 3 guarantees that equation (3.12) defines a valid transition probability. To apply lemma 1, we need to check conditions (2.4) - (2.6). Condition (2.4) always holds, since f is continuous. From equations (3.12) and (3.13), we have

$$[(y_h^+ - y)q_h + (y_h^- - y)(1 - q_h)]/h - \mu(y, t) = 0$$

$$[(y_h^+ - y)^2 q_h + (y_h^- - y)^2 (1 - q_h)]/h - \sigma^2(y, t) = (q_h - \frac{1}{2})\sigma'_y \sigma^2 \sqrt{h} + o(\sqrt{h}).$$

This implies that conditions (2.5) and (2.6) are satisfied. Thus, in applying lemma 1, we establish the following result.

Theorem 1. Let assumptions 1 - 3 hold. Let $Y_k^{(n)}$, $k = 0, 1, \dots, n$, be the binomial Markov chain with lattice generator (3.11) and transition probability (3.12). Define $Y^{(n)}(t) = Y_{[nt]}^{(n)}$. Then $Y^{(n)}(t) \Rightarrow Y(t)$.

In the above theorem, assumption 3 can be replaced by conditions that are easier to verify. Either of the following is sufficient:

- (i) The transformed process $X(t)$ has a locally bounded drift.
- (ii) There exists an $\varepsilon > 0$ such that $\sigma(y, t) \geq \varepsilon$ for all y and $0 < t < T$.

Condition (i) is somewhat weaker than condition (ii). For

example, the geometric Wiener process does not satisfy condition (ii). However, the transformed process $X(t)$, a Brownian motion, satisfies condition (i).

Generally, one can transform the underlying diffusion Y into a new one, X , whose volatility function is either 1 or $a + bX$, where $b \neq 0$. The resulting binomial process will be slightly different, however. If f is the transformation to a diffusion with unity volatility, the corresponding binomial model will be given by equations (3.11) and (3.12). If the transformed process has linear volatility function $a + bX$, then the resulting binomial model is

$$y_h^\pm = f(f^{-1}(y, t) \pm \frac{\ln(1+b\sqrt{h})}{b}, t + h), \quad (3.15)$$

with the same transition probability as in equation (3.12). Using Taylor's expansion for equation (3.15), we have

$$\begin{aligned} y_h^\pm = y \pm \sigma(y, t)\sqrt{h} + \frac{1}{2}[\sigma_y'(y, t) + b]\sigma(y, t)h \\ + \left[\int_0^y \frac{\sigma_t'(y, t)}{\sigma^2(y, t)} dy \right] h + o(h). \end{aligned} \quad (3.16)$$

Generally, a path-independent binomial model for Y would be

$$y_h^\pm = f(f^{-1}(y, t) \pm \sqrt{h} + o^\pm(\sqrt{h}), t+h),$$

where $o^\pm(\sqrt{h})/\sqrt{h} \rightarrow 0$ as $h \rightarrow 0$. (We may even construct examples in which the term $o^\pm(\sqrt{h})$ depends on y .)

Even though the Cox-Rubinstein model (3.1) is computationally simple for diffusions with linear volatility, one may still prefer to transform these diffusions into those with unity volatility. On the other hand, $x_h^\pm = x \pm \sqrt{h}$ may not be the best choice for the transformed process. For example, Trigeorgis (1991) uses $x_h^\pm = x \pm \sqrt{h}(\sqrt{1 + \mu h})$ to achieve stability, where $\mu = r - \sigma^2/2$, r is the risk-free rate, and σ is the volatility of stock returns. Stability can also be achieved through time changes.

4. Singular Diffusions

Many diffusions in financial models have a lower boundary of 0. For example, stock prices and nominal interest rates are always assumed to be nonnegative. This is often modeled by allowing $\sigma(0, t) = 0$. If the drift term $\mu(0, t)$ equals zero as well, state 0 will serve as an absorbing boundary in many cases.⁵ If the drift term is positive at 0, it will pull the process back from zero and is thus considered a reflecting boundary. There are also cases in between these two.

When $\sigma(y, t)$ is very close to zero for a small state y , the up-jump probability $q_h(y, t)$ in equation (3.12) may be pushed out

⁵ The geometric Wiener process is an exception because it has a natural boundary at 0. If the process starts from a positive state, it will never reach this boundary.

of its meaningful range $[0,1]$, and assumption 3 will be violated.⁶

To avoid this problem, we use absorbing or reflecting binomial processes in the approximation. Specifically, we impose an approximating boundary y_t^* for the binomial process $Y^{(n)}$. Let x_t^* be the corresponding approximating boundary for the transformed process $X^{(n)}$. Then $y_t^* = f(x_t^*, t)$. For technical reasons, we may allow $X^{(n)}$ to be slightly below x_t^* on the lattice; thus, $Y^{(n)}$ may move slightly below y_t^* but not below zero.⁷

Generally, y_t^* depends on the number of partitions n , the time t , and the nature of the true boundary 0 . However, in the limit, we require y_t^* to approach zero for large n . In the rest of this section we examine reflecting and absorbing boundaries separately.

4.1. Reflecting Boundary

Assumption 3a. For any $r > 0$, there exists an $N > 0$ such that for any $h = T/n$ with $n > N$, and for any $t \in (0, T)$, there exists an x_t^* such that

⁶ Nelson and Ramaswamy (1990) suggest that the up-jumps at lower states be moved higher (multi-jump) in the lattice to keep the transition probability between 0 and 1. The magnitude of the multiple jump reflects how "strongly" the drift pulls a small state away from zero.

⁷ Actually, the step size h can be controlled so that the binomial process $X^{(n)}$ reaches the boundary x_t^* in exactly an integer number of steps. For details, see the examples in appendix E.

$$0 \leq q_h(f(x,t),t) \leq 1, \quad x_t^* \leq x < r. \quad (4.1)$$

This assumption allows the calculated transition probability in equation (3.12) to exceed 1 for very small states. At any state smaller than x_t^* , the binomial chain cannot jump down any farther. As a result, the first state below x_t^* serves as the reflecting boundary for the approximating binomial chain $X^{(n)}$. The resulting binomial lattice is

$$x_h^+ = x + \sqrt{h} \quad (4.2a)$$

$$x_h^- = \begin{cases} x - \sqrt{h} & \text{if } x > x_t^* \\ x + j\sqrt{h} & \text{if } x \leq x_t^*, \end{cases} \quad (4.2b)$$

where $j \geq 1$ is the smallest odd integer such that

$$\mu(f(x,t),t)h \leq f(x+j\sqrt{h},t+h) - f(x,t). \quad (4.3)$$

The transition probability $q_h(y,t)$ is given by equation (3.12). The adjustment for x_h^- in equation (4.2b) ensures that $q_h(y,t)$ is between 0 and 1, with odd integer j indicating the strength of the reflection. In most cases, $j \leq 3$. When $j = 1$, the binomial process jumps to a higher node in the lattice with probability 1. It is unlikely that $j = 1$ on a binomial lattice. However, we can choose h to make this happen. (See appendix E for details.)

Theorem 2. Let assumptions 1, 2, and 3a hold. Suppose the value of j in assumption 3a is bounded for all n , and let $X_k^{(n)}$, $k = 0, 1, \dots, n$, be the binomial Markov chain with lattice generator (4.2) and transition probability (3.12). Then $Y^{(n)}(t) = f(X_{[nt]}^{(n)}, [nt]) \Rightarrow Y(t)$.

Proof. We need to show that conditions (2.4) - (2.6) hold for all possible states on the lattice. Condition (2.4) holds because f is continuous. By the definition of $q_h(y, t)$, condition (2.5) also holds. From the proof of theorem 1, we know that condition (2.6) holds for all y corresponding to $x > x_t^*$, so we only need verify this condition for $y \leq y_t^*$. From equation (4.3) and using Taylor's expansion, we have

$$\begin{aligned} \mu(y, t)h &\leq f(x+j\sqrt{h}, t+h) - y \\ &= \sigma(y, t)j\sqrt{h} + \frac{1}{2}\sigma_y'(y, t)\sigma(y, t)j^2h + f_t'(x, t)h + o(h). \end{aligned}$$

Thus, $\sigma(y, t) = O(\sqrt{h})$, and

$$\begin{aligned} |\sigma_n^2(y, t) - \sigma^2(y, t)| &\leq \{[f(x+\sqrt{h}, t+h) - y]^2 \\ &\quad + [f(x+j\sqrt{h}, t+h) - y]^2\}/h + \sigma^2(y, t) \\ &\leq (2 + j^2)\sigma^2(y, t) + o(1) \rightarrow 0. \end{aligned}$$

This implies that condition (2.6) holds for $y \leq y_t^*$.

Q.E.D.

4.2. Absorbing Boundary

The absorbing case is relatively simple. Since both the drift and volatility terms vanish at state 0, we need to prescribe state 0 as an absorbing barrier for the approximating binomial process.

Assumption 3b. For any $r > 0$, there exists an $N > 0$ such that for any $h = T/n$ with $n > N$, and for any $t \in (0, T)$,

$$0 \leq q_h(f(x, t), t) \leq 1, \quad 0 \leq x < r \quad (4.4)$$

$$\mu(0, t) = \sigma(0, t) = 0. \quad (4.5)$$

The binomial lattice for the transformed process is

$$x_h^\pm = \begin{cases} x \pm \sqrt{h} & \text{if } x > 0 \\ x & \text{if } x \leq 0. \end{cases} \quad (4.6)$$

Accordingly, for the original process Y ,

$$y_h^\pm = f(x_h^\pm, t). \quad (4.7)$$

The up-jump probability is given by equation (3.12). When $Y^{(n)}$ reaches the absorbing boundary it stays there with probability 1.

Theorem 3. Let assumptions 1, 2, and 3b hold. Let $X_k^{(n)}$, $0 \leq k \leq n$, be the binomial chain defined by equations (4.6) and (2.12). Then $Y^{(n)}(t) = f(X_{[nt]}^{(n)}, [nt]) \Rightarrow Y(t)$.

Proof. As in the proof of theorem 2, we need only check equation

(2.6) at the boundary 0. In fact,

$$\begin{aligned}\sigma_n^2(0, t) - \sigma^2(0, t) &= [f(0, t+h) - f(0, t)]^2/h \\ &= [f'_t(0, t)h + o(h)]^2/h \\ &\rightarrow 0.\end{aligned}$$

Q.E.D.

5. Contingent Claim Approximation

5.1. European Options

Suppose the stock price follows the diffusion process

$$dS(t) = \mu(t, S(t))dt + \sigma(t, S(t))dW(t), \quad S(0) = S_0, \quad (5.1)$$

and the discount bond price $B(t)$ evolves according to the equation

$$dB(t) = r(t, S(t))dt, \quad B(0) = 1. \quad (5.2)$$

Further, assume the stock does not pay dividends. If the terminal payoff of a European contingent claim at maturity T is $g(S(T))$, then at any time $t \leq T$, the discounted terminal payoff is

$$G_{t,T}(S) = \exp\left[-\int_t^T r(u, S(u))du\right]g(S(T)). \quad (5.3)$$

Following Harrison and Kreps (1979) and Harrison and Pliska (1981), there is an equivalent martingale measure Q on $(\Omega, \mathcal{B})$ under which the price of this contingent claim is the expectation of $G_{t,T}(t, S)$. That is,

$$C_{t,T}^E(S) = \mathbb{E}_Q[G_{t,T}(S)]. \quad (5.4)$$

Under measure Q , the stochastic evolution of the stock prices follows the so-called pseudo process, which differs from process (5.1) only in the drift term. Specifically, the stock price process solves

$$dS(t) = r(t, S(t))S(t)dt + \sigma(t, S(t))dW(t), \quad S(0) = S_0. \quad (5.5)$$

Let $\{S^{(n)}\}$ be a sequence of binomial processes that weakly converges to S under Q . Assume, as before, that the time period is evenly divided into n periods of equal length h . For any n , consider a European contingent claim on a stock whose prices follow process $S^{(n)}$. Let $g(S^{(n)}(T))$ be the payoff of such a claim at maturity T , and $r(t, S^{(n)}(t))$ be the instantaneous return on the associated discount bond. Then, this claim can be priced by arbitrage using standard backward recursion on the approximating binomial lattice.

Let $V_k^E(S_k^{(n)})$ be the value of this claim at node $(kh, S_k^{(n)})$ on the binomial lattice, where $S_k^{(n)}$ is the stock price at time kh . After one period, let the binomial chain jump up to $(kh+h, S_k^{(n)+})$ with probability p_k and then jump down to $(kh+h, S_k^{(n)-})$ with probability $1 - p_k$. To eliminate arbitrage, we have

$$V_k^E(S_k^{(n)}) = \exp[-r(kh, S_k^{(n)})] \mathbb{E}_k[V_{k+1}^E(S_{k+1}^{(n)}) | S_k^{(n)}], \quad (5.6)$$

where

$$\mathbb{E}_k[V_{k+1}^E(S_{k+1}^{(n)}) | S_k^{(n)}] = p_k V_{k+1}^E(S_k^{(n)+}) + (1-p_k) V_{k+1}^E(S_{k+1}^{(n)-}).$$

The boundary condition is

$$V_k^E(S_n^{(n)}) = g(S_n^{(n)}).$$

Since $S^{(n)}(t) = S_{[t/h]}^{(n)}$ is Markovian and its sample paths are step functions, an induction argument yields

$$\begin{aligned} V_{t,T}^E(S^{(n)}) &= \exp\left[-\int_t^T r(u, S^{(n)}(u)) du\right] g(S(T)) \\ &= V_{[t/h], [T/h]}^E(S_{[t/h]}^{(n)}). \end{aligned}$$

Therefore, $V_{t,T}^E(S^{(n)}) = \mathbb{E}_Q[G_{t,T}(S^{(n)})]$ is the value of the claim on $S^{(n)}$. If r and g are continuous functions of t and $S(t)$ on $[0, T]$, then $G_{t,T}(S)$ is continuous on $\mathbb{D}^d[t, T]$. Applying lemma 2, we establish the following result.

Theorem 4. Suppose S is the a.e. unique solution of equation (5.1), r and g are continuous in t and $S(t)$ on $[0, T]$, and $\{S^{(n)}\}$ is a sequence of binomial processes that weakly converges to S under measure Q . Then

$$C_{t,T}^E(S^{(n)}) = \mathbb{E}_Q[G_{t,T}(S^{(n)})] \rightarrow C_{t,T}^E(S). \quad (5.7)$$

5.2. Discount Bonds

A discount bond with maturity T can be viewed as a European option that pays \$1 for every state at time T . Assume the instantaneous interest rate $r(t)$ follows

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t), \quad r(0) = r_0. \quad (5.8)$$

Under the local expectations hypothesis,⁸ at any time $t \leq T$, the price of such a discount bond is

$$B_{t,T}(r) = \mathbb{E}_t \exp\left[-\int_t^T r(u)du\right]. \quad (5.9)$$

When the interest-rate process (5.8) is approximated by binomial processes, we can calculate the approximated price of the discount bond on the binomial lattice. Similar to the European options, we have the following convergent theorem for discount bond.

Theorem 4. Suppose r is the a.e. unique solution of equation (5.8) and $\{r^{(n)}\}$ is a sequence of binomial processes that weakly converges to r under measure Q . Then

$$B_{t,T}(r^{(n)}) = \mathbb{E}_t \exp\left(-\int_t^T r^{(n)}(u)du\right) \rightarrow B_{t,T}(r). \quad (5.10)$$

Similarly, $B_{t,T}(r^{(n)})$ can be calculated on the binomial lattice using backward recursion.

⁸ See Ingersoll (1987) for an account of expectations hypotheses.

5.3. American Options

For American options, not all contracts will be held to maturity; early exercise may be optimal. An exercise strategy is best described by a stopping time, since the decision to exercise an option is based only on information available up to that time. Let $\mathcal{T}_{0,T}$ be the class of $\{\mathcal{F}_t\}$ -stopping times with values in $[0, T]$. Following the arbitrage argument of Karatzas (1988), there exists an optimal $\{\mathcal{F}\}$ -stopping time ρ such that the time t price of an American option is

$$C_{t,T}^A(S, \rho) = \sup_{\tau \in \mathcal{T}_{0,T}} \{C_{t,T}^A(S, \tau)\}, \quad (5.11)$$

where

$$C_{t,T}^A(S, \tau) = E_0 \left[\exp \left(- \int_t^\tau r(u, S(u)) du \right) g(S(\tau)) \right] \quad (5.12)$$

and $g(S(u))$ is the immediate payoff if the option is exercised at time u . Actually, we can restrict the optimization over a smaller class of stopping times than the class $\mathcal{T}_{0,T}$. For example, an American put option can be exercised immediately when the stock price falls below a critical boundary. Shiriyayev (1978) has shown that the value $C_{t,T}^A(S, \rho)$ in equation (5.11) will not change if we consider only the class of stopping times $\mathcal{D}_{0,T}$ that take the form

$$\tau_D = \inf\{t \leq T: (t, S(t)) \in D\}, \quad (5.13)$$

where D is a closed subset of $\mathbb{R}^+ \times [0, T]$. From equation (5.13), τ_D

is the first time the process $S(t)$ reaches the stopping region D . Define the continuation region G as $\mathbb{R}^+ \times [0, T] - D$, and suppose the process $S(t)$ starts within G ; that is, $(0, S(0)) \in G$. Then the option is exercised as soon as the stock price reaches the boundary $\partial G = \partial D$. Let ρ be the optimal stopping time in $\mathcal{D}_{0, T}$. If D^* is the corresponding optimal stopping region, then

$$\tau = \inf\{t: \leq T: (t, S(t)) \in D^*\}. \quad (5.14)$$

In binomial approximation, the option can only be exercised at discrete times $t_k = kh$, $k = 0, 1, \dots, n$. Let $\mathcal{J}_{0, T}^{(n)}$ be the subset of $\{\mathcal{F}_t\}$ -stopping times with discrete values ih , $i \leq n$. Then we have the following convergence theorem for American options.

Theorem 6. Suppose S is the a.e. unique solution of equation (5.5) and $S^{(n)}$ is a sequence of processes that weakly converges to S under measure Q . Suppose further that τ is continuous a.e. relative to the measure induced by the limit process $S(t)$, the boundary of the optimal stopping region. Define

$$C_{t, T}^A(S^{(n)}, \rho^{(n)}) = \sup_{\tau \in \mathcal{J}_{0, T}^{(n)}} \mathbb{E}_Q \left\{ \exp \left[- \int_t^{\tau^{(n)}} r(u, S^{(n)}(u)) du \right] g(S^{(n)}(\rho^{(n)})) \right\}. \quad (5.15)$$

Then

$$C_{t, T}^A(S^{(n)}, \rho^{(n)}) \rightarrow C_{t, T}^A(S, \tau). \quad (5.16)$$

Proof. First, we need to discretize the optimal stopping time τ

so we can compare it with $\rho^{(n)}$. Define

$$\tau^{(n)} = \inf\{t = kh: k \leq n, S^{(n)}(t) \leq D^*\}. \quad (5.17)$$

Since $\rho^{(n)}$ is optimal for the price process $S^{(n)}$, we have

$$C_{t,T}^A(S^{(n)}, \rho^{(n)}) \geq C_{t,T}^A(S^{(n)}, \tau^{(n)}) \quad (5.18)$$

for all n . Using Skorokhod embedding (see Kushner [1990]), we can assume that $S^{(n)}$ and S are defined on the same probability space. Since τ is continuous a.e. relative to the measure induced by the limit process $S(t)$, by weak convergence, $\tau^{(n)} \rightarrow \tau$ a.e. Further, since $C_{t,T}^A(S, \tau)$ is continuous in both S and τ , we have

$$C_{t,T}^A(S^{(n)}, \tau^{(n)}) \rightarrow C_{t,T}^A(S, \tau). \quad (5.19)$$

On the other hand, the sequence $\rho^{(n)}$ is tight since $0 < \rho^{(n)} \leq T$. Let ρ be the limit of some convergent subsequence of $\{\rho^{(n)}\}$. Then

$$C_{t,T}^A(S^{(n)}, \rho^{(n)}) \rightarrow C_{t,T}^A(S, \rho). \quad (5.20)$$

Taking the limit in equation (5.18) yields

$$C_{t,T}^A(S, \rho) \geq C_{t,T}^A(S, \tau).$$

However, since τ is optimal under the price process S , we have

$$C_{t,T}^A(S, \rho) = C_{t,T}^A(S, \tau). \quad (5.21)$$

Notice equation (5.21) does not depend on the subsequence. Q.E.D.

For each n , the discrete optimal stopping problem (5.15) can be solved using dynamic programming on the binomial lattice for the approximating process $S^{(n)}$. Let $V_k^A(S_k^{(n)})$ be the value of the American claim on $S^{(n)}$ at node $(kh, S_k^{(n)})$ on the binomial lattice, where $S_k^{(n)}$ is the stock price at time kh . After one period, suppose the binomial chain jumps up to $(kh + h, S_k^{(n)+})$ with probability p_k and then jumps down to $(kh + h, S_k^{(n)-})$ with probability $1 - p_k$. The Bellman equation for the optimization problem in equation (5.15) is then

$$V_k^A(S_k^{(n)}) = \max\{g(S_k^{(n)}), \exp[-r(kh, S_k^{(n)})] \mathbb{E}_k[V_{k+1}^A(S_{k+1}^{(n)}) | S_k^{(n)}]\}, \quad (5.22)$$

where

$$\mathbb{E}_k[V_{k+1}^A(S_{k+1}^{(n)}) | S_k^{(n)}] = p_k V_{k+1}^A(S_k^{(n)+}) + (1-p_k) V_{k+1}^A(S_k^{(n)-}).$$

The boundary condition is

$$V_n^A(S_n^{(n)}) = g(S_n^{(n)}). \quad (5.23)$$

Since $S^{(n)}(t) = S_{[t/h]}^{(n)}$ is Markovian and its sample paths are step functions, an induction argument yields

$$\begin{aligned} V_{t,T}^A(S^{(n)}) &= \exp\left[-\int_t^T r(u, S^{(n)}(u)) du\right] g(S(T)) \\ &= V_{[t/h]}^A(S_{[t/h]}^{(n)}). \end{aligned}$$

A crucial condition in theorem 6 is the continuity of the optimal stopping time τ on the optimal exercising boundary. A sufficient condition for τ to be continuous is that the paths of the diffusion $S(t)$ are tangent to the boundary ∂D with probability 0.⁹ It is also sufficient to have all the points on the boundary ∂D regular for the diffusion $S(t)$. When $\sigma(S(t)) > 0$ on ∂D , a point on ∂D is regular if it can be reached by an open cone (Dynkin [1965]). For American put options on lognormal prices, Van Moerbeke (1976) shows that the optimal boundary is increasing and continuously differentiable in time. Thus, the open cone condition can be easily verified. Most diffusions and their optimal boundaries in financial models fall into this category. (See Kushner [1984 and 1990] for further discussion.)

6. Numerical Examples

In this section, we apply the methods developed in sections 3 and 4 to approximate discount bond and stock option prices. Since all the diffusion processes used here are homogeneous, the time argument t will be dropped whenever appropriate.

⁹ To see this, we need to show that $\forall \varepsilon, \delta > 0, \exists N > 0$ such that

$$P(|\tau - \tau^{(n)}| > \delta) < \varepsilon, \text{ for } n > N.$$

Actually, for any given $\varepsilon, \delta > 0$, if one of the processes S and $S^{(n)}$ hits the boundary K^* first, say at time t , then the other will hit the boundary within time interval $(t, t + \delta)$ with probability $1 - \varepsilon$.

6.1. Bond Pricing

Suppose the instantaneous interest rate follows the mean reverting square root (MRSR) process

$$dy = \kappa(\mu - y) dt + \sigma\sqrt{y} dW, \quad (6.1)$$

where $\kappa, \mu, \sigma > 0$. From Feller's boundary classification, state 0 is an inaccessible boundary when $2\kappa\mu/\sigma^2 \geq 1$, but an accessible (reflecting or absorbing) boundary when $2\kappa\mu/\sigma^2 < 1$.

From equation (3.10), the transformation for this process is

$$f(x) = \begin{cases} \sigma^2 x^2 / 4 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \quad (6.2)$$

Let $\phi = 4\kappa\mu/\sigma^2 - 1$. Then the transformed process $X(t)$ follows

$$dX = \frac{1}{2}(\phi/X - \kappa X)dt + dW. \quad (6.3)$$

In approximating this particular process, we do not need to explicitly specify the approximating boundary x^* .¹⁰ The binomial lattice for $X(t)$ can be constructed as follows:

$$x_h^+ = x + \sqrt{h} \quad (6.4a)$$

$$x_h^- = x + j\sqrt{h}, \quad (6.4b)$$

where $j \geq 1$ is the smallest odd integer such that the transition

¹⁰ It is generally true that x^* can be implicitly determined within the lattice.

probability defined in equation (3.12) does not exceed 1. The following corollary ensures the convergence of this lattice.

Corollary 1. For the MRSR process (6.1), let $X_k^{(n)}$, $k = 0, 1, \dots, n$, be the binomial Markov chain with lattice generator (6.4) and transition probability (3.12). Then $Y^{(n)}(t) = f(X_{[nt]}^{(n)}) \Rightarrow Y(t)$.

Proof. In applying theorem 2, we need to verify assumption 3a. Consider three cases, recalling that $\phi = 4\kappa\mu/\sigma^2 - 1$.

Case 1: $\phi \geq 2 + \kappa h$. Let

$$x^* = \frac{\sqrt{1+\kappa\phi h} - 1}{\kappa\sqrt{h}} \geq \sqrt{h}. \quad (6.5)$$

The only reachable state below x^* on the lattice is still higher than $x^* - \sqrt{h} \geq 0$. Thus $X^{(n)}$ never goes below 0. The down-jump in equation (6.4b) can be rewritten in the form of equation (4.2b).

For $y \geq y^* = f(x^*)$, $j = -1$. From equation (3.12),

$$q_h(y) = \frac{1}{2} + \frac{\phi\sigma^2/4 - \kappa y}{2\sigma\sqrt{y}} \sqrt{h}. \quad (6.6)$$

Clearly, $q_h(y)$ is strictly decreasing for $y \geq y^*$ with $q_h(y^*) = 1$.

For $y < y^*$ and $j \geq 1$, we have from equation (3.12),

$$\begin{aligned} q_h(y) &= \frac{\kappa(\mu-y)h + y - (\sqrt{y} + j\sigma\sqrt{h}/2)^2}{(\sqrt{y} + \sigma\sqrt{h}/2)^2 - (\sqrt{y} + j\sigma\sqrt{h}/2)^2} \\ &= \frac{(\phi - \kappa x^2)\sqrt{h} - j(2x + j\sqrt{h})}{(1-j)[2x + (1+j)\sqrt{h}]} \\ &= \frac{j}{j-1} - \frac{(\phi + j - \kappa x^2)\sqrt{h}}{(j-1)[2x + (1+j)\sqrt{h}]}. \end{aligned} \quad (6.7)$$

Thus, $q_h(y)$ is also strictly increasing for $0 \leq y \leq y^*$ with $q_h(y^*) = 1$ and $q_h(0) = 1 - (\phi-1)/(j^2-1)$. Therefore $j = 3$ is large enough to ensure $0 \leq q(y) \leq 1$ for $0 \leq y \leq y^*$.

We now show that $q_h(y)$ never becomes negative as y increases. For any large number $r > 0$, we can always choose a sufficiently small h such that $q_h(y) > 0$ for $0 \leq y < r$. Actually, the largest state in $\mathcal{Y}^{(n)}$ is $y_{\max} = f(n\sqrt{h}) = f(\sqrt{T/h})$. Then

$$\lim_{h \rightarrow 0} h y_{\max} = \lim_{h \rightarrow 0} h (\sqrt{y_0} + \sigma n \sqrt{h}/2)^2 = \sigma^2 T/4, \text{ and}$$

$$\lim_{h \rightarrow 0} q_n(y_{\max}) = \lim_{h \rightarrow 0} \left[\frac{1}{2} + \frac{\kappa(\mu - y_{\max}) - \sigma^2/4}{2\sigma\sqrt{y_{\max}}} \sqrt{h} \right] = \frac{1}{2} - \kappa\sigma T/8.$$

Thus, as long as $\kappa\sigma T < 16$, there exists an $h^* > 0$ such that $0 \leq q(y) \leq 1$ for all $0 < y < h^*$ and $y \in \mathcal{Y}^{(n)}$.

Case 2: $0 < \phi < 2 + \kappa h$. This is similar to case 1 except that $x_{\min}$ may become negative. The only change in the proof is to set $x = 0$ in equation (6.6) for $x < 0$.

Case 3: $-1 \leq \phi < 0$. In this case, the reflecting barrier is $y^* = 0$. When $X^{(n)}$ first reaches a negative state, say x_{-1} , if we let $j = -1$, then the calculated probability $q(f(x_{-1}))$ in equation (3.12) is still between 0 and 1, since $\mu(0)h = \kappa\mu h \leq \sigma^2 h/4 = f(\sqrt{h})$. From state x_{-1} , $X^{(n)}$ may go down one step further to an even more negative state, say x_{-2} , before it is reflected. Hence, state x_{-2} serves as the reflecting boundary for $X^{(n)}$. Observe that

- a) at x_{-1} , $Y^{(n)}$ jumps to either $f(x_{-1} + \sqrt{h})$ or $f(x_{-2}) = 0$;
- b) at x_{-2} , $Y^{(n)}$ jumps to either $f(x_{-1}) = 0$ or $f(x_{-2} + j\sqrt{h})$;
- c) the drift of $Y(t)$ is the same at states x_{-1} and x_{-2} ; and
- d) the up-jump probability at x_{-1} is between 0 and 1.

From these facts we claim the up-jump probability at x_{-2} is smaller than that at x_{-1} but is still positive.

We now consider the transition probability for $x > 0$,

$$q_h(y) = \begin{cases} \frac{\kappa(\mu-y)h + y}{(\sqrt{y} + \sigma\sqrt{h}/2)^2} & \text{if } 0 \leq x < \sqrt{h} \\ \frac{1}{2} + \frac{\kappa(\mu-y) - \sigma^2/4}{2\sigma\sqrt{y}}\sqrt{h} & \text{if } x \geq \sqrt{h}. \end{cases} \quad (6.8)$$

For $0 \leq x < \sqrt{h}$, we have $0 \leq q_h(y) < (\kappa\mu h + y)/(y + \sigma h^2/4) \leq 1$, since $\sigma\sqrt{h}/2 < \mu$ for small h . For $x \geq \sqrt{h}$, $q_h(y) = \frac{1}{2} + \frac{\phi - \kappa x^2}{4x}\sqrt{h}$ is concave and bounded from above by $1/2$. Actually, for large y the condition $\kappa\sigma T < 16$ guarantees $q_h(y) \geq 0$. Therefore, assumption 3a is satisfied. Q.E.D.

We now turn to approximating discount bond prices. Suppose the local expectations hypothesis holds. Then the time t price of a discount bond that matures at time T is

$$B(t, Y) = \mathbb{E}_t[\exp(-\int_t^T Y(u) du)].$$

Let $\{Y^{(n)}\}$ be the sequence of binomial processes in either corollary 1 or corollary 2. Then the approximated bond price is

$$B(t, Y^{(n)}) = \mathbb{E}_t \left[\exp \left(- \int_t^T Y^{(n)}(u) du \right) \right]. \quad (6.9)$$

Like the European option, the bond price $B(t, Y^{(n)})$ is calculated using backward recursion on the binomial lattice for $Y^{(n)}$. At node (t_k, y) ,

$$B(t_k, y) = [q_h(y)B(t_{k+1}, y_h^+) + (1-q_h(y))B(t_{k+1}, y_h^-)] \exp(-yh). \quad (6.10)$$

The boundary condition is $B(T, y) = 1$.

Table 1.1 shows the approximated prices of a discount bond when the instantaneous interest rate follows the MRSR process (6.1). The first four columns specify the same parameters as in Nelson and Ramaswamy (1990). The volatility σ and the initial interest rate y_0 are annualized, while the maturity T is measured in months. The next four columns display the bond prices obtained using different numbers of partitions in the approximation. The last column contains the theoretical values calculated using the formula of Cox, Ingersoll, and Ross (1985).

Table 1.2 displays similar results for longer maturities under the same setting. These tables clearly illustrate the convergence of the approximated discount bond prices to the corresponding theoretical values for a wide range of case parameters. Compared with Nelson and Ramaswamy's results using the same parameters, our approximations are much more accurate, especially for higher σ , κ , and T values.

Table 1.1. Discount Bond Prices

κ	σ	T	y_0	n				CIR
				5	50	100	200	
.01	.10	1	.05	99.5841	99.5841	99.5840	99.5841	99.5841
.01	.10	1	.11	99.0876	99.0876	99.0876	99.0876	99.0876
.01	.10	6	.05	97.5288	97.5284	97.5283	97.5284	97.5284
.01	.10	6	.11	94.6529	94.6541	94.6542	94.6541	94.6541
.01	.10	12	.05	95.1172	95.1166	95.1167	95.1166	95.1166
.01	.10	12	.11	89.6059	89.6123	89.6127	89.6129	89.6131
.01	.10	60	.05	78.2003	78.3296	78.3373	78.3412	78.3451
.01	.10	60	.11	58.7689	59.0976	59.1167	59.1262	59.1358
.10	.10	1	.05	99.5834	99.5832	99.5832	99.5831	99.5832
.10	.10	1	.11	99.0883	99.0886	99.0885	99.0887	99.0886
.10	.10	6	.05	97.5028	97.4967	97.4964	97.4961	97.4960
.10	.10	6	.11	94.6782	94.6848	94.6851	94.6855	94.6855
.10	.10	12	.05	95.0166	94.9948	94.9937	94.9930	94.9924
.10	.10	12	.11	89.7000	89.7258	89.7272	89.7278	89.7287
.10	.10	60	.05	76.2518	76.0230	76.0114	76.0057	76.0000
.10	.10	60	.11	59.9993	60.4174	60.4401	60.4514	60.4628
.10	.50	1	.05	99.5835	99.5833	99.5834	99.5832	99.5833
.10	.50	1	.11	99.0885	99.0887	99.0888	99.0887	99.0888
.10	.50	6	.05	97.5193	97.5194	97.5193	97.5194	97.5194
.10	.50	6	.11	94.7138	94.7327	94.7338	94.7344	94.7349
.10	.50	12	.05	95.1375	95.1603	95.1619	95.1624	95.1632
.10	.50	12	.11	89.9539	90.0635	90.0697	90.0729	90.0762
.10	.50	60	.05	82.0624	83.3398	83.4165	83.4422	83.4832
.10	.50	60	.11	69.8774	72.2952	72.4238	72.4837	72.5572
.50	.50	1	.05	99.5802	99.5793	99.5793	99.5792	99.5792
.50	.50	1	.11	99.0918	99.0928	99.0928	99.0929	99.0929
.50	.50	6	.05	97.4081	97.3865	97.3854	97.3848	97.3843
.50	.50	6	.11	94.8171	94.8523	94.8542	94.8552	94.8562
.50	.50	12	.05	94.7177	94.6644	94.6618	94.6605	94.6592
.50	.50	12	.11	90.2849	90.4130	90.4199	90.4234	90.4269
.50	.50	60	.05	74.2063	74.7386	74.7885	74.8086	74.8289
.50	.50	60	.11	67.6248	68.4996	68.5585	68.6002	68.6361

n = number of partitions

Interest rate follows equation (6.1) with $\mu = 8\%$

Face value of the bond = \$100

CIR = accurate value derived from Cox, Ingersoll, and Ross (1985)

Table 1.2. Discount Bond Prices (Long Maturities)

κ	σ	T	y_0	n				CIR
				100	200	500	1000	
.01	.10	120	.05	63.8192	63.8402	63.8502	63.8570	63.8613
.01	.10	120	.11	38.8578	38.8930	38.9130	38.9212	38.9283
.01	.10	180	.05	54.8077	54.8574	54.8770	54.8856	54.8943
.01	.10	180	.11	29.0838	29.1448	29.1815	29.1936	29.2059
.01	.10	240	.05	49.0917	49.1546	49.1919	49.2044	49.2173
.01	.10	240	.11	24.0717	24.1594	24.2089	24.2254	24.2420
.01	.10	360	.05	42.1402	42.2215	42.2858	42.3039	42.3243
.01	.10	360	.11	19.3692	19.4833	19.5522	19.5754	19.5987
.10	.10	120	.05	56.7051	56.6985	56.6946	56.6932	56.6919
.10	.10	120	.11	40.0995	40.1218	40.1351	40.1395	40.1440
.10	.10	180	.05	42.2198	42.2178	42.2166	42.2162	42.2159
.10	.10	180	.11	28.2851	28.3115	28.3273	28.3326	28.3379
.10	.10	240	.05	31.4507	31.4557	31.4588	31.4599	31.4610
.10	.10	240	.11	20.5750	20.6030	20.6199	20.6256	20.6312
.10	.10	360	.05	17.4644	17.4826	17.4936	17.4973	17.5010
.10	.10	360	.11	11.2551	11.2848	11.3028	11.3088	11.3147
.10	.50	120	.05	75.1244	75.2436	75.2722	75.3096	75.3333
.10	.50	120	.11	64.7359	64.8698	64.9397	64.9858	65.0224
.10	.50	180	.05	67.7976	68.0587	68.2020	68.2148	68.2741
.10	.50	180	.11	58.2522	58.6563	58.7874	58.8568	58.9177
.10	.50	240	.05	61.5689	61.5859	61.6834	61.7837	61.8842
.10	.50	240	.11	52.8688	53.0933	53.2095	53.2856	53.4032
.10	.50	360	.05	49.6527	50.1658	50.5437	50.7468	50.8428
.10	.50	360	.11	42.2952	43.2952	43.5028	43.5028	43.8750
.50	.50	120	.05	55.6141	55.7323	55.7938	55.8077	55.8290
.50	.50	120	.11	50.9396	51.0150	51.0918	51.1123	51.1349
.50	.50	180	.05	41.4047	41.5117	41.5891	41.6227	41.6571
.50	.50	180	.11	37.8602	37.9391	38.0753	38.1097	38.1539
.50	.50	240	.05	30.6182	30.7976	30.9942	31.0357	31.0827
.50	.50	240	.11	27.9648	28.2506	28.3672	28.4197	28.4687
.50	.50	360	.05	16.7896	17.0698	17.1780	17.2422	17.3053
.50	.50	360	.11	15.4326	15.6090	15.7393	15.7833	15.8499

n = number of partitions

Interest rate follows equation (6.1) with $\mu = 8\%$

Face value of the bond = \$100

CIR = accurate value derived from Cox, Ingersoll, and Ross(1985)

6.2. Stock Options

Consider the CEV stock price process

$$dY = \mu Y dt + \sigma Y^\gamma dW, \quad 0 < \gamma \leq 1. \quad (6.11)$$

When $\frac{1}{2} \leq \gamma < 1$ (which we assume hereafter), state 0 is an absorbing boundary. The transformation function is

$$y = f(x) = \begin{cases} [\sigma(1-\gamma)x]^{1/(1-\gamma)} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases}$$

with inverse $x = g(y) = \frac{y^{1-\gamma}}{\sigma(1-\gamma)}$, ($y \geq 0$). The transformed process $X(t)$ follows the diffusion process

$$dX = [\mu(1-\gamma)X - \frac{1}{2(1-\gamma)X}]dt + dW. \quad (6.12)$$

The approximating binomial processes $X^{(n)}$ and $Y^{(n)}$ are defined by equations (4.6) and (4.7), respectively. Since $-0.5 \leq \gamma < 1$, it can be shown that $q_h(y)$ is increasing for $y \geq 0$.¹¹ For any given

¹¹ When $0 < x \leq \sqrt{h}$, $q_h(y) = (\mu h + 1)y/y^* = (\mu h + 1)[x/(x + \sqrt{h})]^{1/(1-\gamma)}$ is clearly increasing in y . When $x \leq \sqrt{h}$, let $z = (1-\gamma)\sigma\sqrt{h}y^{\gamma-1}$ and $m = 1/(1-\gamma)$; then $q_h(y) = [\mu h + 1 - (1-z)^m]/[(1+z)^m - (1-z)^m]$, and

$$\frac{dq(y)}{dy} = \frac{-(\mu h + 1)[(1+z)^{m-1} - (1-z)^{m-1}] + 2(1+z)^m(1-z)^m \frac{dz}{dy}}{[(1+z)^m - (1-z)^m]^2/m}.$$

Since $m > 1$ and $z > 0$, we have $dz/dy < 0$, $(1+z)^{m-1} - (1-z)^{m-1} > 2$, and $2(1+z)^m(1-z)^m < 2$. This implies $dq(y)/dy > 0$.

number $r > 0$, we can always choose a sufficiently small h such that $q_h(y) > 0$ for all $0 < y < r$. This is true because $q_h(y)$ is decreasing and $\lim_{h \rightarrow 0} q_h(r) = 1/2$. Moreover, let $y_{\max}$ be the largest state on the binomial lattice. Then

$$\lim_{h \rightarrow 0} y_{\max}^{1-\gamma} \sqrt{h} = \lim_{h \rightarrow 0} [y_0^{1-\gamma} + n(1-\gamma)\sigma\sqrt{h}]\sqrt{h} = (1-\gamma)\sigma T, \text{ and}$$

$$\lim_{h \rightarrow 0} q(y_{\max}) = \frac{1 + (1-\gamma)\mu T}{2}.$$

Consequently, $q(y) < 1$ if $(1-\gamma)\mu T < 1$. Assumption 3b is now satisfied. This leads to the following result.

Corollary 2. For the CEV process (6.11), suppose $-0.5 \leq \gamma < 1$. Let $Y_k^{(n)}$, $k = 0, 1, \dots, n$, be the binomial Markov chain with lattice generator (4.7) and transition probability (4.8). Then $Y^{(n)}(t) = f(X_{[nt]}^{(n)}) \Rightarrow Y(t)$.

Having set up a converging binomial lattice for the stock price, options on the stock can be approximated using backward recursion on the lattice, as described in section 5. Table 1.3 shows the approximated values of the call and put options. We fix $\gamma = 0.5$ and set the annual risk-free rate at 5 percent. The parameter σ is standardized such that the initial annual volatility of the stock return is 20 percent. The initial stock price is \$40. The stock prices (X) range from \$35 to \$40, and the maturities are one, four, and seven months. The first three

Table 1.3. Call and Put Option Prices

σ	X	T	Call Prices				Put Prices			
			$n=5$	$n=50$	$n=100$	CR	$n=5$	$n=50$	$n=100$	NR
.2	35	1	5.142	5.150	5.150	5.15	0.000	0.008	0.008	0.008
.2	40	1	1.049	0.998	1.000	1.00	0.900	0.849	0.851	0.842
.2	45	1	0.012	0.018	0.019	0.02	5.000	5.000	5.000	5.000
.3	35	1	5.212	5.232	5.232	5.23	0.070	0.090	0.091	0.088
.3	40	1	1.531	1.455	1.458	1.45	1.383	1.305	1.307	1.296
.3	45	1	0.103	0.141	0.142	0.14	5.032	5.044	5.045	5.042
.4	35	1	5.434	5.412	5.418	5.42	0.294	0.272	0.277	0.270
.4	40	1	2.013	1.911	1.916	1.92	1.865	1.761	1.765	1.751
.4	45	1	0.416	0.386	0.385	0.38	5.286	5.256	5.255	5.245
.2	35	4	5.803	5.782	5.787	5.79	0.245	0.223	0.228	0.223
.2	40	4	2.256	2.159	2.163	2.17	1.661	1.573	1.576	1.570
.2	45	4	0.496	0.450	0.469	0.47	5.060	5.064	5.064	5.060
.3	35	4	6.358	6.319	6.314	6.31	0.801	0.765	0.761	0.751
.3	40	4	3.208	3.060	3.067	3.07	2.617	2.472	2.476	2.470
.3	45	4	1.220	1.194	1.189	1.18	5.696	5.645	5.641	5.628
.4	35	4	6.943	6.976	6.991	6.99	1.386	1.430	1.443	1.430
.4	40	4	4.161	3.963	3.972	3.98	3.572	3.374	3.380	3.374
.4	45	4	1.909	2.001	2.009	2.00	6.390	6.407	6.414	6.345
.2	35	7	6.492	6.436	6.444	6.43	0.522	0.472	0.477	
.2	40	7	3.114	2.992	2.998	3.00	2.077	1.981	1.983	
.2	45	7	1.091	1.041	1.036	1.04	5.256	5.215	5.214	
.3	35	7	7.218	7.247	7.262	7.26	1.255	1.298	1.310	
.3	40	7	4.359	4.170	4.179	4.19	3.330	3.155	3.159	
.3	45	7	2.039	2.130	2.140	2.13	6.124	6.137	6.144	
.4	35	7	8.068	8.248	8.227	8.23	2.172	2.305	2.284	
.4	40	7	5.604	5.350	5.362	5.37	4.581	4.333	4.341	
.4	45	7	3.140	3.311	3.281	3.29	7.131	7.257	7.231	

n = number of partitions

Stock prices follow equation (5.11) with $\gamma = 0.5$

σ = initial annual volatility of stock return

X = strike price

T = maturity in months

Initial stock price = \$40

Annual interest rate = 5 percent

CR = accurate value adopted from Cox and Rubinstein (1985, p. 364)

NR = Nelson and Ramaswamy's (1990) approximations

columns specify case parameters, the next three display call prices for different numbers of partitions in the approximation, and the seventh column reports the theoretical call prices from Cox and Rubinstein (1985, p. 364). Nelson and Ramaswamy (1990) tabulated approximations only for maturities of one and four months, and their results show the same degree of accuracy as ours. For longer maturities, they reported coarse approximations without tabulating the results. The results displayed here for a seven-month maturity clearly illustrate that our approximations not only converge but are also very accurate.

Columns eight through 11 in table 1.3 are the prices of American put options. The last column contains the approximated values obtained by Nelson and Ramaswamy (1990) using finite-difference methods. (Again, they did not tabulate the results for a seven-month maturity.) The trends of our findings for all maturities clearly illustrate that our binomial approximations converge as n increases.

7. Concluding Remarks

We have demonstrated that transformation is a useful tool for simplifying binomial models in diffusion approximations. We have also shown that singular diffusions are better approximated by reflecting or absorbing binomial processes. This is a promising result, and the idea can also be applied to the finite-difference

methods.

An alternative way to achieve computational simplicity within this framework is through lattice adjustment. For one-dimensional diffusions, this may be less efficient than the transformation method. Nonetheless, it may be worthwhile to develop an adjustment scheme for general multidimensional diffusions for which the transformation method fails.

Another contribution of this essay is the convergence result established in approximating American contingent claims. In almost all cases, the optimal early exercise boundaries cannot be analytically solved. However, the approach taken here does require an analytical formula for the boundary. All that is needed is the continuity of the first hitting time with respect to the sample path of the diffusion.

Appendices to Essay 1

A. An Invariant Result

This appendix shows that the probability-symmetric binomial model (3.5) is invariant under time-homogeneous transformations. That is, path-independence cannot be achieved by a time-transformation while probability-symmetry is preserved. To see this, rewrite equation (3.5) as

$$y_h^\pm = y \pm \sigma(y, t)\sqrt{h} + \mu(y, t)h. \quad (\text{A.1})$$

Consider the difference between the two states when one follows the up-then-down path and the other follows the down-then-up path:

$$y^{+-} - y^{-+} = [2\mu'_y\sigma - 2\sigma'_y\mu - \sigma'_{yy}\sigma^2 - 2\sigma'_t]\sqrt{h}^3 + o(h^2). \quad (\text{A.2})$$

A necessary condition for path-independence is

$$2\mu'_y\sigma - 2\sigma'_y\mu - \sigma'_{yy}\sigma^2 - 2\sigma'_t = 0, \quad (\text{A.3})$$

which is equivalent to

$$\mu - \left[\frac{1}{2}\sigma'_y + \int (\sigma'_t/\sigma^2)dy\right]\sigma = 0. \quad (\text{A.4})$$

Suppose a transformation $x = g(y, t)$ is employed. The transformed $X = g(Y, t)$ diffusion then follows

$$dX = M(X)dt + S(X)dW, \quad (\text{A.5})$$

where

$$M(X) = \mu g'_Y + \frac{1}{2} \sigma^2 g''_{YY} + g'_t \text{ and } S(X) = \sigma g'_Y.$$

Note that

$$\begin{aligned} \int \frac{S'_t(x)}{S^2(x)} dx &= \int [(\sigma_t g'_Y + \sigma g''_{tY}) / (\sigma g'_Y)^2] dx \\ &= \int [\sigma_t / \sigma^2 + g''_{tY} / (\sigma g'_Y)] dy \quad (\text{since } dx = g'_Y dy) \\ &= \int (\sigma_t / \sigma^2) dy + g'_t / S - \int g'_t \frac{\partial}{\partial y} \left[\frac{1}{\sigma g'_Y} \right] dy. \end{aligned}$$

We have

$$\begin{aligned} M(x) - \left[\frac{1}{2} S'_X(x) + \int \frac{S'_t(x)}{S^2(x)} dx \right] S \\ &= \mu g'_Y + \frac{1}{2} \sigma^2 g''_{YY} + g'_t - \left[\frac{1}{2} \sigma'_Y g'_Y + \frac{1}{2} \sigma g''_{YY} \right] \sigma \\ &\quad - \left\{ \int (\sigma_t / \sigma^2) dy + g'_t / S - \int g'_t \frac{\partial}{\partial y} \left[\frac{1}{\sigma g'_Y} \right] dy \right\} \sigma g'_Y \\ &= \left\{ \mu / \sigma - \frac{1}{2} \sigma'_Y - \int (\sigma_t / \sigma^2) dy - \int g'_t \frac{\partial}{\partial y} \left[\frac{1}{\sigma g'_Y} \right] dy \right\} \sigma g'_Y. \end{aligned}$$

Therefore, similar to equation (A.4), a necessary condition for the transformed binomial lattice to be path-independent is

$$\mu / \sigma - \frac{1}{2} \sigma'_Y - \int (\sigma_t / \sigma^2) dy - \int g'_t \frac{\partial}{\partial y} \left[\frac{1}{\sigma g'_Y} \right] dy = 0. \quad (\text{A.6})$$

Obviously, under any homogeneous transformation ($g'_t = 0$), conditions (A.4) and (A.6) are the same. That is, the original and transformed processes are path-independent at the same time.

However, a nonhomogeneous transformation may make a difference in this situation.

B. An Adjusted Binomial Lattice

Consider the binomial approximation of diffusion process (2.1). The adjusted binomial lattice (3.8) is obtained by adding an extra term $\lambda(y,t)h$ to the state-symmetric building block (3.3). To determine the local adjustment term $\lambda(y,t)$, calculate the up-down state (y^{+-}) and down-up state (y^{-+}) for the adjusted binomial model (3.8):

$$\begin{cases} y_h^{+-} = y^+ - \sigma(y^+, t)\sqrt{h} + \lambda(y^+, t)h \\ y_h^{-+} = y^- + \sigma(y^-, t)\sqrt{h} + \lambda(y^-, t)h. \end{cases} \quad (\text{B.1})$$

The difference (gap) between these two states is

$$\begin{aligned} y^{+-} - y^{-+} &= [\lambda(y^+, t+h) - \lambda(y^-, t+h)]h + [2\sigma(y^-, t) - \sigma(y^+, t+h) - \sigma(y^-, t+h)]\sqrt{h} \\ &= 2[\lambda'_y(y, t)\sigma(y, t) - \sigma'_y(y, t)\lambda(y, t) - \frac{1}{2}\sigma''_{yy}(y, t)\sigma^2(y, t) - \sigma'_t(y, t)]\sqrt{h}^3 \\ &\quad + o(h^2). \end{aligned} \quad (\text{B.2})$$

Therefore, the difference between the two expected merging states is of order $\sqrt{h}^3$, while that between nonmerging states is of order $\sqrt{h}$. To close this gap, we choose $\lambda(y,t)$ such that the coefficient of $\sqrt{h}^3$ in (B.2) becomes zero. Or, equivalently,

$$\frac{2\lambda'_y(Y, t)\sigma(Y, t) - 2\sigma'_y(Y, t)\lambda(Y, t)}{\sigma^2(Y, t)} = \sigma''_{yy}(Y, t) + \frac{2\sigma'_t(Y, t)}{\sigma^2(Y, t)}. \quad (\text{B.3})$$

Integrate both sides with respect to y and rearrange. Then

$$\lambda(y, t) = \sigma(y, t) \left[\frac{1}{2} \sigma'_y(y, t) + \int \frac{\sigma'_t(y, t)}{\sigma^2(y, t)} dy \right]. \quad (\text{B.4})$$

When $\sigma(y, t) = \sigma(y)$ does not depend on t , we simply choose

$$\lambda(y) = \frac{1}{2} \sigma(y) [\sigma'_y(y) + C], \quad (\text{B.5})$$

where C is a constant. For simplicity, C is set to 0. If we set $C = b$, then the adjusted binomial lattice (3.8) will capture the first three terms in the Taylor expansion of the alternative binomial model (3.16).

C. Pseudo Path-Independent Model

As noted in section 3, the adjusted binomial lattice may not be path-independent. However, under mildly smooth conditions, the gap between the up-then-down state (y^{+-}) and the down-then-up state (y^{-+}) is sufficiently small after the adjustment. By ignoring such minor differences, we obtain a pseudo path-independent lattice. To illustrate this idea, we specify a procedure for reconnecting the nodes as follows.

Pseudo Path-Independent Algorithm:

Step 1. Starting from the initial node y_0 at time 0, branch into two nodes using the adjusted jump scheme in (3.8).

Denote the two nodes in period 1 by $y_{1,0}(h)$ and $y_{1,1}(h)$:

$$\begin{cases} y_{1,0}(h) = y_0(h) + \frac{1}{2}\sigma(y_0)\sigma_Y(y_0)h - \sigma(y_0)\sqrt{h} \\ y_{1,1}(h) = y_0(h) + \frac{1}{2}\sigma(y_0)\sigma_Y(y_0)h + \sigma(y_0)\sqrt{h}. \end{cases} \quad (C.1)$$

Step 2. At the end of period k (or time $t = kh$, $k \geq 1$), there are $k + 1$ nodes $y_{k,j}(h)$, ($0 \leq j \leq k$). Construct the nodes for period $k + 1$ as follows:

$$\begin{cases} y_{k+1,0}(h) = y_{k,0}(h) + \frac{1}{2}\sigma(y_{k,0})\sigma_Y(y_{k,0})h - \sigma(y_{k,0})\sqrt{h} \\ y_{k+1,j+1}(h) = y_{k,j}(h) + \frac{1}{2}\sigma(y_{k,j})\sigma_Y(y_{k,j})h + \sigma(y_{k,j})\sqrt{h}. \end{cases} \quad (0 \leq j \leq k) \quad (C.2)$$

(That is, except for the "bottom" node in period $k+1$, all other nodes are computed from upward moves of the previous period.)

Step 3. In period k , if the process is at state $y_{k,j}$, $0 \leq j \leq k$, then the lattice generator is

$$\begin{array}{ccc} & q_{k,j} & y_{k+1,j+1} \\ & \swarrow & \uparrow \\ y_{k,j} & \text{---} & \\ & \searrow & \downarrow \\ & 1-q_{k,j} & y_{k+1,j} \end{array} \quad (C.3)$$

where

$$q_{k,j} = 0.5 + \frac{\mu(y_{k,j}, t) - \lambda(y_{k,j}, t)}{2\sigma(y_{k,j}, t)}\sqrt{h}. \quad (C.4)$$

Step 4. Repeat steps 2 and 3 until $k = n$.

In the above construction, we ignore the actual down jumps in the adjusted lattice except for the one that is always down. Note that an actual down jump in period k would create an additional k nodes. Let $\{y_{k+1,j}^*(h), j = 0, \dots, k\}$ represent these nodes. Then

$$y_{k+1,j}^*(h) = y_{k,j}(h) + \frac{1}{2}\sigma(y_{k,j})\sigma_Y(y_{k,j})h - \sigma(y_{k,j})\sqrt{h}, \quad 0 \leq j \leq k. \quad (C.5)$$

In general, $y_{k+1,j}^*(h) \neq y_{k+1,j}(h)$. The difference, however, is negligible. Therefore, in the above pseudo path-independent lattice, we ignore the $y^*(h)$ values completely and force the down jumps to reconnect with the up jumps from the immediate state below. Specifically, we bend the downward branches by

$$\Delta y_{k,j}(h) = y_{k+1,j}^*(h) - y_{k+1,j-1}(h), \quad (1 \leq j \leq k) \quad (C.6)$$

such that $k + 1$ pairs of nodes reconnect at time $(k + 1)h$.

In summary, for the pseudo path-independent algorithm, the up and down jumps at node j in period k are

$$y_{k,j}^+(h) = y_{k,j}(h) + \frac{1}{2}\sigma(y_{k,j})\sigma_Y(y_{k,j})h + \sigma(y_{k,j})\sqrt{h} \quad (C.7)$$

$$\begin{aligned} y_{k,j}^-(h) = & y_{k,j-1}(h) + \frac{1}{2}\sigma(y_{k,j-1})\sigma_Y(y_{k,j-1})h \\ & + \sigma(y_{k,j-1})\sqrt{h} + \Delta y_{k,j}(h). \end{aligned} \quad (C.8)$$

Lemma 3. If $\sigma(y,t)$ has locally bounded partial derivatives up to the fourth order in y and up to the second order in t , then

(1) in the adjusted lattice (3.8),

$$\Delta_2(y, h) = y^{+-} - y^{-+} = o(h^2), \text{ and} \quad (\text{C.9})$$

(2) in the pseudo path-independent algorithm,

$$\Delta_{k,j}(h) = o(h), \quad (k > 1, 1 \leq j \leq k). \quad (\text{C.10})$$

Proof. Equation (C.9) follows directly from the discussion in appendix B. To prove equation (C.10), note that

$$\Delta_{k,1}(h) = \Delta_2(y_{k-1,0}, h) = o(h^2). \quad \forall k > 0.$$

Repeatedly using equation (C.9) yields

$$\begin{aligned} \Delta_{k,2}(h) &= \Delta_2(y_{k-2,1}, h) + \Delta_{k-1,1}(h)[1 + O(h)] \\ &= o(h^2) + o(h^2) = 2 o(h^2) \\ &\dots \\ \Delta_{k,j}(h) &= \Delta_2(y_{k-2,j-1}, h) + \Delta_{k-1,j-1}(h)[1 + O(h)] \\ &= o(h^2) + (j-1)o(h^2) = j o(h^2). \end{aligned}$$

Since $j \leq n = T/h$, in the worst case, we have

$$\begin{aligned} \Delta_{k,j}(h) &= n o(h^2) \\ &= o(h), \quad \text{for all } k > 1, \text{ and } j = 1, \dots, k. \end{aligned}$$

Proposition C. Suppose the diffusion equation (2.1) has an a.e. unique solution $Y(t)$ for any given $Y(0)$. Let $Y^{(n)}(k)$ be the binomial Markov chain corresponding to the pseudo path-independent

lattice in (C.7) - (C.8) and the transition probability (C.4). Define $Y^{(n)}(t) = Y_{[nt]}^{(n)}$. Then $Y^{(n)}$ weakly converges to $Y(t)$.

Proof. From lemma 2, we can rewrite equation (C.8) as

$$y_{k,j}^-(h) = y_{k,j}(h) + \frac{1}{2}\sigma(y_{k,j})\sigma_Y(y_{k,j})h + \sigma(y_{k,j})\sqrt{h} + O(\sqrt{h}^3).$$

To simplify notation, we use y^+ , y^- , and y for $y_{k,j}^+(h)$, $y_{k,j}^-(h)$, and $y_{k,j}(h)$, respectively. From equation (C.4), we have $[2q(y,t) - 1]\sigma(y,t) = [\mu(y,t) - \lambda(y,t)]\sqrt{h}$. Then the local drift $\mu_h(y,t)$ and second moment $\sigma_h^2(y,t)$ are as follows:

$$\begin{aligned}\mu_h(y,t) &= \{q(y,t)[y^+ - y] + (1 - q(y,t))[y^- - y]\}/h \\ &= \{q(y,t)[\lambda(y,t)h + \sigma(y,t)\sqrt{h}] \\ &\quad + [1 - q(y,t)][\lambda(y,t)h - \sigma(y,t)\sqrt{h} + O(\sqrt{h}^3)]\}/h \\ &= \{\lambda(y,t)h + [2q(y,t)-1]\sigma(y,t)\sqrt{h} + O(\sqrt{h}^3)\}/h \\ &= \mu(y,t) + O(\sqrt{h}) \\ \sigma_h^2(y,t) &= \{q(y,t)[y^+ - y]^2 + (1 - q(y,t))[y^- - y]^2\}/h \\ &= \{q(y,t)[\lambda(y,t)h + \sigma(y,t)\sqrt{h}]^2 \\ &\quad + [1 - q(y,t)][\lambda(y,t)h - \sigma(y,t)\sqrt{h} + O(\sqrt{h}^3)]^2\}/h \\ &= \{\sigma^2(y,t)h + O(h^2)\}/h \\ &= \sigma^2(y,t) + O(h).\end{aligned}$$

Thus, the local drift $\mu_h(y,t)$ and second moment $\sigma_h^2(y,t)$ converge to the true drift $\mu(y,t)$ and moment $\sigma(y,t)$. From lemma 1, the pseudo path-independent binomial process weakly converges to the

corresponding diffusion limit.

D. Proof of Equations (3.13) and (3.16)

To prove equation (3.13), use the Taylor expansion for (3.11); then

$$y_h^\pm = f(x, t) + \frac{\partial f(x, t)}{\partial x} \sqrt{h} + \left[\frac{1}{2} \frac{\partial^2 f(x, t)}{\partial x^2} + \frac{\partial f(x, t)}{\partial t} \right] h + o(h). \quad (D.1)$$

Note that $f(x, t) = y$ and the two partial derivatives of f with respect to y in the above equation are

$$\frac{\partial f(x, t)}{\partial x} = \sigma(y, t) \quad \text{and} \quad \frac{\partial^2 f(x, t)}{\partial x^2} = \sigma'_y(y, t) \sigma(y, t).$$

To find $\frac{\partial f(x, t)}{\partial t}$, note that $x = g(y, t)$. Thus $0 = \frac{\partial g}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial g}{\partial t}$, and

$$\frac{\partial f(x, t)}{\partial t} = \frac{\partial y}{\partial t} = - \frac{\partial g / \partial t}{\partial g / \partial y} = \sigma(y, t) \int_0^y \frac{\sigma'_t(y, t)}{\sigma^2(y, t)} dy.$$

Substituting these derivatives into equation (D.1) yields equation (3.13).

Next, we prove equation (3.16). In equation (3.9), we choose g such that $X(t)$ has a linear volatility function:

$$\sigma(y) \frac{\partial g}{\partial y} = a + bx \quad (D.2)$$

or

$$\frac{1}{b} \ln(a+bx) = g(y) = \int^y \frac{dZ}{\sigma(Z)}. \quad (D.3)$$

To approximate $X(t)$, use the Cox-Rubinstein binomial model (3.3)

$$x_h^\pm = x \pm (a + bx)\sqrt{h}. \quad (\text{D.4})$$

The corresponding binomial model for $Y(t)$ is

$$y_h^\pm = f(x \pm (a + bx)\sqrt{h}) \quad (\text{D.5})$$

with up-jump probability (3.12). Using the Taylor expansion on the above equation, we obtain equation (3.16). The derivatives used here are

$$\begin{aligned} \frac{\partial f(x)}{\partial x} &= \frac{\sigma(y)}{a+bx}, \text{ and} \\ \frac{\partial^2 f(x)}{\partial x^2} &= \frac{\frac{\partial \sigma(y)}{\partial y}}{a + bx} \frac{\partial f(x)}{\partial x} - \frac{b\sigma(y)}{(a+bx)^2} = \frac{[\sigma'_y(y)+b]\sigma(y)}{(a + bx)^2}. \end{aligned}$$

E. Controlling Step Size

E1. Reflecting Boundary

This appendix shows how to control the partition size so that the approximating binomial chain is a true random walk with a reflecting barrier. Even though the process to be approximated has a reflecting barrier at 0, the barrier for the Markov chain is a small positive state that approaches 0 as the partition size diminishes. Such a barrier can be constructed by solving $q(y) = 1$. Let y^* be the solution. We can choose a partition size h such

a small positive state that approaches 0 as the partition size diminishes. Such a barrier can be constructed by solving $q(y) = 1$. Let y^* be the solution. We can choose a partition size h such that y^* will be a state for the binomial process. Thus, when the process reaches y^* , it can be reflected with probability 1.

This is best explained by way of example. Let $Y(t)$ be the MRSR process (6.1) with $\phi = 4\kappa\mu/\sigma^2 - 1 \geq 0$. The process has a reflecting barrier at 0. The transformed process $X(t)$ is given by equation (6.3). The approximated boundary x^* can be calculated using equation (6.5). Suppose the transformed process $X(t)$ starts from $X(0)$ and hits the boundary x^* in exactly m steps by following an always-down path. That is, $X(0) - m\sqrt{h} = x^*$, which gives

$$\sqrt{h} = \frac{X(0) - x^*}{m}. \quad (\text{E.1})$$

The number of partitions of the time period $[0, T]$ would be the largest integer that is less than or equal to T/h ; i.e., $n = \lfloor T/h \rfloor$. Since T/h may not be an integer, we simply assume that the binomial process stays at the same state on the residual interval $[nh, T]$.

If we choose $x^* = \phi\sqrt{h}/2$, then $\sqrt{h} = \frac{X(0)}{m + \phi/2}$. Starting with $X(0)$, after m steps, the always-down state will be

$$x^* = \phi\sqrt{h}/2, \text{ for } X^{(n)} \quad (\text{E.2})$$

or

If we choose $q(y^*, mh) = 1$, the binomial process will never go down any farther once it reaches y^* . With probability 1, the process $Y^{(n)}$ jumps up from y^* to

$$y^{*+} = (x^* + \sqrt{h})^2 \sigma^2 / 4. \quad (\text{E.5})$$

The local drift and second moment for the discrete process are

$$\mu_h(y^*, mh) = [y^{*+} - y^*] / h = \kappa\mu \text{ and} \quad (\text{E.6a})$$

$$\sigma_h^2(y^*, mh) = [y^{*+} - y^*]^2 / h = (\kappa\mu)^2 h. \quad (\text{E.6b})$$

The drift and variance for the diffusion limit at state y^* are

$$\mu(y^*, mh) = \kappa(\mu - y^*) = \kappa\mu - \kappa(\phi\sigma)^2 h / 16 = \kappa\mu + O(h) \text{ and} \quad (\text{E.7a})$$

$$\sigma^2(y^*, mh) = \sigma^2 y^* = \phi^2 \sigma^4 h / 16 = O(h). \quad (\text{E.7b})$$

For $y > y^*$, we know that $\mu_h(y, t)$ and $\sigma_h(y, t)$ also converge to $\mu(y, t)$ and $\sigma(y, t)$, respectively. We have the following result.

Proposition E1. Suppose the binomial lattice for the MRSR process (6.1) is generated by equations (4.2a) and (4.2b). Let $\phi = 4\kappa\mu/\sigma^2 - 1 > 0$. Suppose the transformed process $X(t)$ starts with $X(0)$ at time $t = 0$ such that $m = \sqrt{n/T} X(0) - \phi/2$ is an integer less than n . Define an approximated boundary y^* as in equation (E.3). Let the transition probability be defined by equation (3.12) when $y > y^*$. At the approximated boundary y^* , set $q(y^*, t) = 1$. Then the resulting binomial process $Y^{(n)}$ weakly converges to $y(t)$.

E2. Absorbing/Reflecting Boundary

Again, we use the MRSR process (6.1) to illustrate our method. Assume $-1 \leq \phi = 4\kappa\mu/\sigma^2 - 1 < 0$. Then $y = 0$ is a sticky boundary. Let y^* be the small state such that one up jump from 0 to y^* with probability 1 matches the local mean exactly with the drift. That is,

$$y^* = \kappa(\mu - 0)h = \frac{\sigma^2(1+\phi)}{4} h, \quad (\text{E.8})$$

or equivalently, $x^* = \sqrt{1+\phi}\sqrt{h}$ for the transformed process. We control the step size h such that if the process starts from $X(0)$ and follows an always-down path, it will hit the small state y^* in exactly m steps. That is, $X(0) - \sqrt{h} = x^*$. For any state y above y^* , $0 \leq q(y, t) \leq 1$. At x^* , $X^{(n)}$ can either jump up to $x^* + \sqrt{h}$ with probability

$$q(y^*, t) = \frac{2(1+\phi)}{2+2\sqrt{1+\phi}+\phi} + o(h), \quad (\text{E.9})$$

or it can jump down to 0 with probability $1 - q(y^*, t)$. The true drift and variance at state y^* are

$$\mu(y^*, mh) = \kappa(\mu - y^*) = \kappa\mu + \kappa \frac{\sigma^2(1+\phi)}{4} h = \kappa\mu + O(h) \text{ and} \quad (\text{E.10a})$$

$$\sigma^2(y^*, mh) = \sigma^2 y^* = \frac{\sigma^4(1+\phi)}{4} h = O(h). \quad (\text{E.10b})$$

The local drift and second moment for the discrete process are

$$\mu_h(y^*, mh) = \{(y^*, t) \left[\frac{\sigma^2(x^* + \sqrt{h})^2}{4} - \frac{\sigma^2 x^{*2}}{4} \right] + [1 - q(y^*, t)] \left[\frac{-\sigma^2 x^{*2}}{4} \right]\} / h$$

$$\begin{aligned}
\mu_h(y^*, mh) &= \{(y^*, t) \left[\frac{\sigma^2(x^* + \sqrt{h})^2}{4} - \frac{\sigma^2 x^{*2}}{4} \right] + [1 - q(y^*, t)] \left[\frac{-\sigma^2 x^{*2}}{4} \right]\} / h \\
&= q(y^*, t) \left[\frac{\sigma^2(x^* + \sqrt{h})^2}{4} \right] - \frac{\sigma^2(x^*)^2}{4} = \kappa\mu \text{ and} \tag{E11.a}
\end{aligned}$$

$$\begin{aligned}
\sigma_h^2(y^*, mh) &= \{q(y^*, t) \left[\frac{\sigma^2(x^* + \sqrt{h})^2}{4} - \frac{\sigma^2 x^{*2}}{4} \right]^2 + [1 - q(y^*, t)] \left[\frac{-\sigma^2 x^{*2}}{4} \right]^2\} / h \\
&= O(h). \tag{E.11b}
\end{aligned}$$

Proposition E2. Suppose the binomial lattice for the MRSR process (6.1) is generated by (6.4). Let $\phi = 4\kappa\mu/\sigma^2 - 1 > 0$. Suppose the transformed process $X(t)$ starts with $X(0)$ such that $m = \sqrt{n/T} X(0) - \sqrt{1+\phi} \sqrt{h}$ is an integer less than n . Define an approximated boundary y^* as in equation (E.8). Let the transition probability be defined by equation (6.5) when $y > y^*$, and let $q(y^*, t)$ be given by equation (E.9). Set $q(0, t) = 0$. Then the resulting binomial process $Y^{(n)}$ weakly converges to $Y(t)$.

ESSAY 2

Approximation of Multidimensional Diffusions and the Term Structure of Interest Rates

1. Introduction

Multivariate models in option pricing arise in many situations. Usually, the value of an option depends either on more than one underlying asset or on one asset whose price is affected by more than one source of uncertainty, such as index options or options on foreign assets facing exchange-rate risk. Another situation is when more than one variable are driven by a single source of uncertainty. For example, in a one-factor term structure model of interest rates, all forward rates are governed by a single Brownian motion. In addition, there are sophisticated models that involve both situations, such as two-factor term structure models.

Discrete multivariate models in option pricing have been studied by, among others, Boyle (1988), and Boyle, Evnine, and Gibbs (1989). Their approach is similar to that of the single variable case. Asset prices are diffusions that can be approximated by discrete processes on a lattice. The transition probabilities of the discrete process can be determined either by arbitrage or by matching the first two moments between the

discrete processes and the risk-neutralized continuous-time processes. The former is possible only when the economy represented by the discrete model is complete.¹ The latter is a mathematical condition for the discrete model to converge to the desired diffusion limit. There are certain advantages to having a complete discrete model, although for valuation purposes market completeness may not be necessary.

Almost all discrete multivariate models require a constant diffusion matrix to obtain computationally simple approximations. This restricts the variables to being normal or lognormal. It is known that for one-dimensional diffusions a transformation can be applied to achieve computational simplicity. (See Nelson and Ramaswamy [1990] and Amin [1991].) For multivariate diffusions this does not always work. The first objective of this essay is to identify the class of multivariate diffusions that can be approximated by simple discrete models.

The second objective is to develop discrete multivariate models for the term structure of interest rates. Existing models in this area have not been very satisfactory. The simple and insightful model of Ho and Lee (1986) assumes a parallel yield curve movement and permits negative interest rates. The

¹ Cox and Huang (1987) show that for an economy with N sources of uncertainty the corresponding discrete model (economy) is complete only if it is $(N+1)$ -nomial. Such models have been developed by He (1990) and Amin (1991).

equilibrium model of Hull and White (1990) requires extensive searching for the market price of risk. Heath, Jarrow, and Morton (1990) exclude negative rates at the expense of introducing path-dependent terms in their lattice.

This paper is organized as follows: Section 2 develops models to approximate multi-dimensional and perfectly correlated diffusions. Section 3 reviews the theory of the term structure of interest rates, emphasizing the arbitrage-free approach. Section 4 develops lattice approximations of specific path-dependent and multi-factor term structure models. Section 5 concludes the essay.

2. Simple Approximation of Multi-Dimensional Diffusions

An n -dimensional diffusion can usually be described as a solution to the following stochastic differential equation (SDE)

$$dY = \mu(t, Y)dt + \sigma(t, Y)dW, \quad Y(0) = Y_0, \quad (2.1)$$

where W is an n -dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, Q)$, $\mu(t, Y)$ is the n -dimensional drift vector, and $\sigma(t, Y)$ is the $n \times n$ diffusion matrix.

When we say a diffusion is approximated using the lattice approach, we mean that a sequence of discrete Markov processes is defined with the aid of lattices, and in the limit the sequence

weakly converges to this diffusion.² As in the one-dimensional case, an n -dimensional diffusion can be approximated by simple lattices if the diffusion matrix $\sigma(t,Y)$ is constant. In this case, the number of states is of the order m^n , where m is the number of time subintervals. For an arbitrary diffusion matrix, the approximating lattice may have approximately $(n+1)^m$ nodes, which requires tremendous computational effort to obtain a good approximation. In this section, we discuss how to construct approximating lattices in which the number of nodes grows polynomially in the number of time subintervals.

2.1. The Transformation Methods

Except for a few special cases, multi-dimensional diffusions cannot be transformed into ones with a constant diffusion matrix. To see this, consider the 2-dimensional homogeneous case of process (2.1). Let $F: \mathbb{R}^2 \mapsto \mathbb{R}$ be twice differentiable. By Ito's lemma, the transformed process $X = F(Y_1, Y_2)$ satisfies

$$dX = m(X_1, X_2)dt + \frac{\partial F}{\partial Y_1} \sigma_1(Y_1, Y_2) dW_1 + \frac{\partial F}{\partial Y_2} \sigma_2(Y_1, Y_2) dW_2, \quad (2.2)$$

where

$$m(X_1, X_2) = \sum_{i=1}^2 \left[\frac{\partial F}{\partial Y_i} \mu_i(Y_1, Y_2) + (1/2) \frac{\partial^2 F}{\partial Y_i^2} \sigma_i^2(Y_1, Y_2) \right]. \quad (2.3)$$

² A brief discussion of weak convergence, diffusion approximation, and computational simplicity is included in Essay 1.

To achieve computational simplicity we need both diffusion terms in equation (2.2) to be constant. That is,

$$\frac{\partial F}{\partial Y_1} \sigma_1(Y_1, Y_2) = C_1 \quad (2.4a)$$

$$\frac{\partial F}{\partial Y_2} \sigma_2(Y_1, Y_2) = C_2, \quad (2.4b)$$

where C_1 and C_2 are constants. Since F is twice differentiable, $\partial^2 F / \partial Y_1 \partial Y_2 = \partial^2 F / \partial Y_2 \partial Y_1$. This leads to

$$C_1 \partial \sigma_1^{-1}(Y_1, Y_2) / \partial Y_2 = C_2 \partial \sigma_2^{-1}(Y_1, Y_2) / \partial Y_1. \quad (2.5)$$

Obviously, condition (2.5) is very strong. It is satisfied if all the components of the diffusion vector are instantaneously independent. Specifically, consider n one-dimensional diffusions defined by the SDE

$$dY_i = \mu_i(t; Y_1, \dots, Y_n) dt + \sigma_i(t, Y_1) dW_i, \quad Y_i(0) = Y_{i0}, \quad (2.6)$$

where W_i , $i = 1, \dots, n$, are standard Brownian motions with constant correlations. The drift term μ_i and diffusion term σ_i satisfy the usual conditions that guarantee equation (2.2) has an a.e. unique solution. We also assume $\sigma_i(t, Y_1)$ does not depend on Y_j for any $j \neq i$. Define transformation functions

$$F_i(t, y_i) = \int_0^{y_i} \frac{dZ}{\sigma_i(t, Z)}. \quad (2.7)$$

The transformed process $X_i = F_i(t, Y_i)$ then satisfies

$$dX_1 = m_1(t; X_1, \dots, X_n)dt + dW_1, \quad X_1(0) = X_{10}, \quad (2.8)$$

where

$$m(t; X_1, \dots, X_n) = \frac{\partial F_1}{\partial Y_1} \mu_1(t; Y_1, \dots, Y_n) + \frac{1}{2} \frac{\partial^2 F_1}{\partial Y_1^2} \sigma_1^2(t, Y_1). \quad (2.9)$$

To approximate $Y_1(t)$, it is sufficient to approximate $X_1(t)$. Except for the transition probabilities, the lattice for $X_1(t)$ is the same as that for an n -dimensional standard Brownian motion. We make the following assumptions on the original and transformed processes.

Assumption 1. The diffusion process (2.1) has an a.e. unique solution $Y(t)$ on $[0, t]$ for any given Y_0 .

Assumption 2. The transformation function $F_1(t, y_1)$ defined by equation (2.7) is strictly increasing for all $t \in [0, T]$.

Assumption 3. The drift term $m_1(t, x_1)$ in equation (2.8) is continuous and locally bounded.

Uniqueness of the solution of equation (2.8) is a direct consequence of the first two assumptions. The third assumption is relatively strong, and it guarantees the stability of most approximation schemes. For diffusions with boundaries, assumption 3 may not hold, and special techniques such as those in Essay 1 are required to obtain stable approximations.

2.2. The Two-Dimensional Case

Simple approximations for the transformed process (2.6) can be obtained using schemes similar to that of Boyle, Evnine, and Bibbs (1989) and He (1991). We illustrate this for the two-dimensional case where the correlation between the Brownian motions W_1 and W_2 is the constant ρ .

2.2.1. A Quadrinomial Lattice

A quadrinomial lattice is one with four jumps at each node. Similar to the 4-jump lattice of Boyle, Evnine, and Gibbs (1989), we construct a quadrinomial lattice as follows. Suppose that at time t , the two diffusions have values X_1 and X_2 , respectively. The four possible pairs $(X_1 \pm \sqrt{h}, X_2 \pm \sqrt{h})$ at time $t + h$ are as shown in figure 2.1.

Figure 2.1 Quadrinomial Lattice for 2-Dimensional Diffusion

The transition probabilities can be solved by matching the local drifts and the correlation; that is

$$q_1 - q_2 - q_3 + q_4 = m_1 \sqrt{h} \quad (2.10a)$$

$$q_1 + q_2 - q_3 - q_4 = m_2 \sqrt{h} \quad (2.10b)$$

$$q_1 - q_2 + q_3 - q_4 = \rho \quad (2.10c)$$

$$q_1 + q_2 + q_3 + q_4 = 1. \quad (2.10d)$$

The last equality comes from the definition of probability. The second moments are automatically matched by definition of the quadrinomial lattice. Solving this system of equations yields

$$q_1 = [1 + \rho + (m_1 + m_2)\sqrt{h}]/4 \quad (2.11a)$$

$$q_2 = [1 - \rho - (m_1 - m_2)\sqrt{h}]/4 \quad (2.11b)$$

$$q_3 = [1 + \rho - (m_1 + m_2)\sqrt{h}]/4 \quad (2.11c)$$

$$q_4 = [1 - \rho + (m_1 - m_2)\sqrt{h}]/4. \quad (2.11d)$$

The lattice is computationally simple since the number of nodes in each time period grows in the order of n^2 . Using lemma 1 of Essay 1, we establish the following result.

Proposition 1. Suppose assumptions 1-3 hold. Then the discrete process generated by the quadrinomial lattice in figure 2.1 with transition probabilities (2.11) weakly converges to the corresponding diffusion limit (2.8) as $h \rightarrow 0$.

2.2.2. A Trinomial Lattice

The straight-forward 4-nomial lattice we discussed above is computationally simple and adequate for the purpose of valuation approximation. Unfortunately, the corresponding discrete process no longer has the usual Arrow-Debreu complete market property. To complete the market, one has to use the more restrictive trinomial process. Subsequently, He (1990) devised such a trinomial model; however, his lattice is not computationally simple when the drift and diffusion terms are not linear. In what follows, we combine the transformation method with the approach of He (1990) to produce a computationally simple trinomial model.

The SDEs that govern the diffusions Y_1 and Y_2 can be rewritten as

$$dY_1 = \mu_1(t; Y_1, Y_2)dt + \sigma_1(t, Y_1)dW_1, \text{ and} \quad (2.12a)$$

$$dY_2 = \mu_2(t; Y_1, Y_2)dt + \sigma_2(t, Y_2)(\rho dW_1 + \sqrt{1-\rho^2}dW_2^*), \quad (2.12b)$$

where W_2^* is a standard Brownian motion independent of W_1 . Then the transformed diffusions X_1 and X_2 satisfy

$$dX_1 = m_1(t; X_1, X_2)dt + dW_1, \text{ and} \quad (2.13a)$$

$$dX_2 = m_2(t; X_1, X_2)dt + (\rho dW_1 + \sqrt{1-\rho^2}dW_2^*). \quad (2.13b)$$

A trinomial lattice can be constructed as shown in figure 2.2. As with the quadrinomial lattice, the second moments and the

Figure 2.2. Trinomial Lattice for 2-Dimensional Diffusion

correlation are automatically matched due to the way the lattice is constructed. The transition probabilities are solved by matching the local drifts. Specifically, we have

$$(q_1 - q_3)\sqrt{3h/2} = m_1 h \quad (2.14a)$$

$$(q_1 - q_3)\rho\sqrt{3h/2} + (q_1 - 2q_2 + q_3)\sqrt{1-\rho^2}\sqrt{h/2} = m_2 h \quad (2.14b)$$

$$q_1 + q_2 + q_3 = 1. \quad (2.14c)$$

Solving this system of equations yields:

$$q_1 = [1 + (\rho m_1 - m_2)\sqrt{h/2}/\sqrt{1-\rho^2} + m_1\sqrt{3h/2}]/3 \quad (2.15a)$$

$$q_2 = [1 - 2(\rho m_1 - m_2)\sqrt{h/2}/\sqrt{1-\rho^2}]/3 \quad (2.15b)$$

$$q_3 = [1 + (\rho m_1 - m_2)\sqrt{h/2}/\sqrt{1-\rho^2} - m_1\sqrt{3h/2}]/3. \quad (2.15c)$$

Clearly, all three probabilities tend to $1/3$ as h diminishes. The following equalities imply that the second moments and the correlation are also matched:

$$\lim_{h \rightarrow 0} [q_1(\sqrt{3h/2})^2 + q_2(-\sqrt{3h/2})^2]/h = 1 \quad (2.16a)$$

$$\begin{aligned} \lim_{h \rightarrow 0} [q_1(\rho\sqrt{3}+\sqrt{1-\rho^2})^2 + q_2(2\sqrt{1-\rho^2})^2 + q_3(-\rho\sqrt{3}+\sqrt{1-\rho^2})^2](\sqrt{h/2})^2/h \\ = [(\rho\sqrt{3}+\sqrt{1-\rho^2})^2 + (2\sqrt{1-\rho^2})^2 + (-\rho\sqrt{3}+\sqrt{1-\rho^2})^2]/6 \\ = 1 \end{aligned} \quad (2.16b)$$

$$\begin{aligned} \lim_{h \rightarrow 0} [q_1(\sqrt{3h/2})(\rho\sqrt{3}+\sqrt{1-\rho^2}) + q_3(-\sqrt{3h/2})(-\rho\sqrt{3}+\sqrt{1-\rho^2})](\sqrt{h/2})/h \\ = [\sqrt{3}(\rho\sqrt{3}+\sqrt{1-\rho^2}) + (-\sqrt{3})(-\rho\sqrt{3}+\sqrt{1-\rho^2})]/6 \\ = \rho. \end{aligned} \quad (2.16c)$$

Similar to proposition 1, we have the following result.

Proposition 2. Suppose assumptions 1-3 hold. Then the discrete process generated by the trinomial lattice in figure 2.2 with transition probabilities (2.15) weakly converges to the corresponding diffusion limit (2.13) as $h \rightarrow 0$.

2.2.3. The Degenerate Case: $\rho = 1$

We now consider the case where both diffusions are driven by a single Brownian motion. The quadrinomial lattice in figure 2.1 is no longer valid. One of the transition probabilities q_2 or q_4 in equations (2.11b) and (2.11c) will be negative since m_1 and m_2 are not always equal. The trinomial lattice in figure 2.2 is also invalid because the probabilities in equations (2.15a-c) are not defined. To overcome these problems, we construct a new trinomial

lattice as shown in figure 2.3.

Figure 2.3 Trinomial Lattice for Degenerate 2-D Diffusion

In the above lattice, $\theta = \text{sign}(m_1 - m_2)$. To match the local drifts, we have

$$(q_1 + \theta q_2 - q_3)\sqrt{h} = m_1 h \quad (2.17a)$$

$$(q_1 - \theta q_2 - q_3)\sqrt{h} = m_2 h \quad (2.17b)$$

$$q_1 + q_2 + q_3 = 1. \quad (2.17c)$$

Solving this system of equations yields

$$q_1 = [1 - |m_1 - m_2|\sqrt{h}/2 + (m_1 + m_2)\sqrt{h}]/2 \quad (2.18a)$$

$$q_2 = |m_1 - m_2|\sqrt{h}/2 \quad (2.18b)$$

$$q_3 = [1 - |m_1 - m_2|\sqrt{h}/2 - (m_1 + m_2)\sqrt{h}]/2. \quad (2.18c)$$

Clearly, the probabilities tend to $1/2$, 0 , and $1/2$, respectively, as h diminishes. The following equations imply that the second moments are matched and the correlation is 1:

$$\lim_{h \rightarrow 0} [q_1 \sqrt{h}^2 + q_2 (\theta \sqrt{h})^2 + q_3 \sqrt{h}^2] / h = 1 \quad (2.19a)$$

$$\lim_{h \rightarrow 0} [q_1 \sqrt{h}^2 + q_2 (-\theta \sqrt{h})^2 + q_3 \sqrt{h}^2] / h = 1 \quad (2.19b)$$

$$\lim_{h \rightarrow 0} [q_1 \sqrt{h} \sqrt{h} + q_2 \theta \sqrt{h} (-\theta \sqrt{h}) + q_3 \sqrt{h} \sqrt{h}] / h = 1. \quad (2.19c)$$

Proposition 3. Suppose assumptions 1-3 hold. Then the discrete process generated by the trinomial lattice in figure 2.3 with transition probabilities (2.18) weakly converges to the corresponding diffusion limit (2.13) with $\rho = 1$ as $h \rightarrow 0$.

2.3. Extension to n Perfectly Correlated Diffusions

We now construct an $(n+1)$ -nomial lattice to approximate n perfectly correlated diffusions Y_i , $i = 1, \dots, n$, simultaneously driven by a common Wiener process $W(t)$. That is,

$$dY_i = \mu_i(t; Y_1, \dots, Y_n) dt + \sigma_i(t, Y_i) dW, \quad Y_i(0) = Y_{i0}. \quad (2.20)$$

Under transformation (2.2), the transformed processes satisfies

$$dX_i = m_i(t; X_1, \dots, X_n) dt + dW, \quad X_i(0) = X_{i0}. \quad (2.21)$$

Again, we need only approximate the X_i 's. To simplify the notation, we use m_i for $m_i(t; X_1, \dots, X_n)$ as long as no confusion arises. Suppose $m_1 \geq m_2 \geq \dots \geq m_n$; otherwise, we can rearrange the order. The $(n+1)$ -nomial lattice is shown in figure 2.4.

The second moments and the correlations are automatically matched independent of the transition probabilities. To match the

Figure 2.4 (n+1)-Nomial Lattice for Degenerate n-D Diffusion

local drifts, we require

$$\begin{aligned}
 q_1 + q_2 + q_3 + \dots + q_n &= 1 \\
 q_1 + q_2 + q_3 + \dots - q_n &= m_1 \sqrt{h} \\
 &\dots \\
 q_1 + q_2 - q_3 + \dots - q_n &= m_{n-1} \sqrt{h} \\
 q_1 - q_2 - q_3 + \dots - q_n &= m_n \sqrt{h}.
 \end{aligned} \tag{2.22}$$

Solving this system of equations yields

$$\begin{aligned}
 q_0 &= (1 - m_1 \sqrt{h})/2 \\
 q_i &= (m_i - m_{i+1}) \sqrt{h}/2, \quad (0 < i < n) \\
 q_n &= (1 + m_n \sqrt{h})/2.
 \end{aligned} \tag{2.23}$$

Proposition 4. Suppose assumptions 1-3 hold. Then the discrete process generated by the $(n+1)$ -nomial lattice in figure 2.4 with transition probabilities (2.23) weakly converges to the corresponding diffusion limit (2.21) as $h \rightarrow 0$.

3. The Theory of the Term Structure of Interest Rates

In this section, we briefly review the expectations hypotheses and the equilibrium and arbitrage-free approaches to study of the term structure of interest rates. This leads to a better understanding of the advantages and limitations of various existing models, and provides the necessary background for applying approximations.

3.1. Notation and Expectations Hypotheses

Consider a continuous economy where discount bonds are traded. The following notation is associated with a discount bond that pays \$1 at maturity date T :

$P(t,T)$ — Price of the bond

$Y(t,T)$ — Yield to maturity

$R(t,T)$ — Return to maturity

$f(t,T)$ — Forward rate.

We then have

$$P(t,T) = \exp\left[-\int_t^T f(t,s)ds\right], \text{ and} \quad (3.1)$$

$$f(t, T) \equiv - \frac{1}{P(t, T)} \frac{\partial P(t, T)}{\partial T}. \quad (3.2)$$

The instantaneous interest rate, or the spot rate, is defined as

$$r(t) \equiv f(t, t) \equiv \lim_{h \rightarrow 0} f(t, t+h). \quad (3.3)$$

A money market account which starts with \$1 at time 0 and grows at rate $r(t)$ has a value $B(t)$ at time t . That is,

$$B(t) = \exp\left\{\int_0^t r(s) ds\right\}. \quad (3.4)$$

The four commonly used expectations hypotheses are:

Unbiased expectation hypothesis:

$$P(t, T) = \exp\left\{-\int_t^T \mathbb{E}_t[r(s)] ds\right\}; \quad (3.5)$$

Return-to-maturity expectations hypothesis:

$$1/P(t, T) = \mathbb{E}_t\left\{\exp\left[\int_t^T r(s) ds\right]\right\}; \quad (3.6)$$

Yield-to-maturity expectations hypothesis:

$$\ln[P(t, T)] = \mathbb{E}_t\left\{-\int_t^T r(s) ds\right\}; \quad (3.7)$$

Risk-neutral (local) expectations hypothesis:

$$P(t, T) = \mathbb{E}_t\left\{\exp\left[-\int_t^T r(s) ds\right]\right\}. \quad (3.8)$$

The expectation operator $\mathbb{E}_t$ is conditioned on all the events

up to time t . It is well-known that in continuous-time only the local expectations hypothesis is consistent in equilibrium, while the other three lead to arbitrage opportunities. The local expectations hypothesis is equivalent to the statement that the relative price $P(t,T)/B(t)$ is a martingale. That is, equation (3.8) is equivalent to

$$P(t,T)/B(t) = E_t\{P(T,T)/B(T)\} \quad (3.9)$$

for any $0 \leq t \leq T$. This can be easily verified by substituting $B(t)$ and $B(T)$ from equation (3.4) into the above equation, and noting the fact that $P(T,T) = 1$.

3.2. Equilibrium Models of Interest Rates

Among early interest-rate models, the one by Vasicek (1977) is still frequently cited in the literature due to its analytical tractability. He assumes the spot rate $r(t)$ evolves according to the equation

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t), \quad (3.10)$$

where κ , θ , and σ are constants and $W(t)$ is a standard Brownian motion. This process is called the Ornstein-Uhlenbeck process which has a known distribution. However, this model permits negative interest rates with positive probability. Alternatively, Cox, Ingersoll, and Ross (1985) suggested the mean-reverting square-root process which satisfies

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r}dW(t). \quad (3.11)$$

This process has gained in popularity because it is nonnegative, mean-reverting, and it has a known distribution. Moreover, the variance increases with the interest rate.

Generally, the interest-rate process in a one-factor model follows a well-defined SDE

$$dr(t) = \mu(t, r)dt + \sigma(t, r)dW(t). \quad (3.12)$$

If we assume the local expectations hypothesis holds, then all discount bonds must be priced according to equation (3.8). Equations (3.8) and (3.12) together imply that the bond price $P(t, T)$ depends only on the current value, not the historical values, of interest rate $r(t)$. This is in subtle contrast to the arbitrage-free approach in the next subsection.

Since investors are not always risk neutral, the bond price in equation (3.8) may not match the observed price. The market price of risk, $\lambda(t)$, should therefore be included in the pricing equation. As a result, equation (3.12) can be rewritten as

$$dr(t) = [\mu(t, r) - \lambda(t)\sigma(t, r)]dt + \sigma(t, r)d\tilde{W}(t), \quad (3.13)$$

where

$$\tilde{W}(t) = W(t) - \int_0^t \lambda(s)ds \quad (3.14)$$

is a standard Brownian motion on $(\Omega, \mathcal{F}, \tilde{Q})$, and $\tilde{Q}$ is an equivalent martingale measure on $(\Omega, \mathcal{F})$ given by

$$d\tilde{Q}/dQ = \exp\left[\int_0^t \lambda(s) dW(s) - \frac{1}{2} \int_0^t \lambda^2(s) du\right]. \quad (3.15)$$

Discount bonds are then priced by

$$P(t, T) = \tilde{\mathbb{E}}_t \left\{ \exp\left[-\int_t^T r(s) ds\right] \right\}, \quad (3.16)$$

where the expectation is taken under probability measure $\tilde{Q}$. In practice, the market price of risk $\lambda(t)$ can be solved by matching the initial term structure. That is, choose $\lambda(t)$ such that

$$P(0, T) = \tilde{\mathbb{E}}_0 \left\{ \exp\left[-\int_0^T r(s) ds\right] \right\} \quad (3.17)$$

for all maturity T . The forward rate implied by the spot-rate process (3.12) is

$$f(t, T) = \tilde{\mathbb{E}}_t \left\{ \exp\left[-\int_t^T r(s) ds\right] r(T) \right\} / \tilde{\mathbb{E}}_t \left\{ \exp\left[-\int_t^T r(s) ds\right] \right\}. \quad (3.18)$$

It is the study of this process that bridges the equilibrium and the arbitrage-free approaches. This, however, is beyond the scope of this essay.

3.3. Arbitrage-Free Models of Interest Rates

3.3.1. The Risk-Neutralized Forward Process

As seen in section 3.2, equilibrium models produce bond prices that explicitly depend on the market price of risk $\lambda(t)$. This is in contrast to the arbitrage-free approach of Heath, Jarrow, and

Morton (1992), which is an extension of Ho and Lee (1986) in the framework of Harrison and Kreps (1979). The arbitrage-free approach takes the current yield curve as input of the model. The forward-rate process $f(t, T)$ is assumed to satisfy an SDE

$$f(t, T) = f(0, T) + \int_0^t \alpha(s, T) ds + \sum_{i=1}^n \int_0^t \sigma_i(s, T) dW_i(s), \quad (3.19)$$

where the initial forward rate $f(0, T)$ can be observed in the market at time 0. Setting $T = t$, we have the spot-rate process

$$r(t) = f(t, t) = f(0, t) + \int_0^t \alpha(v, t) dv + \sum_{i=1}^n \int_0^t \sigma_i(v, t) dW_i(v). \quad (3.20)$$

To avoid arbitrage across bonds of different maturities, Heath, Jarrow, and Morton (1992) showed that the drift term $\alpha(t, T)$ must be constrained to

$$\alpha(t, T) = - \sum_{i=1}^n \sigma_i(t, T) [\phi_i(t) - \int_t^T \sigma_i(t, u) du], \quad (3.21)$$

where $\phi_i(t)$ is the market price of the i th risk and is a predictable process independent of T . Define

$$\tilde{W}_i(t) = W_i(t) - \int_0^t \phi_i(v) dv. \quad (3.22)$$

By the Girsanov theorem (Elliott [1982], Corollary 13.25), $\tilde{W}(t)$ is a standard Brownian motion on $(\Omega, \mathcal{F}, \tilde{Q})$, where $\tilde{Q}$ is an equivalent martingale measure on $(\Omega, \mathcal{F})$ given by

$$d\tilde{Q}/dQ = \exp\left\{ \sum_{i=1}^n \left[\int_0^t \phi_i(v) dW(v) - \frac{1}{2} \int_0^t \phi_i^2(v) dv \right] \right\}. \quad (3.23)$$

The process $\tilde{W}_1(t)$ and measure $\tilde{Q}$ are also referred to as the risk-neutralized Wiener process and the equivalent martingale measure, respectively. Substituting equations (3.21) and (3.22) into equation (3.19), we have

$$f(t, T) = f(0, T) + \sum_{i=1}^n \int_0^t \sigma_i(v, T) \int_v^T \sigma_i(v, u) du dv + \sum_{i=1}^n \int_0^t \sigma_i(v, T) d\tilde{W}_1(v). \quad (3.24)$$

The spot-rate process (3.20) can then be rewritten as

$$r(t) = f(0, t) + \sum_{i=1}^n \int_0^t \sigma_i(v, t) \int_v^t \sigma_i(v, u) du dv + \sum_{i=1}^n \int_0^t \sigma_i(v, t) d\tilde{W}_1(v). \quad (3.25)$$

The main results of Heath, Jarrow and Morton (1992) is stated in the following theorem.

Theorem 1. The value of any interest-rate contingent claim relative to the value of the money market account is a martingale under measure $\tilde{Q}$. Specifically, let $C(t)$ be the price of an interest-rate contingent claim with payoff $C(T)$ at the expiration date T . Then

$$C(t)/B(t) = \tilde{E}[C(T)/B(T) | \mathcal{F}_t], \quad (3.26)$$

where $\tilde{E}$ is the expectation operator under measure $\tilde{Q}$.

Proof. See appendix A.

Theorem 1 ensures that the local expectations hypothesis

holds under measure $\tilde{Q}$. This leads to the bond pricing formula

$$P(t, T) = \tilde{\mathbb{E}}\{\exp[-\int_t^T r(s)ds] \mathcal{F}_t\}. \quad (3.27)$$

Example 1: (Heath, Jarrow, and Morton [1992]) Let the volatility of the forward rate $f(t, T)$ be constant; that is, $\sigma(t, T) = \sigma$. From equation (3.19) the drift is

$$\alpha(t, T) = -\sigma[\phi(t) - \sigma(T-t)].$$

Then

$$\begin{aligned} df(t, T) &= -\sigma[\phi(t) - \sigma(T-t)]dt + \sigma dW(t) \\ &= \sigma^2(T-t)dt + \sigma d\tilde{W}(t), \end{aligned}$$

where $\tilde{W}(t) = W(t) - \int_0^t \phi(s)ds$ is the risk-neutralized standard Brownian motion. Solving this equation, we have

$$f(t, T) = f(0, T) + \sigma^2 t(T-t/2) + \sigma \tilde{W}(t). \quad (3.28)$$

Then the spot rate $r(t) = f(t, t)$ and the discount-bond price are

$$r(t) = f(0, t) + \sigma^2 t^2/2 + \sigma \tilde{W}(t), \text{ and}$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\{-(\sigma^2 tT/2 + \sigma \tilde{W}(t))(T-t)\},$$

respectively. These equations can be rewritten as

$$f(t, T) = f(0, T) - f(0, t) + \sigma^2 t(T-t) + r(t), \text{ and} \quad (3.29)$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\{-[r(t) - f(0, t) + \sigma^2 t(T-t)/2](T-t)\}. \quad (3.30)$$

Applying theorem 1, the time 0 value of a European option expiring at time t , on a discount bond maturing at time T with strike price K , is

$$\begin{aligned} C(0) &= \tilde{\mathbb{E}}\{\max[P(t,T)-K,0]B(0)/B(t)|\mathcal{F}_0\} \\ &= P(0,T)N(d_1) - KP(0,t)N(d_2), \end{aligned} \quad (3.31)$$

where

$$d_1 = [\log \frac{P(0,T)}{KP(0,t)} + \frac{1}{2}\sigma^2(T-t)^2t]/[\sigma(T-t)\sqrt{t}]$$

$$d_2 = d_1 - \sigma(T-t)\sqrt{t}.$$

The bond and option pricing formulae in this example are consistent with the current observed bond prices, and do not explicitly depend on the market price of risk $\phi(t)$. Since the volatility $\sigma(t,T)$ is constant, all uncertainty is suppressed into the Markovian spot-rate process $r(t)$. However, this is not true in general. The spot-rate process is usually path-dependent, which complicates implementation of the model.

3.3.2. The Forward Risk-Adjusted Measure

We have seen that the risk-neutralized martingale measure $\tilde{Q}$ makes the price of any contingent claim, relative to the money market account value, a martingale. We now show that there exists a so-called forward risk-adjusted martingale measure $\tilde{Q}$ under which the price of any contingent claim, relative to the price of a

reference bond, is a martingale. This idea has been explored by Jamshidian (1987) and Ritchken and Sankarasubramanian (1991a). The advantage of this approach is that the discount factor in the pricing formula is factored out of the expectation operator, simplifying contingent-claim valuation in certain circumstances.

Suppose a discount bond with maturity S is used as the reference bond. At time t , the forward price of a discount bond maturing at time T and to be delivered at time S , is

$$F(t, S, T) = \frac{P(t, T)}{P(t, S)} = \exp\left[-\int_t^T f(t, u) du\right], \quad (t \leq S \leq T). \quad (3.32)$$

Define

$$\psi_1(t) = \phi_1(t) - \int_t^S \sigma_1(t, u) du. \quad (3.33)$$

Then

$$\bar{W}_1(t) = W_1(t) - \int_0^t \psi_1(u) du \quad (3.34)$$

is a standard Brownian motion on $(\Omega, \mathcal{F}_t, \bar{Q})$, and $\bar{Q}$ is an equivalent martingale measure on $(\Omega, \mathcal{F})$ given by³

$$d\bar{Q}/dQ = \exp\left\{\sum_{i=1}^n \left[\int_0^t \psi_i(u) dW_i(u) - \frac{1}{2} \int_0^t \psi_i^2(u) du\right]\right\}. \quad (3.35)$$

The process $\bar{W}_1$ is often referred to as the risk-adjusted Wiener process. It reduces to the original Wiener process W_1 when the market price of the i th risk $\phi_1(t)$ is set to $\int_t^S \sigma_1(t, u) du$. The

³ The equivalent martingale measure $\bar{Q}$ depends on S through $\psi_1(t)$.

valuation theorem under the forward risk-adjusted measure is stated as follows. (This can also be found in Ritchken and Sankarasubramanian [1991a].)

Theorem 2. Let $C(t)$ be the price of a European interest-rate contingent claim with payoff $C(S)$ at expiration date S , where $C(S)$ is measurable on $(\Omega, \mathcal{F}_t, \bar{Q})$. Then $C(t)/P(t, S)$ is a $\bar{Q}$ -martingale; that is

$$C(t) = P(t, S) \bar{E}[C(S) | \mathcal{F}_t]. \quad (3.36)$$

In particular, the forward bond price $F(t, S, T)$ is a $\bar{Q}$ -martingale; that is,

$$F(s, S, T) = \bar{E}\{F(t, S, T) | \mathcal{F}_s\}, \quad (s \leq t), \quad (3.37)$$

where $\bar{E}$ is the expectation operator on $(\Omega, \mathcal{F}_t, \bar{Q})$.

Proof. See appendix B.

The forward-rate process can be rewritten as

$$df(t, T) = \sum_{i=1}^n \sigma_i(t, T) \left[\int_S^T \sigma_i(t, u) du + d\bar{W}_i(t) \right]. \quad (3.38)$$

In the special case where $S = T$, we have

$$df(t, S) = \sum_{i=1}^n \sigma_i(t, S) d\bar{W}_i(t). \quad (3.39)$$

This means that the forward rate $f(t, S)$ is also a $\bar{Q}$ -martingale. Although the local expectations hypothesis is no

longer valid under measure $\bar{Q}$, the contingent claim can still be priced with the help of this measure.

For American claims, the discount factor $P(t, \tau^*)$ depends on the optimal exercise date τ^* . Thus, the decomposition cannot be done as easily as in the European case. The problem is further complicated by the fact that the forward risk-adjusted measure $\bar{Q}$ also depends on the optimal exercise date τ^* .

Example 1. (Continued from section 2.3.1.) Substituting $\sigma(t, T) = \sigma$ into equation (3.38), we have

$$df(t, T) = \sigma^2(T-S)dt + \sigma d\bar{W}(t),$$

where $\bar{W} = W - \int_0^t [\phi(u) + \sigma(S-u)]du$ is the forward risk-adjusted standard Brownian motion. Solving this equation yields

$$f(t, T) = f(0, T) + \sigma^2 t(T-S) + \sigma \bar{W}(t). \quad (3.40)$$

Then the spot rate $r(t) = f(t, t)$ and the discount-bond price are

$$r(t) = f(0, t) + \sigma^2 t(t-S) + \sigma \bar{W}(t), \text{ and} \quad (3.41)$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\{-[\sigma^2 t(T+t-2S)/2 + \sigma \bar{W}(t)](T-t)\}, \quad (3.42)$$

respectively. From theorem 2, the time 0 value of a European option expiring at time t , on a discount bond maturing at time T with strike price K , is

$$C(0) = P(0, t) \mathbb{E}\{\max[P(t, T) - K, 0] | \mathcal{F}_0\}. \quad (3.43)$$

Carrying out the expectation in the equation above yields the same call value as in equation (3.31).

4. Approximating the Term Structure

Under the local expectations hypothesis, we need only approximate the risk-neutralized spot-rate process. For equilibrium models, an extensive search for the market price of risk is needed to match the observed bond prices. An alternative is to assume risk neutrality and match the model with the observed market data. This is well-explained in Hull and White (1991b). We now focus our attention to the arbitrage-free approach.

Since spot-rate processes implied by arbitrage-free models are usually non-Markovian (caused by the path-dependent term in the drift), they cannot be approximated by a computationally simple binomial lattice. However, if we define the path-dependent term as a separate variable (process), then it can be put together with the spot-rate process to form a two-dimensional Markovian diffusion. In this section we explain how to use the lattice approach developed in section 2 to approximate this diffusion.

We can avoid path-dependency by assuming a deterministic volatility structure. Such models with deterministic volatility have been studied by Jamshidian (1987, 1989), Ritchken and Sankarasubramanian (1991a,b).

For practical purposes, assume $\sigma_1(t, T)$, the volatility due to

the i th factor, $i = 1, 2, \dots, n$, can be decomposed as follows

$$\sigma_i(t, T) = \sigma_i(t, t)h_i(t, T). \quad (4.1)$$

The first term $\sigma_i(t, t)$ depends only on time t and the spot rate r .

The second term $h_i(t, T)$ is deterministic and can be expressed as

$$h_i(t, T) = \exp\left[-\int_t^T \kappa_i(s) ds\right], \quad (4.2)$$

where $\kappa_i(t)$ is some integrable function. From equation (27) the spot-rate process is

$$\begin{aligned} r(t) = f(0, t) &+ \sum_{i=1}^n \int_0^t \sigma_i^2(v, v) h_i(v, t) \int_v^t h_i(v, u) du dv \\ &+ \sum_{i=1}^n \int_0^t \sigma_i(v, v) h_i(v, t) d\tilde{W}_i(v). \end{aligned} \quad (4.3)$$

The differential form is

$$\begin{aligned} dr(t) = f'_t(0, t)dt &+ \sum_{i=1}^n \left\{ \int_0^t \sigma_i^2(v, v) h_i(v, t) h_i(v, t) dv \right\} dt \\ &- \sum_{i=1}^n \left\{ \int_0^t \sigma_i^2(v, v) h_i(v, t) \kappa_i(t) \int_v^t h_i(v, u) du dv \right\} dt \\ &+ \sum_{i=1}^n \sigma_i(t, t) h_i(t, t) d\tilde{W}_i(t) \\ &+ \sum_{i=1}^n \left\{ \int_0^t \sigma_i(v, v) [-\kappa_i(t) h_i(v, t)] d\tilde{W}_i(t) \right\} dt. \end{aligned} \quad (4.4)$$

In what follows, we treat one-factor and multi-factor models separately.

4.1. Approximating the One-Factor Model

To obtain a one-factor term structure model, we simply drop the index i in the previous equations. From equation (4.4), we have

$$\begin{aligned} dr(t) = & \{f'_t(0,t) + \int_0^t \sigma^2(v,v)h^2(v,t)dv \\ & + \kappa(t)[f(0,t)-r(t)]\}dt + \sigma(t,t)d\tilde{W}(t). \end{aligned} \quad (4.5)$$

Define two auxiliary processes

$$\xi(t) = \int_0^t \sigma^2(v,v)h^2(v,t)dv, \text{ and} \quad (4.6)$$

$$\theta(t) = f(0,t) + [f'_t(0,t) + \xi(t)]/\kappa(t). \quad (4.7)$$

Then

$$d\xi(t) = [\sigma^2(t,t) - 2\kappa(t)\int_0^t \sigma^2(v,v)h^2(v,t)dv]dt. \quad (4.8)$$

Therefore,

$$dr(t) = \kappa(t)[\theta(t)-r(t)]dt + \sigma(t,t)d\tilde{W}(t), \text{ and} \quad (4.9)$$

$$d\xi(t) = [\sigma^2(t,t) - 2\kappa(t)\xi(t)]dt. \quad (4.10)$$

The corresponding forward rate and discount bond price can be written in terms of the spot rate $r(t)$,

$$\begin{aligned} f(t,T) &= f(0,T) + \int_0^t \sigma(v,T) \int_v^T \sigma(v,u)du dv + \int_0^t \sigma(v,T)d\tilde{W}(v) \\ &= f(0,T) + h(t,T) \left\{ \int_0^t [\sigma(v,t) \left(\int_v^t + \int_t^T \right) \sigma(v,u)du] dv + \int_0^t \sigma(v,t)d\tilde{W}(v) \right\} \end{aligned}$$

$$= f(0,T) + h(t,T)[r(t)-f(0,t)] + \xi(t)h(t,T)\int_t^T h(t,u)du \quad (4.11)$$

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left\{-\int_t^T h(t,s)[r(t) - f(0,t) + \xi(t)\int_t^s h(t,u)du]ds\right\}. \quad (4.12)$$

When the spot volatility $\sigma(t,t)$ depends on the spot rate $r(t)$, the auxiliary variable $\xi(t)$ in the drift term may become path-dependent.⁴ We now demonstrate how to use the ideas developed in section 2 to approximate the path-dependent spot-rate process (4.6). First, we specify the two components of $\sigma(t,T)$ in equation (4.1),

$$\sigma(t,t) = \sigma\sqrt{r(t)}, \text{ and} \quad (4.13)$$

$$h(v,t) = e^{-\kappa(t-s)}, \quad (4.14)$$

where σ and κ are positive constants. Equations (4.9) and (4.10) are then reduced to

$$dr(t) = \kappa[\theta(t)-r(t)]dt + \sigma\sqrt{r(t)}d\tilde{W}, \text{ and} \quad (4.15)$$

$$d\xi(t) = [\sigma^2 r(t) - 2\kappa\xi(t)]dt, \quad (4.16)$$

where

$$\xi(t) = \int_0^t \sigma^2 r(s)e^{-\kappa(t-s)}ds, \text{ and} \quad (4.17)$$

⁴ It is possible to have path-independent, non-Gaussian spot-rate processes within the Heath-Jarrow-Morton framework; however, the volatility structure in these cases may be complicated.

$$\theta(t) = f(0, t) + [f'_t(0, t) + \xi(t)]/\kappa. \quad (4.18)$$

We now apply the lattice approximation of section 2.2.3. In figure 2.3, let

$$X_1 = \frac{2\sqrt{r(t)}}{\sigma}, \text{ and} \quad (4.19)$$

$$X_2 = X_1 - \xi(t). \quad (4.20)$$

Then, X_1 and X_2 are perfectly correlated with drifts

$$m_1 = \frac{\kappa[4\theta(t)/\sigma^2 - x^2(t)] - 1}{2x(t)}, \text{ and} \quad (4.21)$$

$$m_2 = m_1 - [\sigma^2 r(t) - 2\kappa\xi(t)], \quad (4.22)$$

respectively. Rewriting the lattice in figure 2.3 in terms of $r(t) = \sigma^2 X_1^2/2$ and $\xi(t) = X_1 - X_2$ yields the following trinomial lattice:

$$\begin{array}{c}
 \begin{array}{c} q_2 \\ q_1 \\ q_0 \end{array} \\
 \begin{array}{c} (r, \xi) \end{array} \left\{ \begin{array}{l} \nearrow \\ \rightarrow \\ \searrow \end{array} \right. \begin{array}{l} (r + \sigma\sqrt{h} + \sigma^2 h/2, \xi) \\ (r + \theta\sigma\sqrt{h} + \sigma^2 h/2, \xi + \theta\sqrt{h}) \\ (r - \sigma\sqrt{h} + \sigma^2 h/2, \xi) \end{array}
 \end{array} \quad (4.23)$$

where

$$\theta = \text{sign}(\sigma^2 r - 2\kappa\xi) \quad (4.24)$$

$$q_0 = \frac{1}{2} \left\{ 1 - \frac{\kappa[\theta(t)r - \sigma^2/2]\sqrt{h}}{\sigma\sqrt{r}} - (1-\theta)|\sigma^2 r - 2\kappa\xi|\sqrt{h} \right\} \quad (4.25a)$$

$$q_1 = |\sigma^2 r - 2\kappa\xi|\sqrt{h} \quad (4.25b)$$

$$q_2 = \frac{1}{2} \left\{ 1 + \frac{\kappa[\theta(t)-r] - \sigma^2/2}{\sigma\sqrt{r}} \sqrt{h} - (1+\theta) |\sigma^2 r - 2\kappa\xi| \sqrt{h} \right\}, \quad (4.25c)$$

for $\xi(t)$ and $r(t)$ on a compact.

The proposed approximation is polynomial in complexity. Since the process becomes singular when $r = 0$, a technique similar to the one in Essay 1 may be used.

4.2. A Multi-Factor Term Structure Model

We now develop an n -factor term structure model. Assume the following deterministic volatility structure

$$\sigma_i(t, T) = \sigma_i(t) h_i(t, T), \quad (4.26)$$

where

$$h_i(t, T) = \exp \left[- \int_t^T \kappa_i(s) ds \right] \quad (4.27)$$

for some integrable function $\kappa_i(t)$, $i = 1, \dots, n$. From equation (3.24), the forward-rate process follows the equation

$$f(t, T) = f(0, T) + \int_0^t \alpha(v, T) dv + \sum_{i=1}^n \left[\int_0^t \sigma_i(v) h_i(v, T) d\tilde{W}_i(v) \right], \quad (4.28)$$

where the $\tilde{W}_i$'s are the risk-neutralized standard Brownian motions, and

$$\alpha(v, T) = \sum_{i=1}^n [\sigma_i^2(v) h_i(v, T) \int_v^T h_i(v, u) du]. \quad (4.29)$$

Define stochastic integrals

$$\tilde{\delta}_1(t) = \int_0^t \sigma_1(v) h_1(v, t) d\tilde{W}_1(v), \quad (4.30)$$

with dynamics

$$d\tilde{\delta}_1(t) = -\kappa(t)\tilde{\delta}_1(t)dt + \sigma_1(t)d\tilde{W}_1(t). \quad (4.31)$$

The normalized process $\tilde{\Delta}_1(t) = [\sigma_1(0)/\sigma_1(t)]\tilde{\delta}_1(t)$ satisfies

$$d\tilde{\Delta}_1(t) = -[\kappa_1(t) + \sigma'_1(t)/\sigma_1(t)]\tilde{\Delta}_1(t)dt + \sigma_1(0)d\tilde{W}_1(t) \quad (4.32)$$

and can be easily approximated using the quadrinomial model of section 2 with $\rho = 0$ and $m_1 = -[\kappa_1(t) + \sigma'_1(t)/\sigma_1(t)]\tilde{\Delta}_1(t)$.

Rewrite equation (4.28) as

$$f(t, T) = f(0, T) + \int_0^t \alpha(v, T)dv + \sum_{i=1}^n [\tilde{\delta}_i(t)h_i(t, T)]. \quad (4.33)$$

Then the spot-rate process and bond price can also be written in terms of $\tilde{\delta}_i(t)$; that is,

$$r(t) = f(0, t) + \int_0^t \alpha(v, t)dv + \sum_{i=1}^n \tilde{\delta}_i(t), \text{ and} \quad (4.34)$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left\{-\int_0^t \int_t^T \alpha(v, u)du dv - \sum_{i=1}^n \tilde{\delta}_i(t) \int_t^T h_i(t, u)du\right\}. \quad (4.35)$$

Therefore, the processes $\tilde{\delta}_i(t)$, $i = 1, \dots, n$, capture all the uncertainty of the term structure. In other words, the spot rate, the forward rate, and the bond price are all deterministic functions of these two processes. We summarize these results in the following theorem.

Theorem 3. (i) Under the assumptions of this section, the forward rate, spot rate, and discount-bond price are

$$f(t, T) = f(0, T) + \sum_{i=1}^n [a_i(t) + b_i(t)g_i(t, T) + \tilde{\delta}_i(t)]h_i(t, T) \quad (4.36)$$

$$r(t) = f(0, t) + \sum_{i=1}^n [a_i(t) + \tilde{\delta}_i(t)] \quad (4.37)$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left\{-\sum_{i=1}^n [a_i(t) + b_i(t)g_i(t, T)/2 + \tilde{\delta}_i(t)]g_i(t, T)\right\}, \quad (4.38)$$

where

$$a_i(t) = \int_0^t \sigma_i^2(v)h_i(v, t)g_i(v, t)dv \quad (4.39)$$

$$b_i(t) = \int_0^t \sigma_i^2(v)h_i^2(v, t)dv \quad (4.40)$$

$$g_i(t, T) = \int_t^T h_i(t, u)du. \quad (4.41)$$

(ii) The time 0 price of a European option expiring at time t , on a discount bond maturing at time T with strike price K , is

$$C(0, t, T) = P(0, T)\Phi(d_1) - KP(0, t)\Phi(d_2), \quad (4.42)$$

where

$$V_P^2(t, T) = \sum_{i=1}^n b_i(t)g_i^2(t, T) \quad (4.43)$$

$$d_1 = \{\log[P(0, T)/KP(0, t)] + V_P^2(t, T)/2\}/V_P(t, T) \quad (4.44)$$

$$d_2 = d_1 - V_P(t, T). \quad (4.45)$$

Proof. See appendix C.

Theorem 3 can also be established under the forward risk-adjusted martingale measure $\bar{Q}$. To see this, define new stochastic integrals

$$\bar{\delta}_1(t) = \int_0^t \sigma_1(v) h_1(v, t) d\bar{W}_1, \quad (4.46)$$

where $\bar{W}_1$ is the risk-adjusted Wiener process defined in equation (3.34), and the reference bond matures at time S . From equation (3.38),

$$f(t, T) = f(0, T) + \sum_{i=1}^n [b_i(t) \int_S^T h_i(t, u) du + \bar{\delta}_i(t)] h_i(t, T). \quad (4.47)$$

Let $S = t$. Then equations (4.36) - (4.38) can be rewritten as

$$f(t, T) = f(0, T) + \sum_{i=1}^n [b_i(t) g_i(t, T) + \bar{\delta}_i(t)] h_i(t, T) \quad (4.48)$$

$$r(t) = f(0, t) + \sum_{i=1}^n \bar{\delta}_i(t) \quad (4.49)$$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left\{-\sum_{i=1}^n [b_i(t) g_i(t, T)/2 + \bar{\delta}_i(t)] g_i(t, T)\right\}. \quad (4.50)$$

Since $P(t, T)$ is lognormal under $\bar{Q}$, the expectation in equation (3.36) can be easily evaluated for a European call option. The result is the same as in equation (4.42).

5. Concluding Remarks

In this essay we extended some of the simple lattice approximation methods for one-dimensional diffusions to higher dimensions. Lattice approximations are also developed for perfectly correlated diffusions, and the methods can be used in path-dependent models.

We also reviewed the theory of the term structure of interest rates, including the equilibrium approach of Cox, Ingersoll, and Ross (1985) and the arbitrage-free approach of Heath, Jarrow, and Morton (1992). Both methods are theoretically sound but difficult to implement. Yet much can be gained by a better understanding of their relationship. In the simplest case, the equivalence of the equilibrium model (Vasicek [1977]) and the arbitrage-free model (Ho and Lee [1986]) is obvious. There is still more to learn about the forward-rate process implied by a general equilibrium model. Also, better tools are needed to approximate the spot-rate process which is usually non-Markovian in an arbitrage-free model.

We demonstrated how to approximate term structures using the lattice approach developed in this essay. Even in a one-factor, arbitrage-free term structure, path-dependence can dramatically complicate the model. In contrast, multi-factor Gaussian models are conceptually simple, and serve as a useful tool for pricing interest-rate contingent claims. The problem of possible negative interest rates, however, remains to be resolved.

Appendices to Essay 2

A. Proof of Theorem 1

Substituting equation (3.19) into equation (3.1) and taking the logarithm yields

$$\begin{aligned}\log P(t, T) &= -\int_t^T f(0, u) du - \int_t^T \int_0^t \alpha(v, u) dv du - \sum_{i=1}^n \int_t^T \int_0^t \sigma_i(v, u) dW_i(v) du \\ &= -\int_t^T f(0, u) du - \int_0^t \int_t^T \alpha(v, u) du dv - \sum_{i=1}^n \int_0^t \int_t^T \sigma_i(v, u) du dW_i(v).\end{aligned}$$

The last equality comes from the Fubini theorem. Then

$$\begin{aligned}d[\log P(t, T)] &= f(0, t) dt - \int_t^T \alpha(t, u) du dt + \int_0^t \alpha(v, t) dv dt \\ &\quad - \sum_{i=1}^n \int_t^T \sigma_i(t, u) du dW_i(t) + \sum_{i=1}^n \int_0^t \sigma_i(v, t) dW_i(v) dt \\ &= r(t) dt - \int_t^T \alpha(t, u) du dt - \sum_{i=1}^n \int_t^T \sigma_i(t, u) du dW_i(t).\end{aligned}$$

Using Ito's rule we have

$$\frac{dP(t, T)}{P(t, T)} = \mu_p(t, T) dt + \sigma_p(t, T) dW(t), \quad (\text{A.1})$$

where

$$\mu_p(t, T) = r(t) - \int_t^T \alpha(t, u) du + \frac{1}{2} \sum_{i=1}^n \left[\int_t^T \sigma_i(t, u) du \right]^2, \text{ and} \quad (\text{A.2})$$

$$\sigma_p(t, T) = - \sum_{i=1}^n \int_t^T \sigma_i(t, u) du. \quad (\text{A.3})$$

Let $Z(t, T) = P(t, T)/B(t)$ be the relative bond price. Using Ito's rule,

$$\frac{dZ(t, T)}{Z(t, T)} = - \sum_{i=1}^n \int_0^t \sigma_i(v, t) d\tilde{W}_i(v). \quad (\text{A.6})$$

Therefore, the relative price $Z(t, T)$ is a martingale on $(\Omega, \mathcal{F}, \tilde{Q})$. Following the argument of Harrison and Kreps (1979), the price of any contingent claim relative to the money market account $B(t)$ is also a martingale. Thus theorem 1 holds. Q.E.D.

B. Proof of Theorem 2

From equations (3.32) and (3.19),

$$\log F(t, S, T) = - \int_S^T f(0, u) du - \int_S^T \int_0^t \alpha(v, u) dv du - \sum_{i=1}^n \int_S^T \int_0^t \sigma_i(v, u) dW_i(v) du,$$

and

$$d [\log F(t, S, T)] = - \int_S^T \alpha(t, u) du dt - \sum_{i=1}^n \int_S^T \sigma_i(t, u) du dW_i(t).$$

Using Itô's lemma, we have

$$\begin{aligned} \frac{dF(t, S, T)}{F(t, S, T)} &= - \int_S^T \alpha(t, u) du dt + \sum_{i=1}^n \frac{1}{2} \left[\int_S^T \sigma_i(t, u) du \right]^2 dt \\ &\quad - \sum_{i=1}^n \int_S^T \sigma_i(t, u) du dW_i(t). \end{aligned} \quad (\text{B.1})$$

From equation (3.21)

$$\begin{aligned}
& \int_S^T \alpha(t, u) du + \sum_{i=1}^n \phi_i(t) \int_S^T \sigma_i(t, u) \\
&= \sum_{i=1}^n \int_S^T \sigma_i(t, u) \int_t^u \sigma_i(t, v) dv du \\
&= \sum_{i=1}^n \int_S^T \sigma_i(t, u) \left[\int_t^S \sigma_i(t, v) dv + \int_S^u \sigma_i(t, v) dv \right] du \\
&= \sum_{i=1}^n \left\{ \int_S^T \sigma_i(t, u) \int_t^S \sigma_i(t, v) dv du + \frac{1}{2} \left[\int_S^T \sigma_i(t, u) du \right]^2 \right\}.
\end{aligned}$$

Substituting the above equation into equation (B.1) gives

$$\frac{dF(t, S, T)}{F(t, S, T)} = \sum_{i=1}^n \int_S^T \sigma_i(t, u) du \left\{ [\phi_i(t) - \int_t^S \sigma_i(t, v) dv] - dW_i(t) \right\}. \quad (B.2)$$

Simplifying the above equation using equations (3.33) and (3.34), we have

$$\frac{dF(t, S, T)}{F(t, S, T)} = - \sum_{i=1}^n \int_S^T \sigma_i(t, u) du d\bar{W}_i(t). \quad (B.3)$$

Therefore, the forward bond price is a $\bar{Q}$ -martingale.

Comparing equations (3.22) and (3.23) with equations (3.34) and (3.35), we have

$$\bar{W}_i(t) = \tilde{W}_i(t) + \int_0^t \int_v^S \sigma_i(v, u) du dv \text{ and}, \quad (B.4)$$

$$d\bar{Q}/d\tilde{Q} = \exp \left\{ \sum_{i=1}^n \int_0^t \int_v^S [-\sigma_i(v, u)] du d\tilde{W}(v) - \frac{1}{2} \int_0^t \left[\int_v^S \sigma_i(v, u) du \right]^2 dv \right\}. \quad (B.5)$$

From equation (3.25), we have

$$\begin{aligned}
\int_0^t r(u) du - \int_0^t f(0, u) du \\
&= \sum_{i=1}^n \int_0^t \int_0^u \sigma_i(v, u) \int_v^u \sigma_i(v, z) dz dv du + \sum_{i=1}^n \int_0^t \int_0^u \sigma_i(v, u) d\tilde{W}_i(v) du \\
&= \sum_{i=1}^n \int_0^t \int_0^t \sigma_i(v, u) \int_v^u \sigma_i(v, z) dz du dv + \sum_{i=1}^n \int_0^t \int_0^t \sigma_i(v, u) du d\tilde{W}_i(v).
\end{aligned} \tag{B.6}$$

The last equality comes from the Fubini theorem. Set $t = S$; then, from equations (B.5), (B.6) and (3.4), and the fact that

$$\int_v^S \sigma_i(v, u) \int_v^u \sigma_i(v, z) dz du = (1/2) \left[\int_v^S \sigma_i(v, u) du \right]^2, \tag{B.7}$$

we have

$$B(S) = B(t) \exp \left[\int_t^S f(t, u) du \right] d\tilde{Q}/d\tilde{Q}. \tag{B.8}$$

Since $d\tilde{Q}/d\tilde{Q}$ is a $\tilde{Q}$ -martingale, $\tilde{E}(d\tilde{Q}/d\tilde{Q}) = 1$. Finally, substituting equation (B.8) into equation (3.34) yields

$$\begin{aligned}
C(t) &= \tilde{E} \left[C(S) \exp \left[- \int_t^S f(t, u) du \right] d\tilde{Q}/d\tilde{Q} \middle| \mathcal{F}_t \right] \\
&= \tilde{E} \left\{ \left[C(S) \exp \left[- \int_t^S f(t, u) du \right] \right] \middle| \mathcal{F}_t \right\} \tilde{E} [d\tilde{Q}/d\tilde{Q} \middle| \mathcal{F}_t] \\
&= P(t, S) \tilde{E} [C(S) \middle| \mathcal{F}_t].
\end{aligned} \tag{B.9}$$

The second equality in equation (B.9) comes from the Radon-Nikodym theorem (Dothan [1990], theorem 9.2). This completes the proof.

C. Proof of Theorem 3

(i) Let $t \leq \tau \leq T$. Then

$$h_1(t, T) = h_1(t, \tau)h_1(\tau, T) \quad (C.1)$$

$$g_1(t, T) = g_1(t, \tau) + h_1(t, \tau)g_1(\tau, T). \quad (C.2)$$

Applying these equalities in the evaluation of the integrals in equations (4.33) - (4.35) yields equations (4.36) - (4.38).

(ii) From equation (3.30), we have

$$C(0, t, T) = \tilde{E}\{\max[P(t, T) - K, 0]B(t) | \mathcal{F}_0\}, \quad (C.3)$$

where $\tilde{E}$ is the expectation operator under the equivalent martingale measure $\tilde{Q}$. We now define a new equivalent martingale measure $\hat{Q}$ given by

$$d\hat{Q}/d\tilde{Q} = \exp\left\{-\sum_{i=1}^n \left[\int_0^t \tilde{\delta}_1(u)du + \int_0^t a_1(u)du\right]\right\}. \quad (C.4)$$

To see that $d\hat{Q}/d\tilde{Q}$ is a martingale, we apply the Fubini theorem to equations (4.30) and (4.39); then

$$\begin{aligned} \int_0^t \tilde{\delta}_1(u)du &= \int_0^t \int_0^u \sigma_1(v)h_1(v, u)d\tilde{W}_1(v)du \\ &= \int_0^t \sigma_1(v) \int_v^t h_1(v, u)dud\tilde{W}_1(v) \\ &= \int_0^t \sigma_1(v)g_1(v, t)d\tilde{W}_1(v) \end{aligned} \quad (C.5)$$

$$\begin{aligned}
\int_0^t a_1(u) du &= \int_0^t \int_0^u \sigma_1^2(v) h_1(v, u) g_1(v, u) dv du \\
&= \int_0^t \sigma_1^2(v) \int_v^t h_1(v, u) g_1(v, u) du dv \\
&= (1/2) \int_0^t \sigma_1^2(v) g_1^2(v, t) dv.
\end{aligned} \tag{C.6}$$

From the Girsanov theorem, the process

$$\hat{W}_1(s) = \tilde{W}_1(s) + \int_0^s \sigma_1(v) g_1(v, t) dv, \quad (s \leq t) \tag{C.7}$$

is a standard Brownian motion under the new measure $\hat{Q}$, and $d\hat{W}(s) = d\tilde{W}(s) + \sigma_1(s) g_1(s, t) ds$. Let

$$\hat{\delta}_1(t) = \int_0^t \sigma_1(s) h_1(s, t) d\hat{W}_1(s). \tag{C.8}$$

Then

$$\begin{aligned}
\hat{\delta}_1(t) &= \tilde{\delta}_1(t) + \int_0^t \sigma_1^2(s) h_1(s, t) g_1(s, t) ds \\
&= \tilde{\delta}_1(t) + a_1(t), \text{ and}
\end{aligned} \tag{C.9}$$

$$d\hat{Q}/d\tilde{Q} = \exp\left[-\sum_{i=1}^n \int_0^t \hat{\delta}_i(u) du\right]. \tag{C.10}$$

To evaluate equation (C.3), first express $P(t, T)$ in terms of $\hat{\delta}_1(u)$

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left\{-\sum_{i=1}^n [b_i(t) g_i(t, T)/2 + \hat{\delta}_i(t)] g_i(t, T)\right\}. \tag{C.11}$$

Note that

$$\begin{aligned}
B(t) &= \exp\left\{\int_0^t f(0,u)du + \int_0^t \tilde{\delta}_1(u)du\right\} \\
&= \frac{d\hat{Q}/d\tilde{Q}}{P(0,t)}.
\end{aligned} \tag{C.12}$$

Let $\hat{\mathbb{E}}$ be the expectation operator under the new measure $\hat{Q}$. Substituting equation (C.12) into equation (C.3) and applying the Radon-Nikodym theorem yields

$$\begin{aligned}
C(0,t,T) &= \tilde{\mathbb{E}}\{\max[P(t,T)-K,0] \frac{d\hat{Q}/d\tilde{Q}}{P(0,t)} | \mathcal{F}_0\} \\
&= \hat{\mathbb{E}}\left\{\frac{\max[P(t,T)-K,0]}{P(0,t)} | \mathcal{F}_0\right\} \tilde{\mathbb{E}}[d\hat{Q}/d\tilde{Q} | \mathcal{F}_0].
\end{aligned} \tag{C.13}$$

Since $\hat{Q}$ is an equivalent martingale measure, $\tilde{\mathbb{E}}[d\hat{Q}/d\tilde{Q} | \mathcal{F}_0] = 1$. From equation (C.11), $\log P(t,T)$ is normally distributed with variance $V_p^2(t,T)$ under measure $\hat{Q}$. Then

$$\hat{\mathbb{E}}\{\max[P(t,T)-K,0] | \mathcal{F}_0\} = \frac{P(0,T)}{P(0,t)} \Phi(d_1) - K\Phi(d_2). \tag{C.14}$$

Substituting equation (C.14) into equation (C.13) yields equation (4.42). Q.E.D.

ESSAY 3

Optimal Bank Portfolio Choice under Fixed-rate Deposit Insurance

1. Introduction

The current system of fixed-rate deposit insurance in the United States gives insured banks the incentive to take on riskier investments than they otherwise would. To relate the cost of deposit insurance to a bank's investment risk, Merton (1977) shows that deposit insurance grants a put option to the insured bank. Under this model, banks tend to take on extremely risky projects to exploit the put option. As a result, fixed-rate deposit insurance is apt to be underpriced for high-risk-taking banks and overpriced for low-risk-taking banks. Implementation of option models for valuing deposit insurance can be found in Marcus and Shaked (1984) and Ronn and Verma (1986).

In reality, not all banks take extreme risks. Being in business is a privilege, which is reflected in a firm's charter value or growth option. Extreme risk-taking may lead a bank into insolvency, forcing it out of business by regulators. The charter value comes from many sources, such as monopoly rents in issuing deposits, economies of scale, superior information in financial

markets, and reputation. The charter, as well as the deposit insurance, can be viewed as an intangible asset to the insured.

Taking into account the charter value, Marcus (1984) shows that banks either minimize or maximize their risk exposure as a result of the trade-offs between the put option value and the charter value. Under a different setting, Buser, Chen, and Kane (1981) show that the trade-offs reestablish an interior solution to the capital structure decision. They also argue that capital requirements and other regulations serve as additional implicit constraints to discourage extreme risk-taking.

Almost all models of deposit insurance assume that banks' asset risk is exogenously given. With the exception of the discussion in Ritchken et al. (1991), the flexibility for banks to dynamically adjust their investment decisions has been mostly ignored. However, their model allows only a finite number of portfolio revisions between audits.

In this essay, we establish a continuous-trading model to identify how an equity-maximizing bank dynamically responds to flat-rate deposit insurance schemes and how this affects the actuarially fair value of deposit insurance. Since investment decisions are carried out by optimizing the investment portfolio, we model the problem as the optimal control of a diffusion process. Upon obtaining the optimal portfolio, the actuarially fair cost of insurance can be easily calculated.

In this model, we use the traditional dynamic programming

approach (Fleming and Rishel [1975])). The limitation of this is that it often reduces the problem to an intractable partial differential equation (PDE) where analytical solutions are rare. Merton's (1971) application to the optimal investment/consumption problems is among the few cases in which analytical solutions are obtained.¹ Fortunately, in our problem the resulting PDE can be explicitly solved provided that the charter value is constant. We assume a lognormal price to warrant an analytical solution, but general price distributions can be easily built into the model.

The dynamic programming procedure can also be carried out numerically by lattice approximation. This is especially attractive when more realistic assumptions are made. As the bank changes its portfolio risk over time, the most common binomial model is no longer path-independent, and the problem size grows exponentially with the number of partitions. This difficulty is resolved by using a trinomial lattice. The lattice is devised to incorporate the decision variable into the transition probabilities rather than into the step size.

This essay is organized as follows. Section 2 formulates the model and summarizes the results under no portfolio revision. Section 3 solves the optimal portfolio problem under continuous portfolio revision. The value of deposit insurance is derived

¹ An alternative technique would be the martingale representation approach of Cox and Huang (1989) and Pliska (1986).

based on the optimal portfolio decisions. Section 4 presents the trinomial approximation of controlled diffusion process. Section 5 extends the model to more general situations, and section 6 concludes the essay. Proof of main results is in the appendix.

2. The Static Model - No Portfolio Revision

Investment Opportunities: Assume that financial markets are complete. The bank can invest in both riskless bonds (earning rate r) and a portfolio of risky securities that follows a geometric Wiener process

$$dS/S = \mu dt + \sigma dW. \quad (1)$$

Capital and Liability: The bank's initial tangible asset $X(0)$ consists of capital $K(0)$ and the deposit base $D(0)$. We assume, for simplicity, no net external cash inflows into the deposit base, no capital injections, and no dividend payments during the time interval $[0, T]$. Deposits earn the riskless rate r because they are fully insured. Let $L(t)$ be the liability at time t ; then

$$L(t) = L(0)e^{rt}. \quad (2)$$

Investment Decisions: Management decides at time zero to put a fraction q of its tangible assets in risky securities and the remaining in riskless bonds. Without portfolio revisions, q is fixed before the audit. The market value of the tangible assets

at time t is

$$X(t) = qX(0)e^{\mu t - \sigma^2 t/2 + \sigma\sqrt{t}\xi} + (1-q)X(0)e^{rt}, \quad (3)$$

where ξ is the standard normal random variable with density and distribution function $n()$ and $N()$, respectively.

Auditing and Closure Rules: The regulator conducts an audit at time T . Assume market-value accounting is adopted. Thus there is no information asymmetry. If the bank is solvent, i.e., the market value of its tangible assets exceeds its liabilities, it claims the residual $X(T) - L(T)$ and keeps its charter. If the bank is insolvent, the regulator takes over and equityholders receive nothing. Let $G(T)$ represent the charter value of a solvent bank at time T . $G(T)$ is assumed to be a fraction of the liability. Define

$$G(t) = fL(t), \quad 0 < f < 1. \quad (4)$$

Let $V(t; q)$ be the equity value at time t under policy q . Then

$$V(T, q) = \begin{cases} X(T) - L(T) + G(T) & \text{if } X(T) > L(T) \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The equity value at time 0 can be obtained by using standard option pricing techniques,

$$V(0, q) = \begin{cases} qX(0)N(d_1) - [L(0) - G(0) - (1-q)X(0)]N(d_2) & \text{if } (1-q)X(0) < L(0) \\ X(0) - L(0) + G(0) & \text{otherwise,} \end{cases} \quad (6)$$

where

$$d_1 = \frac{\ln(qX(0)/[L(0)-(1-q)X(0)] + \sigma^2 T/2}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}.$$

On behalf of the shareholders, management will maximize the equity value by choosing the optimal fraction q^* such that

$$V(0, q^*) = \max_q \{ V(0, q) \}. \quad (7)$$

This optimization problem can be solved analytically. Solvent and insolvent banks are treated separately. Even though an initially insolvent bank would be an unusual case, it is included to complete the analysis. we summarize these results in Theorems 1 and 2.

Theorem 1. For an insolvent bank without portfolio revisions, $q^* = 1$ is optimal. Consequently, the value of deposit insurance is²

$$I(0) = -X(0)N(-d_1^1) + L(0)N(-d_2^1), \quad (8)$$

where

$$d_1^1 = \frac{\ln[X(0)/L(0)] + \sigma^2 T/2}{\sigma\sqrt{T}}$$

$$d_2^1 = d_1^1 - \sigma\sqrt{T}.$$

² The value of deposit insurance always refers to the actuarially fair cost of deposit insurance.

Theorem 2. For a solvent bank without portfolio revision, the optimal policy is

$$q^* = \begin{cases} 1 & \text{if } f < 1-H(m) \\ 0 & \text{if } f \geq 1-H(m), \end{cases}$$

where

$$m = X(0)/L(0)$$

$$H(m) = mN(-d_1^1)/N(-d_2^1).$$

Consequently, the value of deposit insurance is

$$I(0) = \begin{cases} -X(0)N(-d_1^1) + L(0)N(-d_2^1) & \text{if } f < 1-H(m) \\ 0 & \text{if } f \geq 1-H(m). \end{cases} \quad (9)$$

Theorems 1 and 2 show that without revision opportunities between audits, banks always take extreme positions. Insolvent banks have no choice but to gamble. They win everything if their gamble wins, and shift the losses elsewhere if they lose. For solvent banks, trade-offs exists between preserving the charter ($q = 0$) and exploiting the insurance ($q = 1$). With a small charter value, a solvent bank may be better off by taking the riskiest position so as to maximize the value of the deposit insurance. Only solvent banks with a sufficiently large asset-deposit ratio m or a relatively high charter value will invest in riskless bonds.³

The value of insurance for an insolvent bank, or for a

³ This can be shown from the fact that $H(m)$ is an increasing function of m with $H(0) = 0$ and $H(\infty) = 1$.

solvent bank with $f < 1 - H(m)$, is the same as in Merton (1977), where the charter value is zero. When $f \geq 1 - H(m)$, risk-taking is discouraged and the insurance has no intrinsic value.⁴

3. Continuous Portfolio Revision

In this section, we assume that banks can revise their investment portfolios continuously over time at no cost. Let $X(t)$ be the market value of the assets and $q = q(t, X(t))$ be the fraction of risky assets in the portfolio at time $t \in [0, T]$. Then $X(t)$ follows a diffusion process

$$dX(t) = [q\mu + (1-q)r]X(t)dt + q\sigma X(t)dW(t), \quad X(0) = X_0, \quad (10)$$

where $W(t)$ is a standard Brownian motion. The liability and charter value are given by equations (2) and (4), respectively. For valuation purposes, one can substitute μ with r in equation (11). Let $J(t, X(t))$ be the maximum equity value of the bank at time t . Then

$$J(t, X(t)) = \max_q E_t [J(T, X_T) e^{-r(T-t)}]. \quad (11)$$

It has the boundary condition

⁴ To be precise, when $f = 1 - H(m)$, a bank is indifferent between $q = 0$ (preserving the charter) and $q = 1$ (exploiting the insurance). However, the bank's actual decision on q does affect the value of insurance. This discontinuity in the insurance value is one of the drawbacks of static models.

$$J(T, X(T)) = \begin{cases} X(T) - L(T) + G(T) & \text{if } X(T) \geq L(T) \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

We are interested in the maximum equity value $J(0, X(0))$ for any given $X(0) = X_0$ at time zero and the corresponding optimal policy $q^*(t)$ for all $t \in [0, T]$. If the function $J(t, X(t))$ is twice continuously differentiable, then the dynamic programming equation (11) becomes

$$rJ = \max_q \{J_t + rX(t)J_x + \frac{1}{2}[q\sigma X(t)]^2 J_{xx}\}. \quad (13)$$

The optimal value for q in the above equation is either 0 or 1 since the right hand side is a convex function of q . Unfortunately, the function $J(t, X(t))$ is not twice differentiable due to the discontinuity in boundary condition (12); hence, we can not simply use equation (13) to construct the optimal solution for q . However, it can be shown that extreme solutions remain optimal for equation (11) under boundary condition (12). The results are presented in the following theorem.

Theorem 3. Let τ be the solution of the following equation⁵

$$\frac{G(T)}{L(T)} \left[N(\sigma\sqrt{T-\tau}/2) + \frac{2n(\sigma\sqrt{T-\tau}/2)}{\sigma\sqrt{T-\tau}} \right] = 1. \quad (14)$$

Suppose the asset value at time t is $X(t)$. Under the assumptions of section 2 and continuous portfolio revision, the optimal

⁵ If the solution is negative, simply let $\tau = 0$.

decision $q^*(t)$ and the corresponding equity value $J(t, X(t))$ are as follows:

- (1) If $t \in [\tau, T)$ and $X(t) \geq L(t)$, then $q^*(t) = 0$, and

$$J(t, X(t)) = X(t) - [L(t) - G(t)]. \quad (15)$$

- (2) If $t \in [\tau, T)$ and $X(t) < L(t)$, then $q^*(t) = 1$, and

$$J(t, X(t)) = \frac{G(t)}{L(t)} [X(t)N(\gamma_1) + L(t)N(\gamma_2)] \quad (16)$$

where

$$\gamma_1 = \frac{\ln[X(t)/L(t)] + \sigma^2(T-t)/2}{\sigma\sqrt{T-t}}$$

$$\gamma_2 = \gamma_1 - \sigma\sqrt{T-t}.$$

- (3) If $t \in [0, \tau)$, then $q^*(t) = 1$, and

$$\begin{aligned} J(t, X(t)) = \frac{G(t)}{L(t)} [X(t)N(\gamma_1, -\gamma_3, \rho) + L(t)N(\gamma_2, -\gamma_4, \rho)] \\ + X(t)N(\gamma_3) - [L(t) - G(t)]N(\gamma_4) \end{aligned} \quad (17)$$

where

$$\gamma_3 = \frac{\ln[X(t)/L(t)] + \sigma^2(\tau-t)/2}{\sigma\sqrt{\tau-t}}$$

$$\gamma_4 = \gamma_3 - \sigma\sqrt{\tau-t}$$

$$\rho = \sqrt{\tau-t}/\sqrt{T-t}$$

and $N(x, y, \rho)$ is the standard cumulative bivariate normal distribution with correlation coefficient ρ .

In summary, the optimal policy is⁶

$$q^* = \begin{cases} 0 & \text{if } t \in [\tau, T) \text{ and } X(t) \geq L(t) \\ 1 & \text{if } t \in [0, \tau) \text{ or } X(t) < L(t). \end{cases}$$

Theorem 3 clearly illustrates the trade-offs between preserving the charter and exploiting the deposit insurance. As we will show later, the deposit insurance is essentially a put option on the bank's assets that matures at the time of the audit. The longer the time before an audit, the higher the value of the deposit insurance. Prior to time τ , the deposit insurance is more valuable than the fixed charter value, and shareholders exploit the deposit insurance by choosing $q = 1$. After time τ , since the audit is near, the deposit insurance is less valuable than the charter, and shareholders will do their best to ensure that the market value of the bank's assets remains above the solvency curve $L(t)$ in order to preserve its charter. Figure 1 shows this optimal policy where the riskless rate is set to zero.

The critical time τ is uniquely determined by equation (14) for $0 \leq f \leq 1$. To see this, rewrite equation (14) with $\beta = \sigma\sqrt{T-\tau}/2$:

$$N(\beta) + \frac{e^{-\beta^2/2}}{\sqrt{2\pi}\beta} = 1/f. \quad (18)$$

⁶ Actually, when $t \in [\tau, T)$ and $X(t) > L(t)$, any q is optimal as long as q is set at 0 if and when $X(t)$ hits the solvency curve $L(t)$.

Since the left-hand side of (18) decreases from $+\infty$ to 1 as β goes from 0 to $+\infty$, a positive β is uniquely determined. It can be shown that τ is increasing in σ and decreasing in f . That is, the higher the volatility of the risky asset and the lower the charter, the longer the bank sets $q = 1$. However, τ does not depend on the the riskless interest rate r .

Figure 3.1. Optimal Portfolio Policies

As an example, consider an audit period of one year. Suppose the volatility of the risky assets is $\sigma = 10$ percent annually, and the charter value is $f = 10$ percent of the deposit base. Solving equation (14) yields $\tau = 0.293$. If f drops to 5 percent, τ will increase to 0.834. If there is no charter at all, τ equals T , the audit date.

To obtain the value of the deposit insurance $I(0)$, note that the equity value comes from three sources: namely, the initial capital $K(0)$, the deposit insurance $I(0)$, and the charter value $G(0)$. That is,

$$J(0, X(0)) = K(0) + I(0) + G(0)P\{X(T) \geq L(T)\}, \quad (19)$$

where $P\{X(T) \geq L(T)\}$ is the probability that the bank passes the audit. Following the same argument as in the proof of Theorem 3, we have

$$P\{X(T) \geq L(T)\} = \begin{cases} 1 & \text{if } \tau=0, X(0) \geq L(0) \\ \frac{X(0)}{L(0)} N(\gamma_1) + N(\gamma_2) & \text{if } \tau=0, X(0) < L(0) \\ N(\gamma_4) + \frac{X(0)}{L(0)} N(\gamma_1, -\gamma_3, \rho) + N(\gamma_2, -\gamma_4, \rho) & \text{if } \tau>0, \end{cases}$$

where the γ 's are evaluated at time $t = 0$. Substituting this into equation (19), we have the actuarially fair value of deposit insurance for a bank with continuous revision opportunities

$$I(0) = \begin{cases} 0 & \text{if } \tau=0, X(0) \geq L(0) \\ L(0) - X(0) & \text{if } \tau=0, X(0) < L(0) \\ -X(0)N(-\gamma_3) + L(0)N(-\gamma_4) & \text{if } \tau>0, \end{cases} \quad (20)$$

where γ_3 and γ_4 are evaluated at time $t = 0$.

This insurance value can be viewed as a put option on the bank's assets with maturity τ instead of T . This clearly explains the impact of the charter value and the continuous portfolio revision on the value of deposit insurance. Since $\tau < T$ as long as $f > 0$, the deposit insurance is less valuable in the presence

of charter value. Compared to the static model, the insurance value in equation (20) is continuous in terms of charter value and capital-asset ratio. Even for very highly capitalized banks, as long as $\tau > 0$, the insurance has a positive value.

4. Trinomial Approximation

For general terminal payoff functions other than the one in equation (5), analytical solutions may not always exist, and numerical procedures must be used to solve the optimal portfolio problem. Without portfolio revisions, a simple binomial model can be used to approximate the bank's asset value. However, when the portfolio is revised, the resulting lattice becomes path-dependent.

To see this, partition the audit period $[0, T]$ into n subintervals of equal length $h = T/n$. The asset portfolio may be revised at discrete decision points $t_i = ih$, $i = 0, 1, \dots, n-1$. Let $q(t_i, X(t_i))$ be the revised fraction of risky investments at time t_i if the market value of the bank's assets is $X(t_i)$. Let q be initially set to q_0 . The portfolio is revised at time t_1 by changing q_0 to q_1 at the up state and q_2 at the down state, respectively. The two-period binomial lattice looks like

where

$$u_i = (1-q_i)e^{rh} + q_i \exp(\sigma\sqrt{h})$$

$$d_i = (1-q_i)e^{rh} + q_i \exp(-\sigma\sqrt{h})$$

for $i = 0, 1, 2$.

Clearly, the lattice is path-dependent if $u_0 d_1 \neq d_0 u_2$. To overcome this difficulty, a path-independent lattice is first set up as if there is no portfolio revision. Then, when the portfolio is revised to a new q value at a revision point, one changes only the transition probabilities such that the drift and variance terms match locally. This suggests adding one more degree of freedom to the lattice. Consider the following trinomial lattice when the asset value at time t_i is X_i :

$$\begin{array}{c}
 p_1 \\
 \nearrow \\
 X_i \text{ --- } p_0 \text{ --- } X_i \exp(rh) \\
 \searrow \\
 p_{-1}
 \end{array}
 \begin{array}{l}
 X_i \exp(rh + \sigma\sqrt{h}) \\
 \\
 X_i \exp(rh - \sigma\sqrt{h}).
 \end{array}
 \quad (21a)$$

The transition probabilities are set to

$$p_1 = \frac{q^2}{2} \left(1 - \frac{\sigma\sqrt{h}}{2}\right) \quad (21b)$$

$$p_0 = 1 - q^2 \quad (21c)$$

$$p_{-1} = \frac{q^2}{2} \left(1 + \frac{\sigma\sqrt{h}}{2}\right). \quad (21d)$$

Obviously, $\sum_j p_j = 1$. The first and second local moments are

$$\begin{aligned} \mu_h(X_1, t_1) &= \sum_j p_j [\exp(rh + j\sigma\sqrt{h}) - 1] X_1 / h \\ &= rX_1 + O(\sqrt{h}), \end{aligned} \quad (22)$$

$$\begin{aligned} \sigma_h^2(X_1, t_1) &= \sum_j p_j [X_1 \exp(rh + j\sigma\sqrt{h}) - X_1]^2 / h \\ &= (q\sigma X_1)^2 + O(\sqrt{h}). \end{aligned} \quad (23)$$

As $h \rightarrow 0$, these moments converge to the true mean and variance of the diffusion process $X(t)$ in equation (10). This ensures that the trinomial process converges to the process $X(t)$ in distribution.

To find the optimal policy q^* , a dynamic programming procedure can be applied to the trinomial lattice. At the end-nodes, payoff values are given. Working backward, at any node X_1 , an optimal policy $q_1^*(h)$ and equity value can be easily obtained. Under certain smoothness conditions on the payoff function, as $h \rightarrow 0$, $q_1^*(h)$ will converge to the optimal policy q^* . The optimal policy of Theorem 3 can be easily confirmed using this procedure.

5. A Second Look at the Charter Value

In sections 2 and 3, Marcus's (1984) specification of the charter value is adopted. The bank either retains or loses the full charter value depending on whether or not it is solvent at the audit date. This corresponds to the terminal payoff curve OBCD in figure 2. However, despite its simplicity, this specification is far from realistic.

Figure 3.2. Alternative Payoff Functions

For example, regulators may, for economic or political reasons, choose to inject additional funds into a slightly insolvent bank rather than simply to close it. Thus, the payoff curve OBCD in figure 2 should stretch further to the left. As for

below the liability value just before the audit, it would be to the bank's advantage to inject additional funds in order to preserve the charter. It may do so as long as the charter value exceeds the liability minus asset value. This suggests the payoff curve OAD of a call option with strike price $L(T) - G(T)$. In this case, the charter can be viewed as part of the bank's tangible assets.

However, when a bank is close to insolvency, it may face financial distress or bankruptcy costs, which would decrease the charter value. Usually the charter value depends not only on the size of the deposit base, but also on the soundness of the bank (such as the capital-deposit ratio). When this ratio drops below a certain level, a regulatory tax is likely to be charged (Buser, Chen, and Kane [1981]). Therefore, a more reasonable payoff function would be somewhat like the OEFD curve in figure 3.2. For a highly capitalized bank, the charter value is proportional to the deposit base (the F-D segment). As the bank lowers its capital, the charter-deposit ratio decreases (the E-F segment). If the capital is too low, the charter value is zero (the O-E segment).

After the payoff curve is specified, one can use the trinomial approximation of section 4 to calculate the present value of bank equity and the actuarially fair price of deposit insurance. For demonstration purposes, suppose the payoff curve has the following form:

$$V(T) = \begin{cases} \{1 - \delta \exp[-(\frac{M(T)}{fL(T)})^\alpha]\}M(T) & \text{if } M(T) \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (24)$$

where $M(T) = X(T) - (1-f)L(T)$, $1 \geq \delta \geq 0$, and $\alpha \geq 0$. The parameter δ measures the relative magnitude of bankruptcy costs, and α represents the bank's risk attitude toward bankruptcy.

This payoff function contains many interesting special cases. When $f = 0$, it reduces to the case of Merton (1977). When $\delta = 0$, it reduces to the OAD curve in figure 2 where there is no bankruptcy cost present. When $\alpha = +\infty$ and $\delta = 1$, it reduces to that of Marcus (1984), which corresponds to the OBCD payoff curve in figure 2.

Figure 3.3. Some Specific Payoff Functions

Figure 3.3 shows the payoff function (24) for $\alpha = 1, 2, 3$ and ∞ , while $\delta = 1$. The corresponding optimal policies are shown in figure 4, where the other parameters are $T = 1$, $r = 0$, $X_0 = L_0 = 100$, $\sigma = 0.1$, and $f = 0.05$.

Figure 3.4. Optimal Policies Under the Payoffs in Figure 3.3

All of the optimal policies are similar to the one in Theorem 3. Banks initially choose $q = 1$. After a critical time τ , there is a critical curve $K(t)$. If asset value $X(t)$ is above $K(t)$, $q = 0$ is optimal; otherwise $q = 1$ is optimal. In contrast to Theorem 3, the critical curve $K(t)$ is no longer a straight line. It is interesting to note that the larger the value of α , the lower the

critical curve $K(t)$, because banks are more risk-averse to bankruptcy as α increases.

6. Concluding Remarks

In this essay we develop a stochastic control model to analyze the investment decisions of a bank whose deposits are fully insured under a fixed-rate insurance premium. We show how banks dynamically adjust their investment portfolios in response to market information and how this flexibility affects both investment decisions and the value of deposit insurance. The optimal portfolio problem is solved analytically assuming lognormal asset price and constant charter value. For general payoff patterns, an efficient numerical procedure is presented.

Under continuous portfolio revision we show that, before some critical time τ , the bank always takes the riskiest position regardless of its solvency situation. The bank acts cautiously only between time τ and the audit date T . The deposit insurance remains a put option, but with maturity τ instead of T . This critical time τ depends on the charter value, on the volatility of the risky assets, and on the time between audits.

In the regulators' point of view, to protect the insurance fund, regulation should be designed to give banks an incentive to lower the critical time τ . For example, during periods of

economic turbulence, as the volatility of risky assets is high, more frequent audits may be worthwhile. Since the value of deposit insurance is a decreasing function of the charter value, the model also predicts that lowering the charter value would be counterproductive.

Appendix to Essay 3

Proof of Theorem 1. Since $X(0) < L(0)$, from equation (6) we have

$$\partial V / \partial q = X(0) [N(d_1) - N(d_2)] + \frac{X(0)G(0)[L(0) - X(0)]n(d_1)}{[L(0) - (1-q)X(0)]^2 \sigma \sqrt{T}} > 0.$$

The equity value V is increasing in q ; $q^* = 1$ is optimal. Q.E.D.

Proof of Theorem 2. For a solvent bank, when $q \leq q_{\min} \equiv 1 - L(0)/X(0)$, the riskless bonds alone will be enough to pay off the obligation at time T , and the bank will pass the audit with certainty. In this case, $V(0, q) = X(0) - [L(0) - G(0)]$.

When $q > q_{\min}$, we have $L(0) - (1-q)X(0) > 0$, and

$$\frac{\partial^2 V}{\partial q^2} = \frac{X(0)[L(0) - X(0)]^2 n(d_1)}{[L(0) - (1-q)X(0)]^3 q \sigma \sqrt{T}}$$

$$\{L(0) - (1-q)X(0) + G(0)\left[\frac{2q}{L(0) - X(0)} - \frac{n(d_1)}{\sigma \sqrt{T}}\right]\}$$

A close look at the right-hand side reveals that $\partial^2 V / \partial q^2$ changes sign at most once on $[q_{\min}, 1]$. Let Q be the sign-changing point; then $V(0, q)$ is flat on $[0, q_{\min}]$, concave on $[q_{\min}, Q]$, and convex on $[Q, 1]$. Noting that $\partial V / \partial q = 0$ at $q_{\min}$, $V(0, q)$ has no maximum on the open interval $(0, q_{\min})$. Thus, the optimal policy q^* is either 1 or any value in $[0, q_{\min}]$. From equations (6) and (7)

$$V(0, q^*) = \max \{ V(0, 0), V(0, 1) \}$$

$$= \max \{ X(0) - [L(0) - G(0)], X(0)N(d_1^1) - [L(0) - G(0)]N(d_2^1) \}.$$

This leads to equation (10).

Q.E.D.

To prove Theorem 3, a few lemmas are necessary. Lemma 1 is an adaptation of Fleming and Rishel (1975, p. 124, Theorem V.5.1). Lemma 2 is a classic result (Bhattacharya and Waymire [1990, p. 32]). In the rest of the proof, we use the shorthand notations J and f for $J(t, X(t))$ and $f(s; t, X(t))$, respectively, as long as no confusion arises.

Lemma 1. (Sufficient optimality condition for discounted stochastic dynamic programming) Let $X(t)$ be a diffusion on $[0, T]$

$$dX(t) = \mu(X)dt + \sigma(X)dW(t), \quad X(0) = X_0, \quad (\text{A.1})$$

where μ and σ satisfy the linear growth and the Lipschitz conditions. Let $M(t, X)$ and $J(T, X)$ be continuous and satisfy the polynomial growth condition. Let $J(t, X)$ solve the following dynamic programming equation

$$rJ = \max \{ J_t + \mu(X)J_x + \frac{1}{2}[\sigma(X)]^2 J_{xx} + M(t, X) \} \quad (\text{A.2})$$

with boundary value $J(T, X(T))$. If $J(t, X)$ is twice differentiable for $t \in [0, T]$ and continuous for $t \in [0, T]$, then

$$J(t, X(t)) \geq \mathbb{E}_q \left[\int_0^T M(s, X(s)) ds + J(T, X(T)) \right] \quad (\text{A.3})$$

for any admissible policy q .

Lemma 2. Let $X(t)$ be a Brownian motion with drift μ and diffusion coefficient σ . Let T_z be the first time the process reaches level z conditioned on $X(0) = x$. Then the probability density and distribution functions of T_z are

$$f(t; x, z) = \frac{(z-x)}{\sqrt{2\pi}\sigma t^{3/2}} \exp\left[-\frac{(z-x-\mu t)^2}{2\sigma^2 t}\right], \quad t > 0, \quad (\text{A.4})$$

$$F(t; x, z) = \int_0^t f(s; x, z) ds = N\left(\frac{x-z+\mu t}{\sigma\sqrt{t}}\right) - e^{2(z-x)\mu/\sigma^2} N\left(\frac{x-z-\mu t}{\sigma\sqrt{t}}\right). \quad (\text{A.5})$$

Lemma 3. The functional $J(t, X(t))$ and the policy q^* defined in Theorem 3 is optimal if

(1) when J_{xx} is continuous at $(t, X(t))$, the maximizing q is

$$q^* = \begin{cases} 1 & \text{if } J_{xx} \geq 0 \\ 0 & \text{if } J_{xx} < 0 \end{cases} \quad (\text{A.6})$$

and

$$rJ = J_t + rX(t)J_x + \frac{1}{2}[\sigma X(t)]^2 J_{xx} \quad \text{if } q^* = 1 \quad (\text{A.7})$$

$$rJ = J_t + rX(t)J_x \quad \text{if } q^* = 0 \quad (\text{A.8})$$

(2) when J has a jump at $(t, X(t))$,

$$J_x^+ < J_x^- \quad (\text{A.9a})$$

$$q^* = 0 \quad (\text{A.9b})$$

where $J_x^\pm = J_x(t, X(t) \pm 0)$.

Proof: Part (1) follows immediately from Lemma 1 with $\mu(X) = rX$. To show part (2), note that $J(t, X(t))$ is twice differentiable except when $X(t) = L(t)$ and $t \in (\tau, T]$ where $J_X^- > J_X^+ = 1$ and $J_{XX}^- = J_{XX}^+ = 0$; when $t \in (\tau, T]$, $J(t, X(t))$ is convex for $X(t) \leq L(t)$ and linear for $X(t) \geq L(t)$ (see the proof of Theorem 3). To apply Lemma 1, add a smoothing term P^ε to J such that $J^\varepsilon(t, X(t)) = J(t, X(t)) + P^\varepsilon(t, X(t))$ is twice differentiable, convex for $X(t) \leq L(t)$, and concave for $X(t) \geq L(t)$ for $t \in (\tau, T)$ and for any small number $\varepsilon > 0$. For example, one such P^ε is

$$P^\varepsilon = \begin{cases} -\varepsilon \pi \frac{\Delta J^L}{2} & \text{if } X(t) > L(t) + \varepsilon \pi, t \in (\tau, T) \\ -[X(t) - L(t) + \sin(\frac{X(t) - L(t)}{\varepsilon})\varepsilon] \frac{\Delta J^L}{2} & \text{if } 0 \leq X(t) - L(t) \leq \varepsilon \pi, t \in (\tau, T) \\ 0 & \text{otherwise,} \end{cases}$$

where $\Delta J^L = J_X^+(t, L(t)) - J_X^-(t, L(t))$. Define

$$\phi(P^\varepsilon) = -rP^\varepsilon + P_t^\varepsilon + rX(t)P_X^\varepsilon.$$

Then for any admissible policy q ,

$$\begin{aligned} & -rJ^\varepsilon + J_t^\varepsilon + rX(t)J_X^\varepsilon + \frac{1}{2}[qX(t)\sigma]^2 J_{XX}^\varepsilon - \phi(P^\varepsilon) \\ & \leq -rJ^\varepsilon + J_t^\varepsilon + rX(t)J_X^\varepsilon + \frac{1}{2}[q^*X(t)\sigma]^2 J_{XX}^\varepsilon - \phi(P^\varepsilon) \\ & = 0, \end{aligned}$$

where q^* is the policy in Theorem 3. Therefore, $J^\varepsilon(t, X(t)) = J(t, X(t)) + P^\varepsilon(t, X(t))$ is the solution of the dynamic programming equation

$$rJ^E = \max_q \{J_t^E + rX(t)J_x^E + \frac{1}{2}[qX(t)\sigma]^2 J_{xx}^E - \phi(P^E)\}$$

for $t \in (\tau, T-\delta)$ and any small $\delta > 0$. Applying Lemma 1, we have

$$\begin{aligned} J(t, X(t)) &= J^E(t, X(t)) - P^E(t, X(t)) \\ &\geq \mathbb{E}_q \left[-\int_t^{T-\delta} \phi(P^E) dt + J^E(T-\delta, X(T-\delta)) \right] - P^E(t, X(t)) \\ &= \mathbb{E}_q J(T-\delta, X(T-\delta)) - \mathbb{E}_q \left[\int_t^{T-\delta} \phi(P^E) dt \right] \\ &\quad + \mathbb{E}_q [P^E(T-\delta, X(T-\delta))] - P^E(t, X(t)). \end{aligned}$$

Let $\varepsilon, \delta \rightarrow 0$. The last three terms on the right-hand side all go to zero. Then $J(t, X(t)) \geq \mathbb{E}_q [J(T, X(T))]$ for any q . This implies $J(t, X(t))$ and q^* are optimal. Q.E.D.

Proof of Theorem 3. We need to show that the value function $J(t, X(t))$ and the corresponding policy q^* satisfy the conditions in Lemma 3.

Case 1. Let $t \in (\tau, T]$ and $X(t) \geq L(t)$. When $X(t) > L(t)$, $q^*(t) = 0$ and $J(t, X(t))$ in equation (15) together satisfy the conditions (A.6) and (A.8) in Lemma 3. When $X(t) = L(t)$, as we will show later, J_x is not continuous in $X(t)$. However, from Lemma 3, $q^* = 0$ is optimal if $J_x^+ < J_x^-$. Since $J_x^+ = 1$, we need only to show that $J_x^- > 1$ at $X(t) = L(t)$.

First note that $J(t, X(t))$ is continuous at $X(t) = L(t)$. In fact, as $X(t) \uparrow L(t)$, $\gamma_1 \rightarrow -\sigma\sqrt{T-t}/2$ and $\gamma_2 \rightarrow \sigma\sqrt{T-t}/2$ in equation

(16). Further manipulation yields $J(t, X(t)) \rightarrow G(t) = J(t, X(t))$. Now differentiate $J(t, X(t))$ in equation (16), and let $X(t) \uparrow L(t)$. Then

$$\begin{aligned} J_x(t, X(t)) &= \frac{G(t)}{L(t)} [N(\gamma_1) + X(t)n(\gamma_1)\frac{\partial \gamma_1}{\partial X} + L(t)n(\gamma_2)\frac{\partial \gamma_2}{\partial X}] \\ &\rightarrow \frac{G(t)}{L(t)} [N(\sigma\sqrt{T-t}/2) + \frac{2n(\sigma\sqrt{T-t}/2)}{\sigma\sqrt{T-t}}]. \end{aligned} \quad (\text{A.10})$$

Since $\partial J_x^- / \partial t = -\frac{G(t)}{L(t)} \frac{n(\sigma\sqrt{T-t}/2)}{\sigma(T-t)^{3/2}} < 0$, J_x^- is strictly increasing in t . Noting that $J_x^- = 1$ at $t = \tau$, we have $J_x^- > 1$ for $t \in (\tau, T]$.

Case 2. Let $t \in (\tau, T]$ and $X(t) < L(t)$. Differentiating equation (16), and noting that $X(t)n(\gamma_1) = L(t)n(\gamma_2)$, we have

$$\begin{aligned} J_{xx} &= \frac{\partial^2}{\partial X^2} [X(t)N(\gamma_1) + L(t)N(\gamma_2)] \frac{G(t)}{X(t)} \\ &= \frac{2f(T; t, X(t))G(t)}{\sigma^2 X^2(t)} \geq 0. \end{aligned} \quad (\text{A.11})$$

J_{xx} is obviously continuous. To show that $q^*(t) = 1$ is optimal, we need only to check that condition (A.7) in Lemma 3 is satisfied. Toward this goal, let $Y(t) = \ln[X(t)] - rt$; then

$$dY(t) = -(\sigma^2/2)dt + \sigma dW(t), \quad Y(0) = \ln[X(0)].$$

The first passage times are the same for the geometric Wiener process $X(t)$ to reach $L(s)$ given $X(t)$ at time t and for the Brownian motion $Y(t)$ to reach $\ln[L(s)] - rs$ given $Y(t) = \ln[X(t)]$

- rt at time t . From Lemma 2, the density function of this first passage time is

$$f(s; t, X(t)) = \frac{\ln[L(t)/X(t)]}{\sqrt{2\pi\sigma}(s-t)^{3/2}} \exp\left\{-\frac{[\ln(\frac{L(t)}{X(t)}) + \frac{1}{2}\sigma^2(s-t)]^2}{2\sigma^2(s-t)}\right\} \quad (s > t).$$

It is easy to show that $J(t, X(t)) = \int_t^T G(t)f(s; t, X(t))ds$. Since the density function f satisfies the backward Kolmogorov equation

$$f_t + rX(t)f_x + \frac{1}{2}[X(t)\sigma]^2 f_{xx} = 0, \quad (\text{A.12})$$

condition (A.7) can be easily checked:

$$\begin{aligned} & -rJ + J_t + rX(t)J_x + \frac{1}{2}[X(t)\sigma]^2 J_{xx} \\ &= -rG(t) \int_t^T f ds + [rG(t) \int_t^T f ds + G(t) \int_t^T f_t ds] \\ & \quad + rX(t)G(t) \int_t^T f_x ds + \frac{1}{2}[X(t)\sigma]^2 G(t) \int_t^T f_{xx} ds \\ &= G(t) \int_t^T \{f_t + rX(t)f_x + \frac{1}{2}[X(t)\sigma]^2 f_{xx}\} ds \\ &= 0. \end{aligned}$$

Case 3: Let $t \in [0, \tau]$. We first show that $J(t, X(t))$ in equation (17) is the risk-neutral value of a contingent claim with terminal value $J(\tau, X(\tau))$ at time τ . To see this, let

$$J(t, X(t)) = e^{-r(\tau-t)} \int_{-\infty}^{\infty} J(\tau, X(\tau)) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz, \quad (\text{A.13})$$

where $X(\tau) = X(t)e^{(r-\sigma^2/2)(\tau-t)+\sigma(\tau-t)^{1/2}z}$. Substituting equations (15) and (16) into equation (A.13),

$$J(t, X(t)) = e^{-r(\tau-t)} \int_{-\gamma_4}^{\infty} \{X(\tau) - [L(\tau) - G(\tau)]\} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz \\ + e^{-r(\tau-t)} \int_{-\infty}^{-\gamma_4} \frac{G(\tau)}{L(\tau)} [X(\tau)N(\gamma_1) + L(\tau)N(\gamma_2)] \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz,$$

where γ_1 and γ_2 are evaluated at τ rather than at t . Carrying out the above integrations gives equation (17). From equation (A.13) we have

$$\frac{1}{2}[X(t)\sigma]^2 J_{xx}(t, X(t)) = e^{-r(\tau-t)} \int_{-\infty}^{\infty} \frac{1}{2}[X(\tau)\sigma]^2 J_{xx}(\tau, X(\tau)) \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz.$$

Since $J_{xx}(\tau, X(\tau)) \geq 0$ from cases 1 and 2, $J_{xx}(t, X(t)) \geq 0$.

Now we need only to check condition (A.7) in order to show $q^*(t) = 1$ is optimal. Let $p = p(\tau, y; t, X(t))$ be the density function of the lognormal price $X(\tau)$ conditioned on $X(t)$. Rewrite equation (17) as

$$J(t, X(t)) = e^{-r(\tau-t)} \int_0^{\infty} J(\tau, y) p(\tau, y; t, X(t)) dy. \quad (A.14)$$

Then equation (A.7) can be established by the fact that $p(\tau, y; t, X(t))$ satisfies the backward Kolmogorov equation (A.10).

References

- Amin, K., 1991, "On the Computation of Continuous Time Option Prices Using Discrete Approximations," *Journal of Financial and Quantitative Analysis* 26, 4, 477-495.
- Bhattacharya, R.N., and E.C. Waymire, 1990, *Stochastic Processes with Applications*, John Wiley & Sons, New York.
- Billingsley, P., 1968, *Convergence of Probability Measures*, John Wiley & Sons, New York.
- Black F. and M. Schole, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economics*, 81, 637-659.
- Boyle, P., 1988, "A lattice Framework for Option Pricing with Two State Variables," *Journal of Financial and Quantitative Analysis* 23, 1-12.
- Buser, S., A. Chen, and E. Kane, 1981, "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital," *Journal of Finance*, 36, 51-60.
- Boyle, P., J. Evnine, and S. Gibbs, 1989, "Numerical Evaluation of Multivariate Contingent Claims," *The Review of Financial Studies* 2, 241-251.
- Cox, J. C., J. E. Ingersoll Jr., and S. A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53, 385-407.
- Cox, J. C., S. Ross, and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229-263.
- Cox, J. C. and M. Rubinstein, 1985, *Options Markets*, Prentice-Hall, Englewood Cliffs, N.J.
- Dothan, M., 1990, *Prices in Financial Markets*, Oxford University Press, New York.
- Dynkin, E. B., 1965, *Markov Processes*, Springer-Verlag, Berlin.

- Elliott, R., 1982, *Stochastic Calculus and Applications*, Springer-Verlag, New York.
- Ethier, S. N., and T. G. Kurtz, 1986, *Markov Processes: Characterization and Convergence*, John Wiley & Sons, New York.
- Fleming, W.H., and R.W. Rishel, 1975, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York.
- Ingersoll, J. E., Jr., 1987, *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa, New Jersey.
- Harrison, J. M. and D. M. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381-408.
- Harrison, J. M. and S. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11, 215-260.
- He, H., 1991, "Convergence from Discrete to Continuous Time Financial Models," *Review of Financial Studies*, 3, 523-546.
- Heath, D., R. Jarrow, and A Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica*, 60, forthcoming.
- Heath, D., R. Jarrow, and A Morton, 1990, "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation," *Journal of Financial and Quantitative Analysis* 25, 4, 419-440.
- Ho, T. S. and S. Lee, 1986, "Term Structure Movements and Pricing of Interest Rate Contingent Claims," *Journal of Finance* 41, 1011-1028.
- Hull, J. and A. White, 1990a, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis* 25, 1, 87-99.
- Hull, J. and A. White, 1990b, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25, 1, 87-99.
- Jamshidian, F., 1987, "Pricing of Contingent Claims in the One-Factor Term Structure Model," *Working Paper*, Merrill Lynch Capital Markets.

- Jamshidian, F., 1989, "The Multifactor Gaussian Interest Rate Model and Implementation," *Working Paper*, Merrill Lynch Capital Markets.
- Karatzas, I., 1988, "On the Pricing of American Options," *Applied Mathematics and Optimization*, 17, 37-60.
- Kushner, H. J., 1984, *Approximation and Weak Convergence Methods for Random Processes*, MIT Press, Cambridge, Mass.
- Kushner, H. J., 1990, "Numerical Methods for Stochastic Control Problems in Continuous Time," *SIAM Journal of Control and Optimization*, 28, 999-1048.
- Marcus, A.J., 1984, "Deregulation and Bank Financial Policy," *Journal of Banking and Finance*, 8, 557-565.
- , and I. Shaked, 1984, "The Valuation of FDIC Deposit Insurance Using Option-pricing Estimates," *Journal of Money, Credit, and Banking*, 16, 446-460.
- Merton, R.C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373-413.
- , 1973, "The Theory of Rational Option Pricing," *The Bell Journal of Economics and Management Science* 4, 141-183.
- , 1977, "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees," *Journal of Banking and Finance*, 1, 3-11.
- Nelson, D.B., 1990, "ARCH Models as Diffusion Approximation," *Journal of Econometrics*, 45, 7-38.
- Nelson, D.B. and K. Ramaswamy, 1990, "Simple Binomial Processes as Diffusion Approximations in Financial Models," *Review of Financial Studies*, 3, 393-430.
- Ritchken P. and L. Sankarasubramanian, 1991a, "Stochastic Process of the Yield Curve and the Pricing of Interest Rate Claim," *Working Paper*, Case Western Reserve University.
- Ritchken P. and L. Sankarasubramanian, 1991b, "A Multi-Factor Model of the Quality Option in Treasury Futures Contracts," *Working Paper*, Case Western Reserve University.

- Ritchken, P., J. Thomson, R. DeGennaro, and A. Li, 1991, "The Asset Flexibility Option and the Value of Deposit Insurance," *Proceedings of the Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago, May.
- Ronn, E.I., and A.K. Verma, 1986, "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model," *Journal of Finance*, 41, 871-895.
- Sharpe, W. F., 1978, *Investment*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Shiryayev, A. N., 1978, *Optimal Stopping Rules*, Springer-Verlag, New York.
- Trigeorgis, L., 1991, "A Log-Transformed Binomial Numerical Analysis Method for Valuing Complex Multi-Option Investments," *Journal of Financial and Quantitative Analysis*, 26, 309-326.
- Van Moerbeke, P., 1976, "On Optimal Stopping and Free Boundary Problems," *Archive for Rational Mechanics and Analysis*, 60, 101-148.
- Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure", *Journal of Finance* 32, 177-188.