
Gerald W. Buetow, Jr., CFA
BFRC Services, LLC

James Sochacki
James Madison University

Term-Structure Models Using Binomial Trees



The Research Foundation of AIMR™

Research Foundation Publications

Active Currency Management

by Murali Ramaswami

Common Determinants of Liquidity and Trading

by Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam

Company Performance and Measures of Value Added

by Pamela P. Peterson, CFA, and David R. Peterson

Controlling Misfit Risk in Multiple-Manager Investment Programs

by Jeffery V. Bailey, CFA, and David E. Tierney

Corporate Governance and Firm Performance

by Jonathan M. Karpoff, M. Wayne Marr, Jr., and Morris G. Danielson

Country Risk in Global Financial Management

by Claude B. Erb, CFA, Campbell R. Harvey, and Tadas E. Viskanta

Country, Sector, and Company Factors in Global Equity Portfolios

by Peter J.B. Hopkins and C. Hayes Miller, CFA

Currency Management: Concepts and Practices

by Roger G. Clarke and Mark P. Kritzman, CFA

Earnings: Measurement, Disclosure, and the Impact on Equity Valuation

by D. Eric Hirst and Patrick E. Hopkins

Economic Foundations of Capital Market Returns

by Brian D. Singer, CFA, and Kevin Terhaar, CFA

Emerging Stock Markets: Risk, Return, and Performance

by Christopher B. Barry, John W. Peavy III, CFA, and Mauricio Rodriguez

The Founders of Modern Finance: Their Prize-Winning Concepts and 1990 Nobel Lectures

Franchise Value and the Price/Earnings Ratio
by Martin L. Leibowitz and Stanley Kogelman

Global Asset Management and Performance Attribution

by Denis S. Karnosky and Brian D. Singer, CFA

Interest Rate and Currency Swaps: A Tutorial

by Keith C. Brown, CFA, and Donald J. Smith

Interest Rate Modeling and the Risk Premiums in Interest Rate Swaps

by Robert Brooks, CFA

The International Equity Commitment

by Stephen A. Gorman, CFA

Investment Styles, Market Anomalies, and Global Stock Selection

by Richard O. Michaud

Long-Range Forecasting

by William S. Gray, CFA

Managed Futures and Their Role in Investment Portfolios

by Don M. Chance, CFA

The Modern Role of Bond Covenants

by Ileen B. Malitz

Options and Futures: A Tutorial

by Roger G. Clarke

Risk Management, Derivatives, and Financial Analysis under SFAS No. 133

by Gary L. Gastineau, Donald J. Smith, and Rebecca Todd, CFA

The Role of Monetary Policy in Investment Management

by Gerald R. Jensen, Robert R. Johnson, CFA, and Jeffrey M. Mercer

Sales-Driven Franchise Value

by Martin L. Leibowitz

Time Diversification Revisited

by William Reichenstein, CFA, and Dovalee Dorsett

The Welfare Effects of Soft Dollar Brokerage: Law and Economics

by Stephen M. Horan, CFA, and D. Bruce Johnsen

Term-Structure Models Using Binomial Trees

To obtain the *AIMR Product Catalog*, contact:
AIMR, P.O. Box 3668, Charlottesville, Virginia 22903, U.S.A.
Phone 434-951-5499; Fax 434-951-5262; E-mail info@aimr.org
or
visit AIMR's World Wide Web site at www.aimr.org
to view the AIMR publications list.

The Research Foundation of The Association for Investment Management and Research™, the Research Foundation of AIMR™, and the Research Foundation logo are trademarks owned by the Research Foundation of the Association for Investment Management and Research. CFA®, Chartered Financial Analyst™, AIMR-PPS®, and GIPS® are just a few of the trademarks owned by the Association for Investment Management and Research. To view a list of the Association for Investment Management and Research's trademarks and a Guide for the Use of AIMR's Marks, please visit our Web site at www.aimr.org.

© 2001 The Research Foundation of the Association for Investment Management and Research

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the copyright holder.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional service. If legal advice or other expert assistance is required, the services of a competent professional should be sought.

ISBN 0-943205-53-0

Printed in the United States of America

November 2001

Editorial Staff

Roger S. Mitchell
Book Editor

Lisa S. Medders
Assistant Editor

Jayne M. Dudley
Production Manager

Donna C. Hancock
Production Coordinator

Kelly T. Bruton/Lois A. Carrier
Composition

Mission

The Research Foundation's mission is to identify, fund, and publish research that is relevant to the AIMR Global Body of Knowledge and useful for AIMR member investment practitioners and investors.

Biographies

Gerald W. Buetow, Jr., CFA, is currently president of BFRC Services, LLC, a quantitative financial consulting firm in Charlottesville, Virginia. Previously, he was the Wheat First Professor of Finance and director of the quantitative finance program at James Madison University. He was also lead quantitative researcher for Prudential Investment's Quantitative Investment Management Group, where he managed enhanced index funds and developed structured securities. Mr. Buetow has recently worked on developing financial modeling software for risk management and contingent claim valuation. He has published numerous articles in academic and practitioner journals as well as in various edited works. Mr. Buetow also recently completed a book with Frank J. Fabozzi titled *Valuation of Interest Rate Swaps and Swaptions*. He holds a B.S. in electrical engineering, an M.S. in finance from the University of Texas at Dallas, and a Ph.D. in finance and econometrics from Lehigh University and has extensive training in nuclear engineering. Mr. Buetow can be contacted via his Web site at www.BFRCServices.com.

James Sochacki is an associate professor and director of the Center for Computational Mathematics and Modeling in the mathematics department at James Madison University. He has spent several years working with the quantitative finance program at James Madison University and taught mathematical finance several times. Mr. Sochacki developed many of the applied and computational mathematics courses at James Madison University. He has published articles on scientific and financial computation and modeling. Mr. Sochacki has also directed and consulted on many scientific and financial projects in academia, industry, and business. He holds a Ph.D. in applied mathematics from the University of Wyoming.

Contents

| | |
|---|------|
| Foreword | viii |
| Preface | x |
| Chapter 1. Interest Rate Modeling..... | 1 |
| Chapter 2. Models for the Short Rate | 11 |
| Chapter 3. The Binomial Trees | 34 |
| Chapter 4. Examples of the Binomial Models | 57 |
| Chapter 5. Conclusion..... | 80 |
| Appendix A. Variable Time Step | 81 |
| Appendix B. Continuous Compounding of Interest Rates | 83 |
| Appendix C. Implementation and Calibration Issues for the Binomial Models..... | 84 |
| References | 91 |
| Selected AIMR Publications | 94 |

Foreword

Term-structure models are essential for the valuation of interest rate dependent claims. Although term-structure experts have produced a variety of useful models, they involve complex mathematics, which limits their accessibility to investment practitioners who are not engaged in this area of specialization. Moreover, the original “journal” versions of these models and their subsequent descriptions in text books often abstract from many important details necessary for implementation. These circumstances make it difficult for investors to compare the prices of interest rate dependent claims, to assess the appropriateness of alternative term-structure software products, and to build their own term-structure models.

With this monograph, Gerald W. Buetow, Jr., CFA, and James Sochacki go a long way toward ameliorating this problem. They begin with a concise but hardly superficial overview of interest rate modeling, and they introduce the binomial tree framework. Having thoroughly prepared the reader, they next present the five most important no-arbitrage term-structure models:

- **Ho-Lee Model.** This model was the first no-arbitrage term-structure model. It assumes constant and identical volatility for all spot and forward rates and does not incorporate mean reversion.
- **Hull-White Model.** This model extends the Ho-Lee model to allow for mean reversion.
- **Kalotay-Williams-Fabozzi Model.** This model assumes a lognormal distribution and eliminates the problem of negative short rates, which can occur with the Ho-Lee and Hull-White models.
- **Black-Karasinski Model.** An extension of the Kalotay-Williams-Fabozzi Model, this model controls the growth in the short rate.
- **Black-Derman-Toy Model.** This model permits independent and time-varying spot-rate volatilities.

Buetow and Sochacki explain clearly and precisely the strengths and weaknesses of each model and how these models relate to one another. In addition, they generate binomial trees for each of the models to demonstrate the implications of their different features. Perhaps of most interest to the investment practitioner, they provide all of the detailed information necessary

to implement these models, along with an easy-to-follow numerical example. For the first time in a single source, Buetow and Sochacki remove the opaqueness of the important no-arbitrage term-structure models, rendering them available for use by investment practitioners. The Research Foundation is pleased to present *Term-Structure Models Using Binomial Trees*.

Mark Kritzman, CFA
Research Director
The Research Foundation of the
Association for Investment Management and Research

Preface

This monograph was a by-product of the Fixed-Income Specialization Project (FISP) sponsored by AIMR. Throughout the FISP and the corresponding application development, called the Fixed-Income Specialization Application Project (FISAP), it became evident that most users of interest rate models do not have a very good understanding of them. This is largely due to the lack of available literature in this area. More specifically, existing readings require the reader to have an advanced understanding of mathematics. The demand for a readable source was evident from our discussions with both the participants and other investment professionals. This monograph is our attempt to meet that demand. While mathematics is unavoidable in this area, we have tried to concentrate more on the finance so that the reader will not get lost in the algebra. We have also tried to outline the specifics required if an analyst needs to implement his or her own model. We hope this monograph provides an intelligible and useful guide on term-structure models for readers, thus eliminating the frustration often associated with this area.

In Memory of

Captain William F. Burke, Jr.
Engine Company 21, FDNY

A Leader, Comedian, Friend, and Hero

You paid the supreme sacrifice to ensure the safety of your men and countless others on the morning of September 11, 2001. Men of your courage, leadership, kindness, intelligence, and humor are far too rare in this world. I am forever indebted to you for helping me navigate the labyrinth of adolescence and helping me solve the numerous personal and professional dilemmas that come with age. The current value of that counsel has no metric. My memories of our friendship are more precious than a pure arbitrage. You will be missed but never forgotten.

GWB

1. Interest Rate Modeling

Many models are available for modeling the short rate—that is, the forward rate for the short time period—and many papers and books have been written on these models. Models for the short rate come in various forms: No-arbitrage, equilibrium, Markov, and non-Markov are some examples. These frameworks often involve intimidating mathematics. For the practitioner, the no-arbitrage models seem to be the most widely accepted, but unfortunately, implementing these models can be challenging. This monograph explains many of the financial and mathematical features of no-arbitrage models.

Using any of these models on security or contingent-claim valuation is not as straightforward as it first may appear. What makes implementation even more difficult is the lack of guidance in the extant literature. The original articles assume, when it comes to actual implementation, that the user has a lot of mathematical knowledge of the models. That is, when a user actually sits down to generate the appropriate code to implement the model, the articles are found wanting. Jamishidian (1991), Rebonato (1996), Tuckman (1996), Bjerksund and Stensland (1996), Sundaresan (1997), and Hull (2000) offer some additional insight into implementation. Still, these studies do not explain either the equations or the properties of the models in enough detail to give the user the information needed to implement the models properly. We hope to fill that need with this work.

In this monograph, we develop the models and the mathematical equations and approximations for them, and we give examples of the models. We then present the strengths and weaknesses of the models and indicate how the models can be used to model the short rate.

First, we develop the terminology and notation used in this monograph. Because interest rates are stated at certain times and act over a period of time, we need to formulate how time is measured. Calendar time is measured in years from a certain starting time. That is, if we start measuring interest rates on January 15, 1970, then $t = 0$ on January 15, 1970, $t = 1$ on January 15, 1971, $t = 2$ on January 15, 1972, and so on. For example, the interest rate may be effective at $t = 1$ for a time period of six months. We can divide the calendar time into increments called “time steps.” If the time step is six months, or 0.5 years, then there are two time steps in $t = 1$ and four time steps in $t = 2$. We number the time steps consecutively. The time steps do not have to be of the same length. For example, the first time step could be six months, the second

could be five months, and the third could be four months. In that case, $t = 1$ is in the third time step.

We always choose $t_0 = 0$ as the starting time for our interest rate models. If the calendar times for modeling the interest rate are chosen to be $0 = t_0 < t_1 < t_2 < \dots < t_n$, then the time steps are calculated to be

$$\begin{aligned}\tau_1 &= t_1 - t_0, \\ \tau_2 &= t_2 - t_1, \dots,\end{aligned}$$

and

$$\tau_n = t_n - t_{n-1}.$$

If the time steps are constant, then $\tau_1 = \tau_2 = \dots = \tau_n = \tau$, and we calculate that

$$\begin{aligned}t_1 &= t_0 + \tau, \\ t_2 &= t_1 + \tau = t_0 + 2\tau, \dots,\end{aligned}$$

and

$$t_n = t_{n-1} + \tau = t_0 + n\tau.$$

If the time steps are not constant, then

$$\begin{aligned}t_1 &= t_0 + \tau_1, \\ t_2 &= t_1 + \tau_2 \\ &= t_0 + \tau_1 + \tau_2, \dots,\end{aligned}$$

and

$$\begin{aligned}t_n &= t_{n-1} + \tau_n \\ &= t_0 + \tau_1 + \tau_2 + \dots + \tau_n.\end{aligned}$$

We present the models for a constant time step and treat the nonconstant, or variable, time steps in Appendix A. We use τ to denote a constant time step.

In addition to a formula for measuring time, we need to define our notation for the spot rate. The spot rate is the rate that corresponds to a particular point in time measured from $t = 0$. It is compounded at a certain time step for a fixed number of years. We use R_1 for the spot rate at time $t = 0$ for one time period, R_2 for the spot rate at time $t = 0$ for two time periods, and so on. That is, R_n is the spot rate at time $t = 0$ for n time periods.¹ The spot rates used in this presentation were obtained from zero-coupon bonds. Such information can be

¹For a complete discussion of the spot rate, see Tuckman, or Sundaresan, or Buetow and Fabozzi 2001.

found in any number of sources, including the *Wall Street Journal*. The spot rates up to time $t = t_n = n\tau$ compounded at given time periods are referred to as the term structure of interest rates.

Forward Rates

Suppose we invest A_2 dollars at $t = 0$ for two time periods; at the end of the second time period, $t = t_2 = 2\tau$, we have

$$P_2 = A_2(1 + R_2\tau)^2 \quad (1)$$

or

$$A_2 = \frac{P_2}{(1 + R_2\tau)^2}. \quad (2)$$

That is, A_2 is the present value of P_2 dollars two time periods from now under a spot rate of R_2 compounded every period. (We treat the case of continuous compounding in Appendix B.)

The discount factor is

$$\frac{1}{(1 + R_2\tau)^2}.$$

Now, suppose we want to invest A_2 dollars in one time period and then invest the proceeds, P_1 , for another period. At $t = 0$, we invest A_2 dollars for one time period. At the end of the time period, we have

$$P_1 = A_2(1 + R_1\tau). \quad (3)$$

That is,

$$A_2 = \frac{P_1}{1 + R_1\tau}. \quad (4)$$

Now, we invest these P_1 dollars for another time period. We want to determine the interest rate, r_1 , that lands us at the end of this process with the same P_2 dollars as in the first scenario. We have from Equation 1 and Equation 3 that

$$\begin{aligned} P_1(1 + r_1\tau) &= P_2 \\ &= A_2(1 + R_2\tau)^2. \end{aligned} \quad (5)$$

Using Equation 2 and Equation 4 with Equation 5, we have

$$\begin{aligned}
 A_2 &= \frac{P_2}{(1 + R_2\tau)^2} \\
 &= \frac{P_1}{1 + R_1\tau} \\
 &= \frac{P_2/(1 + r_1\tau)}{1 + R_1\tau}.
 \end{aligned} \tag{6}$$

The interest rate r_1 that gives us the same present value under both scenarios (i.e., makes Equations 5 and 6 true) is called “the forward rate for the short time period” or “the short rate at the first time period.” The forward rate at $t = 0$ is r_0 , and $r_0 = R_1$. In particular, r_n is the short rate n time periods in the future.

Under the usual market assumptions, if

$$\frac{P_2}{(1 + r_1\tau)} > A_2(1 + R_1\tau)$$

or

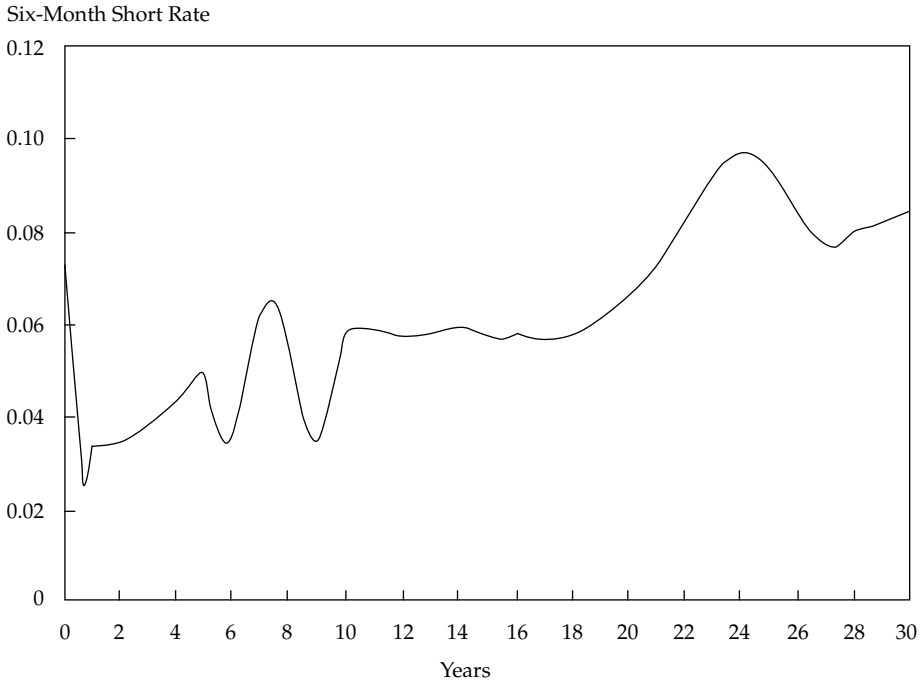
$$\frac{P_2}{(1 + r_1\tau)} < A_2(1 + R_1\tau),$$

then an arbitrage opportunity exists. That is, a risk-free profit is available without any initial investment. Because the models we are considering are no-arbitrage models, Equation 6 must be satisfied. In particular, from Equation 6, we must have

$$\begin{aligned}
 \frac{1}{(1 + R_2\tau)^2} &= \frac{1/(1 + r_1\tau)}{1 + R_1\tau} \\
 &= \frac{1/(1 + r_1\tau)}{1 + r_0\tau}.
 \end{aligned} \tag{7}$$

Because the short rate is the forward rate at a given t for a short time period, we can denote it as $r(t)$. That is, $r(t)$ is the short (forward) rate at time t for a short time period, τ . This short time period for the short rate is usually fixed over calendar time. In the models we consider, the short rates are usually for a fixed six-month period.

Financial analysts are interested in determining possible short rates at future points in time. For example, they are interested in knowing what the six-month short rate will be seven years from now. Suppose the graph in **Figure 1** is a plot of the six-month (fixed-time-period) short rate for a calendar

Figure 1. Six-Month Short Rate (Forward Rate) for 30 Years

time of 30 years. The graph indicates that the six-month short rate at Year 7 is 6 percent, which we would write as $r(7) = 0.06$.²

The graph in Figure 1 shows that the short rate can have significant up and down movement. This movement is called “the volatility of the short rate.” Figure 1 also shows that the volatility is a function of the calendar time, t . Therefore, a model of the interest rate must be able to incorporate volatility. The models we present incorporate the volatility of the short rate, which we call “the local volatility” to distinguish it from yield volatility.

The no-arbitrage models we discuss attempt to answer the following question: Can a model be developed that incorporates the parameters affecting the short rate that will produce graphs consistent with the actual short rate? An elementary knowledge of calculus and probability is all that is necessary to learn and understand the processes that are commonly used to model the short rate within the no-arbitrage framework. The model must make intuitive sense.

²For a semiannual model, when $t = 7$ years and $\tau = 0.5$ years, there are 14 periods.

The models we present are some of the most commonly used models in the finance literature. Our objective is for readers to understand the models so that they can implement them properly. Once analysts understand the appropriate use of the models, they can apply them to contingent-claim valuation, as in Buetow and Fabozzi (2001), or to risk-metric computation, as in Buetow and Johnson (2000).

Introduction to Binomial Models

Consider the following simple example of a volatile short rate for a two-period bond that pays \$100 at the end of the second time period. Because many unpredictable events can change market conditions (including short rates and spot rates), the spot rates at time $t = \tau$ will be different from the spot rates at time $t = 0$. Suppose we believe that at time $t = \tau$, the short rate for the next time period can have only one of two possible values, r_u or r_d , with $r_u > r_d$. If there is a probability q that the short rate will be r_u , then there is a probability $1 - q$ that the short rate will be r_d (because the total of the probabilities must be 1).

With these conditions in mind, from Equation 2, the present value of the bond is

$$A_2 = \frac{\$100}{(1 + R_2\tau)^2},$$

and from Equation 4, the present value at time period τ is either

$$p_u = \frac{\$100}{1 + r_u\tau}$$

or

$$p_d = \frac{\$100}{1 + r_d\tau}.$$

Therefore, the present value at $t = 0$ is either

$$\frac{p_u}{1 + r_0\tau}$$

or

$$\frac{p_d}{1 + r_0\tau}.$$

Thus, within the no-arbitrage framework, either

$$\begin{aligned}
A_2 &= \frac{\$100}{(1 + R_2\tau)^2} \\
&= \frac{p_u}{1 + r_0\tau} \\
&= \frac{\$100/(1 + r_u\tau)}{1 + r_0\tau}
\end{aligned} \tag{8a}$$

or

$$\begin{aligned}
A_2 &= \frac{\$100}{(1 + R_2\tau)^2} \\
&= \frac{p_d}{1 + r_0\tau} \\
&= \frac{\$100/(1 + r_d\tau)}{1 + r_0\tau}.
\end{aligned} \tag{8b}$$

Equations 8a and 8b are similar to Equation 7. A_2 is now set equal to the expected value of Equations 8a and 8b, which gives

$$\begin{aligned}
A_2 &= \frac{\$100}{(1 + R_2\tau)^2} \\
&= q \frac{p_u}{1 + r_0\tau} + (1 - q) \frac{p_d}{1 + r_0\tau} \\
&= q \frac{\$100/(1 + r_u\tau)}{1 + r_0\tau} + (1 - q) \frac{\$100/(1 + r_d\tau)}{1 + r_0\tau}
\end{aligned}$$

or, after dividing by \$100 (the value of the bond),

$$\frac{1}{(1 + R_2\tau)^2} = q \frac{1/(1 + r_u\tau)}{1 + r_0\tau} + (1 - q) \frac{1/(1 + r_d\tau)}{1 + r_0\tau}.$$

Clearing fractions leads to

$$\frac{1}{(1 + R_2\tau)^2} (1 + r_0\tau)(1 + r_d\tau)(1 + r_u\tau) = q(1 + r_d\tau) + (1 - q)(1 + r_u\tau).$$

Finally, subtracting the right-hand side from both sides produces

$$\frac{1}{(1 + R_2^\tau)^2} (1 + r_0^\tau)(1 + r_d^\tau)(1 + r_u^\tau) - q(1 + r_d^\tau) - (1 - q)(1 + r_u^\tau) = 0 \quad (9a)$$

or

$$A_2(1 + r_0^\tau)(1 + r_d^\tau)(1 + r_u^\tau) - q(1 + r_d^\tau) - (1 - q)(1 + r_u^\tau) = 0. \quad (9b)$$

Equation 9b is the fundamental equation used in no-arbitrage binomial models to determine the possible short rates. To use Equation 9b, one has to model r_u and r_d .

In the following chapters, we discuss five models that give formulas for r_u and r_d . The models are the Ho–Lee (HL; Ho and Lee 1986), Hull–White (HW; Hull and White 1994), Kalotay–Williams–Fabozzi (KWF;³ Kalotay, Williams, and Fabozzi 1993), Black–Karasinski (BK; Black and Karasinski 1991), and Black–Derman–Toy (BDT; Black, Derman, and Toy 1990). These models build on Equation 9b. Historically, the HL model was introduced first, then the BDT model, then the BK model, then the KWF model, and then the HW model. We, however, present the models in such a way that each one arises mathematically from the previous model. Thus, for example, a computer program designed to implement the HL model could easily be adapted for the other models.

In Chapter 3, we explain binomial short-rate models and their corresponding price trees. We cover many of the commonly used models for several reasons. First, each model has different properties, and these properties must be understood before the model is applied for valuation purposes. Second, the models require different input parameters. These inputs must also be properly understood before the model is implemented. Third, all the models have limitations that need to be appreciated prior to their use. Finally, interpreting outputs (values or risk metrics) from the model is impossible unless the underlying model and its corresponding characteristics are understood.⁴

Our goal for this monograph is to help readers gain an appreciation for the complexity and power of binomial models. But keep in mind that, although binomial models require far less computing than higher-order tree models, such as trinomial and tetranomial models, these higher-order models do have some advantages, which the reader should also appreciate after completing this monograph. Fortunately, despite the interest rate lattice used in the higher-order models, the mechanics are similar to the mechanics of the

³The actual KWF model is a slight variation of the model we present. Other names are also associated with this model.

⁴Many other reasons exist, but they do not need to be listed here.

binomial models, so the concepts developed here are easily extrapolated to the other lattice structures.

What we develop in this monograph can also be straightforwardly modified to work for trinomial and tetranomial models. A trinomial model would have three possible values for the short rate r_1 —say, r_u , r_m , and r_d —with the respective probabilities q_u , q_m , and q_d (where $q_u + q_m + q_d = 1$). Equation 9b would become

$$A_2(1 + r_0\tau)(1 + r_d\tau)(1 + r_m\tau)(1 + r_u\tau) - q_u(1 + r_d\tau)(1 + r_m\tau) - q_m(1 + r_d\tau)(1 + r_u\tau) - q_d(1 + r_m\tau)(1 + r_u\tau) = 0. \quad (10)$$

We briefly discuss trinomial models at the end of Chapters 3 and 4.

The models we present are the basic models. Advanced readers may suggest that the more sophisticated models are more useful. Undoubtedly, they are. But the more complicated models cannot be properly understood without an understanding of the fundamental models. Moreover, despite the refinements of the sophisticated models, the models we discuss are still widely used by practitioners.

Uses of Term-Structure Models

At this point, the reader may think that binomial models for the short rate are severely limited, but they are not. Binomial models exhibit many of the desirable properties of the short rate. Additionally, they are commonly used as valuation models for both bonds and fixed-income contingent claims.⁵

The concepts presented in this monograph have far-reaching applications throughout the investment arena. Almost every interest rate product in existence can be analyzed by using the models presented, so to reach an accurate conclusion about an investment, investment professionals must understand these models. Securities that particularly lend themselves to analysis by way of lattice models are those with option characteristics, such as swap options, caps, floors, callable bonds, putable bonds, bonds with sinking-fund provisions, bond options, options on bond futures, capped floaters, and credit options. Even securities without embedded options, however, can be analyzed within this framework as part of credit-spread or scenario analysis. The applicability of the models within the fixed-income area is almost endless.

Another important application of these models for the investment professional is the measurement of interest rate sensitivity. Specifically, effective duration, effective convexity, partial duration, key-rate duration, option-

⁵Two good sources on how to use short-rate binomial models and their corresponding price trees are Fabozzi (2000) and Buetow and Fabozzi.

adjusted spread, slope duration, and curvature duration are easily evaluated with these models. Before the proper interpretation of the model's output, however, the model generating the metrics must be understood because each model generates different values. We cannot overemphasize this point. An investment professional must be very comfortable with the characteristics of these models before interpreting any of the valuation or sensitivity results.

2. Models for the Short Rate

Short-term interest rates change through time, so we model the short-term rate using a mathematical process. Specifically, we use differential equations or differential processes to capture the dynamics of the short-term interest rate. The model must also incorporate empirical and intuitive characteristics of interest rates, which the models presented in this chapter attempt to do.

The models we consider have the general form

$$df[r(t)] = \{\theta(t) + \rho(t)g[r(t)]\} dt + \sigma[r(t), t]dz, \quad (11)$$

where f and g are suitably chosen functions, θ (which we will show is the drift of the short rate) is determined by the market, and ρ (the tendency to an equilibrium short rate) can be chosen by the user of the model or dictated by the market. The term σ is the local volatility of the short rate. In Equation 11, $dz = \varepsilon\sqrt{dt}$, where $\varepsilon \sim N(0,1)$ and $N(0,1)$ is the normal distribution with a mean of 0 and a standard deviation of 1.⁶ This component is incorporated in the differential equation to model the randomness (volatility) of interest rates.⁷

Differential equations with this type of term are called stochastic differential equations (SDEs). The models based on Equation 11 are known as “one-factor models” because they model only the short rate (one factor).

The right-hand side of Equation 11 is the sum of two components. The first component is contained in the braces and is multiplied by the term dt . This component is the expected or average change in rates over a short time period, dt . This component is where certain characteristics of interest rates, such as mean reversion (MR), are incorporated. Interest rates are expected to change by what this term suggests. The second component models the randomness (or risk) of interest rates. This component is a product of the terms $\sigma[r(t), t]$ and dz . This component dictates the distribution characteristics of interest rates. Depending on the model, interest rates are either normally or lognormally distributed. As explained, on average, this component is zero, which is why the first component, the expected value, needs to be thoroughly understood.

⁶The variable z follows a Wiener process. This means $dz = \varepsilon\sqrt{dt}$ and that the values of Δz for any two different time intervals Δt are independent.

⁷Throughout the manuscript, we will take liberties with terminology, but this will not detract from the financial implications of the development.

We will not study the technical issues of SDEs in this monograph, but we will use Equation 11 to develop five one-factor models for the short rate. To study Equation 11 in full detail, one needs a good background in stochastic calculus.⁸

Because analytic (or closed-form solutions) to Equation 11 are not feasible even when stochastic calculus is used, numerical approximations to Equation 11 that incorporate the stochastic term $\sigma[r(t), t]dz$ are usually applied. Numerical calculations can be done only at discrete calendar times; therefore, calendar times at which the short rate will be calculated must be chosen. In this monograph, the chosen times are denoted as indicated in Chapter 1, and the times are usually chosen so that the time steps will match the time period for the spot rates of the term structure. This approach makes it easier to incorporate the no-arbitrage condition (bond prices from the model must match the bond prices from the term structure of spot rates). We later explain, however, that some models do not allow the time step to match the time periods for the spot rates. In those cases, the analyst has to use interpolation to match the model time steps with the time periods of the term structure.

We let $r(t_k)$ denote the exact solution to Equation 11 at time $t = t_k$. Because we cannot (in general) generate an analytic form for $r(t_k)$, we develop a numerical (or discrete) process that will approximate $r(t_k)$. We denote this approximation by r_k . That is, our numerical algorithm will generate a value r_k that approximates $r(t_k)$. We write $r_k \approx r(t_k)$.

We now use calculus to develop a numerical approximation for Equation 11. We have that

$$dr \approx \Delta r = r(t_{k+1}) - r(t_k) \approx r_{k+1} - r_k \quad (12a)$$

when

$$dt \approx \Delta t = t_{k+1} - t_k \quad (12b)$$

is close to zero. In many cases, Δr is a fairly good approximation of dr , even if dt is different from zero. (We will use r_k both as $r(t_k)$ and as the approximation of $r(t_k)$. The context will make clear whether we are using the exact value or the approximation.) We will use τ for Δt . If the time steps vary, τ will also depend on k . We use Equations 12 on Equation 11 to develop our numerical approximation for Equation 11.

⁸Determining solutions to Equation 11 is generally difficult. Stochastic calculus must be used to obtain closed-form solutions. A good textbook on the mathematical theory behind Equation 11 is Neftci (1996). For example, to focus attention on the first term of the SDEs (Equation 11), we state that we are taking the expectation of the SDE. We are not really applying the expectations operator to both sides of the equation. We are simply eliminating the second term of the SDE to analyze the first component of the equation. If we did apply the expectations operator, the only difference to our analysis would be a scaling factor; the dynamics do not change. Applying the operator unnecessarily complicates the analysis.

As stated earlier, because the expected value of the random term is zero, we analyze the properties of Equation 11 with only the first component included. In other words, we set the second component equal to its expected value of zero and study the resulting equation. Most of the properties of this resulting equation (without the stochastic term) will apply to Equation 11 (with the stochastic term). A basic understanding of calculus is all that is needed to study this equation.

Inputs into the model must satisfy the mathematical properties of both the SDEs and their approximations. We assume the time step is constant and make adjustments when it is not. The five models are now developed.

Ho-Lee Model

In the HL model or process $f(r) = r$, $g(r) = 0$, and $\rho = 0$ in Equation 11. The HL process is, therefore, given by

$$dr = \theta dt + \sigma dz. \quad (13)$$

Because $dz = \varepsilon \sqrt{dt}$ is normally distributed, the HL process models the short rate as normal. The solution to Equation 13, assuming $r(0) = r_0$ is given by

$$r(t) = r_0 + \int_0^t \theta ds + \int_0^t \sigma dz, \quad (14a)$$

where the integral involving σ is a stochastic integral. If θ is constant, the solution can be expressed as

$$r(t) = r_0 + \theta t + \int_0^t \sigma dz. \quad (14b)$$

Equation 14b shows that the HL process models an interest rate that can change proportionally with time t through the constant of proportionality θ and a random disturbance determined by σ . That is, the larger the magnitude of θ , the larger the average change in the short rate over time. This effect is why θ is called the drift in the short rate. Also, the smaller θ is, the larger the influence of the random disturbance is. A shortcoming of the HL process is that the short rate can be negative. Hull shows that θ is related to the slope of the term structure.

A numerical approximation for Equation 13 can be obtained by using Equations 12a and 12b. Letting $t_k = k\tau$ and $r_k \approx r(k\tau)$ gives

$$r_{k+1} - r_k = \theta_k \tau + \sigma_k \Delta z_k$$

as an approximation for Equation 13 or

$$r_{k+1} = r_k + \theta_k \tau + \sigma_k \Delta z_k, \quad (15)$$

where Δz_k is a numerical (discrete) approximation of dz . Because $dz = \varepsilon \sqrt{dt}$, we can further approximate Equation 15 by

$$r_{k+1} = r_k + \theta_k \tau + \sigma_k \varepsilon_k \sqrt{\tau}, \quad (16)$$

where ε_k is a random number generated from a normal distribution $N(0,1)$.

Equation 16 is the form of the expression that is used for r_{k+1} to build the HL binomial tree. In the HL process, Equation 16 is used for the formula for r_u and r_d in Equation 9. We address the issues of this approximation for the binomial tree in the next chapter.

We first consider the expected value solution to Equation 16 when θ is constant. Equation 16 under these requirements is

$$r_{k+1} = r_k + \tau \theta, \quad (17a)$$

and the solution is given by

$$r_k = c + k\delta, \quad (17b)$$

where c and δ are constants. In particular, $c = r_0 = R_1$ and $\delta = \theta\tau$. The last equation shows that the mean short rate in the HL process increases or decreases at a constant rate θ over time, depending on the sign of θ . As a matter of fact, Equation 17b shows that the short rate grows without bound if $\theta > 0$ and decreases without bound (i.e., becomes extremely negative) if $\theta < 0$. The binomial tree examples considered in Chapter 4 illustrate this effect.

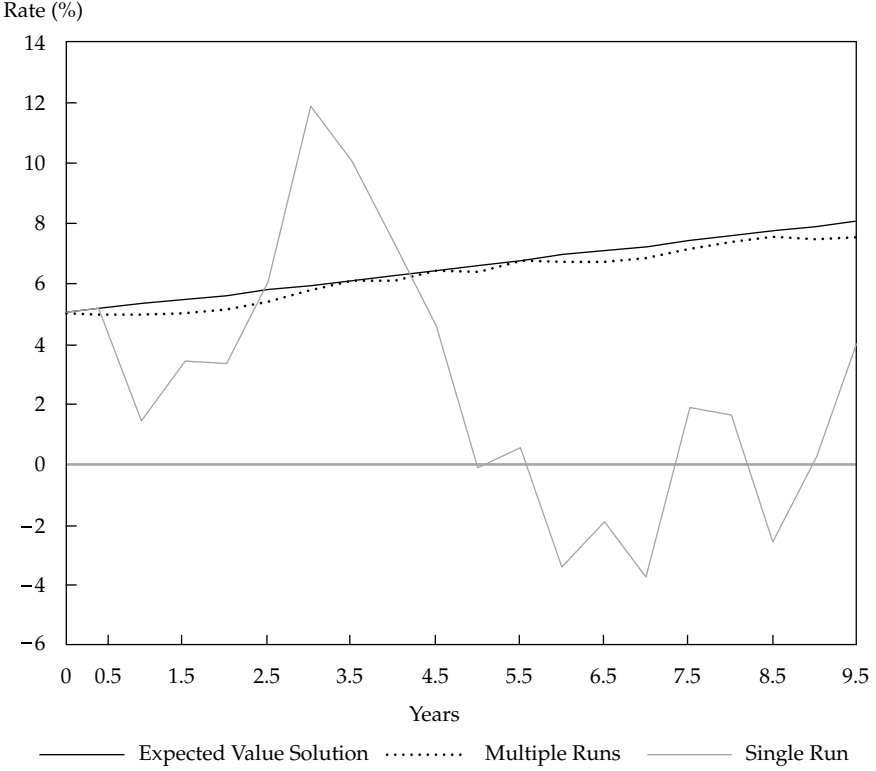
For now, look at solutions to Equation 13 and Equation 16. We use a random number generator to model the normal distribution. **Figure 2** depicts a Ho–Lee example with $\sigma = 10$ percent averaged over 1,000 simulations (the multiple runs line) and the corresponding expected value solution from Equations 16 and 17b with $c = r_0 = R_1 = 0.05$. The slope of the curve for the expected value solution is the drift, θ . As expected, the expected value solution curve and the curve for the average over 1,000 simulations have the same general trend. Clearly, the single random run is highly volatile. For the numerical approximations using Equation 16, we used $\tau = 0.5$. (Unless specified otherwise, we use $\tau = 0.5$ in all our numerical approximations.)

Hull–White Model

In the HW model or process, $f(r) = r$, $g(r) = r$, and $\rho = -\phi$. Therefore, the stochastic process for the HW model for the short rate is

$$dr = (\theta - \phi r)dt + \sigma dz. \quad (18)$$

The short-rate process in the HW models is seen to be normal as in the HL process. We consider the case where the parameters θ and ϕ are constant over time. Note that if $\phi = 0$, the HW process reduces to the HL process. (The HW process will be similar to the HL process, if ϕ is close to zero.) The introduction

Figure 2. Example Ho-Lee Solutions

of ϕ in the HW model is an attempt to incorporate mean reversion and to correct for the uncontrolled growth (or decline) in the HL model; ϕ is the mean reversion rate. Note that, despite this term, the HW model can still result in negative interest rates.

The expected value of Equation 18 leads to the ordinary differential equation

$$dr = (\theta - \phi r)dt, \quad (19)$$

whose solution is given by

$$r(t) = \frac{\theta}{\phi} + ce^{-\phi t}, \quad (20)$$

where

$$c = r_0 - \frac{\theta}{\phi}. \quad (21)$$

If $\phi > 0$, we see from Equation 20 that

$$\begin{aligned}\lim_{t \rightarrow \infty} r(t) &= \frac{\theta}{\phi} \\ &= \mu.\end{aligned}$$

Therefore, for positive mean reversion ($\phi > 0$), the HW process will converge to the short rate, $\mu = \theta/\phi$. Thus, the term μ is called the target or long-term mean rate. For negative mean reversion ($\phi < 0$) the short rate grows exponentially over time.

Equation 21 indicates that c in the HW process gives the distance of the initial short rate from the target short rate. Therefore, the closer the initial short rate is to the target rate, the smaller c will be and the smaller changes in $r(t)$ will be through time. In particular, Equation 21 shows that if $r(0) = \mu$, then $c = 0$, and from Equation 20 that $r(t) = \mu$ for all t . That is, under this scenario, the short rate is always the target rate.

Consider the following examples that give qualitative information for c and μ . Factoring ϕ in Equation 18 leads to

$$dr = \phi(\mu - r)dt + \sigma dz,$$

the expected value of which leads to

$$dr = \phi(\mu - r)dt.$$

We see that if $r > \mu$, then dr is negative and r will decrease, and if $r < \mu$, then dr is positive and r will increase. That is, r will approach the target rate μ . The larger ϕ is, the faster this approach is to the target rate μ . This is why ϕ is called the mean reversion or mean reversion rate. It regulates how fast the target rate is reached.

Since the target rate μ is equal to θ/ϕ , we can solve for the drift, θ , or the mean reversion, ϕ . That is,

$$\theta = \mu\phi \tag{22}$$

or

$$\phi = \frac{\theta}{\mu}. \tag{23}$$

It is seen from Equation 22 and Equation 23 that there is a strong relationship between the drift and mean reversion that can be used to reach any desired target rate. How large the mean reversion should be is an important financial question. Equations 22 and 23 can be used to set target

rates. Equations 20 and 21 allow one to determine how long it takes to reach the target rate. We consider examples of these properties below.

Using Equations 12a and 12b to approximate Equation 18 gives us

$$r_{k+1} = r_k + (\theta_k - \phi_k r_k)\tau + \sigma_k \varepsilon_k \sqrt{\tau}. \quad (24)$$

If θ and ϕ are constant, then the solution to the expected value of Equation 24 has the form

$$r_k = \alpha \beta^k + \gamma.$$

To determine α , β , and γ , we substitute this form for r_k into Equation 24 without the stochastic component and obtain that $\beta = (1 - \phi\tau)$, $\gamma = \theta/\phi = \mu$, and $\alpha = r_0 - \mu$. Therefore,

$$r_k = \alpha(1 - \phi\tau)^k + \frac{\theta}{\phi}. \quad (25)$$

Note that if $0 < \phi\tau < 2$, then $-1 < 1 - \phi\tau < 1$ and

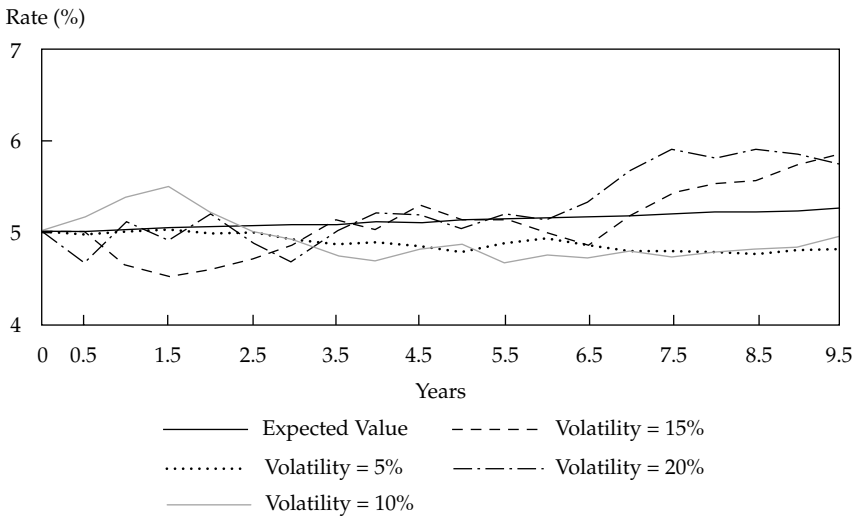
$$\begin{aligned} \lim_{k \rightarrow \infty} r_k &= \frac{\theta}{\phi} \\ &= \mu, \end{aligned}$$

which is the same result that is obtained from Equation 20 for the HW SDE under the same scenario. As will be seen later, the condition $0 < \phi\tau < 2$ can be maintained in modeling the short rate. Equation 24 is the formula that will be used for r_u and r_d in Equation 9 for the HW process.

Figure 3 illustrates the results for five HW examples, with $r(0) = 5$ percent; $\phi = 0.02$ (denoted as “mean reversion” in Figure 3 and other figures); $\mu = 6.5$ percent; and $\sigma = 5$ percent, 10 percent, 15 percent, and 20 percent averaged over 1,000 simulations. Figure 3 also includes the expected value solution ($\sigma = 0$) from Equations 20 and 21, with $c = -0.015$. The drift from Equation 22 is given by $\theta = \mu\phi = (0.065)(0.02) = 0.0013$. Again, as expected, the average over 1,000 simulations is close to the expected value solution. Note that the graphs have not reached the target rate of 6.5 percent, because a mean reversion of 0.02 is small.

Figure 4 presents five HW examples with $r(0) = 0.05$; $\phi = 0, 0.01, 0.02, 0.05$, and 0.15294 ; and $\sigma = 10$ percent. The results presented are an average of 1,000 simulations. Because $r(0) = 0.05$ and $\mu = 0.065$ in all the examples, c is the same for each case and is equal to -0.015 (from Equation 21). The values for θ for each case can be calculated from Equation 22. From Equations 20 and 21, ϕ can be used to determine the time of convergence to the target rate.

Figure 3. Example Hull-White Solutions
(mean reversion equals 0.02; volatility varies)



We use Equations 20 and 21 to calculate the time it takes to get within 5 percent of the target rate for each value of ϕ , as shown in **Table 1**. Choosing a meaningful value for ϕ is critical in modeling interest rates. For example, in Figure 4, mean reversion of 0.15294 is shown because this value for ϕ is necessary to get within 5 percent of the target rate after 10 years. Figure 4 shows this effect, and the curve with $\phi = 0$ grows as expected. The value of ϕ clearly has a strong impact on the convergence to the target rate.

As pointed out earlier, if the initial rate and target rate were closer, the convergence time would also drop. Table 1 shows the drift values and convergence times when the target rate is 5.75 percent. Table 1 clearly demonstrates that the relationship between $r(0)$ and μ significantly affects the convergence time.

Kalotay-Williams-Fabozzi Model

For the KWF process, $f(r) = \ln(r)$ (where \ln is the natural logarithm with base $e \approx 2.718$), $g(r) = 0$, and $\rho = 0$ in Equation 11. Thus, the differential process is

$$d\ln(r) = \theta dt + \sigma dz. \quad (26a)$$

This model is directly analogous to the HL model. Setting $u = \ln(r)$ in equation 26a, the HL process (Equation 13) for u is

$$du = \theta dt + \sigma dz. \quad (26b)$$

Figure 4. Example Hull-White Solutions
(mean reversion varies; volatility equals 10 percent)

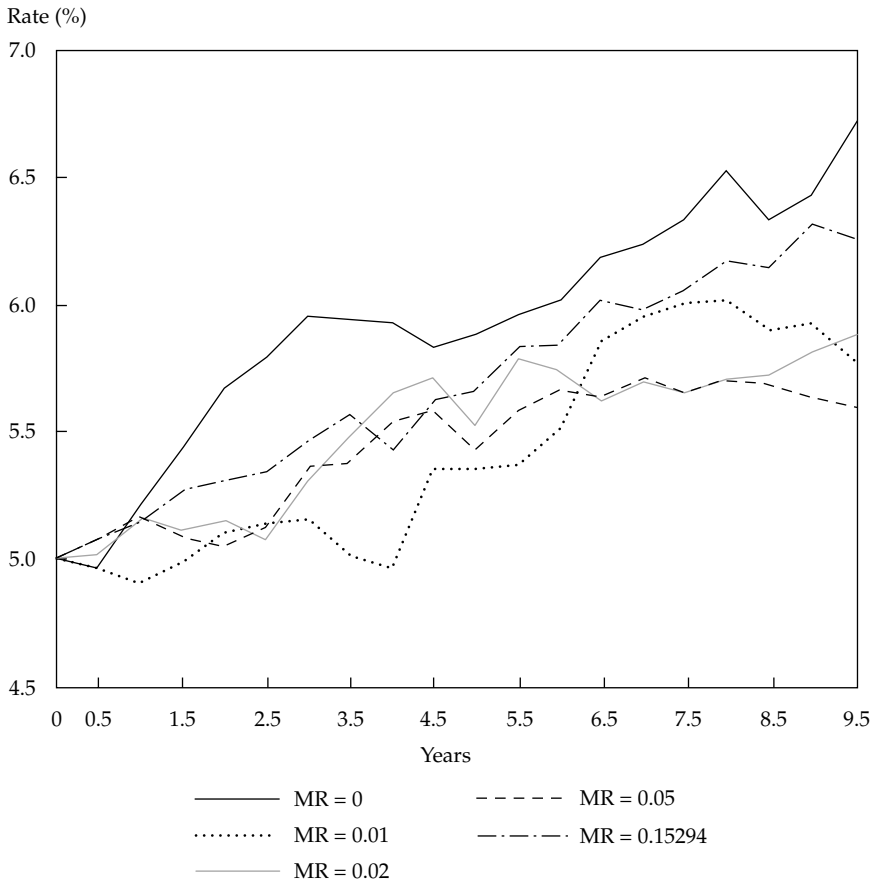


Table 1. Determining the Convergence Time to Target Rate for the Hull-White Model

| Mean Reversion (ϕ) | Target Rate (μ) = 6.5% | | Target Rate (μ) = 5.75% | |
|---------------------------|------------------------------|---|-------------------------------|---|
| | Drift Rate (θ) | Time to Converge to Target Rate (years) | Drift Rate (θ) | Time to Converge to Target Rate (years) |
| 0.01 | 0.00065 | 152.94 | 0.000575 | 95.86 |
| 0.02 | 0.0013 | 76.47 | 0.0015 | 47.94 |
| 0.05 | 0.00325 | 30.59 | 0.002875 | 19.18 |
| 0.15294 | 0.00994 | 10.0 | 0.00879 | 6.27 |

Because u follows a normal process, $\ln(r)$ follows a normal process, and so r follows a lognormal process. Because u is the same as the HL and HW processes, u can become negative, but $u = \ln(r)$ so $r = e^u$; thus, r is always positive. Therefore, the KWF model eliminates the problems of negative short rates that occur with the HL and HW models.

Taking the expectation of Equation 26a results in

$$d\ln(r) = \theta dt,$$

and for Equation 26b

$$du = \theta dt.$$

The expected value in Equation 14a yields

$$\begin{aligned}\ln[r(t)] &= u \\ &= u(0) + \int_0^t \theta ds,\end{aligned}$$

and because $u(0) = \ln[r(0)] = \ln(r_0)$,

$$\ln[r(t)] = \ln(r_0) + \int_0^t \theta ds.$$

Taking the exponential of both sides of this equation gives

$$r(t) = r_0 e^{\int_0^t \theta ds}, \quad (27)$$

showing that $r(t) > 0$ because $r(0) > 0$. Therefore, for constant θ , if $\theta > 0$, the short rate in the KWF process grows without bound, and if $\theta < 0$, the short rate in the KWF process decays to 0.

From Equation 16, the discrete approximation of Equation 26b is

$$u_{k+1} = u_k + \theta_k \tau + \sigma_k \varepsilon_k \sqrt{\tau}, \quad (28a)$$

and the exponential of this equation gives the discrete approximation of Equation 26a:

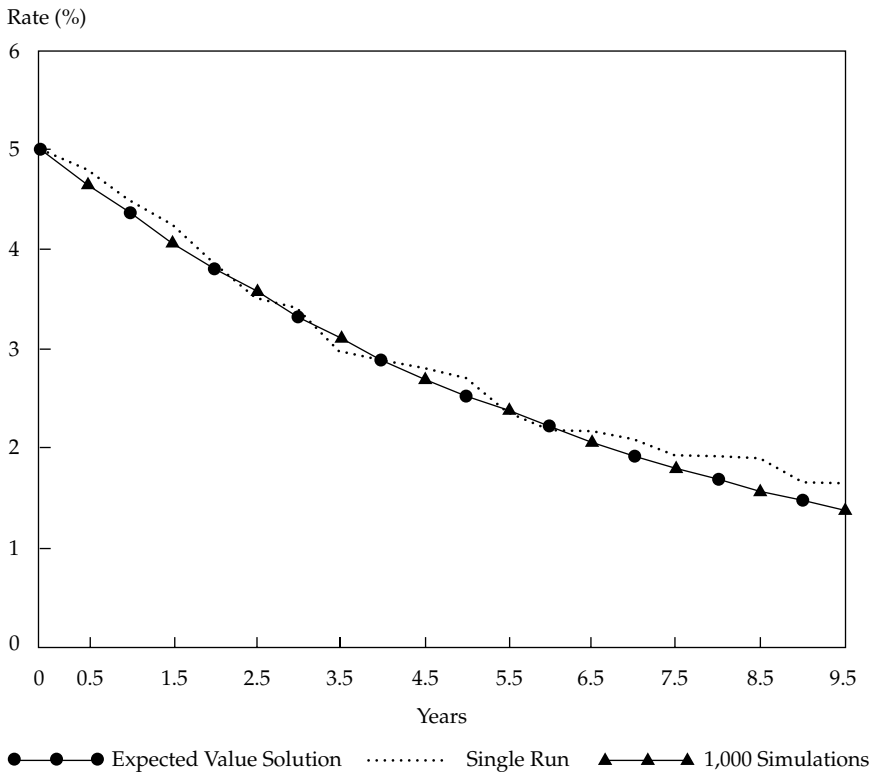
$$r_{k+1} = r_k e^{\theta_k \tau + \sigma_k \varepsilon_k \sqrt{\tau}}. \quad (28b)$$

From Equation 28b, we see that the numerical approximation of the expected value for Equation 26a has similar properties to the solution to the KWF SDE. That is, for constant θ , if $\theta > 0$, the short rate grows without bound, and if $\theta < 0$, the short rate decays to zero. Equation 28b is the formula we will use for r_u and r_d in Equation 9 for the KWF process.

Figure 5 is similar to Figure 2 and presents a KWF model example with $\sigma = 10$ percent for one random run, $\sigma = 10$ percent averaged over 1,000 random simulations, and the expected value solution from Equations 27, with $c = \ln r(0) = \ln(0.05) = -2.995$. The value we use for the drift is -0.136684 . Note that the KWF random run follows the expected value solution more closely than the HL example did in Figure 2. This pattern is also exhibited in the binomial tree models in Chapter 3.

The actual KWF model is a variation of the model we have presented. Because the actual KWF model does not explicitly incorporate the drift, the model does not always have a solution for the binomial tree. Our presentation of the KWF model, however, is easily modified to obtain the actual KWF binomial tree.

Figure 5. Example Kalotay–Williams–Fabozzi Solutions



Black-Karasinski Model

In the BK model, we set $f(r) = \ln(r)$, $\rho = -\phi$, and $g(r) = \ln(r)$ in Equation 11 to obtain the SDE

$$d\ln(r) = [\theta - \phi \ln(r)]dt + \sigma dz. \quad (29a)$$

We now work with Equation 29a using Equation 18 for the HW process in a manner similar to the way we used results from the HL process to develop the KWF process. Setting $u = \ln(r)$ in Equation 29a, we obtain

$$du = (\theta - \phi u)dt + \sigma dz, \quad (29b)$$

which is the HW process for u . Again, note that u has all the same properties as r in the HW model. Because $r = e^u$ in the BK process, $r > 0$ in the BK process. This difference is the advantage the BK model has over the HW model. Thus, the BK process is an extension of the KWF process, just as the HW process is an extension of the HL process. The main difference is that the BK process is a lognormal extension of the lognormal KWF process. As a matter of fact, if $\phi = 0$, the BK process reduces to the KWF process. The introduction of ϕ by Black and Karasinski controls the growth of the short rate in the KWF process.

From Equation 20, we have

$$u(t) = \frac{\theta}{\phi} + ce^{-\phi t},$$

and after taking exponentials,

$$\begin{aligned} r(t) &= e^{u(t)} \\ &= e^{(\theta/\phi) + ce^{-\phi t}}. \end{aligned} \quad (30)$$

For $\phi < 0$, we see that r grows without bound and that for $\phi > 0$,

$$\begin{aligned} \lim_{t \rightarrow \infty} r(t) &= e^{\theta/\phi} \\ &= \mu. \end{aligned}$$

The target rate for the BK process is the exponential of the target rate for the HW process.

As in the HW process from Equation 30 (or Equations 20 and 21), we see that

$$c = \ln(r_0) - \frac{\theta}{\phi} \quad (31)$$

in the BK process. Again, the closer the initial rate is to the target rate, the faster the BK process converges to the target rate. From Equations 30 and 31, we see that if the initial short rate is the target rate, then $r(t) = \mu$ for all t in the BK process, which is analogous to the result we found for the HW process.

As we did for the HW SDE, we present some examples that give qualitative information about c and μ for the BK model. Given the target rate $\mu = e^{\theta/\phi}$, we can solve for the drift or the mean reversion similarly to Equations 22 and 23 in the HW model. As before, we have

$$\theta = \phi \ln(\mu) \quad (32)$$

and

$$\phi = \frac{\theta}{\ln(\mu)}. \quad (33)$$

Later, Equations 32 and 33 will be used to set target rates in the BK models so that we can compare the target rates of Figure 4 with the HW model. Equations 30 and 31 are used to determine how long it takes to reach the target rate.

We discretize $u = \ln(r)$ in Equation 29b, just as we did for the HW SDEs, and then let $r = e^u$. This approach is analogous to the way we used the HL discrete process to get the KWF discrete process. The equations corresponding to Equation 24 are

$$u_{k+1} = u_k + (\theta_k - \phi_k u_k)\tau + \sigma_k \varepsilon_k \sqrt{\tau} \quad (34a)$$

or, after taking the exponential of both sides of Equation 34a,

$$r_{k+1} = r_k e^{[\theta_k - \phi_k \ln(r_k)]\tau + \sigma_k \varepsilon_k \sqrt{\tau}}. \quad (34b)$$

Similar to Equation 25, if θ and ϕ are constant, the solution to the expected value in Equation 34b is

$$r_k = e^{\alpha(1-\phi\tau)^k + (\theta/\phi)}. \quad (35)$$

Note that from Equation 35,

$$\begin{aligned} \lim_{k \rightarrow \infty} r_k &= e^{\theta/\phi} \\ &= \mu \end{aligned}$$

for $0 < \phi\tau < 2$. This result is similar to what we obtain from Equation 25 for the HW SDEs. Equation 34b is the formula we use for r_u and r_d in Equation 9 for the BK process within the binomial framework.

Figure 6 corresponds to Figure 3 and presents five BK solutions with $r(0) = 0.05$; $\phi = 0.02$; $\mu = 6.5$ percent; $\sigma = 5$ percent, 10 percent, 15 percent, and 20 percent averaged over 1,000 simulations; and the expected value solution from Equations 30 and 34 with $c = \ln(r_0) - \ln(\mu) = -0.26236$. From Equation 32, the drift is -0.05467 . Note that after 10 years, the target rate of 6.5 percent has not been reached.

Figure 7 presents five BK examples with $r(0) = 0.05$; $\phi = 0, 0.01, 0.02, 0.05$, and 0.1632 ; and $\sigma = 10$ percent. Figure 7 for the BK SDE is analogous (under the u substitution) to Figure 4 for the HW SDEs. The results presented are an average of 1,000 simulations. The drifts are chosen from Equation 32 to match the target rate of 6.5 percent. Note that only the graph with $\phi = 0.1632$ is approaching the target rate after 10 years because 0.1632 is the only value for ϕ that has the BK SDE short rate get within 5 percent of the target rate

Figure 6. Example Black–Karasinski Solutions
(mean reversion equals 0.02; volatility varies)

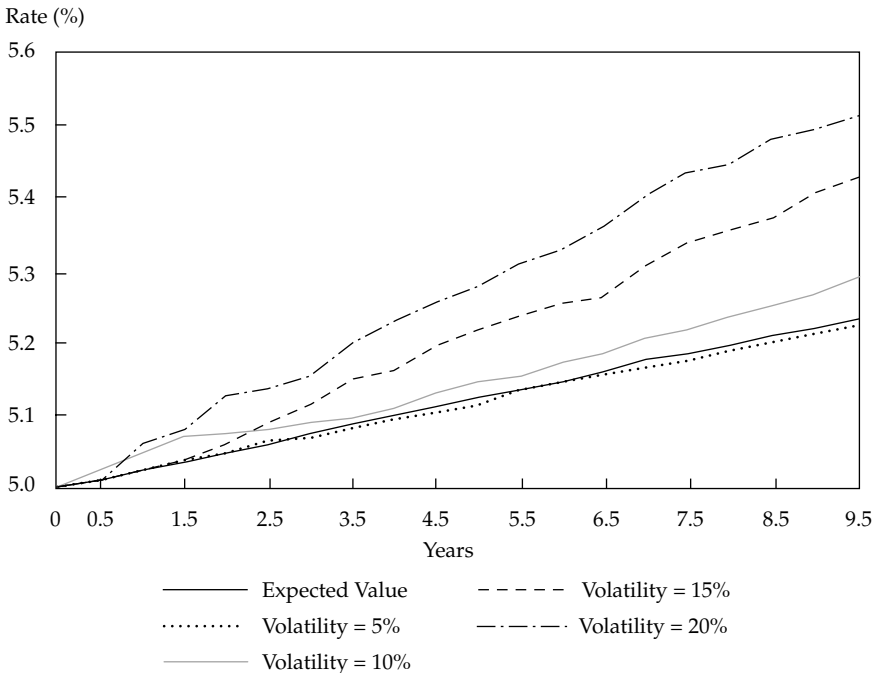
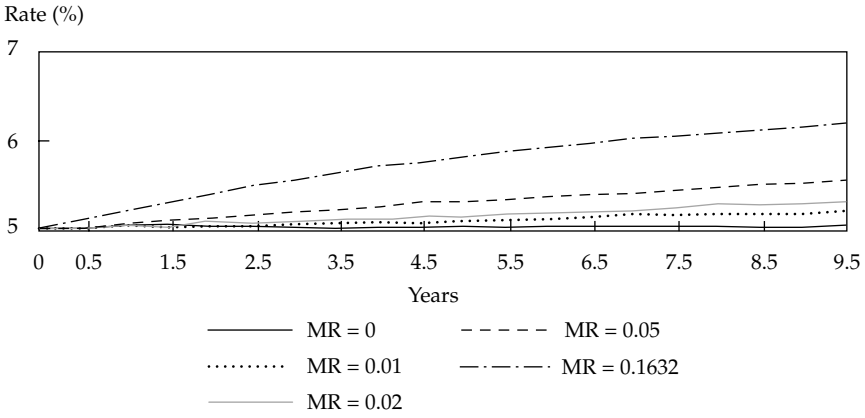


Figure 7. Example Black–Karasinski Solutions
(mean reversion varies; volatility equals 10 percent)



after 10 years. The value is determined from Equations 30 and 31. As in the HW examples, a target rate closer to 5 percent would greatly reduce the convergence time to the target rate for the mean reversion values. As in the HW SDE, it is crucial to choose the mean reversion so that convergence to the target rate is achieved at the desired time.

Note that the BK models seem to be more stable than the HW models. We will compare our binomial models with these results in Chapter 3. These give us a base for comparison in the sections that follow.

Black–Derman–Toy Model

The BDT model is a lognormal model with mean reversion, but the mean reversion is endogenous to the model. That is, the mean reversion is determined from the input parameters to the model. Historically, the BDT model came before the BK model, but we can now use the intuition developed from the HL, HW, KWF, and BK models to help with the development of the BDT model. The BDT model is the most complicated of the five models.

The equation describing the interest rate dynamics in the BDT model has $f(r) = \ln(r)$ and $g(r) = \ln(r)$ in Equation 11, as in the BK model. Therefore, the short rate in the BDT model follows the lognormal process:

$$d\ln(r) = [\theta(t) + \rho(t)\ln(r)]dt + \sigma(t)dz.$$

In the BDT model, however, $\rho(t) = d/dt \ln[\sigma(t)] = \sigma'(t)/\sigma(t)$, giving

$$d\ln(r) = \left[\theta(t) + \frac{\sigma'(t)}{\sigma(t)} \ln(r) \right] dt + \sigma(t) dz. \quad (36a)$$

Making the substitution $u = \ln(r)$ leads to

$$du = \left[\theta(t) + \frac{\sigma'(t)}{\sigma(t)} u \right] dt + \sigma(t) dz. \quad (36b)$$

Notice the similarity between Equation 36 and Equation 29 of the BK model. We expect $\sigma'(t)/\sigma(t)$ to behave similarly to $-\phi(t)$ in the BK model.⁹ This expression should give mean reversion in the short rate when it is negative. That is, we expect that, if $\sigma'(t) < 0$ (implying $\sigma(t)$ is decreasing), then the BDT model will give mean reversion. On the other hand, when $\sigma'(t) > 0$ (implying $\sigma(t)$ is increasing), the short rates in the BDT model will grow with no mean reversion. If $\sigma(t)$ is constant in the BDT model, then $\sigma'(t) = 0$, so $\rho = 0$ and Equation 36a becomes the KWF model (Equation 26). Therefore, we study only the case of varying local volatility for the BDT model.

We now consider the expected value of Equation 36. Under these conditions, Equation 36 becomes

$$\begin{aligned} d\ln(r) &= du \\ &= \left[\theta(t) + \frac{\sigma'(t)}{\sigma(t)} u \right] dt \\ &= \left[\theta(t) + \frac{\sigma'(t)}{\sigma(t)} \ln(r) \right] dt. \end{aligned} \quad (37)$$

Solving this equation for u , as we did in the KWF and BK models, gives us

$$u(t) = \left[\frac{u(0)}{\sigma(0)} + \int_0^t \frac{\theta(s)}{\sigma(s)} ds \right] \sigma(t),$$

and substituting $u = \ln(r)$ gives

$$\begin{aligned} r(t) &= e^{\left[\frac{\ln(r_0)}{\sigma_0} + \int_0^t \frac{\theta(s)}{\sigma(s)} ds \right] \sigma(t)} \\ &= e^{\frac{\sigma(t) \ln(r_0)}{\sigma_0}} e^{\sigma(t) \int_0^t \frac{\theta(s)}{\sigma(s)} ds}, \end{aligned}$$

or

$$r(t) = r_0 e^{\frac{\sigma(t) - \sigma_0}{\sigma_0} \ln(r_0)} e^{\sigma(t) \int_0^t \frac{\theta(s)}{\sigma(s)} ds}. \quad (38)$$

⁹This is what is meant by mean reversion being endogenous to the BDT model. Mean reversion is obtained from $\sigma(t)$ rather than the user dictating ϕ .

Note that the expected value of the BDT short rate depends on the local volatility. If the local volatility has a decreasing structure, the first exponential term in Equation 38 will have a negative exponent and will cause a decrease in the short rate—and vice versa if the local volatility has an increasing structure. It is important to note that mean reversion in the BDT model comes from the local volatility structure.

Now consider some specific cases for Equation 38. If θ is constant, then Equation 38 becomes

$$r(t) = r_0 e^{\frac{\sigma(t) - \sigma_0}{\sigma_0} \ln(r_0)} e^{-\theta \sigma(t) \int_0^t \frac{1}{\sigma(s)} ds}. \quad (39)$$

Because the volatility of short rates is usually small, the term $1/\sigma(s)$ can be a large number, and integrating over such a function will give a large value. Therefore, the exponent in the last term of Equation 39 has the possibility of becoming large, and even if the market volatility is small, the rates in the BDT model can become unbounded. We now consider some specific cases for the local volatility in the BDT model.

Suppose

$$\frac{\sigma'(t)}{\sigma(t)} = a,$$

where a is a constant, thus

$$\sigma(t) = \sigma(0)e^{at} = \sigma_0 e^{at},$$

and under this condition, the solution to Equation 37 is given by

$$u = \left[u(0) + \frac{\theta}{a} \right] e^{at} - \frac{\theta}{a}.$$

Since $u(0) = \ln(r_0) < 0$ (because $0 < r < 1$), the expression $u(0) + \theta/a$ could be positive or negative, depending on the sign and magnitude of the drift θ . This solution shows that if $a > 0$ and $u(0) + \theta/a < 0$, then $u \rightarrow -\infty$ and thus $r = e^u \rightarrow 0$. It also shows that if $a > 0$ and $u(0) + \theta/a > 0$, then $u \rightarrow \infty$ and $r = e^u \rightarrow \infty$. That is, for $a > 0$, the BDT short rate could blow up or decay to zero. If $a < 0$, then $u \rightarrow -\theta/a$, which implies that $r \rightarrow e^{-\theta/a}$. That is, if the local volatility decays exponentially over time, the BDT short rate will approach a target rate. The target rate will depend on the magnitude and sign of θ and a . We will address this issue when we study the numerical approximations to the BDT model.

Now, consider the case where the local volatility is linear. That is,

$$\sigma'(t) = m$$

and

$$\sigma(t) = mt + \sigma_0,$$

where m is a constant. Using Equation 39 gives

$$r(t) = r_0 e^{(mt/\sigma_0)\ln(r_0)} e^{(\theta t + \theta \sigma_0/m)\ln[(mt + \sigma_0)/\sigma_0]}$$

for a linear local volatility, if θ is constant. This solution is interesting. Because $\ln(r_0) < 0$, we see that, if $m < 0$, the first exponential term in the solution will increase. But the second term contains $\ln[(mt + \sigma_0)/\sigma_0]$, which is negative if m is negative. Therefore, if θ is positive, the second term will decrease, and if θ is negative, the second term will increase. Therefore, r can grow without bound or tend to a target rate. A similar situation exists for $m > 0$. Therefore, for linear local volatility, the short rate in the BDT model can grow without bound for negative or positive θ . If $m < 0$, it is possible that $\sigma(t) = mt + \sigma_0 < 0$. In this case, \ln is undefined and the BDT short rate does not exist. Therefore, we suggest that if the BDT model is used to model a T -year bond with a linearly decreasing volatility structure, the slope $m < 0$ should satisfy

$$\begin{aligned} \frac{\sigma(T)}{\sigma_0} &= \frac{mT + \sigma_0}{\sigma_0} \\ &= 1 + \frac{mT}{\sigma_0} > 0.5. \end{aligned}$$

These types of problems have to be dealt with when considering numerical solutions (e.g., binomial models) for the BDT model, and examples are shown in Chapter 4.

We now consider some numerical solutions to the BDT process. To discretize Equation 36a for the BDT model, we start off again by approximating du in Equation 36b by using Equations 12a and 12b to get

$$u_{k+1} = u_k + (\theta_k + \rho_k u_k)\tau + \sigma_k \varepsilon_k \sqrt{\tau}, \quad (40)$$

where

$$\rho_k = \frac{\sigma'_k}{\sigma_k}.$$

We approximate this term by

$$\frac{(\sigma_{k+1} - \sigma_k)/\tau}{\sigma_k}.$$

That is, we approximate σ'_k by a discrete approximation of the derivative. We now have

$$u_{k+1} = u_k + \left(\theta_k + \frac{(\sigma_{k+1} - \sigma_k)^{\prime\tau}}{\sigma_k} u_k \right) \tau + \sigma_k \varepsilon_k \sqrt{\tau}$$

or

$$\begin{aligned} u_{k+1} &= u_k \left(1 + \frac{(\sigma_{k+1} - \sigma_k)^{\prime\tau}}{\sigma_k} \right) + \theta_k \tau + \sigma_k \varepsilon_k \sqrt{\tau} \\ &= \frac{\sigma_{k+1}}{\sigma_k} u_k + \theta_k \tau + \sigma_k \varepsilon_k \sqrt{\tau}. \end{aligned} \quad (41)$$

In the expected value case, Equation 41 leads to

$$u_{k+1} = \frac{\sigma_{k+1}}{\sigma_k} u_k + \theta_k \tau. \quad (42)$$

We consider some of the solutions to Equation 42. Equation 42 gives

$$\begin{aligned} u_1 &= \frac{\sigma_1}{\sigma_0} u_0 + \theta_0 \tau \\ u_2 &= \frac{\sigma_2}{\sigma_1} u_1 + \theta_1 \tau = \frac{\sigma_2}{\sigma_1} \left(\frac{\sigma_1}{\sigma_0} u_0 + \theta_0 \tau \right) + \theta_1 \tau = \frac{\sigma_2}{\sigma_0} u_0 + \frac{\sigma_2}{\sigma_1} \theta_0 \tau + \theta_1 \tau \\ &\vdots \\ u_k &= \frac{\sigma_k}{\sigma_{k-1}} u_{k-1} + \theta_{k-1} \tau = \frac{\sigma_k}{\sigma_0} u_0 + \sum_{j=1}^{k-1} \left(\frac{\sigma_k}{\sigma_j} \theta_{j-1} \tau \right) + \theta_{k-1} \tau. \end{aligned}$$

We see that u and thus $\ln(r)$ depend heavily on the local volatility. In particular, if

$$\frac{\sigma_{k+1}}{\sigma_k} = \alpha,$$

where α is a constant, then

$$u_k = \alpha^k u_0 + \sum_{j=0}^{k-1} \alpha^j \theta_{k-j-1} \tau.$$

The exponential of this equation gives

$$r_k = r_0 e^{(\alpha^k - 1) \ln(r_0)} e^{\sum_{j=0}^{k-1} \alpha^j \theta_{k-j-1} \tau}.$$

This equation is interesting because $\ln(r_0) < 0$. If $\alpha > 1$, the first exponential term decreases. When $\theta < 0$, the second exponential term also decreases and the BDT short rate should approach a target rate. Conversely, when $\theta > 0$, the second exponential term increases. In this case, we can approach a target rate or the second term can dominate. If $\alpha < 1$, then a similar situation arises. Therefore, in order to get meaningful numerical results for the BDT short rates, we strongly recommend that α be close to 1 and that the term structure of spot rates not have too large a slope.

Taking the exponential of Equation 40 gives

$$r_{k+1} = r_k e^{[\theta_k + \rho_k \ln(r_k)] \tau + \sigma_k \varepsilon_k \sqrt{\tau}}. \quad (43)$$

This expression will be used for r_u and r_d in Equation 9 to generate the binomial BDT short rates.

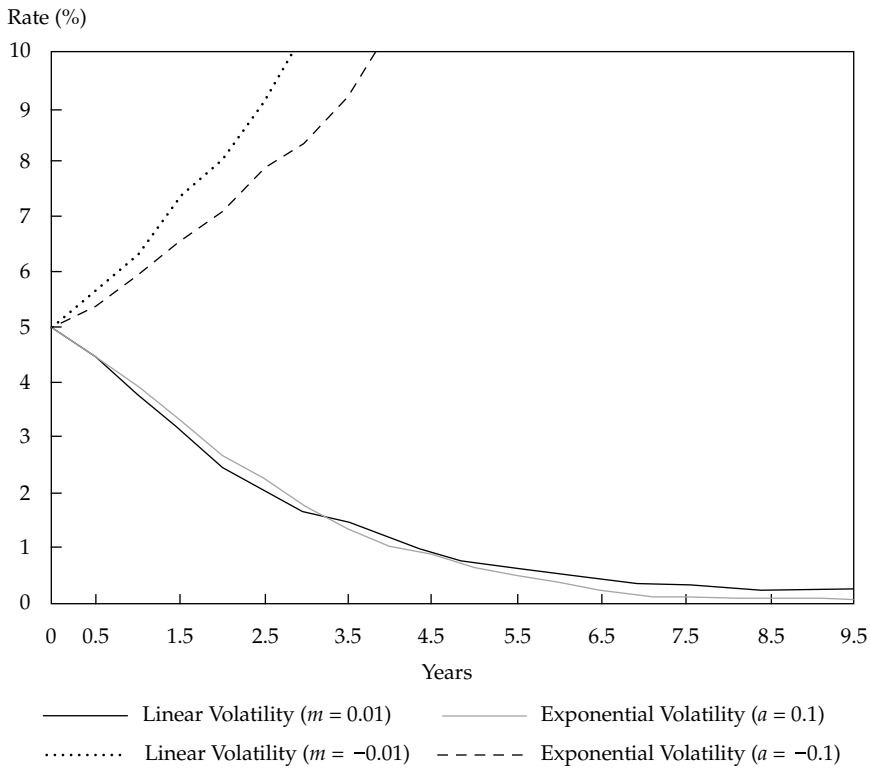
Figure 8 presents BDT solutions with $r(0) = 0.05$ and $\theta = -0.05467$, as in the BK examples, with an exponentially increasing, an exponentially decreasing, a linearly increasing, and a linearly decreasing volatility all starting at 10 percent for a single simulation path. **Figure 9** presents the average of 1,000 simulations using the same inputs. Figures 8 and 9 illustrate that for the BDT process, the single simulation exhibits behavior similar to that of the average from multiple simulations. Figures 8 and 9 show that the solutions decrease when the volatility is exponentially increasing because $a = 0.1 > 0$ and

$$\begin{aligned} u(0) + \frac{\theta}{a} &= \ln(0.05) + \frac{-0.05467}{0.1} \\ &= -3.54 < 0 \end{aligned}$$

In the figures, the solutions appear to grow without bound when the volatility is exponentially decreasing because the target rate is $e^{-\theta/a} = e^{0.05467/-0.1} = 0.5789$. We leave the linear examples for the reader to analyze.

Figure 10 presents the same results from Figure 8 except that $\theta = -0.25$. For the decreasing exponential volatility structure, the target rate of $e^{-\theta/a} = e^{0.25/-0.1} = 0.0821$ is approached, but the decreasing linear volatility structure

Figure 8. Example Black–Derman–Toy Solutions
(single run; various volatilities)



does not approach a target rate. Increasing volatility structures still have decreasing rates. (These phenomena are also discussed for the BDT binomial examples in Chapter 4.) The BDT model is obviously complicated, but our rules of thumb should help interpret results from the model.

The expected value analysis presented in this chapter is important because the corresponding properties will also hold under the more general case, which includes the stochastic component. Consequently, the properties presented within this section will also hold under more general circumstances. The random runs for the lognormal models are closer to their averages than are the random runs for the normal models. These properties will also be important in studying and using the binomial trees. In fact, we use what has been learned in this chapter to test and analyze our binomial models in Chapter 4. The discrete approximations we have developed for the five models are used to build the binomial models in Chapter 3.

Figure 9. Example Black–Derman–Toy Solutions
(1,000 runs; various volatilities)

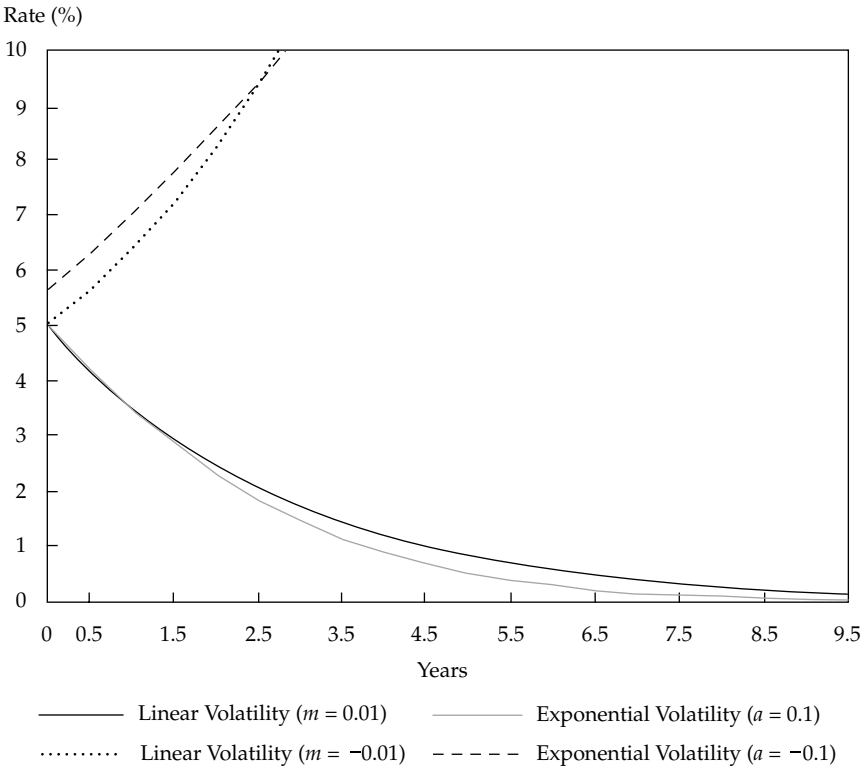
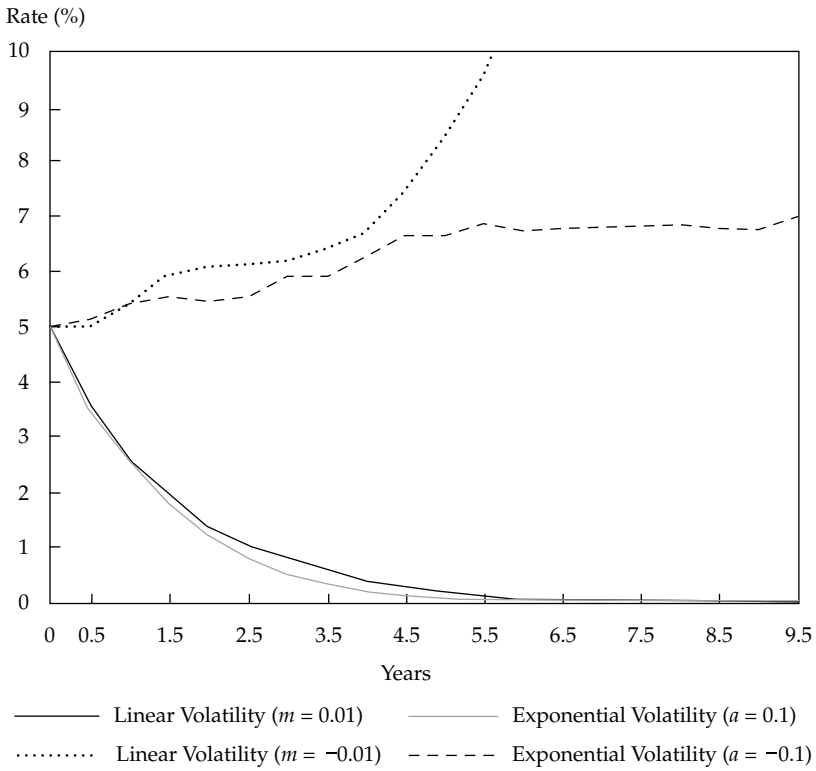


Figure 10. Example Black–Derman–Toy Solutions
(1,000 runs; various volatilities)



3. The Binomial Trees

For each of the five models that we studied and developed in Chapter 2, we now determine formulas giving the two possible values r_u and r_d for r_1 in Equation 9. The two values, r_u and r_d , for the short rate, r_1 , are called the up-state, or up-move, and the down-state, or down-move, respectively. This is our starting point for the algorithm that develops the short rates for the binomial trees for the Ho–Lee (HL; Equations 13 and 16), Hull–White (HW; Equations 18 and 24), Kalotay–Williams–Fabozzi (KWF; Equations 26 and 28), Black–Karasinski (BK; Equations 29 and 34), and Black–Derman–Toy (BDT; Equations 36, 40, and 43) models.

Before we give the formulas for the up-state, r_u , and the down-state, r_d , in Equation 9, we develop the basic structure of a binomial tree. Following this, we give the anticipated formulas for r_u and r_d in Equation 9 for each of the five models. Then, we show how to modify these formulas to build an algorithm to generate the formulas for all the short rates that are used to develop the binomial tree. During this process, we also show some of the properties that the numerical solutions to the binomial tree have to satisfy. These are in addition to the properties we developed in Chapter 2.

A schematic of how we move from r_0 to the up-state, r_u , and the down-state, r_d , is shown in **Figure 11**. This schematic is called a binomial tree, or binomial lattice, because there are only two possible states emanating from r_0 . One branch leads to the up-state, r_u , and the other to the down-state, r_d ; r_0 , r_u , and r_d are also referred to as nodes in the tree.

To make this more concrete, a numerical example of a tree is presented in **Figure 12**. We also present the corresponding price tree. The price tree is based on a zero-coupon bond with a par value of \$100. Recall that the rates act for a period of time and the prices are at the given time. Figure 12 contains the results for a two-period binomial interest rate tree in which the short rates are semiannual.

Figure 11. Example Binomial Tree

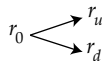


Figure 12. Numerical Example of a Binomial Tree



To calculate the prices in Figure 12, one uses backward induction (the two previous adjacent nodes) and the discount factors. We use $q = 0.5$, and because we are assuming semiannual rates, we have $t_1 = 0.5$ and $\tau = \tau_1 = 0.5$ (six months). For example, to compute \$96.70 in the price tree of Figure 12, we calculate

$$\frac{q(100) + (1 - q)100}{1 + r_u \tau} = \frac{0.5(100) + (1 - 0.5)(100)}{[1 + (0.0683)(0.5)]} \quad (44a)$$

$$= 96.70,$$

and to get the price of the bond (the first node of the price tree in Figure 12), we calculate

$$\frac{q(96.70) + (1 - q)(97.49)}{1 + r_0 \tau} = \frac{0.5(96.70) + (1 - 0.5)(97.49)}{[1 + (0.0605)(0.5)]} \quad (44b)$$

$$= 94.24.$$

In order to model short rates for more than one time period, we have to add more branches to the tree in Figure 11. If we are at node r_u , the short rate for the next time period should be able to be one of two values, which we will denote as r_{uu} and r_{ud} . The rate r_{uu} is the up-move for the short rate r_u at time t_2 , and r_{ud} is the down-move for the short rate r_u at time t_2 . Similarly, if we are at node r_d , then the short rate for the next time period will be one of two values, which are denoted as r_{du} and r_{dd} . The rate r_{du} is the up-move for the short rate r_d at time t_2 , and r_{dd} is the down-move for the short rate r_d at time t_2 .

To determine r_{uu} , r_{ud} , r_{du} , and r_{dd} in a manner similar to that used for determining r_u and r_d , we require formulas and/or equations that can be solved that are similar to those used for r_u and r_d . Such formulas will give us an algorithm for calculating possible values for the short rates at the n given calendar times, $t_1 < t_2 < \dots < t_n$.

Notice that from r_0 , we went to two possible values for the next short rate of r_u and r_d . That is, we increased the number of possible short rates by a

factor of two, for which we need two formulas. From r_u and r_d , we went to two possible values for each of these short rates at the next time, and thus, we must have four formulas. In other words, we again increased the number of formulas by a factor of two. If we continued this process, we would increase the number of formulas by a factor of two to eight formulas, as seen in **Figure 13**. If we keep doubling the previous number of possible short rates, we soon will have too many short rates to be computed. For example, after 10 time steps, we would have $2^{10} = 1,024$ formulas to calculate, and at 30 time steps, we would have $2^{30} \approx 10^9$ formulas to calculate, which would be impractical even for today's supercomputers. We would be unable to build a 30-year one-year-time-step binomial tree model. Such a scenario is clearly unacceptable.

To reduce the number of formulas (and thus calculations), we have to put a restriction on the algorithm. The restriction we implement is called the *recombination condition*. This condition simply requires that an up-move from a down-state equal a down-move from an up-state. For example,

$$r_{ud} = r_{du}.$$

The recombination condition reduces the number of possible short rates at t_2 from 4 to 3 and at t_3 from 8 to 4. The recombination condition is used at each interior node of the tree. This requirement will permit an increase of only one node from one time step to the next. So, for a 10-period tree, we only need 10 equations; and for the 30-period tree, we only need 30 equations. This number of equations is much more manageable. **Figure 14** shows a four-period short-rate binomial tree incorporating the recombination criterion just described.

Now, a more efficient method is needed for subscribing our nodes in Figure 14, or the number of subscripts will become ridiculous. We use two subscripts. The first subscript denotes the time step, and the second subscript denotes the level on the tree at that time step. That is, node $r_{k,j}$ is the short

Figure 13. Expansion of Binomial Short-Rate Tree without Recombination

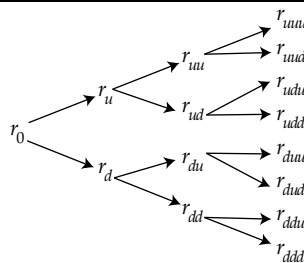
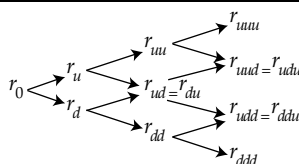


Figure 14. Four-Period Short-Rate Binomial Tree with Recombination Condition



rate at time t_k at level j . For example, $r_{0,1} = r_0$, $r_{1,1} = r_u$, and $r_{1,2} = r_d$. **Figure 15** is the same binomial tree as in Figure 14 except that it uses the new subscripts. All nodes along the uppermost branch of the tree have the form $r_{k,1}$. This set of nodes is called the up-state of the tree. For example, in Figure 15, the up-state of the tree consists of the nodes $r_{0,1}$, $r_{1,1}$, $r_{2,1}$, and $r_{3,1}$. The up-state of the tree represents the largest possible values at time t_k for the short rates of the tree. A graph of these values versus time gives the upper range for the short rates. Similarly, the nodes along the lowest branch of the tree have the form $r_{k,k+1}$. This set of nodes is called the down-state of the tree. For example, in Figure 15, the down-state of the tree has the nodes $r_{0,1}$, $r_{1,2}$, $r_{2,3}$, and $r_{3,4}$. The down-state of the tree represents the smallest possible values at time t_k for the short rates of the tree. A graph of these values versus time gives the lower range for the short rates.

The interest rate tree in Figure 15 spans four time steps, and each $r_{k,j}$ represents a forward rate that spans a short period of time. An up-move from node $r_{k,j}$ is a move to node $r_{k+1,j}$, and a down-move is a move to node $r_{k+1,j+1}$. As mentioned earlier, it is convenient to work with the price tree derived from the corresponding discount factors. The corresponding price tree is presented in **Figure 16**. As noted earlier, the subscripts on $p_{k,j}$ correspond to those on $r_{k,j}$. Notice that the price tree has one more set of nodes than the corresponding interest rate tree. Because interest rates span time and prices do not, a four-time-step interest rate tree spans a five-time-step price tree.

Figure 15. Four-Period Short-Rate Binomial Tree with Recombination Condition and Only Two Subscripts

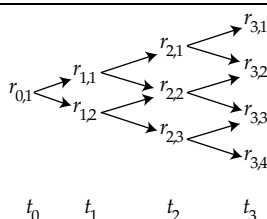


Figure 16. Corresponding Price Tree for Four-Period Short-Rate Binomial Tree

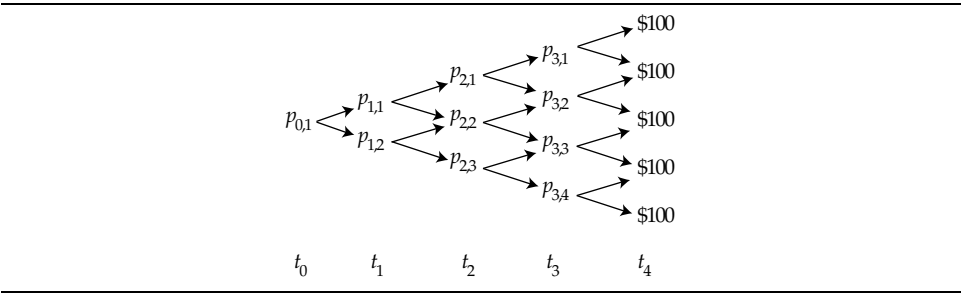


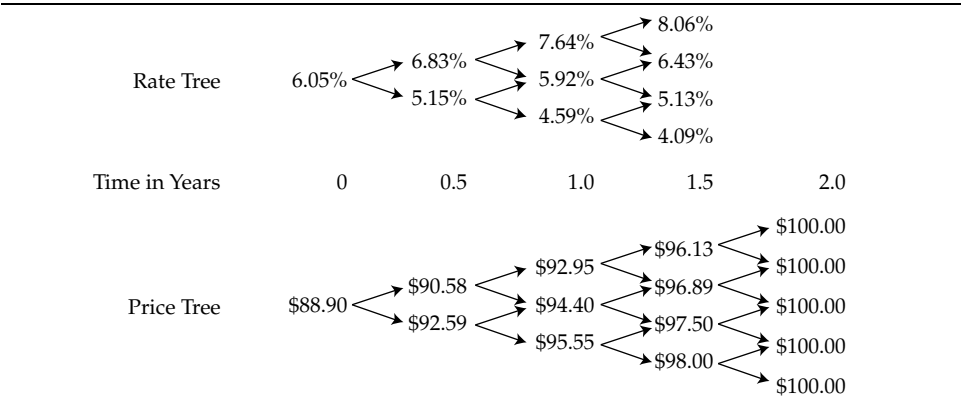
Figure 17 shows an example of a four-period short rate tree with period $\tau = 0.5$ years and its corresponding five-period pricing tree with numerical values. The pricing trees are derived just like the pricing trees in Figure 12. The reader can use Equation 44 with $q = 0.5$, $\tau = \tau_1 = \tau_2 = \tau_3 = 0.5$ and backward induction to verify these prices. In particular, we note that

$$p_{k,j} = \frac{q(p_{k+1,j}) + (1 - q)p_{k+1,j+1}}{1 + r_{k,j}\tau}. \tag{45}$$

Therefore, to get the prices at time t_k , we first need to calculate all the prices at time t_{k+1} (which is why it is called backward induction.) To build a binomial price tree from a binomial rate tree, one programs Equation 45.

The separation between the up-state and down-state rates, as shown in **Figure 18**, is a spread of possible future short rates. Notice that this spread

Figure 17. Four-Period Short-Rate Tree and Corresponding Five-Period Pricing Tree



increases with time. A small spread indicates that our model either has a low volatility or high mean reversion. Only the HW, BK, and BDT models include mean reversion.

The algorithms for generating the short rates for each of the five models are now developed in a slightly different order than we presented in Chapter 2. The reason for this difference is that, within the binomial lattice framework, the HW and BK binomial trees must have a variable time step. We first present the formulas for the up-state, r_u , and the down-state, r_d , in Equation 9 for each of the five models of Chapter 2.

Formulas for the Up-State and Down-State

In this section, we derive the necessary formulas required to solve for r_u and r_d within the one-period binomial framework. Specifically, a quadratic equation is developed for each model; the solution to this equation is used to compute r_u and r_d for each model within the binomial framework.

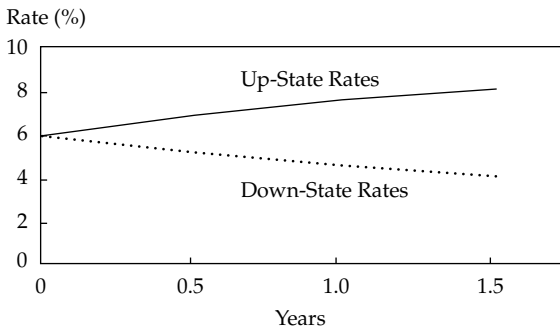
Ho-Lee Formulas. From the discrete approximation Equation 16 for the HL SDE Equation 13, we use

$$\begin{aligned} r_u &= r_{1,1} \\ &= r_0 + m_0\tau + \sigma_0\sqrt{\tau} \end{aligned} \quad (46a)$$

for the up-move from r_0 and

$$\begin{aligned} r_d &= r_{1,2} \\ &= r_0 + m_0\tau - \sigma_0\sqrt{\tau} \end{aligned} \quad (46b)$$

Figure 18. Up-State and Down-State Rates for Tree Shown in Figure 17



for the down-move from r_0 . We are using m_0 for the numerical approximation to the drift value θ_0 . Notice that

$$r_u - r_d = 2\sigma_0\sqrt{\tau}.$$

That is, the volatility, σ , determines the spread in the up-state and the down-state.

Substituting the formulas for the up-state, r_u , and the down-state, r_d , from Equation 46 into Equation 9 gives the following quadratic equation for m_0 for the HL binomial tree

$$\begin{aligned} A_2(1 + r_0\tau)[1 + (r_0 + m_0\tau - \sigma_0\sqrt{\tau})\tau][1 + (r_0 + m_0\tau + \sigma_0\sqrt{\tau})\tau] \\ - q[1 + (r_0 + m_0\tau - \sigma_0\sqrt{\tau})\tau] - (1 - q)[1 + (r_0 + m_0\tau + \sigma_0\sqrt{\tau})\tau] = 0. \end{aligned} \quad (47)$$

Many authors suggest using the Newton–Raphson method (Burden and Faires 1998) on this equation to solve for m_0 . This method involves differentiating this equation with respect to m_0 . To avoid this differentiation, we use the bisection method for m_0 . See Burden and Faires for a complete description of the bisection method.

Hull–White Formulas. Following the same development used for the HL process, we use the discrete approximation Equation 24 for the HW SDE Equation 18 to get

$$\begin{aligned} r_u &= r_{1,1} \\ &= r_0 + m_0\tau - \phi_0 r_0\tau + \sigma_0\sqrt{\tau} \end{aligned} \quad (48a)$$

for the up-move from r_0 and

$$\begin{aligned} r_d &= r_{1,2} \\ &= r_0 + m_0\tau - \phi_0 r_0\tau - \sigma_0\sqrt{\tau} \end{aligned} \quad (48b)$$

for the down-move from r_0 . Notice again that

$$r_u - r_d = 2\sigma_0\sqrt{\tau}.$$

Substituting the formulas for the up-state, r_u , and the down-state, r_d , from Equation 48 into Equation 9 gives the following quadratic equation for m_0 for the HW binomial tree

$$\begin{aligned}
 & A_2(1 + r_0\tau)[1 + (r_0 + m_0\tau - \phi_0 r_0\tau - \sigma_0\sqrt{\tau})\tau][1 + (r_0 + m_0\tau - \phi_0 r_0\tau + \sigma_0\sqrt{\tau})\tau] \\
 & - q[1 + (r_0 + m_0\tau - \phi_0 r_0\tau - \sigma_0\sqrt{\tau})\tau] \\
 & - (1 - q)[1 + (r_0 + m_0\tau - \phi_0 r_0\tau + \sigma_0\sqrt{\tau})\tau] = 0.
 \end{aligned} \tag{49}$$

We again solve this equation using the bisection method.

Kalotay–Williams–Fabozzi Formulas. From the discrete approximation Equation 28b for the KWF SDE Equation 26a, we use

$$\begin{aligned}
 r_u &= r_{1,1} \\
 &= r_0 e^{m_0\tau + \sigma_0\sqrt{\tau}}
 \end{aligned} \tag{50a}$$

for the up-move from r_0 and

$$\begin{aligned}
 r_d &= r_{1,2} \\
 &= r_0 e^{m_0\tau - \sigma_0\sqrt{\tau}}
 \end{aligned} \tag{50b}$$

for the down-move from r_0 for the KWF process. That is, we use an exponential of the HL formulas (Equation 46).

Notice that

$$\frac{r_u}{r_d} = e^{2\sigma_0\sqrt{\tau}}$$

or, after taking logarithms,

$$\ln(r_u) - \ln(r_d) = 2\sigma_0\sqrt{\tau}.$$

That is, the volatility determines the spread in the natural logarithm of the up-state and the natural logarithm of the down-state. This formula is the log-normal version of the spread for the normal HL and HW binomial trees.

Substituting the formulas for the up-state, r_u , and the down-state, r_d , from Equation 50 into Equation 9 gives the following quadratic equation for m_0

$$\begin{aligned}
 & A_2(1 + r_0\tau)\left[1 + \left(r_0 e^{m_0\tau - \sigma_0\sqrt{\tau}}\right)\tau\right]\left[1 + \left(r_0 e^{m_0\tau + \sigma_0\sqrt{\tau}}\right)\tau\right] \\
 & - q\left[1 + \left(r_0 e^{m_0\tau - \sigma_0\sqrt{\tau}}\right)\tau\right] - (1 - q)\left[1 + \left(r_0 e^{m_0\tau + \sigma_0\sqrt{\tau}}\right)\tau\right] = 0.
 \end{aligned} \tag{51}$$

We also solve this equation using the bisection method.

Black–Karasinski Formulas. Because the BK process adds mean reversion to the KWF process, from Equation 50 we have

$$\begin{aligned} r_u &= r_{1,1} \\ &= r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau + \sigma_0 \sqrt{\tau}} \end{aligned} \quad (52a)$$

for the up-move from r_0 and

$$\begin{aligned} r_d &= r_{1,2} \\ &= r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau - \sigma_0 \sqrt{\tau}} \end{aligned} \quad (52b)$$

for the down-move from r_0 . We could also use Equations 29a and 34b. Again,

$$\frac{r_u}{r_d} = e^{2\sigma_0 \sqrt{\tau}}$$

and

$$\ln(r_u) - \ln(r_d) = 2\sigma_0 \sqrt{\tau}.$$

Substituting the formulas for the up-state, r_u , and the down-state, r_d , from Equation 52 into Equation 9 gives the following quadratic equation for m_0

$$\begin{aligned} A_2(1 + r_0 \tau) &\left[1 + \left(r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau - \sigma_0 \sqrt{\tau}} \right) \tau \right] \left[1 + \left(r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau + \sigma_0 \sqrt{\tau}} \right) \tau \right] \\ &- q \left[1 + \left(r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau - \sigma_0 \sqrt{\tau}} \right) \tau \right] \\ &- (1 - q) \left[1 + \left(r_0 e^{m_0 \tau - \phi_0 \ln(r_0) \tau + \sigma_0 \sqrt{\tau}} \right) \tau \right] = 0. \end{aligned} \quad (53)$$

This equation can also be solved using the bisection method.

Black–Derman–Toy Formulas. The BDT model is similar to the KWF and BK formulas because the BDT process is lognormal. We use

$$\begin{aligned} r_u &= r_{1,1} \\ &= r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}} \end{aligned} \quad (54a)$$

for the up-move from r_0 and

$$\begin{aligned} r_d &= r_{1,2} \\ &= r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}} \end{aligned} \quad (54b)$$

for the down-move from r_0 . That is, we use the same formula as in the KWF formulas (Equation 50), but now we have

$$m_0 = \theta_0 - \frac{\sigma(0)}{\sigma'(0)} \ln(r_0).$$

(One could have also used Equation 36a and Equation 43.) We still have that

$$\frac{r_u}{r_d} = e^{2\sigma_0 \sqrt{\tau}}$$

and

$$\ln(r_u) - \ln(r_d) = 2\sigma_0 \sqrt{\tau}.$$

Substituting the formulas for the up-state, r_u , and the down-state, r_d , from Equation 54 into Equation 9 gives the following quadratic equation for m_0

$$\begin{aligned} A_2(1 + r_0 \tau) \left[1 + \left(r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}} \right) \tau \right] \left[1 + \left(r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}} \right) \tau \right] \\ - q \left[1 + \left(r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}} \right) \tau \right] - (1 - q) \left[1 + \left(r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}} \right) \tau \right] = 0. \end{aligned} \quad (55)$$

As in all the other models, we again solve this equation using the bisection method for m_0 . We will obviously get the same answer as in the KWF model, but the meaning for m_0 is different. In the KWF model, m_0 is the drift, whereas in the BDT model, m_0 involves the drift, the volatility, and the initial spot rate. We will see below how local volatility is incorporated into the BDT model for the remaining short rates in the BDT binomial tree.

We note that we are able to solve for m_0 in Equations 47, 49, 51, 53, and 55 using the bisection method. Also, notice that all of these equations have a similar form. That is, once we had the HL Equation 47, the others were slight modifications in a sequential manner. In the next section, we focus on an algorithm that generates all the short rates for each of our five binomial trees. As in the development of Equations 47, 49, 51, 53, and 55, the HL binomial tree serves as the base case, and we modify this base algorithm to get the others.

Algorithms for Generating Binomial Trees

In the last section, we developed the equations that produce the first two possible short rates in the binomial tree for each of our five models. In this section, we show how to develop algorithms that generate all the short rates for the binomial trees. Because the variable time step creates difficulties in the binomial models, we present the development of the HL, KWF, and BDT algorithms first because they do not require a variable time step. We then show why the HW and BK binomial models require a varying time step. In Appendix A, we outline an algorithm that will generate short rates for these two binomial models.

First, we present some notation. We let $R_1, R_2, R_3, \dots, R_n$ be the term structure of the spot rates, and we let A_k represent the price of the k -period zero-coupon bond determined by R_k . That is,

$$A_k = \frac{1}{(1 + R_k \tau)^k}.$$

In particular, the price tree of the n -period bond we are modeling must give

$$A_n = \frac{1}{(1 + R_n \tau)^n}$$

because of the no-arbitrage condition. Recall that the no-arbitrage condition requires that the price of the bond through the short rates (i.e., the model) must equal the price of the bond through the spot rates (i.e., the market). This check can be done on binomial tree algorithms to make sure they are correct. In particular, for straight bonds, the HL, KWF, and BDT binomial trees should give the same price for the same term structure, regardless of the volatility, σ .¹⁰

Ho-Lee Formulas for Short Rates. Equation 16 is used to build the algorithm for generating the short rates for the HL binomial tree. We develop the equations for the short rates using the recombination condition to determine the conditions on the parameters in the HL model.

In the HL binomial tree, the expressions for $r_{k,j}$ that correspond to Equation 46 are

$$r_{k+1,j} = r_{k,j} + m_k \tau + \sigma_k \sqrt{\tau} \quad (56a)$$

¹⁰The same is true for the HW and BK binomial trees for the constant time step. Interpolation will give round-off errors in the variable time step case.

for an up-move and

$$r_{k+1,j+1} = r_{k,j} + m_k \tau - \sigma_k \sqrt{\tau} \quad (56b)$$

for a down-move. Using Equation 46 or 56 gives

$$r_{1,1} = r_0 + m_0 \tau + \sigma_0 \sqrt{\tau}$$

and

$$r_{1,2} = r_0 + m_0 \tau - \sigma_0 \sqrt{\tau}.$$

Equation 56 also gives

$$\begin{aligned} r_{2,1} &= r_{1,1} + m_1 \tau + \sigma_1 \sqrt{\tau}, \\ r_{2,2} &= r_{1,1} + m_1 \tau - \sigma_1 \sqrt{\tau} = r_{1,2} + m_1 \tau + \sigma_1 \sqrt{\tau}, \end{aligned} \quad (57)$$

and

$$r_{2,3} = r_{1,2} + m_1 \tau - \sigma_1 \sqrt{\tau}.$$

The two values for $r_{2,2}$ in Equation 57 arise from the recombination condition. (A down-move from $r_{1,1}$ must equal an up-move from $r_{1,2}$.) We solve this equation for σ_1 . Because

$$r_{1,1} + m_1 \tau - \sigma_1 \sqrt{\tau} = r_{1,2} + m_1 \tau + \sigma_1 \sqrt{\tau},$$

we have

$$\sigma_1 = \frac{r_{1,1} - r_{1,2}}{2\sqrt{\tau}}.$$

Now using $r_{1,1}$ and $r_{1,2}$ from above gives

$$\begin{aligned} \sigma_1 &= \frac{2\sigma_0 \sqrt{\tau}}{2\sqrt{\tau}} \\ &= \sigma_0. \end{aligned}$$

The recombination condition has put the stringent condition on the HL binomial tree that σ is constant over time. That is, $\sigma_k = \sigma$ for $1 \leq k \leq n$ in the HL binomial model. We must use this condition in Chapter 4, when we build the HL binomial tree.

For the HL binomial (recombining) tree, it is required that

$$r_{k+1,j} = r_{k,j} + m_k \tau + \sigma \sqrt{\tau} \quad (58a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} + m_k \tau - \sigma \sqrt{\tau} \quad (58b)$$

for a down-move. Also note that

$$r_{k+1,j} - r_{k,j} = m_k \tau + \sigma \sqrt{\tau}$$

and

$$r_{k+1,j} - r_{k+1,j+1} = 2\sigma \sqrt{\tau}.$$

That is, the HL binomial short rates change by the drift and a constant (random) disturbance, and the difference between short rates between two adjacent levels at any time is also a constant. Specifically, the spread in the up-state and the down-state grows linearly over time because

$$r_{k,1} - r_{k,k+1} = 2k\sigma \sqrt{\tau}.$$

We now derive the equation that the short rates $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$ must satisfy. We do this again by applying the no-arbitrage requirement. From the corresponding prices for $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$, we have

$$\begin{aligned} A_3 &= \frac{1}{(1 + R_3 \tau)^3} \\ &= \frac{qp_{1,1} + (1-q)p_{1,2}}{1 + r_0 \tau} \\ &= \frac{q \left[\frac{qp_{2,1} + (1-q)qp_{2,2}}{1 + r_{1,1} \tau} \right] + (1-q) \left[\frac{qp_{2,2} + (1-q)qp_{2,3}}{1 + r_{1,2} \tau} \right]}{1 + r_0 \tau}, \end{aligned}$$

which, after clearing fractions, simplifies to

$$\begin{aligned} &A_3(1 + r_0 \tau)(1 + r_{1,1} \tau)(1 + r_{1,2} \tau) - q^2(1 + r_{1,2} \tau)p_{2,1} \\ &- q(1-q)[(1 + r_{1,1} \tau) + (1 + r_{1,2} \tau)]p_{2,2} - (1-q)^2(1 + r_{1,1} \tau)p_{2,3} = 0. \end{aligned}$$

Substituting for $p_{2,1}$, $p_{2,2}$, and $p_{2,3}$ with the discount factors $p_{2,k} = 1/(1 + r_{2,k} \tau)$ for $1 \leq k \leq 3$ and again clearing fractions, we have

$$A_3(1+r_{0\tau})(1+r_{1,1\tau})(1+r_{1,2\tau})(1+r_{2,1\tau})(1+r_{2,2\tau})(1+r_{2,3\tau}) \\ -q^2(1+r_{1,2\tau})(1+r_{2,2\tau})(1+r_{2,3\tau})-q(1-q)[(1+r_{1,1\tau})+(1+r_{1,2\tau})] \\ \times (1+r_{2,1\tau})(1+r_{2,3\tau})-(1-q)^2(1+r_{1,1\tau})(1+r_{2,1\tau})(1+r_{2,2\tau})=0. \quad (59)$$

We now substitute in the values for $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$ from Equation 57 or Equation 58 to get a cubic equation in m_1 . We again use the bisection method to find m_1 that makes economic sense and determine $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$ from Equation 57 or Equation 58, using the fact that we already determined $r_{1,1}$ and $r_{1,2}$ in Equation 47.

We now give the algorithm for the HL short rates. Note that Equation 47 can be written as

$$c_{1,1}(1+r_{1,1\tau})(1+r_{1,2\tau})+c_{1,2}(1+r_{1,2\tau})+c_{1,3}(1+r_{1,1\tau})=0 \quad (60)$$

and that Equation 59 can be written as

$$c_{2,1}(1+r_{2,1\tau})(1+r_{2,2\tau})(1+r_{2,3\tau})+c_{2,2}(1+r_{2,2\tau})(1+r_{2,3\tau}) \\ +c_{2,3}(1+r_{2,1\tau})(1+r_{2,3\tau})+c_{2,4}(1+r_{2,1\tau})(1+r_{2,2\tau})=0. \quad (61)$$

We now introduce some variables that will help to generate the coefficients for the polynomials that determine the interest rates at time period k or time $t_k = k\tau$. This process is done in two steps. The first step is to notice how the coefficients, $c_{k,j}$, are related to the interest rates at the previous time periods. Note that if $a_{1,1} = 1 + r_{0\tau}$, $a_{1,2} = -1$, and $a_{1,3} = -1$, then $c_{1,1} = A_2 a_{1,1}$, $c_{1,2} = q a_{1,2}$, and $c_{1,3} = (1-q) a_{1,3}$. In the next step, we have $b_{1,1} = a_{1,2}(1+r_{1,2\tau})$, $b_{1,2} = a_{1,3}(1+r_{1,1\tau})$, $a_{2,1} = (1+r_{0\tau})(1+r_{1,1\tau})(1+r_{1,2\tau})$, $a_{2,2} = b_{1,1}$, $a_{2,3} = b_{1,1} + b_{1,2}$, and $a_{2,4} = b_{1,2}$. After substituting back, $c_{2,1} = A_3 a_{2,1}$, $c_{2,2} = q^2 a_{2,2}$, $c_{2,3} = q(1-q) a_{2,3}$, and $c_{2,4} = (1-q)^2 a_{2,4}$. Let $b_{2,1} = a_{2,2}(1+r_{2,2\tau})(1+r_{2,3\tau})$, $b_{2,2} = a_{2,3}(1+r_{2,1\tau})(1+r_{2,3\tau})$, and $b_{2,3} = a_{2,4}(1+r_{2,1\tau})(1+r_{2,2\tau})$, and continue the process. We can generalize this process for $k \geq 2$. Let $a_{k,1} = \prod_{n=0}^{k-1} \prod_{j=1}^{n+1} (1+r_{n,j\tau})$, $a_{k,2} = b_{k-1,1}$, $a_{k,j} = b_{k-1,j-2} + b_{k-1,j-1}$ for $j = 3, \dots, k+1$, $a_{k,k+2} = b_{k-1,k}$, and $c_{k,1} = A_{k+1} a_{k,1}$, $c_{k,j+1} = q^{k-j+1} (1-q)^{j-1} a_{k,j+1}$ for $j = 1, \dots, k+1$. Then $b_{k,j} = a_{k,j+1} \prod_{\substack{n=1 \\ n \neq j}}^{k+1} (1+r_{k,n\tau})$ for $j = 1, \dots, k+1$. This process will generate the coefficients $c_{k,j}$ for $j = 1, \dots, k+2$ in the k th degree polynomial

$$f_k = c_{k,1} \prod_{j=1}^{k+1} (1+r_{k,j\tau}) + \sum_{j=1}^{k+1} c_{k,j+1} \prod_{\substack{n=1 \\ n \neq j}}^{k+1} (1+r_{k,n\tau}). \quad (62)$$

Solving $f_k = 0$ determines m_{k-1} and thus $r_{k,j}$ for $j = 1, \dots, k + 1$ at the k th period. Again the bisection method will converge quickly because there is usually only one root, m_{k-1} , of Equation 62 between -1.0 and 1.0 .

Kalotay–Williams–Fabozzi Formulas for Short Rates. The reason we do not substitute the formulas for the short rates in Equations 60–62 is that these equations will also generate the formulas for the short rates for the KWF binomial tree. For this model,

$$r_{k+1,j} = r_{k,j} e^{m_k \tau + \sigma_k \sqrt{\tau}} \quad (63a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} e^{m_k \tau - \sigma_k \sqrt{\tau}} \quad (63b)$$

for a down-move. These equations correspond to Equations 50. From Equation 63 or 50, we have

$$r_{1,1} = r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}}$$

and

$$r_{1,2} = r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}}.$$

From Equation 63, the short rates for the next time period are

$$\begin{aligned} r_{2,1} &= r_{1,1} e^{m_1 \tau + \sigma_1 \sqrt{\tau}}, \\ r_{2,2} &= r_{1,1} e^{m_1 \tau - \sigma_1 \sqrt{\tau}} = r_{1,2} e^{m_1 \tau + \sigma_1 \sqrt{\tau}}, \end{aligned}$$

and

$$r_{2,3} = r_{1,2} e^{m_1 \tau - \sigma_1 \sqrt{\tau}}.$$

It is simple to show that the same steps used in the HL formulas lead to a KWF binomial recombination condition that gives the same requirement as the HL binomial recombination condition. That is, the volatility must be constant.

Thus, we have

$$r_{k+1,j} = r_{k,j} e^{m_k \tau + \sigma \sqrt{\tau}} \quad (64a)$$

and

$$r_{k+1,j+1} = r_{k,j} e^{m_k \tau - \sigma_k \sqrt{\tau}} \quad (64b)$$

and

$$\frac{r_{k,1}}{r_{k,k+1}} = e^{2k\sigma\sqrt{\tau}}.$$

Therefore, we see that the spread between the up- and down-most states in the KWF binomial lattice grows exponentially with time. To generate these KWF short rates, we replace the short rates in Equations 60–62 with the formulas given by Equation 64 and solve Equations 60–62 for m_{k-1} using the bisection method. Therefore, the basic program for the HL binomial short rates is easily modified to handle the KWF binomial short rates.

Black–Derman–Toy Formulas for Short Rates. In the BDT model, the expressions for $r_{k,j}$ corresponding to Equation 54 are

$$r_{k+1,j} = r_{k,j} e^{m_{k,j} \tau + \sigma_k \sqrt{\tau}} \quad (65a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} e^{m_{k,j} \tau - \sigma_k \sqrt{\tau}} \quad (65b)$$

for a down-move. For example, we get Equation 54

$$r_{1,1} = r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}}$$

and

$$r_{1,2} = r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}},$$

and, for $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$,

$$\begin{aligned} r_{2,1} &= r_{1,1} e^{m_{1,1} \tau + \sigma_1 \sqrt{\tau}}, \\ r_{2,2} &= r_{1,1} e^{m_{1,1} \tau - \sigma_1 \sqrt{\tau}} = r_{1,2} e^{m_{1,2} \tau + \sigma_1 \sqrt{\tau}}, \end{aligned}$$

and

$$r_{2,3} = r_{1,2} e^{m_{1,2} \tau - \sigma_1 \sqrt{\tau}}. \quad (66)$$

The second line in Equation 66 is the recombination condition, which leads to

$$r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}} e^{m_{1,1} \tau - \sigma_1 \sqrt{\tau}} = r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}} e^{m_{1,2} \tau + \sigma_1 \sqrt{\tau}}.$$

Simplifying gives

$$2\sigma_0 \sqrt{\tau} + (m_{1,1} - m_{1,2})\tau = 2\sigma_1 \sqrt{\tau}$$

or

$$\sigma_1 = \frac{2\sigma_0 \sqrt{\tau} + (m_{1,1} - m_{1,2})\tau}{2\sqrt{\tau}}.$$

We see that if $m_{1,1} = m_{1,2}$, then $\sigma_1 = \sigma_0$. Because we want the local volatility to vary over time in the BDT model, the drift m must be a function of both time and level. In building the BDT binomial model, the drift must have the form $m_{k,j}$.

Consider another consequence of the recombination requirement. Equation 66 also gives

$$\frac{r_{2,1}}{r_{2,2}} = e^{2\sigma_1 \sqrt{\tau}}$$

or

$$\sigma_1 = \frac{\ln(r_{2,1}) - \ln(r_{2,2})}{2\sqrt{\tau}} \quad (67)$$

and

$$\frac{r_{2,1}}{r_{1,1}} = e^{m_{1,1} \tau} e^{\sigma_1 \sqrt{\tau}}$$

or

$$m_{1,1} = \frac{\ln(r_{2,1}) - \ln(r_{1,1}) - \sigma_1 \sqrt{\tau}}{\tau}. \quad (68)$$

Similarly,

$$m_{1,2} = \frac{\ln(r_{2,2}) - \ln(r_{1,2}) - \sigma_1 \sqrt{\tau}}{\tau}. \quad (69)$$

We see that from $r_{k,j}$ for $1 \leq j \leq k+1$, we can determine the local volatility, σ_{k-1} , and the drift rates, $m_{k-1,j}$, for $1 \leq j \leq k$. The general formulas are the obvious extensions of Equations 67–69.

As pointed out in the discussion of the up-state/down-state formulas for the Black–Derman–Toy model, another equation is needed in order to incorporate yield volatility, s_k . Yield volatility is incorporated by using the market yield volatility structure, s , the up-state and down-state yields, y_u and y_d , and prices for time period k . The conditions are

$$y_u = y_d e^{2s_{k-1}\sqrt{\tau}} \quad (70)$$

and

$$A_{k+1} = \frac{qp_{k,u} + (1-q)p_{k,d}}{1+r_0\tau}, \quad (71)$$

where $p_{k,u} = 1/(1+y_{k,u}\tau)^k$ is the up-price and $p_{k,d} = 1/(1+y_{k,d}\tau)^k$ is the down-price. This incorporation of yield volatility is illustrated in **Figure 19**.

Consider the case for $k = 2$. Combining Equations 70–71 and the up- and down-prices leads to

$$\begin{aligned} A_3(1+r_0\tau) \left((1+y_{2,d}\tau e^{2s_1\sqrt{\tau}})^2 (1+y_{2,d}\tau)^2 \right. \\ \left. - q(1+y_{2,d}\tau)^2 - (1-q)(1+y_{2,d}\tau e^{2s_1\sqrt{\tau}})^2 \right) = 0. \end{aligned} \quad (72)$$

Notice that Equation 72 is similar to Equations 55 and 60. The solution for $y_{k,d}$ is obtained from Equation 72 using the bisection method and the fact that $0 < y_{k,d} < 1$.

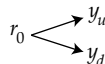
We now determine the BDT rates $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$. We use $p_{2,u}$ to solve for the up-state BDT rates and $p_{2,d}$ to solve for the down-state rates. The no-arbitrage equations are

$$p_{2,u}(1+r_{1,1}\tau)(1+r_{2,1}\tau)(1+r_{2,2}\tau) - q(1+r_{2,2}\tau) - (1-q)(1+r_{2,1}\tau) = 0 \quad (73a)$$

and

$$p_{2,d}(1+r_{1,2}\tau)(1+r_{2,2}\tau)(1+r_{2,3}\tau) - q(1+r_{2,3}\tau) - (1-q)(1+r_{2,2}\tau) = 0, \quad (73b)$$

Figure 19. Binomial Tree Incorporating Volatility



and from Equation 65,

$$\frac{r_{2,1}}{r_{2,2}} = \frac{r_{2,2}}{r_{2,3}}$$

or

$$r_{2,2}^2 = r_{2,1}r_{2,3}. \quad (73c)$$

Equations 73a, 73c, and 73b are solved for the rates $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$ by the Newton–Raphson method (in this order for ease of computation). We then get the drifts and the local volatility from Equations 67–69.

The process for obtaining the BDT binomial short rates is to first determine the up-price, $p_{k,u}$, and the down-price, $p_{k,d}$, from Equation 72. Equation 72 is easily extended for all k . These two equations for the up-price and down-price are solved using the bisection method. Equation 73 is also easily extended. At time $t_k = k\tau$, we will have two equations similar to Equations 73a and 73b for the up-price and the down-price from the no-arbitrage requirement and $k-1$ equations similar to Equation 73c. Because Equations 73a, 73c, and 73b are a system of $k+1$ nonlinear equations for the BDT short rates, the bisection method cannot be used. Instead, the Newton–Raphson method for systems of nonlinear equations must be used.

Hull–White Formulas for Short Rates. In the HW binomial tree, the expressions for $r_{k,j}$ that correspond to Equation 48 are

$$r_{k+1,j} = r_{k,j} + m_k\tau - \phi_k r_{k,j}\tau + \sigma_k\sqrt{\tau} \quad (74a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} + m_k\tau - \phi_k r_{k,j}\tau - \sigma_k\sqrt{\tau} \quad (74b)$$

for a down-move. For example, Equation 48 is

$$r_{1,1} = r_0 + m_0\tau - \phi_0 r_0\tau + \sigma_0\sqrt{\tau}$$

and

$$r_{1,2} = r_0 + m_0\tau - \phi_0 r_0\tau - \sigma_0\sqrt{\tau},$$

and the equations for the HW short rates are

$$r_{2,1} = r_{1,1} + m_1\tau - \phi_1 r_{1,1}\tau + \sigma_1\sqrt{\tau},$$

$$r_{2,2} = r_{1,1} + m_1\tau - \phi_1 r_{1,1}\tau - \sigma_1\sqrt{\tau} = r_{1,2} + m_1\tau - \phi_1 r_{1,2}\tau + \sigma_1\sqrt{\tau},$$

and

$$r_{2,3} = r_{1,2} + m_1\tau - \phi_1 r_{1,2}\tau - \sigma_1\sqrt{\tau}. \quad (75)$$

The recombination condition is

$$(r_{1,1} - r_{1,2})(1 - \phi_1\tau) = 2\sigma_1\sqrt{\tau}.$$

Using $r_{1,1}$ and $r_{1,2}$ from above gives

$$2\sigma_0\sqrt{\tau}(1 - \phi_1\tau) = 2\sigma_1\sqrt{\tau}, \quad (76)$$

which we write as

$$(1 - \phi_1\tau) = \frac{\sigma_1}{\sigma_0}.$$

This says that ϕ_1 must be chosen to satisfy

$$\phi_1 = \frac{1 - (\sigma_1/\sigma_0)}{\tau}. \quad (77)$$

Equation 77 indicates that if $\sigma_1 = \sigma_0$, then $\phi_1 = 0$. This is clearly not acceptable. In order to incorporate mean reversion for a constant local volatility, the time step must be allowed to vary with k in Equation 74. This gives us

$$r_{k+1,j} = r_{k,j} + m_k\tau_k - \phi_k r_{k,j}\tau_k + \sigma_k\sqrt{\tau_k} \quad (78a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} + m_k\tau_k - \phi_k r_{k,j}\tau_k - \sigma_k\sqrt{\tau_k} \quad (78b)$$

for a down-move. For example, Equation 48 is

$$r_{1,1} = r_0 + m_0\tau_0 - \phi_0 r_0\tau_0 + \sigma_0\sqrt{\tau_0}$$

and

$$r_{1,2} = r_0 + m_0\tau_0 - \phi_0 r_0\tau_0 - \sigma_0\sqrt{\tau_0},$$

and the equations for the HW short rates for the next time period are

$$\begin{aligned} r_{2,1} &= r_{1,1} + m_1\tau_1 - \phi_1 r_{1,1}\tau_1 + \sigma_1\sqrt{\tau_1}, \\ r_{2,2} &= r_{1,1} + m_1\tau_1 - \phi_1 r_{1,1}\tau_1 - \sigma_1\sqrt{\tau_1} = r_{1,2} + m_1\tau_1 - \phi_1 r_{1,2}\tau_1 + \sigma_1\sqrt{\tau_1}, \end{aligned}$$

and

$$r_{2,3} = r_{1,2} + m_1\tau_1 - \phi_1 r_{1,2}\tau_1 - \sigma_1\sqrt{\tau_1}. \quad (79)$$

The recombination condition is

$$(r_{1,1} - r_{1,2})(1 - \phi_1\tau_1) = 2\sigma_1\sqrt{\tau_1}.$$

Using $r_{1,1}$ and $r_{1,2}$ from above gives

$$2\sigma_0\sqrt{\tau_0}(1 - \phi_1\tau_1) = 2\sigma_1\sqrt{\tau_1},$$

which upon some algebraic manipulation becomes

$$\sigma_0\sqrt{\tau_0}\phi_1\tau_1 + \sigma_1\sqrt{\tau_1} - \sigma_0\sqrt{\tau_0} = 0.$$

This is a quadratic equation in $\sqrt{\tau_1}$ whose solution is

$$\tau_1 = \tau_0 \frac{4(\sigma_0/\sigma_1)^2}{\left[1 + \sqrt{1 + 4(\sigma_0/\sigma_1)^2\tau_0\phi_1}\right]^2}. \quad (80)$$

For example, if $\sigma_1 = \sigma_0$, $\tau_0 = \tau = 0.5$, and the mean reversion is $\phi_1 = 0.01$, then $\tau_1 \cong 0.495$. From Equation 80, it is seen that increasing ϕ will decrease the time step, τ_1 . We can also use Equation 80 to solve for the mean reversion, ϕ_1 , or the local volatility, σ_1 . Doing this gives us

$$\phi_1 = \frac{\sigma_0\sqrt{\tau_0} - \sigma_1\sqrt{\tau_1}}{\sigma_0\sqrt{\tau_0}\tau_1} \quad (81)$$

and

$$\sigma_1 = \frac{\sigma_0\sqrt{\tau_0}(1 - \phi_1\tau_1)}{\sqrt{\tau_1}}. \quad (82)$$

In order to keep the time step constant at τ and let the local volatility vary, then, according to Equation 81, the mean reversion must be

$$\phi_1 = \frac{\sigma_0 - \sigma_1}{\sigma_0\tau}. \quad (83)$$

If the mean reversion is to be chosen and the time step is to be kept constant at τ , then, according to Equation 82, the local volatility must be

$$\sigma_1 = \sigma_0(1 - \phi_1\tau). \quad (84)$$

In Equations 80, 83, and 84, it is easily shown that the subscript 0 can be replaced by $k - 1$, the subscript 1 can be replaced by k , and the subscript 2 by $k + 1$. Using these replacements in Equation 80 gives

$$\tau_{k+1} = \tau_k \frac{4(\sigma_k/\sigma_{k+1})^2}{\left[1 + \sqrt{1 + 4(\sigma_k/\sigma_{k+1})^2 \tau_k \phi_{k+1}}\right]^2} \quad (85)$$

for the variable time step. Similarly, Equation 83 becomes

$$\phi_{k+1} = \frac{\sigma_k - \sigma_{k+1}}{\sigma_k \tau}, \quad (86)$$

and Equation 84 is now

$$\sigma_{k+1} = \sigma_k(1 - \phi_{k+1}\tau). \quad (87)$$

Black–Karasinski Formulas for Short Rates. The BK model expressions for $r_{k,j}$ that correspond to Equation 52 are

$$r_{k+1,j} = r_{k,j} e^{m_k \tau_k - \phi_k \ln(r_{k,j}) \tau_k + \sigma_k \sqrt{\tau_k}} \quad (88a)$$

for an up-move and

$$r_{k+1,j+1} = r_{k,j} e^{m_k \tau_k - \phi_k \ln(r_{k,j}) \tau_k - \sigma_k \sqrt{\tau_k}} \quad (88b)$$

for a down-move. Showing that the recombination conditions for the BK binomial tree are the same as for the HW recombination condition is straightforward. In order to obtain Equations 80–87, simply follow the steps for the HL recombination condition, the steps for the BDT recombination condition, or the steps for the HW recombination condition with Equation 88. Consequently, the BK binomial algorithm is obtained by modifying the HW binomial algorithm, which is obtained by modifying the HL binomial algorithm. Because the BK SDE is a lognormal process, one can build a BK binomial tree that parallels the BDT binomial tree. The variable time step must be incorporated in these binomial trees (as addressed in Appendix A). If one always chooses the mean reversion and the input volatility to satisfy Equation 86 or 87, then the time step is always constant. Such a choice, however, places a strong restriction on the HW and BK models. These two equations show that

an increasing input volatility structure requires a negative mean reversion to get a constant time step.

Another problem with a variable time step is that market pricing data are difficult to acquire at the varying time steps. Therefore, we have to interpolate the known market prices to get the values at the variable time step in order to implement the no-arbitrage condition. Additionally, the short rates given by the binomial model will be at the variable time steps, and this tree will have to be interpolated to get the short rate at the fixed time steps. Consequently, meeting the no-arbitrage condition is more difficult. We iterate on the HW and BK solutions until a time step is achieved that almost matches the time steps of the known market data. In the next chapter, we look at some examples that highlight these difficulties.

Trinomial Framework. One way to avoid the issues of a variable time step is to use a trinomial tree for the HW and BK SDEs. Adding a third rate allows for another degree of freedom and releases the requirement of the variable time step. The trinomial trees are based on the trinomial version of Equation 10 at the end of Chapter 1.

The nodes in the trinomial tree will be labeled $r_{k,j}$ as they are in the binomial tree. The rates in the trinomial tree must follow the expected value and the variance of the HW and BK SDEs. The trinomial tree must recombine just as the binomial tree did, and the recombination conditions are derived in a similar manner, except now three moves are possible instead of two. The trinomial tree should also give a spread in the rates as the binomial tree did. The price tree is built in an analogous manner from the trinomial rate tree as the binomial price tree was built (Equation 45), except now three prices and three probabilities exist.

The nodes in the HW trinomial tree would be denoted

$$r_u = r_{k+1,j} = r_{k,j} + \delta_{k,j} + \alpha_k$$

for an up-move,

$$r_m = r_{k+1,j+1} = r_{k,j} + \delta_{k,j}$$

for a middle-move, and

$$r_d = r_{k+1,j+2} = r_{k,j} + \delta_{k,j} - \alpha_k$$

for a down-move, where $\delta_{k,j}$ would approximate the expected value component, α_k would approximate the distributional component of the HW SDE (Equation 18), and the index j would run from 1 to $2k + 1$. The trinomial trees would be developed using these equations in a manner analogous to the development in this chapter. The BK trinomial tree would use the exponentiation of this equation as it did in the binomial case presented in this chapter.

4. Examples of the Binomial Models

This chapter presents some specific examples of binomial lattices. Obviously, all of the interesting phenomena that occur in these models cannot be covered. So, we provide some examples that highlight the more salient financial and mathematical properties of the processes introduced in Chapters 2 and 3. We also offer some examples to show why trinomial models for the Hull–White (HW) and Black–Karasinski (BK) processes are necessary. Finally, we set $q = 0.5$.¹¹

In this chapter, all short rates are assumed to be semiannual rates. That is, we use $\tau = 0.5$ as our fixed time step. The bonds modeled are five-year bonds. Thus, there are 10 time periods. Time is always measured in years.

Because the Ho–Lee (HL), Kalotay–Williams–Fabozzi (KWF), and Black–Derman–Toy (BDT) models can use the constant time step $\tau = 0.5$, we consider examples of these models first. We analyze the models using the same inputs so that we can compare the differences between the models. For this reason, we have more than one model, which allows the user to select the one that either best captures a particular interest rate environment or can most easily be used to evaluate securities (e.g., caps and floors can be priced using the implied cap curve so that only models that allow for control over local volatility make sense). Next, we explain similar models for the HW and BK processes that explicitly incorporate mean reversion. Because the HL and KWF processes can be obtained from the HW and BK processes, respectively, by setting $\phi = 0$, the same is true of the algorithms for generating the binomial trees. We do these comparisons for completeness and to show the development process involved in binomial models. The chapter concludes with examples of the BK model using Equation 62, the BDT method using Equation 73, and a comparison of the HW and BK binomial models. The latter comparison is accomplished by incorporating a constant time step through Equation 86. That is, we choose ϕ in the BK process to guarantee a constant time step.

¹¹See Jarrow (1996) for a more thorough explanation of the use of risk neutrality in term-structure modeling. More rigorous development can be found in Harrison and Pliska (1981) and Harrison and Kreps (1979). Realize that $q = 0.5$ is not a necessary or sufficient condition for risk neutrality. In order to get the expected value and variance in the binomial trees, however, q must be 0.5.

Table 2 and **Table 3** show the input data for our examples. Table 2 shows the three term structure of spot rates. FTS is a flat term structure of 5 percent. ITS is a linearly increasing term structure with $R(0) = 0.05$ and a slope of 0.005. DTS is a decreasing term structure with $R(0) = 0.05$ and a slope of -0.005 . We have chosen $R(0) = 0.05$ so that a comparison can be made between the binomial trees with the examples presented in Chapter 2 for each of the five models. Because the HL and KWF models require a constant volatility, we define the volatility in the examples. The volatility structures in Table 2 are for the BDT, HW, and BK models. LDV20 is a linearly decreasing volatility starting at 20 percent with a slope of -0.025 per year. LIV is a linearly increasing volatility starting at 10 percent with a slope of 0.005 per year. LDV is a linearly decreasing volatility starting at 10 percent with a slope of -0.005 per year. EIV is an exponentially increasing volatility given by $0.1e^{0.1t}$. EDV is an exponentially decreasing volatility given by $0.1e^{-0.1t}$.

Users of the HW and BK models are free to choose the volatility and the mean reversion, but then the time steps have to be given by Equation 85. In Table 3, we have presented what ϕ has to be in order to keep the time step constant at $\tau = 0.5$ in the HW and BK models for each of LDV20, LIV, LDV, EIV, and EDV. In Table 3, MRLDV20 (mean reversion with linearly decreasing volatility starting at 20 percent) corresponds to LDV20. The other mean reversion structures in Table 3 correspond to the other volatility structures (e.g., MREIV is the mean reversion structure required for the exponentially increasing volatility structure, etc.). Recall that ϕ has to be negative for an

Table 2. The Term Structures and Volatility Structures for the Binomial Examples

| Time (years) | Spot Rates (percent) | | | Volatilities (percent) | | | | |
|-----------------|-------------------------|------|------|---------------------------|---------|---------|---------|---------|
| | FTS | ITS | DTS | LDV20 | LIV | LDV | EIV | EDV |
| 0.0 | 5.0 | 5.00 | 5.00 | 20.0000 | 10.0000 | 10.0000 | 10.0000 | 10.0000 |
| 0.5 | 5.0 | 5.25 | 4.75 | 19.8750 | 10.2500 | 9.7500 | 10.5127 | 9.5123 |
| 1.0 | 5.0 | 5.50 | 4.50 | 19.7500 | 10.5000 | 9.5000 | 11.0517 | 9.0484 |
| 1.5 | 5.0 | 5.75 | 4.25 | 19.6250 | 10.7500 | 9.2500 | 11.6183 | 8.6071 |
| 2.0 | 5.0 | 6.00 | 4.00 | 19.5000 | 11.0000 | 9.0000 | 12.2140 | 8.1873 |
| 2.5 | 5.0 | 6.25 | 3.75 | 19.3750 | 11.2500 | 8.7500 | 12.8403 | 7.7880 |
| 3.0 | 5.0 | 6.50 | 3.50 | 19.2500 | 11.5000 | 8.5000 | 13.4986 | 7.4082 |
| 3.5 | 5.0 | 6.75 | 3.25 | 19.1250 | 11.7500 | 8.2500 | 14.1907 | 7.0469 |
| 4.0 | 5.0 | 7.00 | 3.00 | 19.0000 | 12.0000 | 8.0000 | 14.9182 | 6.7032 |
| 4.5 | 5.0 | 7.25 | 2.75 | 18.8750 | 12.2500 | 7.7500 | 15.6831 | 6.3763 |

Table 3. The Mean Reversions Required to Keep a Constant Time Step of 0.5 Years for the Volatility Structures in Table 2 (percent)

| Time (years) | MRLDV20 | MRLIV | MRLDV | MREIV | MREDV |
|-----------------|---------|-------|-------|--------|-------|
| 0.0 | 1.00 | 5.00 | 5.00 | -10.25 | 9.75 |
| 0.5 | 1.25 | -5.00 | 5.00 | -10.25 | 9.75 |
| 1.0 | 1.26 | -4.88 | 5.13 | -10.25 | 9.75 |
| 1.5 | 1.27 | -4.76 | 5.26 | -10.25 | 9.75 |
| 2.0 | 1.27 | -4.65 | 5.41 | -10.25 | 9.75 |
| 2.5 | 1.28 | -4.55 | 5.56 | -10.25 | 9.75 |
| 3.0 | 1.29 | -4.44 | 5.71 | -10.25 | 9.75 |
| 3.5 | 1.30 | -4.35 | 5.88 | -10.25 | 9.75 |
| 4.0 | 1.31 | -4.26 | 6.06 | -10.25 | 9.75 |
| 4.5 | 1.32 | -4.17 | 6.25 | -10.25 | 9.75 |

increasing volatility structure. We present some examples with these mean reversions. It is interesting to note that for an exponentially changing volatility structure, the time step is constant for a constant mean reversion. One can use Equations 86 and 87 to find this relationship. This sort of analysis can be used to model the implied volatility from cap prices, which is known as the cap curve.

Examples for HL, KWF, and BDT Processes

Figure 20 presents four examples of the HL process that verify the results we developed in Chapter 2 and presented in Figure 2. Figure 20 also includes a price tree. Panel A of Figure 20 shows a binomial tree for the FTS with a constant volatility of 1 percent. Negative interest rates are possible, even with this small volatility, because, as shown by Equation 58 in Chapter 3, the spread in the HL binomial short rates at $t = t_k$ is given by

$$r_{k,1} - r_{k,k+1} = 2k\sigma\sqrt{\tau}.$$

For example, in Panel A of Figure 20 at $t = t_7 = 3.5$ years,

$$\begin{aligned} r_{7,1} - r_{7,8} &= 0.1002 - 0.0012 \\ &= 0.099 \approx 2(7)(0.01)\sqrt{0.5} \approx 0.098995. \end{aligned}$$

The difference is caused by the round-off in the results presented in Panel A. For the FTS with a constant volatility of 10 percent, Panel B of Figure 20 shows that the magnitude of the rates is greater and that there are more negative rates, as expected (and as demonstrated in Figure 2). In Panel C of Figure 20,

Figure 20. HL Binomial Trees

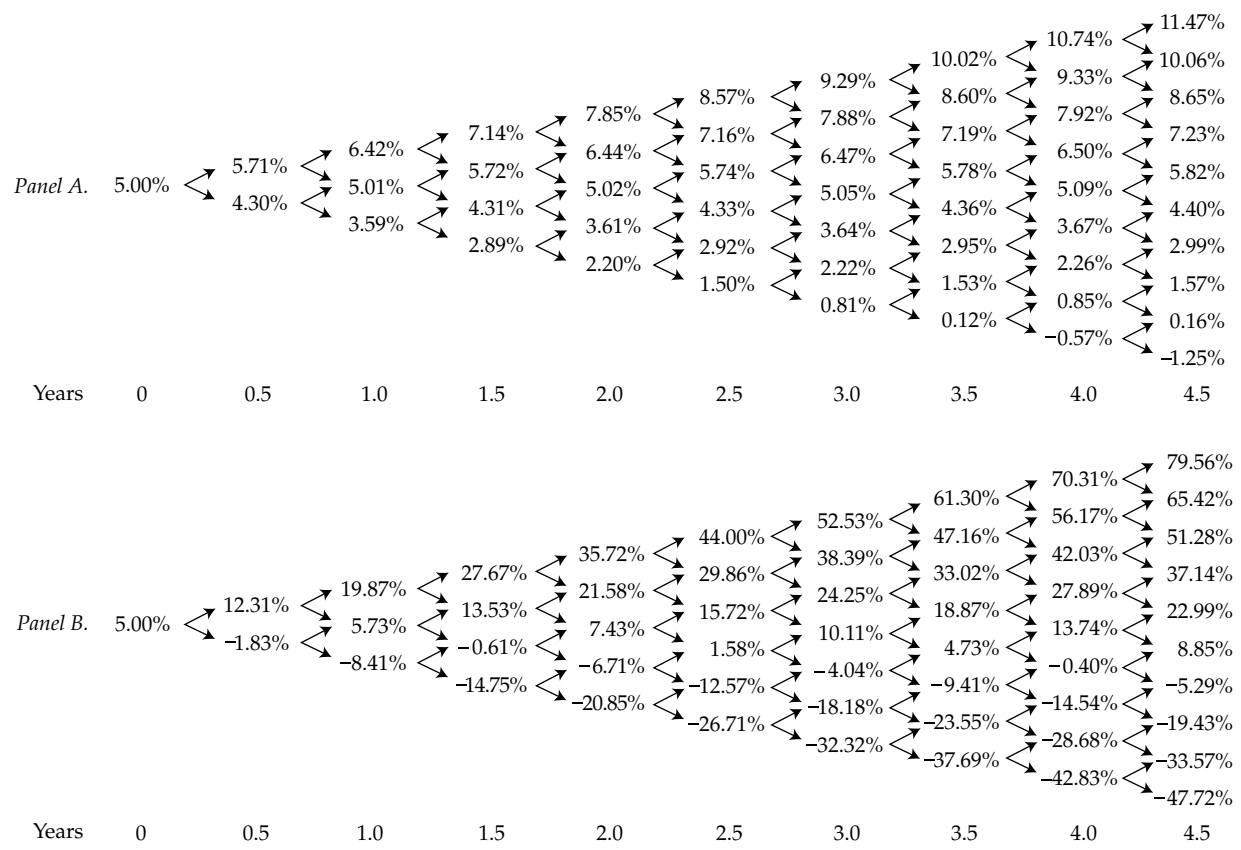


Figure 20. HL Binomial Trees

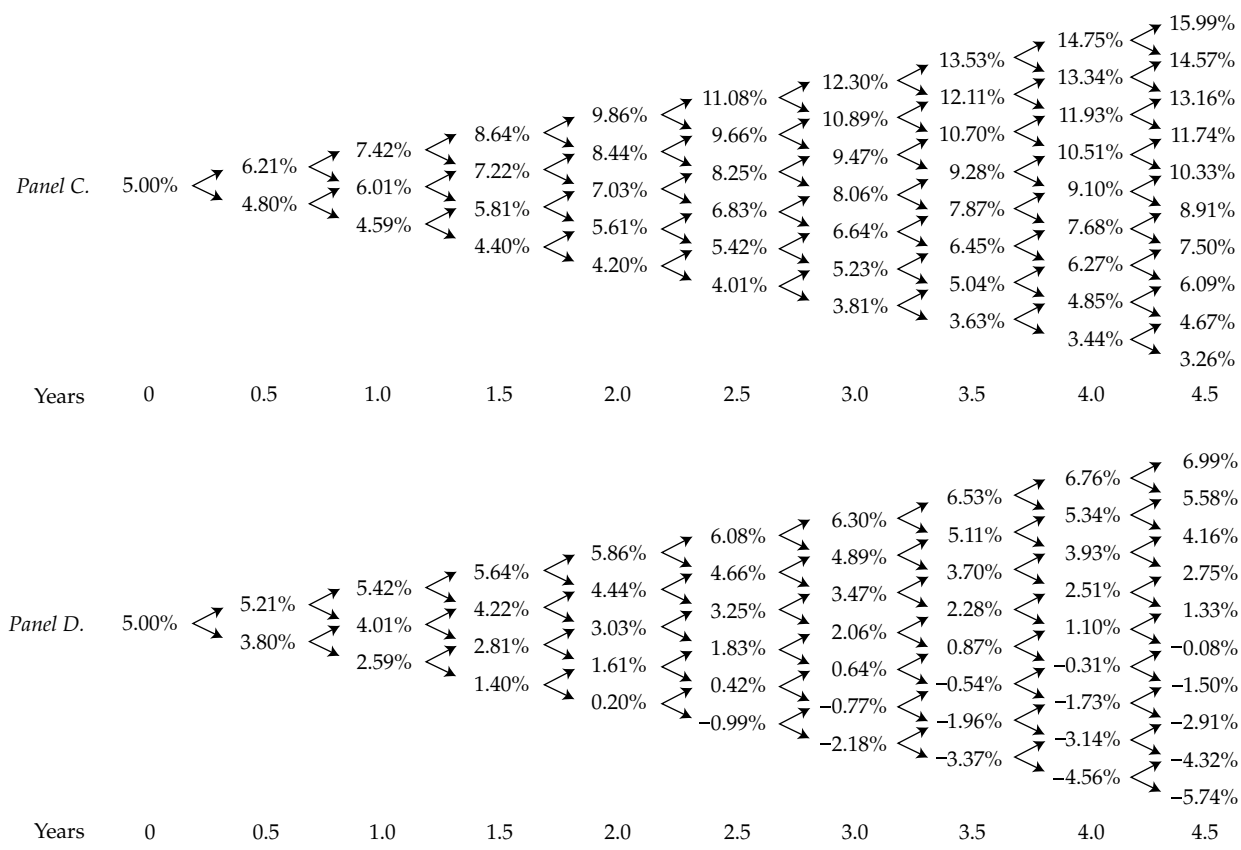
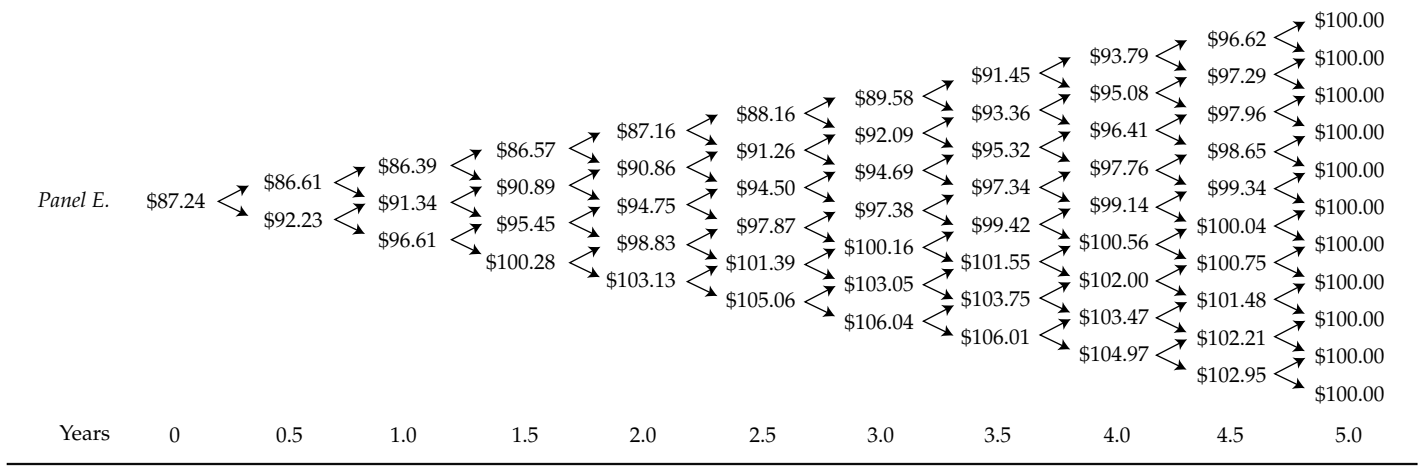


Figure 20. HL Binomial Trees



we have a binomial tree for the ITS with a constant volatility of 1 percent, and we see that, as expected from the results in Chapter 2, there are no negative interest rates. Because the drift is related to the slope and the slope is positive, we expect the rates to grow in the HL process. Panel D of Figure 20, which is a binomial tree for the DTS with a constant volatility of 1 percent, has more negative interest rates than Panel A because Panel D has a negative drift and thus the rates should decrease over time. As pointed out in Chapter 3, the negative rates are not completely detrimental, because we average over the rates to get the price of the bond. As a matter of fact, as long as the HL rates satisfy

$$(1 + r_{k,j\tau}) > 0,$$

they can be used to price bonds and contingent claims. Panel E of Figure 20 presents the corresponding price tree for Panel D, which shows that the price of the bond is given by the final spot rate. That is, the price of the bond is

$$\frac{100}{[1 + R(10)\tau]^{10}} = \frac{100}{[1 + (0.0275)(0.5)]^{10}} \approx \$87.24.$$

Figure 21 presents the lognormal version of the results in Figure 20. That is, it presents KWF binomial trees. We see that the rates are positive but that they grow over time. But the growth and the spread in the rates are much smaller than in the HL case, even for the higher volatilities. Panel A of Figure 21 presents the binomial tree for the same data used to create Panel B of Figure 20. In Panel B of Figure 21, a constant volatility of 20 percent is used, leading to an increase in the value of the up-state rates and a decrease in the value of the down-state rates. Remember that the spread in the rates follows from Equation 64 in Chapter 3. That is,

$$\frac{r_{k,1}}{r_{k,k+1}} = e^{2k\sigma\sqrt{\tau}}.$$

Panel C of Figure 21 presents the results for the ITS with a volatility of 10 percent, and the rates are larger throughout the lattice. Panel D of Figure 21 presents the results for the DTS, with rates that are significantly smaller and approaching zero. These trees verify the results developed in Chapter 2. Note that the price tree (Panel E) corresponding to the binomial tree in Panel D in Figure 21 gives the same price as the HL price tree in Panel E of Figure 20, as guaranteed by the no-arbitrage property. Also note that the prices in the KWF tree are all under \$100, and the prices in the KWF tree have much less spread than the prices in the HL price tree.

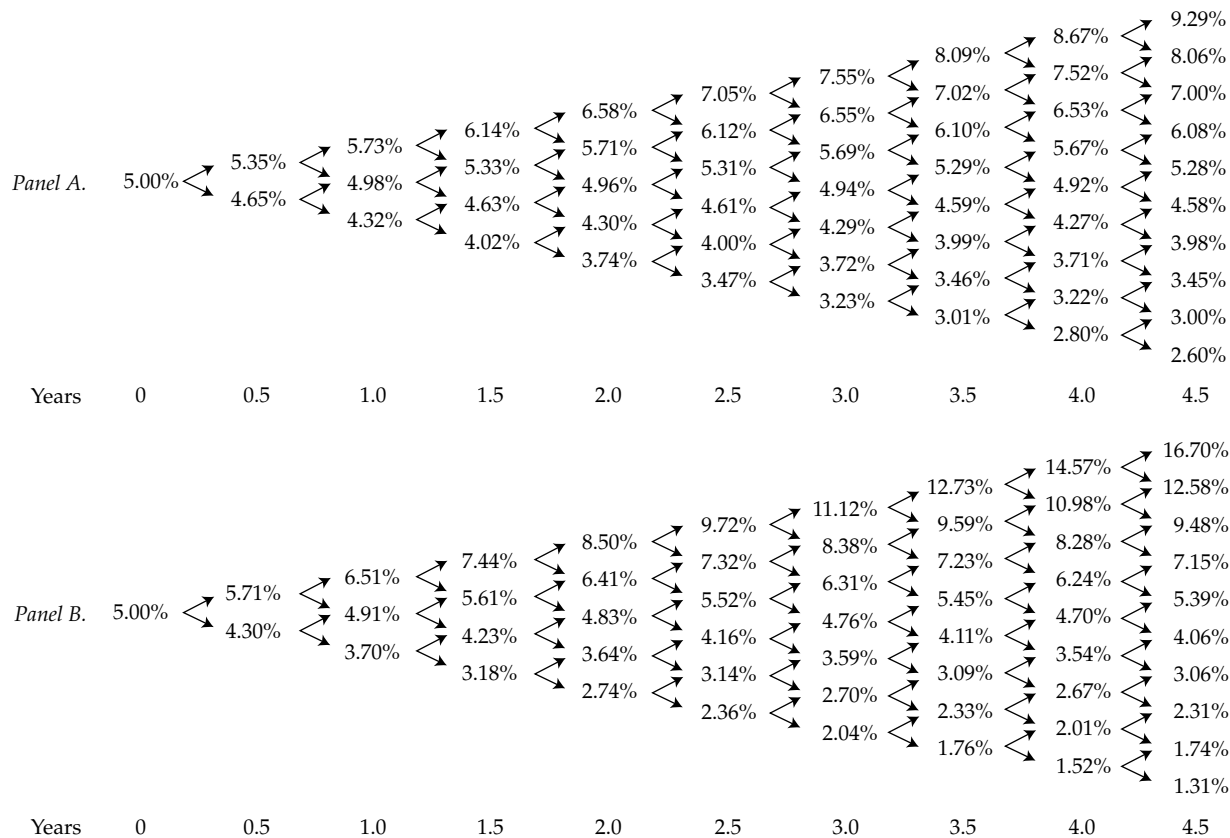
Figure 21. KWF Binomial Trees

Figure 21. KWF Binomial Trees

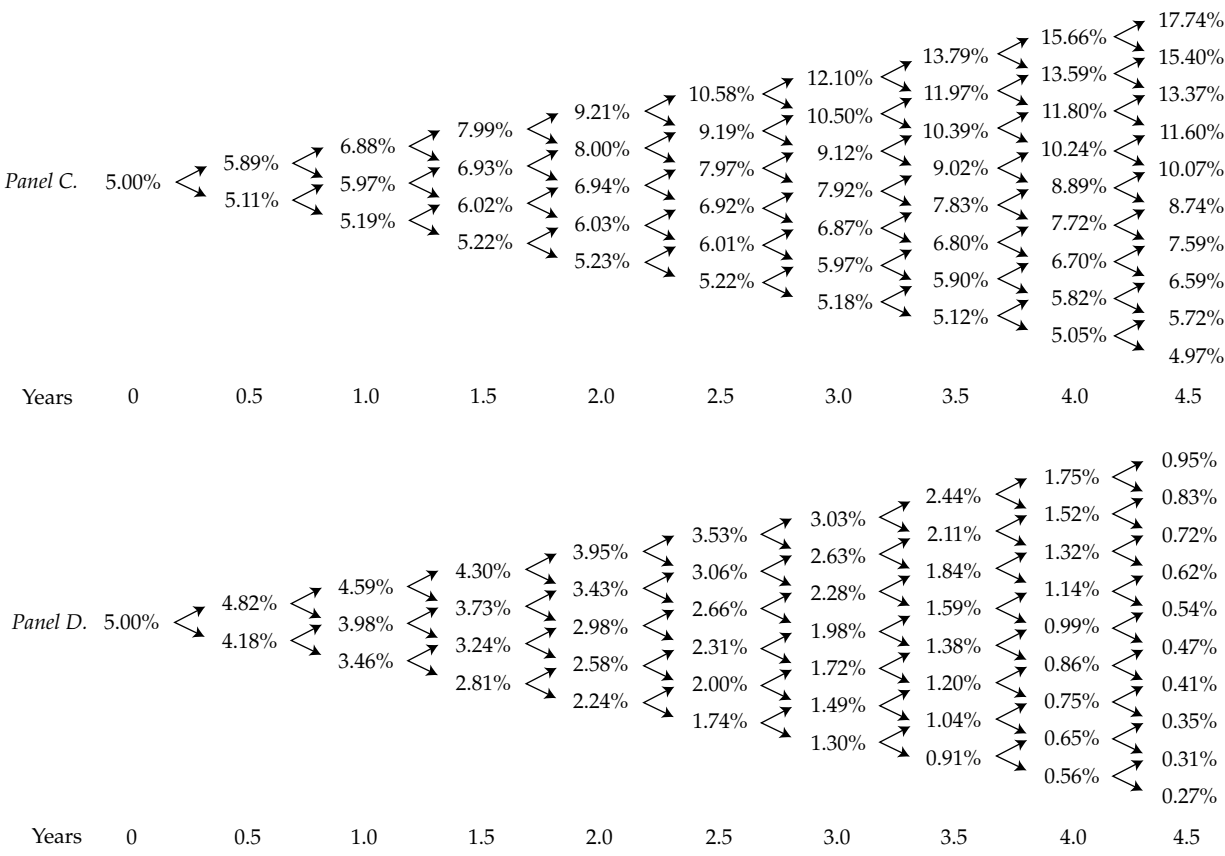
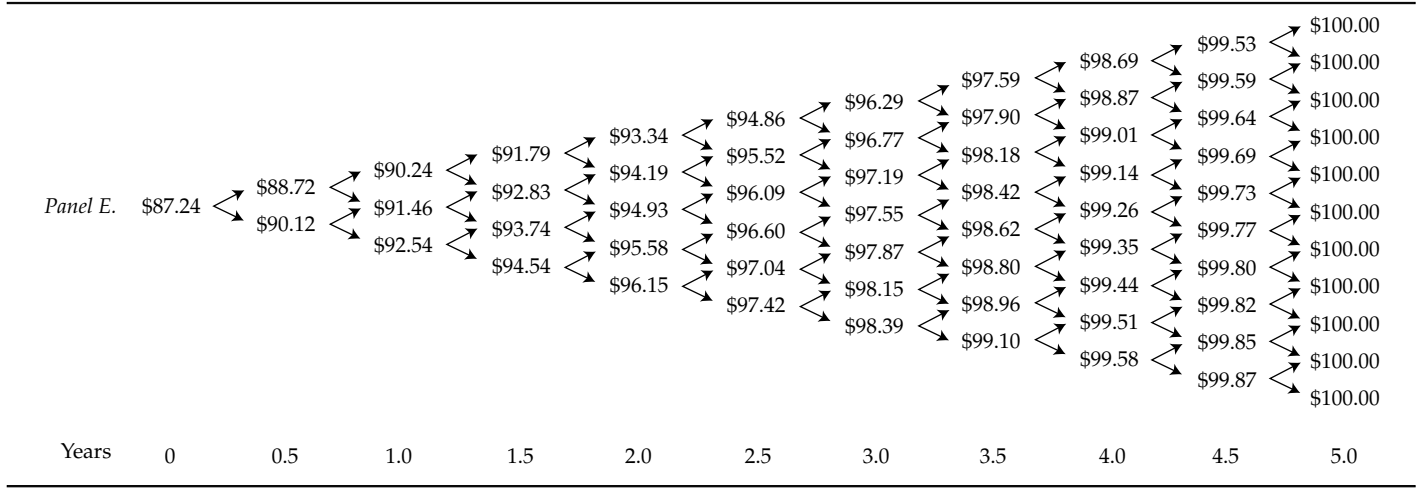


Figure 21. KWF Binomial Trees



The difference in the size of the spread between the HL and the KWF models is a result of the distributional properties of the two models. The short rate in the HL model is normally distributed, allowing for negative rates and, subsequently, a wider spread between the down- and up-states. Conversely, because the short rates in the KWF model are lognormally distributed, not allowing for negative rates, the spread is smaller. It is the difference in the rate spreads that leads to the wider price spreads. The distributional properties are evident by analyzing the figures more closely. The HL rate lattices are symmetric around the center nodes, whereas the KWF rate lattices are asymmetric and skewed to higher rates around the center nodes. This difference is the result of the different distributions implied by the stochastic differential equations that are used to implement the binomial models, as explained in Chapter 2.

Figure 22 presents the BDT binomial trees resulting from the input data in Table 2. These BDT binomial trees verify the results presented in Chapter 2 and demonstrated in Figures 8–10. Panel A of Figure 22 presents the results for the DTS with a volatility starting at 20 percent and decreasing at 0.005 per year. The rates and the spread in the rates decrease significantly because of the large decrease in the volatility structure, as pointed out in Chapter 2. Note that Panel B of Figure 22, which presents the corresponding price tree, gives the same price as the HL and KWF trees for the DTS, as guaranteed by the no-arbitrage property. The three price trees are for significantly different volatilities. We see that the prices in the BDT price tree are more in line with the prices in the KWF tree, as expected from the fact that both models are lognormal. Also, note the asymmetric and skewed nature of the rate spreads, which is similar to the KWF because of the lognormality.

Panel C of Figure 22 presents a BDT binomial tree for the ITS with a volatility starting at 10 percent and decreasing linearly at 1 percent per year. Note the significant decrease in the spread of the rates over the KWF binomial tree in Panel C of Figure 21. This result is expected because there is mean reversion from the decreasing volatility structure. Panel D presents the BDT binomial tree for the ITS with a volatility starting at 10 percent and increasing linearly at 1 percent per year. Note the increase in the rates and the spread of the rates as expected. In Panel E, the BDT binomial tree is for the ITS with EIV, and in Panel F, the binomial tree is for the ITS with EDV. The change in the rates and the spread of the rates is as expected. For the DTS, the BDT process cannot give short rates for EDV, a possibility that was pointed out in Chapter 2.

The BDT and the KWF models will be almost identical when the volatility structure is flat (horizontal) because of the lognormal nature of both models,

Figure 22. BDT Binomial Trees

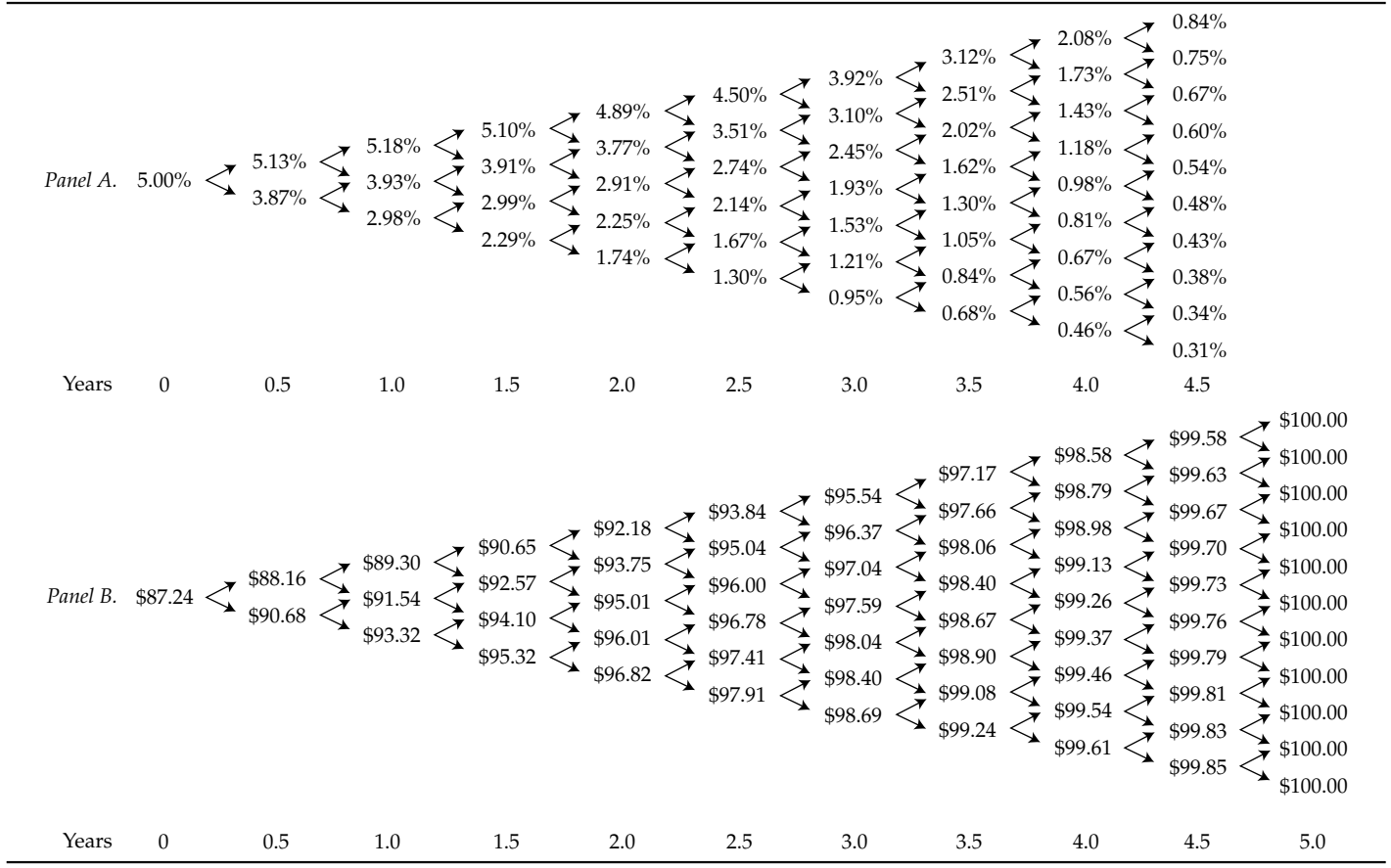


Figure 22. BDT Binomial Trees

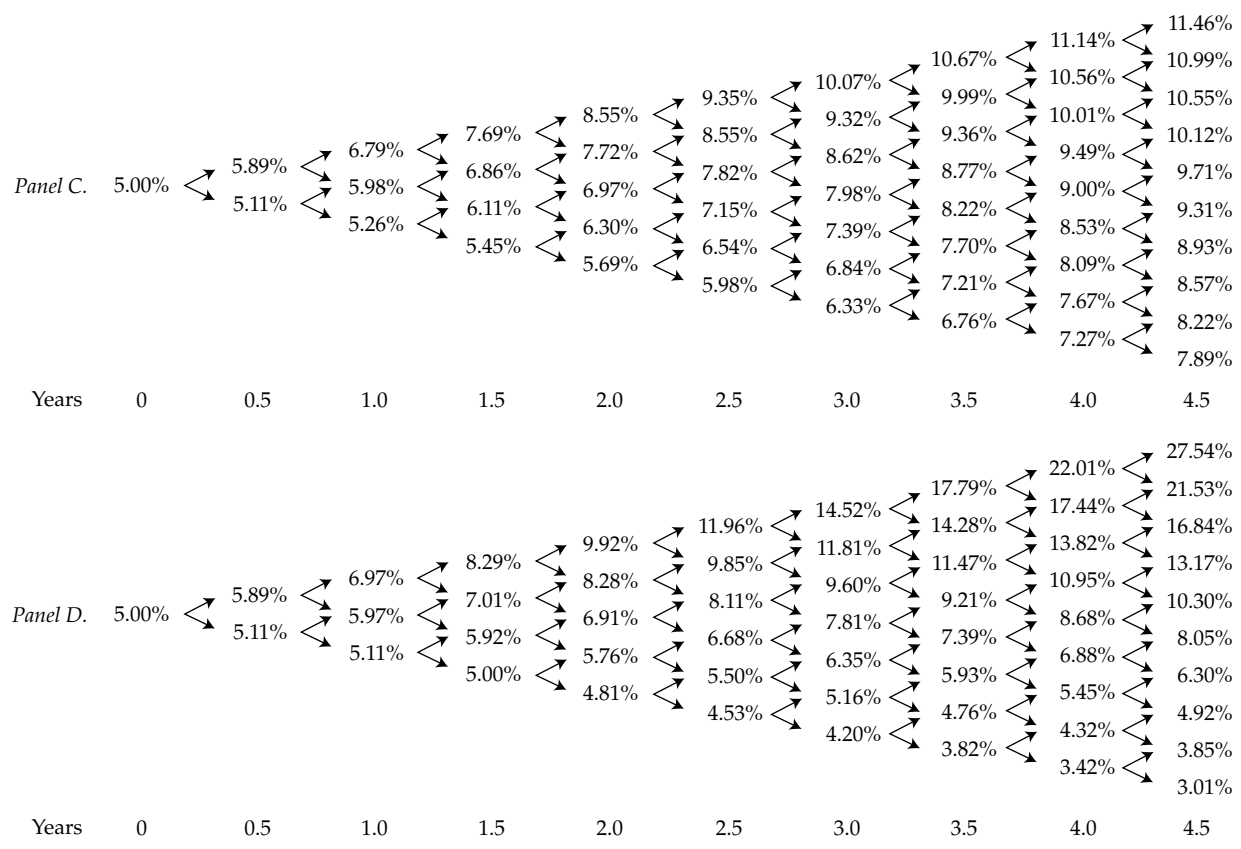
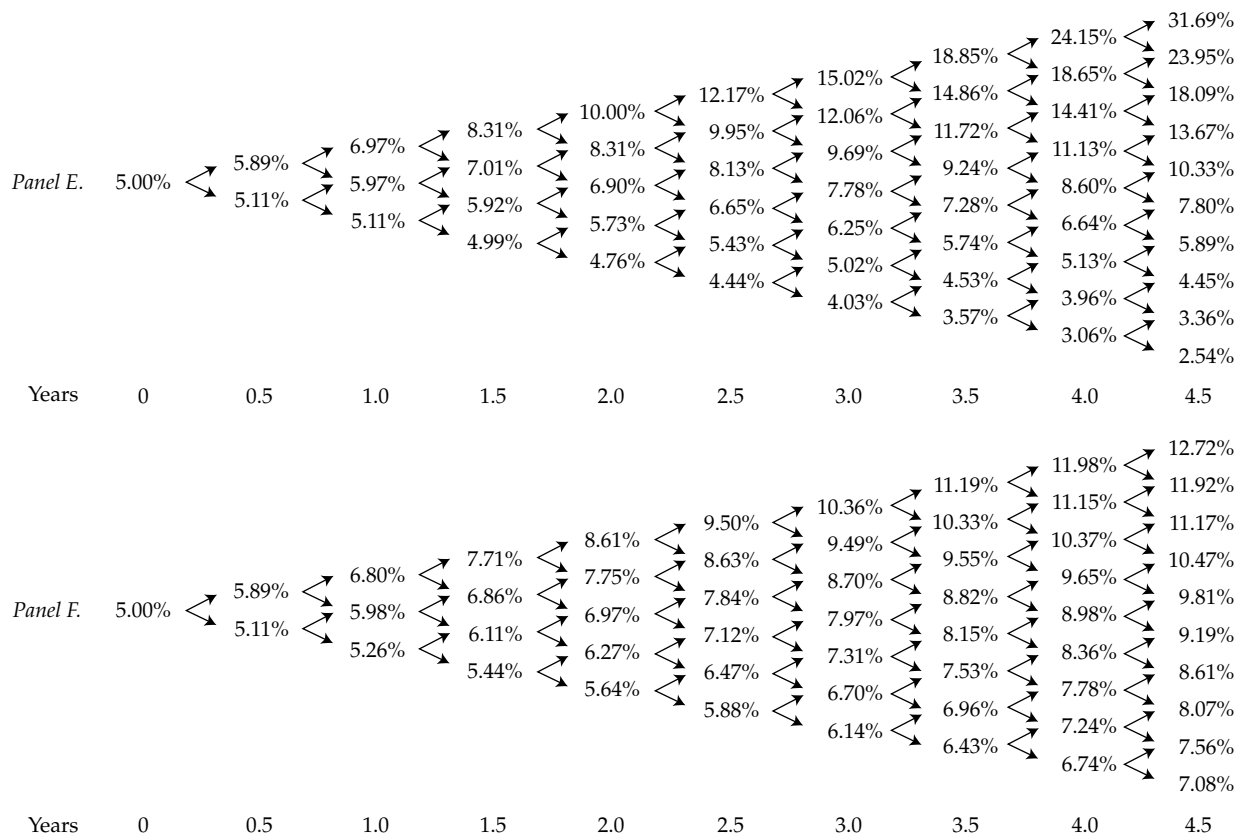


Figure 22. BDT Binomial Trees

as shown in Chapter 2. In fact, the only difference between the two models is the mean reversion implicitly incorporated into the BDT through the shape of the volatility structure. So, only when the volatility structure has significantly changed through time do the two models give very different results.

Examples for HW and BK Processes

Exhibit 1 and **Exhibit 2** present the up-state and down-state short rates for the HW and BK binomial-tree methods, respectively, for $\phi = 0, 0.05, 0.10$, and 0.15 (denoted by MR in the Exhibits) using the modification of Equation 62 for a variable time step. Also presented is the cubic interpolation of the up- and down-states that fit the constant-time-step time periods. **Figure 23** and **Figure 24** show the graphs of the up-state and down-state short rates for the HW and BK binomial trees, respectively. Notice that the size of the time steps decreases with time for a fixed ϕ . Also, note that the size of the time steps decreases with increasing ϕ . In Figure 4 of Chapter 2, the mean reversion in the HW process does not become apparent until ϕ is close to 0.16 . Similar results are presented for the BK process in Figure 7 of Chapter 2.

The binomial tree results presented in Exhibits 1 and 2 and Figures 23–26 should be considered and used as only first approximations for binomial trees to the HW and BK processes. These trees are actually accurate approximations only for small mean reversion in which the time step is almost constant. These first approximations are intended to illustrate the difficulties of the inclusion of mean reversion. They show that large mean reversion leads to a significantly decreasing time step, but they also show how mean reversion affects the rates in the binomial tree (especially in the spread of the rates). For the mean reversion values presented in the examples, iteration on the binomial tree is required to get more accurate approximations of the HW and BK SDEs. The binomial trees presented for these two processes show that an algorithm can be developed for building accurate HW and BK binomial trees. These difficulties (resulting mostly from the variable time step) are the reasons why trinomial lattices are usually developed for the HW and BK processes.

Despite these drawbacks, binomial models for the HW and BK processes do have financially meaningful applications. Consider the following examples of how these two models can be used effectively in the binomial framework. **Figure 25** presents the BK binomial tree obtained from using the BDT method of Chapter 3 on the BK short rates for the FTS and a constant volatility of 10 percent and $\phi = 5$ percent. Comparing the spreads against the BDT lattices shows that there is mean reversion in this binomial tree. Our studies show that the BK process, using the BDT method for solving the binomial

Exhibit 1. Up-State and Down-State Short Rates for the HW Binomial Tree Method*MR = 0*

| | | | | | | | | | | |
|--------------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.050 | 0.123 | 0.199 | 0.277 | 0.357 | 0.440 | 0.525 | 0.613 | 0.703 | 0.796 |
| Down-state | 0.050 | -0.018 | -0.084 | -0.148 | -0.209 | -0.267 | -0.323 | -0.377 | -0.428 | -0.477 |

MR = 0.05

| | | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 0.9765 | 1.4315 | 1.8669 | 2.2844 | 2.6852 | 3.0708 | 3.4422 | 3.8004 | 4.1463 |
| Up-state | 0.05 | 0.1202 | 0.1934 | 0.2602 | 0.3273 | 0.3931 | 0.4575 | 0.5207 | 0.5825 | 0.643 | 0.7022 |
| Down-state | 0.05 | -0.018 | -0.076 | -0.136 | -0.19 | -0.24 | -0.288 | -0.332 | -0.375 | -0.416 | -0.454 |

Cubic int

| | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.05 | 0.1202 | 0.197 | 0.2705 | 0.3481 | 0.4276 | 0.509 | 0.5922 | 0.6771 | 0.7636 |
| Down-state | 0.05 | -0.018 | -0.079 | -0.144 | -0.206 | -0.266 | -0.324 | -0.382 | -0.438 | -0.494 |

MR = 0.10

| | | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 0.9555 | 1.3737 | 1.7602 | 2.1194 | 2.4549 | 2.7696 | 3.066 | 3.346 | 3.6113 |
| Up-state | 0.05 | 0.1175 | 0.1892 | 0.2469 | 0.3047 | 0.3599 | 0.4124 | 0.4625 | 0.5104 | 0.5563 | 0.6004 |
| Down-state | 0.05 | -0.017 | -0.069 | -0.126 | -0.175 | -0.219 | -0.261 | -0.3 | -0.336 | -0.371 | -0.404 |

Cubic int

| | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.05 | 0.1175 | 0.1957 | 0.2652 | 0.3415 | 0.4195 | 0.4997 | 0.5818 | 0.6659 | 0.7521 |
| Down-state | 0.05 | -0.017 | -0.075 | -0.143 | -0.205 | -0.266 | -0.328 | -0.39 | -0.453 | -0.516 |

Exhibit 1. Up-State and Down-State Short Rates for the HW Binomial Tree Method
(continued)

| | | | | | | | | | | | |
|------------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| <i>MR = 0.15</i> | | | | | | | | | | | |
| Time (years) | 0.0 | 0.5 | 0.9366 | 1.324 | 1.672 | 1.9878 | 2.2768 | 2.5432 | 2.7902 | 3.0205 | 3.2361 |
| Up-state | 0.05 | 0.1151 | 0.1857 | 0.2358 | 0.2869 | 0.3348 | 0.3796 | 0.4217 | 0.4614 | 0.499 | 0.5347 |
| Down-state | 0.05 | -0.017 | -0.063 | -0.118 | -0.163 | -0.203 | -0.24 | -0.274 | -0.306 | -0.337 | -0.366 |
| Cubic int | | | | | | | | | | | |
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 | |
| Up-state | 0.05 | 0.1151 | 0.1946 | 0.2608 | 0.3366 | 0.4148 | 0.4956 | 0.5793 | 0.6659 | 0.7559 | |
| Down-state | 0.05 | -0.017 | -0.071 | -0.141 | -0.204 | -0.269 | -0.334 | -0.402 | -0.472 | -0.545 | |

Exhibit 2. Up-State and Down-State Short Rates for the BK Binomial Tree Method*MR = 0*

| | | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.050 | 0.054 | 0.057 | 0.061 | 0.066 | 0.070 | 0.075 | 0.081 | 0.087 | 0.093 |
| Down-state | 0.050 | 0.046 | 0.043 | 0.040 | 0.037 | 0.035 | 0.032 | 0.030 | 0.028 | 0.026 |

MR = 0.05

| | | | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Time (years) | 0.000 | 0.500 | 0.976 | 1.431 | 1.867 | 2.284 | 2.685 | 3.071 | 3.442 | 3.800 | 4.146 |
| Up-state | 0.050 | 0.056 | 0.060 | 0.063 | 0.067 | 0.070 | 0.074 | 0.078 | 0.082 | 0.086 | 0.090 |
| Down-state | 0.050 | 0.049 | 0.045 | 0.042 | 0.039 | 0.037 | 0.035 | 0.033 | 0.031 | 0.029 | 0.028 |

Cubic int

| | | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Time (years) | 0.000 | 0.500 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.050 | 0.056 | 0.06 | 0.064 | 0.068 | 0.072 | 0.077 | 0.082 | 0.088 | 0.094 |
| Down-state | 0.050 | 0.049 | 0.045 | 0.042 | 0.038 | 0.036 | 0.033 | 0.031 | 0.028 | 0.026 |

MR = 0.10

| | | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 0.9555 | 1.3737 | 1.7602 | 2.1194 | 2.4549 | 2.7696 | 3.066 | 3.346 | 3.6113 |
| Up-state | 0.05 | 0.0588 | 0.0609 | 0.0633 | 0.0659 | 0.0687 | 0.0714 | 0.0742 | 0.0771 | 0.08 | 0.0829 |
| Down-state | 0.05 | 0.051 | 0.0465 | 0.043 | 0.0401 | 0.0377 | 0.0356 | 0.0338 | 0.0323 | 0.0308 | 0.0296 |

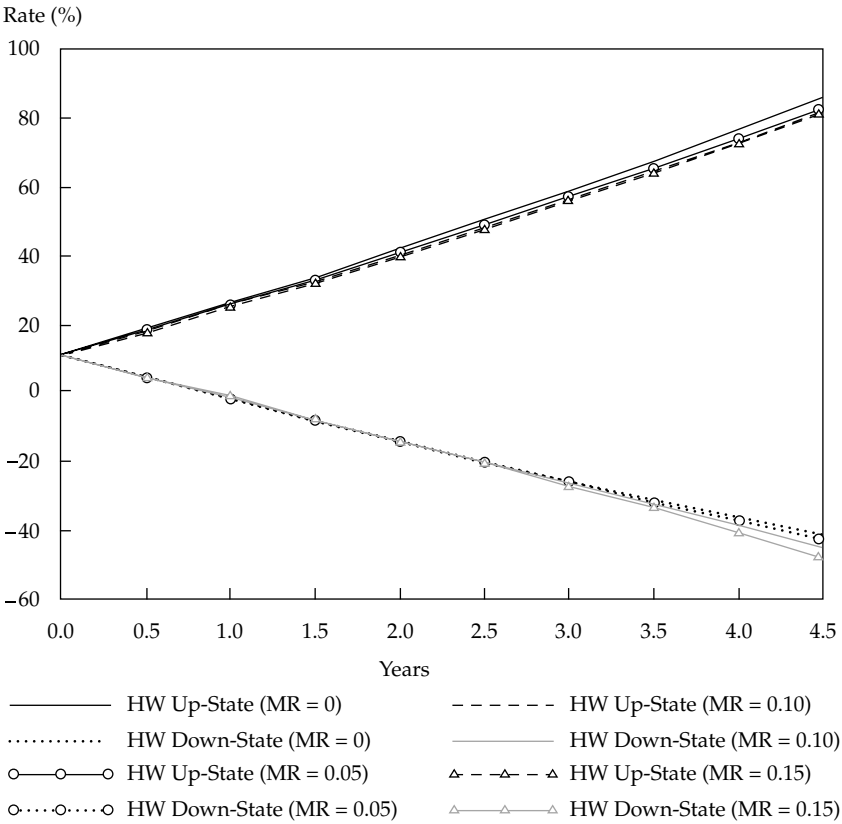
Cubic int

| | | | | | | | | | | |
|--------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 |
| Up-state | 0.05 | 0.0588 | 0.0611 | 0.0642 | 0.0677 | 0.0718 | 0.0764 | 0.0816 | 0.0874 | 0.0939 |
| Down-state | 0.05 | 0.051 | 0.046 | 0.042 | 0.0385 | 0.0354 | 0.0326 | 0.0301 | 0.0278 | 0.0257 |

Exhibit 2. Up-State and Down-State Short Rates for the BK Binomial Tree Method (continued)
(continued)

| | | | | | | | | | | | |
|------------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| <i>MR = 0.15</i> | | | | | | | | | | | |
| Time (years) | 0.0 | 0.5 | 0.9366 | 1.324 | 1.672 | 1.9878 | 2.2768 | 2.5432 | 2.7902 | 3.0205 | 3.2361 |
| Up-state | 0.05 | 0.0613 | 0.0614 | 0.0629 | 0.0647 | 0.0667 | 0.0689 | 0.0711 | 0.0734 | 0.0757 | 0.0779 |
| Down-state | 0.05 | 0.0532 | 0.0471 | 0.0433 | 0.0403 | 0.038 | 0.0361 | 0.0345 | 0.0331 | 0.0319 | 0.0308 |
| Cubic int | | | | | | | | | | | |
| Time (years) | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 | |
| Up-state | 0.05 | 0.0613 | 0.0614 | 0.0638 | 0.0668 | 0.0708 | 0.0754 | 0.0809 | 0.0872 | 0.0949 | |
| Down-state | 0.05 | 0.0532 | 0.0463 | 0.0418 | 0.038 | 0.0348 | 0.032 | 0.0295 | 0.0272 | 0.0251 | |

Figure 23. Up-State and Down-State Short Rates for the HW Binomial Tree



trees, gives a more accurate first approximation of mean reversion than the BK method presented in Chapter 3.

Figure 26 presents the HW and BK binomial trees for the FTS with LDV and MRLDV so that the time step remains constant at 0.5 years. Comparing the interest rate spreads in the binomial tree in Panel A of Figure 26 with the binomial tree in Panel B of Figure 20 and the corresponding data in Exhibit 1 shows that there is mean reversion in this example. Similarly, the binomial tree in Panel B of Figure 26 shows mean reversion when compared with Figure 24 and the corresponding data in Exhibit 2 and also with Figure 25. It is also interesting to point out that the BDT binomial tree will give the same results as Panel B of Figure 26 for this volatility structure.

Figure 24. Up-State and Down-State Short Rates for the BK Binomial Tree

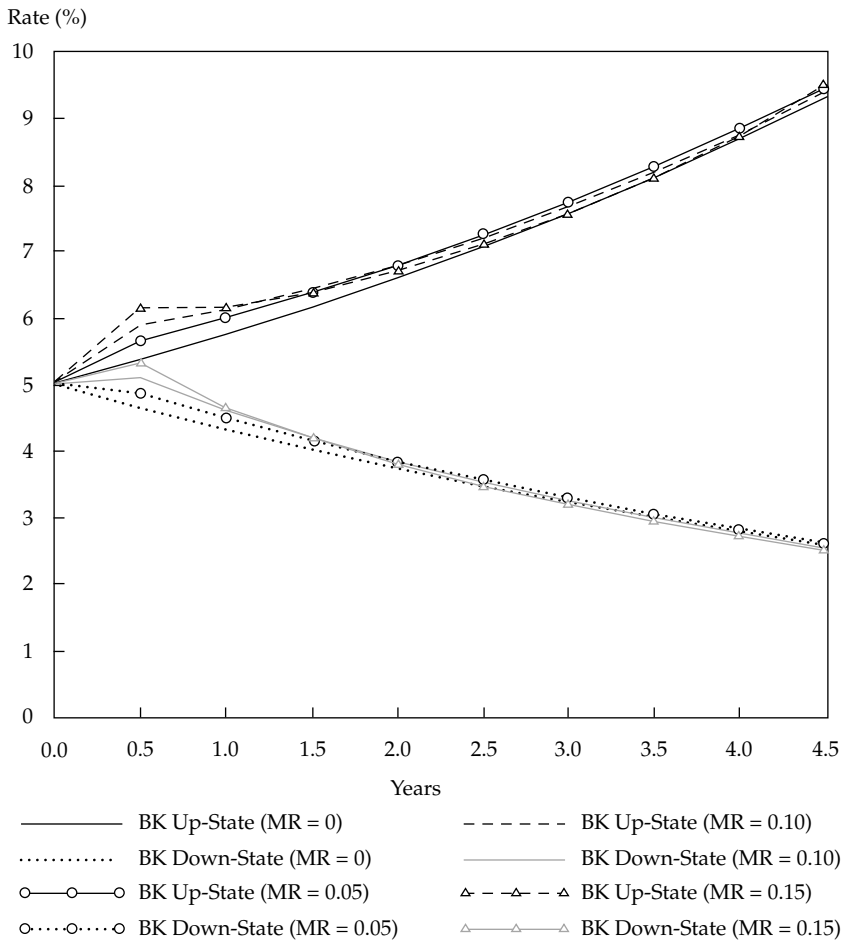


Figure 25. Example of a BK Binomial Tree

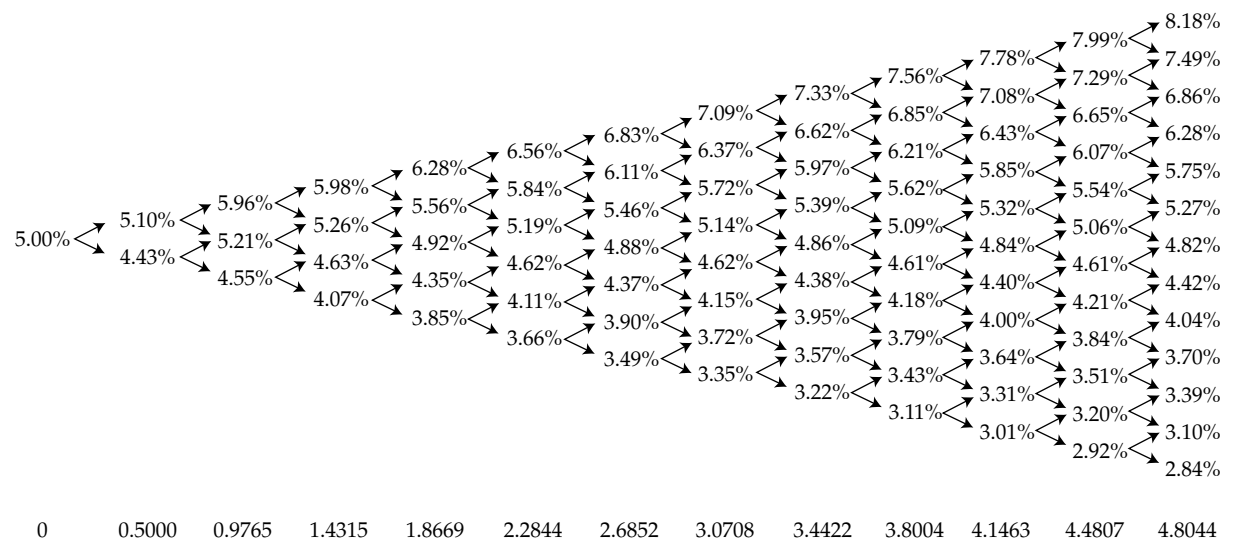
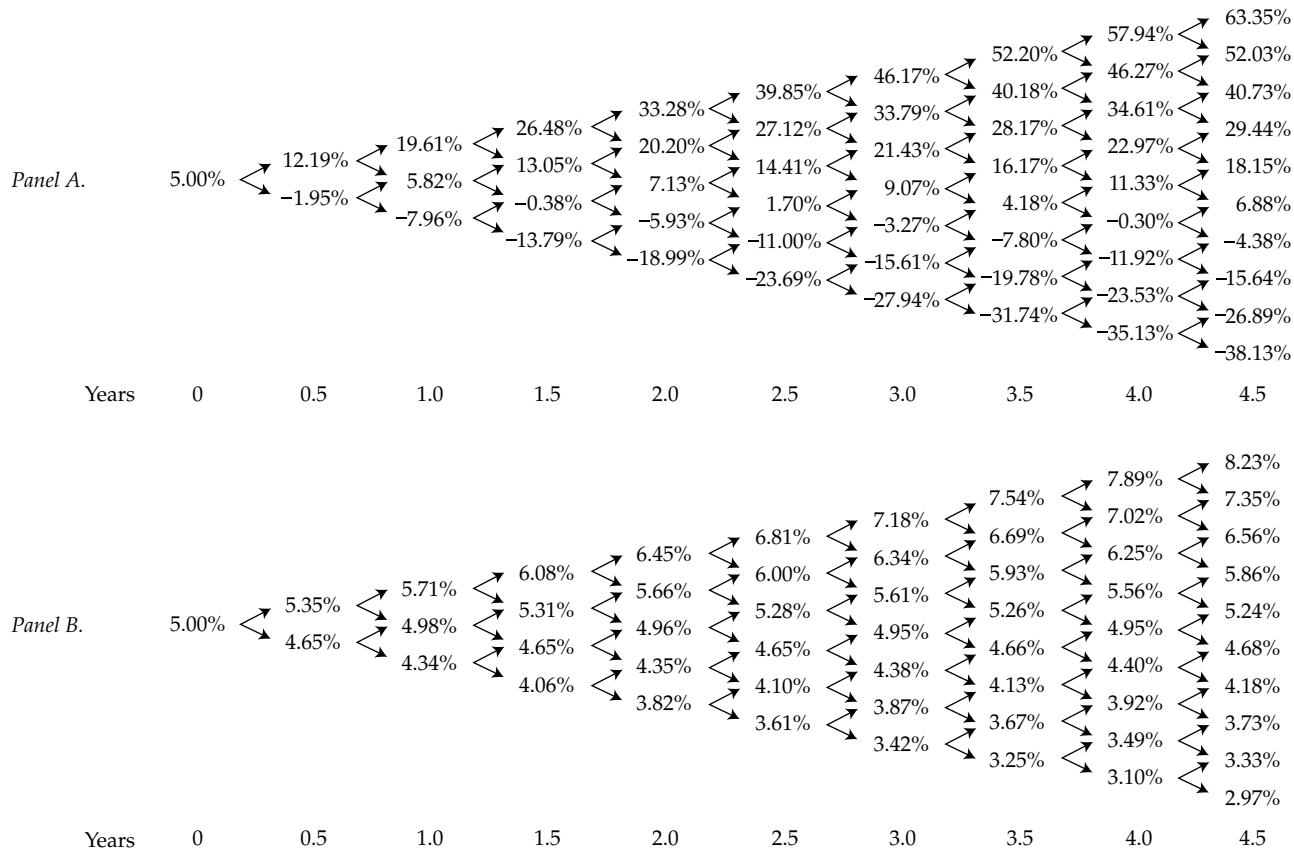


Figure 26. Examples of HW and BK Binomial Trees with Flat Term Structures



5. Conclusion

We have presented five models, each with different strengths and weaknesses. The Ho–Lee (HL) and Hull–White (HW) models are based on normally distributed rates, as is evident in the symmetry of the resulting binomial lattices. They both allow negative interest rates. The HW model incorporates mean reversion in an attempt to prevent negative rates. The degree of freedom used by the inclusion of the mean reversion term forces the HW model to have a changing time step within the binomial framework.

The Kalotay–Williams–Fabozzi (KWF), Black–Karasinski (BK), and Black–Derman–Toy (BDT) models are all based on lognormally distributed rates. This basis is evident in the asymmetry and skewness of the resulting binomial lattices. Accordingly, none of these allows negative rates, which is an improvement over the HW and the HL models. The KWF model is the simplest of these three models and is really no more than a lognormal version of the HL model. The BK model is analogous to the HW model in that it includes a mean reversion term. Similarly, the degree of freedom used by the inclusion of the mean reversion term forces the BK model to have a changing time step within the binomial framework. To avoid using up a degree of freedom on mean reversion, the BDT implicitly incorporates mean reversion through the volatility structure, which allows for a constant time step while including mean reversion. The incorporation of mean reversion, however, makes the BDT model far more difficult than the other models and also limits the interest rate scenarios that allow for a solution in the BDT framework.

Because each model has strengths and weaknesses and slightly different characteristics, the characteristics must be thoroughly understood before the models are implemented. Otherwise, valuation and risk-measurement results cannot be analyzed for different dealers (see Buetow, Hanke, and Fabozzi, forthcoming 2001). Moreover, knowledge of the various characteristics of each model is critical for the evaluation of software packages.

Appendix A. Variable Time Step

Because we are measuring calendar time by

$$0 = t_0 < t_1 < t_2 < \dots, t_n,$$

the time steps are

$$\tau_1 = t_1 - t_0, \tau_2 = t_2 - t_1, \dots, \tau_n = t_n - t_{n-1}.$$

If the time steps are not constant (they vary), then

$$t_1 = t_0 + \tau_1, t_2 = t_1 + \tau_2 = t_0 + \tau_1 + \tau_2, \dots,$$

and

$$t_n = t_{n-1} + \tau_n = t_0 + \tau_1 + \tau_2 + \dots + \tau_n.$$

In this case, the discount factor for the zero-coupon bond is given by

$$A_k = \frac{1}{(1 + \bar{R}_k \tau)^{t_k/\tau}},$$

where τ is the time period corresponding to the spot rates contained in the term structure and \bar{R}_k is the spot rate during the time periods containing the time t_k . That is, there is an l such that $l\tau \leq t_k \leq (l+1)\tau$, and \bar{R}_k is the spot rate during this time period.

This adjustment must also be made for the short rates. For example, if we are using Equation 62 for the Hull–White (HW) and Black–Karasinski (BK) binomial short rates with a variable time step, then we have to replace the $(1 + r_{k,j}\tau)$ terms by $(1 + r_{k,j}\tau)^{\tau_k/\tau}$ and the $(1 + r_{k,n}\tau)$ terms by $(1 + r_{k,n}\tau)^{\tau_k/\tau}$. If we are using the Black–Derman–Toy (BDT) binomial short rate method for the BK binomial short rates with a variable time step, then we must replace Equation 72 with

$$\begin{aligned} A_{k+1}(1 + r_0\tau) \left(1 + \bar{y}_{k,d}\tau e^{2\bar{s}_{k-1}\sqrt{\tau}}\right)^{t_k/\tau} (1 + \bar{y}_{k,d}\tau)^{t_k/\tau} \\ - q(1 + \bar{y}_{k,d}\tau)^{t_k/\tau} - (1 - q) \left(1 + \bar{y}_{k,d}\tau e^{2\bar{s}_{k-1}\sqrt{\tau}}\right)^{t_k/\tau} = 0, \end{aligned}$$

where \bar{s}_{k-1} is the interpolated value from the inputted market-yield volatilities, and use $p_{k,u} = 1 / (1 + \bar{y}_{k,d}\tau)^{t_k/\tau}$ as the up-price and $p_{k,d} = 1 / (1 + \bar{y}_{k,d}\tau)^{t_k/\tau}$ as the down-price. In Equations 73a and 73b, we replace $(1 + r_{k,j}\tau)$ terms with $(1 + r_{k,j}\tau)^{\tau/\tau}$ terms. We use the Newton–Raphson method to solve for these BK short rates, just as for the BDT binomial short rates.

The Newton–Raphson method can be used to solve the single equation $f(w) = 0$ for w . An initial guess for w is made. Call this guess w_0 . The Newton–Raphson algorithm is defined by

$$w_{m+1} = w_m - \frac{f(w_m)}{f'(w_m)}.$$

This method (usually) converges quickly to w . If you have a system of equations, the Newton–Raphson method is applied to each equation in the system. See Burden and Faires for the details.

Appendix B. Continuous Compounding of Interest Rates

In Equation 1 of the manuscript, $P_2 = A_2(1 + R_2\tau)^2$, and in Equation 2, $A_2 = P_2/(1 + R_2\tau)^2$. These equations are the formulas for interest being compounded once every time period. That is, the interest is compounded at $t_1 = \tau$ and again at $t_2 = 2\tau$. If the interest were being compounded continuously, the corresponding equations would be

$$\begin{aligned} P_2 &= A_2 e^{R_2 t_2} \\ &= A_2 e^{2R_2 \tau} \end{aligned} \tag{B1}$$

and

$$\begin{aligned} A_2 &= P_2 e^{-R_2 t_2} \\ &= P_2 e^{-2R_2 \tau}. \end{aligned} \tag{B2}$$

Therefore, if we want to use continuous compounding, we replace the discount factors in the text, $1/(1 + r_{m,n}\tau)$, by the discount factors $e^{-r_{m,n}\tau}$ and we replace the $1 + r_{m,n}\tau$ expressions by $e^{r_{m,n}\tau}$.

We would still use the bisection method to solve for the short rates, $r_{k,j}$, except in the Black–Derman–Toy (BDT) methods, where we would use the Newton–Raphson method. Significant differences in these two forms of compounding will show up only for long-period bonds, high interest rates, or high-premium bonds.

Appendix C. Implementation and Calibration Issues for the Binomial Models

This appendix presents an overview of implementation and calibration issues. In particular, we present a few numerical examples so that the reader can obtain a better appreciation of the algorithm required to properly implement no-arbitrage interest rate models. We also offer an overview of calibration techniques commonly employed to adapt the models to various market information. This discussion coverage is by no means exhaustive.

Implementation Examples

Equation 9 is used to find the first two rates in the binomial model for each of the five models presented. The inputs to the Ho–Lee (HL) and Kalotay–Williams–Fabozzi (KWF) binomial models are the term structure of zero rates and the short-rate volatility. The short-rate volatility can be obtained in various ways. Nonconstant volatilities are at the heart of calibration issues and are presented in the next section of the appendix.

To illustrate the implementation of these models, we assume a semiannual periodicity with $r_0 = R_1 = 3.5$ percent, $R_2 = 4.25$ percent, and $R_3 = 5.5$ percent. We also assume a constant short-rate volatility of 5 percent ($\sigma = 5$ percent). We now consider an example for each of the HL and KWF binomial trees using these data.

The two rates emanating from r_0 are found from Equations 46a and 46b,

$$r_u = r_{1,1} = 0.035 + 0.5m_0 + 0.05\sqrt{0.5}$$

for the up-move and

$$r_d = r_{1,2} = 0.035 + 0.5m_0 - 0.05\sqrt{0.5}$$

for the down-move. Now, we substitute these two rates, R_2 , and $q = 0.5$ into Equation 9a to get the rather ugly expression

$$\begin{aligned}
 & \frac{1}{[1 + (0.0425)(0.5)]^2} \\
 & \times [1 + (0.035)(0.5)][1 + (0.035 + 0.5m_0 - 0.05\sqrt{0.5})(0.5)] \\
 & \times [1 + (0.035 + 0.5m_0 + 0.05\sqrt{0.5})(0.5)] \\
 & - 0.5[1 + (0.035 + 0.5m_0 + 0.05\sqrt{0.5})(0.5)] \\
 & - 0.5[1 + (0.035 + 0.5m_0 - 0.05\sqrt{0.5})(0.5)] = 0,
 \end{aligned}$$

which upon simplification (and roundoff) becomes

$$(0.976)(1 + 0.25m_0)(1.035 + 0.25m_0) - 0.5(1 + 0.25m_0) - 0.5(1.035 + 0.25m_0) = 0.$$

This is a quadratic equation that has the following two solutions:

$$m_0 = 0.03127$$

and

$$m_0 = -4.071.$$

The economically feasible solution is obviously $m_0 = 0.03127$. That is, the first drift value in this HL binomial model is 0.03127, or 3.127 percent. Substituting this drift value into the formulas gives $r_u = 8.6$ percent and $r_d = 1.53$ percent. Also note that $r_u - r_d = 2\sigma\sqrt{\tau} = 2(0.05)\sqrt{0.5} = 0.07$. Because $r_{1,1} = 8.60$ percent and $r_{1,2} = 1.53$ percent, we get (from Equation 57)

$$r_{2,1} = r_{1,1} + m_1\tau + \sigma\sqrt{\tau} = 0.086 + 0.5m_1 + 0.035,$$

$$r_{2,2} = r_{1,1} + m_1\tau - \sigma\sqrt{\tau} = 0.086 + 0.5m_1 - 0.035,$$

and

$$r_{2,3} = r_{1,2} + m_1\tau - \sigma\sqrt{\tau} = 0.0153 + 0.5m_1 - 0.035.$$

We combine these with $P_3 = 1/(1 + R_3\tau)^3 = 0.9218$ and substitute into Equation 59. The bisection method is used on the resulting cubic expression to obtain the feasible solution, $m_1 = 0.0628$, which results in $r_{2,1} = 15.28$ percent, $r_{2,2} = 8.20$ percent, and $r_{2,3} = 1.13$ percent. We leave it to the reader to show that $r_{2,1} - r_{2,3} = 4\sigma\sqrt{\tau}$.

We now use these data to generate the KWF rates. The two rates emanating from r_0 are found from Equation 50,

$$\begin{aligned}
 r_u &= r_{1,1} \\
 &= r_0 e^{m_0 \tau + \sigma_0 \sqrt{\tau}} \\
 &= 0.035 e^{0.5 m_0 + 0.035}
 \end{aligned}$$

for the up-move and

$$\begin{aligned}
 r_d &= r_{1,2} \\
 &= r_0 e^{m_0 \tau - \sigma_0 \sqrt{\tau}} \\
 &= 0.035 e^{0.5 m_0 - 0.035}
 \end{aligned}$$

for the down-move. Substituting into Equation 9a, as we did in the HL model, we get

$$\begin{aligned}
 &(0.976)(1 + 0.0175 e^{0.5 m_0 + 0.035})(1 + 0.0175 e^{0.5 m_0 - 0.035}) \\
 &\quad - 0.5(1 + 0.0175 e^{0.5 m_0 - 0.035}) - 0.5(1 + 0.0175 e^{0.5 m_0 + 0.035}) = 0.
 \end{aligned}$$

Again, using the bisection method, we find the first drift, $m_0 = 0.7133$, which gives $r_u = 5.18$ percent and $r_d = 4.83$ percent. We again see that $\ln(r_u/r_d) = 2\sigma\sqrt{\tau} = 2(0.05)\sqrt{0.5} = 0.07$ (or $r_u/r_d = e^{2\sigma\sqrt{\tau}}$), as shown in Chapter 3. Because $r_{1,1} = 5.18$ percent and $r_{1,2} = 4.83$ percent, we get (from Equation 64)

$$\begin{aligned}
 r_{2,1} &= 0.0518 e^{0.5 m_1 + 0.035}, \\
 r_{2,2} &= 0.0518 e^{0.5 m_1 - 0.035},
 \end{aligned}$$

and

$$r_{2,3} = 0.0483 e^{0.5 m_1 - 0.035}.$$

We combine these with $P_3 = 1/(1 + R_3\tau)^3 = 0.9218$ and substitute into Equation 59. The bisection method is used on the resulting cubic expression to obtain the feasible solution, $m_1 = 0.9436$. This results in $r_{2,1} = 8.60$ percent, $r_{2,2} = 8.01$ percent, and $r_{2,3} = 7.47$ percent. We leave it to the reader to show $r_{2,1}/r_{2,3} = e^{4\sigma\sqrt{\tau}}$.

For both models, this propagation process is continued until we span the desired maturity. The equations that are solved at each period are polynomials that are easily solved using the bisection technique. A robust algorithm will

incorporate the simple algebraic structure of these polynomials. The following six-step procedure outlines the implementation process.

1. Obtain the term structure of zero rates.
2. Determine a constant volatility from market conditions.
3. Obtain the first drift, m_0 , for $r_{1,1}$ and $r_{1,2}$, using Equation 9 and the bisection method, and obtain $r_{1,1}$ and $r_{1,2}$.
4. Generate the equations for $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$.
5. Generate the polynomials at each time step using Equation 62.
6. Repeat Step 5, propagating forward in the lattice.

The Hull–White (HW) and Black–Karasinski (BK) binomial tree approximations presented in Chapters 3 and 4 follow the outline above, except that the time step must vary in order to have enough degrees of freedom to solve the equations. The details of the variable time step are discussed in Appendix A. Equations 48a and 48b are substituted into Equation 9 for the HW model, and Equations 52a and 52b are substituted into Equation 9 for the BK model. The tree evolves in a similar manner as in the HL and KWF models (except that Equations 78 are used for the HW model and Equations 88 are used for the BK model).

We now present the implementation procedure for the Black–Derman–Toy (BDT) model. The inputs for the BDT model are the term structure of zero rates and the yield volatility corresponding to each time step of the input vectors.

The BDT model uses the same term structure as in the HL and KWF examples. For the yield volatilities, we use $s_0 = \sigma_0 = 0.05$ and let $s_1 = 0.06$. Because the first two rates in the binomial tree for the BDT model are the same as the KWF model, we use $r_{1,1} = 5.18$ percent and $r_{1,2} = 4.83$ percent to generate $r_{2,1}$, $r_{2,2}$, and $r_{2,3}$ for the BDT model.

Equation 70 gives

$$\begin{aligned} y_u &= y_d e^{2s_1\sqrt{\tau}} \\ &= y_d e^{2(0.06)\sqrt{0.5}} \\ &= y_d e^{0.08485}, \end{aligned}$$

and Equation 72 results in

$$\begin{aligned} &0.9218[1 + (0.035)(0.5)](1 + 0.5y_d e^{0.08485})^2(1 + 0.5y_d)^2 \\ &- 0.5(1 + 0.5y_d)^2 - 0.5(1 + 0.5y_d e^{0.08485})^2 = 0. \end{aligned}$$

The feasible solution to this equation is $y_d = 0.0623197$ and $y_u = 0.678385$. Therefore, $p_d = 0.94048$ and $p_u = 0.93546$. Equations 73a, 73b, and 73c become, respectively,

$$0.93546[1 + (0.0518)(0.5)](1 + 0.5r_{2,1})(1 + 0.5r_{2,2}) - 0.5(1 + 0.5r_{2,2}) - 0.5(1 + 0.5r_{2,1}) = 0,$$

$$0.94048[1 + (0.0483)(0.5)](1 + 0.5r_{2,2})(1 + 0.5r_{2,3}) - 0.5(1 + 0.5r_{2,3}) - 0.5(1 + 0.5r_{2,2}) = 0.$$

and

$$r_{2,2}^2 = r_{2,1}r_{2,3}.$$

Using the Newton–Raphson method, the feasible solution to this system is $r_{2,1} = 8.80$ percent, $r_{2,2} = 8.01$ percent, and $r_{2,3} = 7.29$ percent. This input information will result in a local volatility. Specifically, from Equation 67, the local volatility is $\sigma_1 = [\ln(0.088) - \ln(0.0801)] / (2\sqrt{0.5}) = 6.64$ percent. This can be used as the local volatility in the HW or BK binomial models.

The following seven-step process outlines the implementation procedure for the BDT binomial model.

1. Obtain the term structure of zero rates.
2. Determine the yield volatility from market conditions for each time step.
3. Obtain the first drift, m_0 , for $r_{1,1}$ and $r_{1,2}$ using Equation 9a and the bisection method, and obtain $r_{1,1}$ and $r_{1,2}$.
4. Using Equation 70 through Equation 72, determine y_u, y_d and then p_u, p_d .
5. Set up Equation 73 for the rates at each time step in the binomial tree.
6. Use the Newton–Raphson method to find the feasible rates for the equations derived in Step 5.
7. Repeat Steps 5 and 6, propagating forward in the lattice.

If the user wants to incorporate yield volatility in the BK model and use the above outline for the BK model, the process is similar. But the variable time step outlined in Appendix A must be used.

Calibration Issues

For the HL and KWF models, a constant short-rate volatility is used throughout the interest rate lattice. Because the volatility is constant, an estimate of the appropriate current short-rate volatility is usually used. The model explicitly assumes that this volatility is the same throughout the lattice. Numerous econometric techniques exist to estimate this parameter. Regardless of the sophistication of the modeling technique used, experience plays a critical role.

The inputs to the HW and BK models are ϕ and σ . Analytic solutions presented in Chapter 2 for the ordinary differential equations (ODEs) can be used to generate a first approximation for the forward rates, and then the input can be “adjusted” according to some minimization-of-error formula for the expected forward rates by using the market prices of forward contracts. One can also use the discrete approximations to the SDEs presented in Chapter 2 to generate many first approximate solutions with varying inputs and use the inputs that give the results closest to known market conditions.

Another calibration procedure is suggested in Black and Karasinski (1991). They suggest starting with constant inputs, building a short-term binomial tree, pricing the options from these inputs (as was shown in Chapter 4), and then comparing these prices with prices given by Black’s formula for options (Black 1976). Iterate on this procedure until the two models are consistent. That is, if the price of the option is off, calibrate the parameters until a desirable result is reached. A minimization-of-error method is often employed. Hull (2000) also suggests this approach.

A common minimization-of-error method is obtained from using the market prices of interest rate caps. A cap is a portfolio of short-term interest rate options called caplets. These prices are used to compute the implied volatility for each caplet. The resulting vector of short-rate volatilities is often called the cap curve. These implied volatilities are then used as the short-rate volatilities throughout the interest rate lattice. The interest rate lattice is then calibrated to retain the no-arbitrage property and ensure that the local volatilities throughout the tree match the cap curve.

Following the presentation in Hull, a caplet can be exercised at time t_k with payoff at time t_{k+1} given by

$$A\tau_{k+1}\max(\rho_k - \rho_x, 0),$$

where A is the principal, ρ_k is the forward rate at time t_k for the period $\tau_{k+1} = t_{k+1} - t_k$ at time t_k , and ρ_x is the option exercise forward rate at time t_k for the period τ_{k+1} at time t_k . If ρ_k is lognormal with volatility, σ_k , the value of this option according to Black’s formula is easily shown to be

$$c = A\tau_{k+1}p_{k+1}[r_k N(d_1) - \rho_x N(d_2)],$$

where

$$d_1 = \frac{\ln(r_k) - \ln(\rho_x) + (1/2)\sigma_k^2 t_k}{\sigma_k \sqrt{t_k}}$$

and

$$d_2 = d_1 - \sigma_k \sqrt{t_k},$$

where r_k is the forward rate and p_{k+1} is the discounted zero coupon price. Because we know the value of the option and all of the other input information, we can solve each of these for the implied volatility, σ_k .¹ This gives a local volatility that can be used in the BDT or BK binomial models. Black and Karasinski also discuss some of these issues in using the cap curve to calibrate input parameters. Calibrating models to the cap curve requires an additional degree of freedom because it necessitates an additional input into the model. This added input often leads to the use of trinomial models instead of binomial models. The BDT model can be used, however, if we replace the yield volatility curve with the cap curve as an input. This approach retains the fixed-time increment and allows for the inclusion of the cap curve. The solution methodology outlined in the previous section is then modified slightly so that the cap curve results, instead of the yield volatilities, are incorporated at each time step. This method will sometimes result in situations for which no solution exists. The trinomial approach is a more robust framework if the user wants to calibrate the interest rate lattice to the cap curve.

¹ We actually use the Newton–Raphson method to compute the volatility because no analytic expression exists for σ_k .

References

- Backus, D., S. Foresi, and C. Telmer. 2000. "Discrete-Time Models of Bond Pricing." In *Advanced Fixed-Income Valuation Tools*. Edited by Narasimhan Jegadeesh and Bruce Tuckman. New York: John Wiley & Sons.
- Bjerkstrand, P., and G. Stensland. 1996. "Implementation of the Black–Derman–Toy Interest Rate Model." *Journal of Fixed Income*, vol. 6, no. 2 (September):67–75.
- Black, F. 1976. "The Pricing of Commodity Contracts." *Journal of Financial Economics*, vol. 3:167–179.
- Black, F., and P. Karasinski. 1991. "Bond and Option Pricing When Short Rates Are Lognormal." *Financial Analysts Journal*, vol. 47, no. 4 (July/August):52–59.
- Black, F., E. Derman, and W. Toy. 1990. "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options." *Financial Analysts Journal*, vol. 46, no. 1 (January/February):33–39.
- Brennan, M., and E. Schwartz. 1979. "A Continuous Time Approach to the Pricing of Bonds." *Journal of Banking and Finance*, vol. 3, no. 2:133–156.
- . 1982. "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency." *Journal of Financial and Quantitative Analysis*, vol. 17, no. 3:301–329.
- Buetow, G., and F. Fabozzi. 2001. *Valuation of Interest Rate Swaps and Swaptions*. New Hope, PA: Frank J. Fabozzi Associates.
- Buetow, G., and R. Johnson. 2000. "A Primer on Effective Duration and Convexity." In *Professional Perspectives on Fixed Income Portfolio Management*, vol. 1. Edited by Frank J. Fabozzi. New Hope, PA: Frank J. Fabozzi Associates.
- Buetow, G., B. Hanke, and F. Fabozzi. Forthcoming 2001. "The Impact of Different Interest Rate Models on Effective Duration." *Journal of Fixed Income*.
- Burden, R., and J.D. Faires. 1998. *Numerical Methods*. Pacific Grove, CA: Brooks/Cole Publishing.
- Cox, J., J. Ingersoll, and S. Ross. 1985. "A Theory of the Term Structure of Interest Rates." *Econometrica*, vol. 53:385–407.
- Cox, J., S. Ross, and M. Rubinstein. 1979. "Option Pricing: A Simplified Approach." *Journal of Financial Economics*, vol. 7, no. 3:229–264.

Fabozzi, Frank. 1998. *Valuation of Fixed Income Securities and Derivatives*. 3rd ed. New Hope, PA: Frank J. Fabozzi Associates.

———. 2000. *Fixed-Income Analysis*. New Hope, PA: Frank J. Fabozzi Associates.

Harrison, J., and D. Kreps. 1979. "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory*, vol. 20:381–408.

Harrison, J., and S. Pliska. 1981. "Martingales and Stochastic Integrals in the Theory of Continuous Trading." *Stochastic Processes and Their Applications*:215–260.

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*, vol. 25, no. 4:419–440.

———. 1992. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology." *Econometrica*, vol. 60, no. 1:77–105

Ho, T., and S.B. Lee. 1986. "Term Structure Movements and Pricing Interest Rate Contingent Claims." *Journal of Finance*, vol. 41, no. 5 (December):1011–29.

Hull, J. 2000. *Options, Futures, and Other Derivatives*. 4th ed. Upper Saddle River, NJ: Prentice Hall.

Hull, J., and A. White. 1990. "Pricing Interest Rate Derivative Securities." *Review of Financial Studies*, vol. 3, no. 4:573–592.

———. 1993. "One-Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities." *Journal of Financial and Quantitative Analysis*, vol. 28, no. 2:235–254.

———. 1994. "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models." *Journal of Derivatives*, vol. 2, no. 1 (Fall):7–16.

———. 1996. "Using Hull–White Interest Rate Lattices." *Journal of Derivatives*, vol. 3, no. 3 (Spring):26–36.

Jamishidian, F. 1991. "Forward Induction and Construction of Yield Curve Diffusion Models." *Journal of Fixed Income*, vol. 1, no. 1 (June):62–74.

Jarrow, R. 1996. *Modelling Fixed Income Securities and Interest Rate Options*. New York: McGraw-Hill.

- Kalotay, A., G. Williams, and F. Fabozzi. 1993. "A Model for the Valuation of Bonds and Embedded Options." *Financial Analysts Journal*, vol. 49, no. 3 (May/June):35–46.
- Longstaff, F. 1989. "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates." *Journal of Financial Economics*, vol. 23, no. 2:195–224.
- . 1992. "Multiple Equilibria and Term Structure Models." *Journal of Financial Economics*, vol. 32, no. 3:333–344.
- Longstaff, F., and E. Schwartz. 1992. "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model." *Journal of Finance*, vol. 47, no. 4:1259–82.
- Neftci, Salih N. 1996. *An Introduction to the Mathematics of Financial Derivatives*. San Diego, CA: Academic Press.
- Rebonato, R. 1996. *Interest-Rate Option Models*. New York: John Wiley & Sons.
- Sundaresan, Suresh. 1997. *Fixed Income Markets and Their Derivatives*. Cincinnati, OH: Southwest College Publishing.
- Tuckman, B. 1996. *Fixed Income Securities Tools for Today's Markets*. New York: John Wiley & Sons.
- Vasicek, O. 1977. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, vol. 5, no. 2:177–188.