

Abstract

CROTTY, MICHAEL THOMAS. Assessing the Effects of Variability in Interest Rate Derivative Pricing. (Under the direction of Peter Bloomfield.)

Interest rate derivatives are financial instruments similar in spirit to stock options. The value of such a derivative depends on the level of a particular interest rate at a designated time. One type of interest rate derivative is an interest rate cap, which pays a notional amount multiplied by the amount by which a specified interest rate exceeds the strike rate at periodic intervals until maturity. Many models exist for pricing interest rate derivatives. This research uses the Hull-White model (Hull, 2003) for the short rate, implemented with a trinomial tree. This model can be used to determine pricing for interest rate derivatives.

The first part of this research explores various spline methods for estimating the term structure of the zero rate curve, one of the inputs into the Hull-White trinomial tree. The zero rate is the interest rate that would be earned on a bond that has no intermediate coupon payments and pays face value at maturity. The zero rate curve is modeled with splines extending an approach developed by Fisher et al. (1994).

The second part of this research uses bootstrap techniques to determine the variability of the zero rate curve at various maturities and to propagate the variability of the zero rate curve into the Hull-White pricing model. The latter bootstrap includes the entire method of derivative pricing from the calculation of a zero curve using a set of bond prices through entering that zero curve into the pricing model, and on to the pricing of a derivative.

ASSESSING THE EFFECTS OF VARIABILITY IN INTEREST
RATE DERIVATIVE PRICING

BY

MICHAEL T. CROTTY

A DISSERTATION SUBMITTED TO THE GRADUATE FACULTY OF
NORTH CAROLINA STATE UNIVERSITY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

STATISTICS

RALEIGH, NORTH CAROLINA

2 NOVEMBER 2006

APPROVED BY:

DR. PETER BLOOMFIELD (CHAIR)

DR. DAVID DICKEY

DR. SUJIT GHOSH

DR. DENIS PELLETIER

Dedication

To my loving wife Niamh.

Without your support and encouragement, this never would have been possible.

Biography

Michael Crotty was born in Moline, Illinois, on January 31, 1980 to Thomas and Dawn Crotty. He has two younger brothers, Kevin and Matt. He grew up in Rock Island, Illinois, before moving to Raleigh, North Carolina, in early 1991. He graduated from Sanderson High School in 1998. He then attended North Carolina State University and majored in statistics. During his undergraduate years, he met his wife Niamh through the Catholic Campus Ministry leadership team. They married on October 4, 2003 in Winston-Salem, North Carolina, and they are expecting their first child in 2007.

Michael was awarded the SAS Institute Scholarship for undergraduate studies in statistics at North Carolina State University. This scholarship included summer internships at SAS Institute in Cary, North Carolina. He graduated as a valedictorian in 2001. In August of that year, he began the graduate program at NC State, having been awarded the Gertrude Cox Fellowship for his first year. In May 2003, he completed his Masters in Statistics degree, and he began coursework for his PhD in Statistics in the fall of 2003. He also began working at this time with Dr. Peter Bloomfield on research in financial statistics. During his graduate career, he spent many semesters as a teaching assistant for Bill Hunt's undergraduate statistical research class. This allowed him to meet many statisticians and researchers in the US Environmental Protection Agency and the NC Department of Environment and Natural Resources.

Outside of statistics, Michael's interests are traveling, sports and spending time with family and friends. He has been to 40 of the 50 states and many European countries. He enjoys playing tennis, racketball, basketball, golf, and card games with friends. He is a lifelong fan of the Chicago Cubs and Wolfpack athletics.

Acknowledgements

This research was supported primarily through the VIGRE grant awarded to the statistics department at North Carolina State University through the National Science Foundation. The department also provided office space, computers, travel funding, and a helpful staff. My committee of Dave Dickey, Sujit Ghosh, Denis Pelletier, and my advisor, Peter Bloomfield, have been encouraging and helpful throughout the process of completing this research.

Much of my education was also funded by SAS Institute; I am grateful to the many people at SAS with whom I had the opportunity to work. I also appreciate the confidence Bill Swallow had in me to recruit me into the field of statistics (and away from the blue university down the road). I have also benefitted greatly from working with Bill Hunt, both as an undergraduate student and as a teaching assistant for his undergraduate research program. In this vein, there are many other teachers and professors through the years that have helped shape me as a person and as a scholar.

My fellow students in the program have been an important source of moral support. Whether it was working together on homework assignments in the first few years or being a sounding board for ideas with research progress, many students have helped me through this process. I would especially like to thank Clay, Joe, Darryl, Steve, Hugh, Jimmy, Ross, Kirsten, Lavanya, and Matt Gribbin.

Sometimes, the best way to get away from the stress of graduate school was to literally get away from it. That's where friends outside of the department really came through for me over the years. Tuesdays dinners, trips to the beach, late nights playing

cards, and going to wine and beer festivals around North Carolina – all these things helped lower my stress level during graduate school. To Chris, Katie, Pepper, Mary, Cathy, Kyle, Kristin, Carl, and Melissa, thank you so much!

All of my family has been extremely encouraging and supportive throughout the entire process of my education, especially the past few years. My extended family members on all sides were always checking in on how I was progressing and always proud of my accomplishments. Both of my immediate families have been very supportive. Thanks go out to all my siblings – Matt, Kevin, Cormac, Padraig, and Sinead – and my parents-in-law, Pat and Teresa.

I am also very appreciative of the support given to me throughout everything I have done in life from my parents, Tom and Dawn. They have provided so much support through the years and helped me become the person I am today.

Last, and *certainly* not least, I would like to thank my wife Niamh. She has been by my side through all of my graduate school years. She has been there to celebrate the triumphs and been a source of strength for me to keep going through the frustrating times. The best thing I found during my years at North Carolina State was, quite simply, the love of my life. Thank you, Niamh, for putting up with me going through school. You're the best!

Table of Contents

List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Background	2
1.1.1 Interest Rate Model Research	2
1.1.2 Spline Research	4
1.1.3 Bootstrap Research	6
1.2 Literature Review	7
1.2.1 Interest Rate Model Literature	8
1.2.2 Spline Literature	14
1.2.3 Bootstrap Literature	16
1.3 Hull-White Yield Curve Model	19
1.3.1 Model Description and Discretization	20
1.3.2 Pricing Derivatives Using the Trinomial Tree	22
2 Modeling the Term Structure	26
2.1 Description of the Data and Notation	27
2.2 Comparison of Two Cubic Spline Bases	28
2.3 Effects of Changing the Timescale	30

2.4	Incorporating Weights into the Spline Fit	38
2.4.1	Optimal k Values	42
2.5	Comparing the Spline Methods	44
2.6	Discussion of the Treasury Premium	47
2.7	Bypassing the Zero Curve Spline	49
2.8	Background Calculations	49
2.8.1	Calculations for $l(\tau)$	50
2.8.2	Calculations for $z(\tau)$	53
2.8.3	Calculations for $\alpha(\tau)$	54
2.8.4	Derivation of FNZ GCV Criterion	57
3	Bootstrap Analysis	59
3.1	Description of the Method	60
3.2	Bootstrap Results	62
3.2.1	Discount Curve Spline Bootstrap Results	62
3.2.2	Caplet Pricing Bootstrap Results	63
3.3	Bootstrap Analysis of the $\alpha(\tau)$ Spline	71
	Bibliography	77
	Appendix	80
A	Trinomial Tree Construction	81

B λ Comparison Tables	87
---	----

List of Tables

1.1	Zero Rates Used in Pricing Example	23
2.1	Spline Comparison (14 Jun 2005)	29
2.2	Spline Comparison (28 Feb 2006)	29
2.3	SSE Comparison (14 Jun 2005)	34
2.4	Mean and Standard Deviation Comparison (14 Jun 2005)	35
2.5	SSE Comparison (28 Feb 2006)	36
2.6	Mean and Standard Deviation Comparison (28 Feb 2006)	37
2.7	Updated Entries of Fisher et al. (1994) Table	54
B.1	λ Values in Weighted Case (14 Jun 2005)	87
B.2	λ Values in Weighted Case (28 Feb 2006)	89

List of Figures

1.1	Trinomial Tree Illustration	24
2.1	Standard Timescale (14 Jun 2005)	31
2.2	Square Root Timescale (14 Jun 2005)	32
2.3	Zero Curve Comparison (14 Jun 2005)	33
2.4	Standard Timescale (28 Feb 2006)	36
2.5	Square Root Timescale (28 Feb 2006)	37
2.6	Zero Curve Comparison (28 Feb 2006)	38
2.7	k Value Comparison (14 Jun 2005)	43
2.8	k Value Comparison (28 Feb 2006)	44
2.9	Yield Curve Estimation Comparison (14 Jun 2005)	45
2.10	Yield Curve Estimation Difference Plot (14 Jun 2005)	46
2.11	Yield Curve Estimation Comparison (28 Feb 2006)	47
2.12	Yield Curve Estimation Difference Plot (28 Feb 2006)	48
3.1	Discount Curve Bootstrap Results (14 Jun 2005)	63
3.2	Discount Curve Bootstrap Results (28 Feb 2006)	64
3.3	Caplet Pricing Bootstrap Means (outside knots)	65
3.4	Caplet Pricing Bootstrap Coefficients of Variation (outside knots)	66
3.5	Caplet Pricing Bootstrap Means (inside knots)	67
3.6	Caplet Pricing Bootstrap Coefficients of Variation (inside knots)	68
3.7	Caplet Pricing Bootstrap Means (outside knots)	69
3.8	Caplet Pricing Bootstrap Coefficients of Variation (outside knots)	70

3.9	Caplet Pricing Bootstrap Means (inside knots)	71
3.10	Caplet Pricing Bootstrap Coefficients of Variation (inside knots)	72
3.11	Caplet Pricing Bootstrap Means (outside knots)	73
3.12	Caplet Pricing Bootstrap Coefficients of Variation (outside knots)	74
3.13	Caplet Pricing Bootstrap Means (outside knots)	75
3.14	Caplet Pricing Bootstrap Coefficients of Variation (outside knots)	76
A.1	Trinomial Tree Branching	82

Chapter 1

Introduction

This chapter will provide background of the research areas involved in this dissertation and a brief literature review of each. In Section 1.3, we discuss the Hull-White model for short term interest rates as well as various aspects of the trinomial tree implementation of this model used in this research. There is also a discussion of using the output of the trinomial tree to price derivatives.

The remainder of this dissertation is arranged as follows. In Chapter 2, we describe the bond data used in the subsequent analysis and discuss the procedure used to spline the interest rate. There are also sections in this chapter related to the details of fitting a spline to the interest rate term structure and a discussion of the treasury premium's relationship to using government bond prices to spline the risk-free interest rate. In Chapter 3, we describe the methods used to bootstrap the trinomial tree pricing process and present the results of the bootstrap analysis.

There are two appendices at the conclusion of this dissertation. Appendix A provides a detailed description of the construction of the trinomial trees used in this research. Appendix B provides tables of results too lengthy to include in the main body of Chapter 2.

1.1 Background

This section will cover background material in three different areas of research: a history of interest rate models, an account of the role of smoothing spline methods in our research, and a description of the statistical bootstrap.

1.1.1 Interest Rate Model Research

Research in the field of interest rate modeling is important for pricing interest rate derivatives. An interest rate derivative is a financial instrument in which the payoff is determined by a future level of a specified interest rate. Many types of interest rate derivatives exist; the three most popular are bond options, interest rate swap options, and interest rate caps/floors/collars. A bond option is an option where the underlying asset is a bond. An interest rate swap option, or swaption, is an option to enter into an interest rate swap where a specified interest rate is exchanged for a floating interest rate.

An interest rate cap is a derivative that pays a notional amount multiplied by the excess of a specified interest rate over the agreed-upon strike rate (also called the *cap rate*). The payment is structured as a stream of periodic payments, called caplets. The period length between payments is called the *tenor* of the agreement; generally, this length is three months (or a quarter of a year). An interest rate floor is similar to a cap, except that it has a floor rate, which is a strike rate where the payoff depends on the excess of the specified interest rate below the strike rate. An interest rate collar is a derivative combining a cap and a floor to give protection against interest rates going too high or too low.

Chapter 1. Introduction

Terminology in the field of interest rate modeling differs greatly. In Yan (2001), the author makes reference to the array of terminology:

The term structure of interest rates is also called “the yield curve of zero-coupon bonds.” With this correspondence in mind, I use “term structure” and “yield curve” interchangeably in this article. In market jargon, however, the yield curve may refer to yields to maturity of on-the-run coupon bonds.

The development of the primary models in this field is covered thoroughly in the literature review later in this introduction. Here we present a general roadmap of the model types. The first models for the interest rate dealt with the short rate, which is the annualized interest rate at which money can be borrowed for a very short time. The first two classes of short rate models developed were equilibrium models and no-arbitrage models. In an equilibrium model, the current term structure of the interest rate is produced as an output of the model. In a no-arbitrage model, the current term structure is an input of the model. Therefore, no-arbitrage models are able to reproduce the current term structure exactly.

There are many models beyond the short rate models initially developed. The next major set of models are called Heath-Jarrow-Morton (HJM) models. In an HJM model, the instantaneous forward rate is modeled instead of the short rate. The most recent developments in this field have been market models and infinite-dimensional models, both of which are discussed in further detail in the literature review.

As an example of a short rate model to motivate the problem of modeling the term structure of interest rates, we now present the Vasicek model in a general setting. In the framework presented above, the Vasicek model is an equilibrium model. The model is a stochastic process that contains both drift and standard deviation terms,

Chapter 1. Introduction

and mean reversion is incorporated into the model. This means that the process tends to have negative drift when rates are high and positive drift when rates are low. This corresponds to empirical results from the marketplace. Three constants, a , b , and σ , in the model must be calibrated using the prices of financial instruments in the marketplace. The model then uses the process $dr = a(b - r)dt + \sigma dz$ to model the short rate r .¹ The short rate r is pulled back towards the long-term mean b at a rate of a , and a normally distributed stochastic term σdz is added to the mean reversion. In this example, the instantaneous drift is $a(b - r)$ and the instantaneous standard deviation is σ . This model is presented in detail in Hull (2003), who gives formulae for pricing interest rate derivatives using the Vasicek model.

Further details of the progress of interest rate modeling are presented in Section 1.2.1, and Section 1.3 contains the technical details of the no-arbitrage class Hull-White model for the short rate.

1.1.2 Spline Research

Spline methods were first introduced by I. J. Schoenberg (Schoenberg, 1946); in the 1960s and 1970s, a number of researchers including J. H. Ahlberg, Christian Reinsch, and Grace Wahba built upon the ideas of Schoenberg. A brief review of the literature as it relates to this research is included in Section 1.2.2. This section will give a brief non-technical introduction to smoothing splines.

A spline allows a user to fit a smooth curve to a set of data. Spline methods come in many variations, but the general concept consists of dividing the range of the data into

¹This process is similar to a process known as an Ornstein-Uhlenbeck (O-U) process in other research disciplines. See pages 271 and 461 of Neftci (2000). An O-U process is equivalent to the Vasicek model with $b = 0$.

Chapter 1. Introduction

sections through the use of “knots”; a knot is a point that divides two such sections of the data. Between knot points, a smooth curve is fit to the data, and conditions are met so that the curves in each section of the data connect at each knot point. Generally, derivative conditions must be met at the knot points to ensure smoothness in the spline. The result is a smooth curve that fits the data according to a criterion defined by the spline method itself.

Since this research involves cubic splines, we will briefly describe the procedure for cubic spline construction. First, knot points are chosen; these can be chosen to be equally spaced or not, according to the situation. (In this research, we choose the knots points using a quantile approach so that the knot points are closer together where the data is more dense and farther apart where the data is more sparse.) In each section of the data between knot points, a cubic function is fit to the data. These cubic functions must meet each other at the knot points such that their first and second derivatives are continuous, ensuring a smooth spline function over the entire data set.

Historically, splines were developed outside of the mathematical arena. They were first used by draftsmen and the aviation industry. The use by draftsmen is described at the beginning of Ahlberg et al. (1967):

For many years, long, thin strips of wood or some other material have been used much like French curves by draftsmen to fair in a smooth curve between specified points. These strips or splines are anchored in place by attaching lead weights called “ducks” at points along a spline. By varying the points where the ducks are attached to the spline itself and the position of both the spline and the duck relative to the drafting surface, the spline can be made to pass through the specified points provided a sufficient number of ducks are used.

From this practice, Schoenberg (1946) developed the mathematical spline.

Chapter 1. Introduction

Finally, we mention the smoothing spline, which is a variation of a cubic spline in that it adds a penalty term to the minimization criterion. This term penalizes excess roughness (non-smoothness) in the spline fit. In general, the more knot points that are used, the less smooth the fit will be. The use of smoothing splines allows the user to use a large number of knot points and then smooth out the resulting fit as needed. The method used to determine the value of the coefficient on the penalty term is generalized cross validation (GCV). Generalized cross validation is a variation of ordinary cross validation developed by Craven and Wahba (1979) for the case of spline fitting. Ordinary cross validation for splines involved far too much computational time, and GCV was able to shorten that time. More discussion of smoothing splines and GCV is found in Section 1.2.2 and Chapter 2.

1.1.3 Bootstrap Research

The statistical bootstrap was introduced in 1979 by Bradley Efron. He introduced it in light of the much older jackknife procedure, and detailed how the jackknife can be thought of as a linear expansion method for approximating the bootstrap. The statistical jackknife and statistical bootstrap are resampling methods developed to estimate the bias and standard error of a statistic. The jackknife was developed first. The method of the jackknife is to successively remove one observation from a sample and calculate the statistic of interest from this “jackknife sample.” There are n such jackknife samples in a sample of size n observations. The jackknife bias and standard error can then be computed using formulae provided in Chapter 11 of Efron and Tibshirani (1993).

Chapter 1. Introduction

In contrast to the jackknife, the statistical bootstrap resamples from a sample with replacement. In the bootstrap world, the sample that has been collected is treated as a substitute for the overall population, and another sample of the same size is taken from the initial sample with replacement. This process is repeated many times, in such a way that B independent bootstrap samples are taken, each consisting of n observations selected with replacement from the initial sample of n values. For each of the B bootstrap samples, the statistic of interest is computed, and the standard error is estimated by the sample standard deviation of the B replications. The bootstrap method allows estimation of standard errors with fewer distributional assumptions being made regarding the data in question. Further details of the computations for the single variable bootstrap will be provided in the beginning of Chapter 3.

The bootstrap has also been used for more complicated data structures than univariate data. Chapters 8 and 9 in Efron and Tibshirani (1993) discuss the use of bootstrap methods in two-sample problems, time series problems, and regression problems.

1.2 Literature Review

This section reviews the literature in interest rate modeling, spline methods, and bootstrap methods. Two primary sources, the textbook *Options, Futures, and Other Derivatives* (Hull, 2003) and a Federal Reserve working paper entitled *Fitting the term structure of interest rates with smoothing splines* (Fisher et al., 1994), are briefly described in this introductory section. As in the Background section, the literature review will be organized into three subsections corresponding to the three areas of research we built upon in the current research.

Chapter 1. Introduction

Hull (2003) is a textbook designed for students and financial practitioners familiar with finance and statistics concepts. Chapter 23 of Hull (2003) discusses numerous models for the short rate, including the Hull-White model. A section in this chapter describes the implementation of the trinomial tree that we used in the course of this research.

Fisher et al. (1994) is a working paper published through the Federal Reserve that describes a method to spline the term structure of interest rates using the market prices of treasury bonds. We used this approach in our current research to provide an input into the trinomial tree method for pricing interest rate derivatives.

1.2.1 Interest Rate Model Literature

Pricing models for interest rate derivatives date back to the 1980s when these instruments were first developed in the marketplace. The earliest approach was to modify the Black-Scholes (1973) model for pricing general options. However, Hull (2003) mentions that this model is appropriate when interest rates are assumed to be constant or deterministic. Hull cites two questions that arise when interest rates are assumed to be stochastic in this model: first, the Black-Scholes approach ignores that interest rates are stochastic when discounting back to present value and second, this approach ignores the fact that the forward price is not the same as the futures price. In light of these concerns, Hull shows later that setting the futures price equal to the forward price discounted at today's T-year maturity zero rate is correct. This result shows that this approach "therefore has a sounder theoretical basis and wider applicability than is sometimes supposed (Hull, 2003, Section 22.1)." For pricing interest rate derivatives,

Chapter 1. Introduction

interest rates must be assumed to be stochastic because constant or deterministic interest rates would make option pricing a trivial exercise.

For these reasons, a rich literature now suggests new models for the term structure of interest rates. These include, in particular, two classes of “classic” models: equilibrium models and no-arbitrage models. Both of these classes will be described in this introduction. The most recent developments beyond these two classes of models will also be reviewed. Special attention will be paid to the no-arbitrage class of term structure models since it contains the Hull-White model, which is of primary focus in this dissertation.

Equilibrium Models

The earliest departures from the Black-Scholes approach were the class of equilibrium models, which use assumptions regarding economic variables to describe the evolution of the zero curve through time. In one-factor equilibrium models, all the rates move in the same direction, but not necessarily by the same amount (Hull, 2003). The Rendleman and Bartter model is the simplest one-factor equilibrium model. The process for the short-term interest rate is modeled like a stock price using geometric Brownian motion without mean reversion (Rendleman and Bartter, 1980). Hull (2003) notes that this is less than ideal since interest rates have historically reverted back to a long-term mean. The Vasicek model, which incorporated mean reversion, addressed this concern. However, this model had a downside as well since it allowed negative interest rates. Although most research seems to agree that negative interest rates are inappropriate in term structure models, Black (1995) has a brief discussion of this issue. Black notes that since the short rate is bound by zero, it is itself like an option. He states that the

Chapter 1. Introduction

real interest rate may become negative, but the short rate cannot since individuals will not keep their money in a negative interest-bearing instrument. They will instead put their money in currency and receive an interest rate of zero percent.

The most popular of the one-factor equilibrium models is the Cox, Ingersoll, and Ross model (denoted CIR model). Since it uses a square root process in the volatility term, the model does not allow negative interest rates. This also implies that as interest rates increase, the standard deviation of changes in the interest rate will increase as well. The CIR model also incorporates mean-reversion like the Vasicek model (Hull, 2003). Of these three models, the CIR model has received the most attention in the term structure modeling literature. However, the Vasicek model is the one that was extended by Hull and White to develop the Hull-White model. Section 1.3 contains more detailed explanations of the Hull-White model, including the formulae behind it.

Two-factor equilibrium models receive only a brief discussion in Hull (2003) and in Hunt and Kennedy (2004). Hull (2003) mentions two models, Brennan-Schwartz and Longstaff-Schwartz, that both have a second factor that allows for more flexibility in modeling the term structure. Hunt and Kennedy (2004) note that two-factor models are rarely used in derivative pricing primarily because of the difficulty in implementation.

No-arbitrage Models

An alternative approach to equilibrium models is the class of no-arbitrage models, which contains the Ho-Lee and Hull-White models. Equilibrium models do not necessarily match the current term structure, but no-arbitrage models are designed so that today's term structure is matched exactly. The current term structure is an out-

Chapter 1. Introduction

put of an equilibrium model, but it is an input to a no-arbitrage model (Hull, 2003). Ho and Lee (1986) presented the first no-arbitrage model. Their model was initially described as a binomial tree of bond prices that included parameters to account for the standard deviation of the short rate and for the market price of risk of the short rate. Hull (2003) presents a continuous time limit of the discrete time process initially presented by Ho and Lee. In Hull and White (1990), the Vasicek model was extended so that it took the current term structure as an input to the model. (Hunt and Kennedy (2004) call the Hull-White model the Vasicek-Hull-White model since it is an extension of the Vasicek model.) The Hull-White model also improved upon the Ho-Lee model by incorporating mean-reversion, which the Ho-Lee model lacks. Both of these models take the initial zero curve as an input, but neither model requires that the zero curve be differentiable. Some no-arbitrage models can be implemented in a discrete-time sense through the use of trees. Section 1.3.1 details the construction of a trinomial tree in the case of the Hull-White model.

No-arbitrage models, like equilibrium models, began with only one function of time, but by making a second parameter also a function of time, no-arbitrage models can be more closely fitted to current market data of actively traded options. Hull (2003) briefly discuss two such extensions, the Black-Derman-Toy binomial tree procedure and the Black-Karasinski model. The downside to these models is that they can produce future volatility term structures that are quite different from the current volatility term structure.

All of these models also involve calibration to determine the values of their parameters. Hull (2003) gives a general description of the calibration procedure. The parameters are determined from market data on actively traded options using a goodness-of-fit

Chapter 1. Introduction

measure such as minimizing the sum of squared differences between the market price and the model price of each instrument being used for calibration.

Pricing interest rate derivatives using the interest rate models discussed thus far has two main limitations. First, they only have one source of uncertainty. Second, they do not offer complete freedom in choosing the volatility structure. More recent efforts have addressed these issues by including a second factor (source of uncertainty). These efforts include models developed by Duffie and Kan (1994) and Hull and White (1994). The two-factor Hull-White model still matches the initial term structure exactly, but since it has a second source of uncertainty, it allows for more variation in the term structure to better capture the differing amounts of volatility in the forward rate throughout the term to maturity.

More Recent Models

Another approach to modeling the yield curve involves modeling the processes that are followed by instantaneous forward rates. The primary model in this approach is the Heath-Jarrow-Morton (HJM) model. The HJM model is a generalization of the Ho-Lee model to multiple factors over continuous time. The HJM model takes the initial forward rate curve as given and describes its evolution across time using a continuous time stochastic process (Heath et al., 1992).

Hull (2003) notes that one negative aspect of the HJM model framework is the difficulty of calibrating the model to the prices of actively traded instruments. Yan (2001) also points out that since HJM models use continuous compounding of the instantaneous forward rate curve, specifying a lognormal process in the HJM framework is no longer possible. In light of these issues, the most recent work in interest rate

Chapter 1. Introduction

models has been on two fronts: LIBOR (London Interbank Offered Rate) or market models and so-called “infinite-dimensional” models. Yan (2001) claims that the market models enjoy the most widespread use in the financial industry.

Market models deal directly with observable rates in the marketplace, and they give the user complete freedom in specifying the term structure of the volatility. Brace et al. (1997), Jamshidian (1997), and Miltersen et al. (1997) all developed models of the forward rate that are used by traders in practice. Miltersen et al. (1997) present a model that provides closed form solutions for interest rate derivative prices that coincide with results obtained from a modified Black-Scholes formula. They also show that by specifying a particular volatility structure, the lognormal assumption on forward rates is consistent with the HJM model. Hull (2003) points out that both HJM and LIBOR models have the disadvantage of not being easily implemented using recombining trees; instead, Monte Carlo simulation must be used to implement these models. (Recombining trees are binomial or trinomial trees where the paths taken by going down and then up or up and then down lead to the same node. Recombining trinomial trees have $2n + 1$ nodes at timestep n , whereas non-recombining trinomial trees have 3^n nodes at timestep n .)

The final area of interest rate models is the class of “infinite-dimensional” models; this area also goes by the terms “random-field” or “stochastic-string” models (Yan, 2001). In recent years, a number of developments in this area have been presented. In Sornette (1998), the term structure is formulated in such a way that the forward curve is viewed as the deformation of a “string.” The article details the analogy of the forward curve in terms of string theory in physics. Goldstein (2000) uses random fields to model the term structure of interest rates. This approach is deemed advantageous

Chapter 1. Introduction

since it does not require recalibration, it provides a parsimonious description of term structure dynamics, and it naturally accounts for the fact that the best hedge for a bond is one of a similar maturity. Goldstein also points out that random field models of the term structure are generalizations of finite-factor models, but they are consistent with the current yield curve and term structure innovation. Yan (2001) recognizes that infinite-dimensional models provide straightforward pricing of bond options. Finally, Izzi and Racheva (2002) present a new model based on the assumption that the short rate follows a mean-reverting jump-diffusion process that is the sum of one Brownian motion and a finite number of Poisson processes. However, this model requires the development of a recombining hexanomial tree that fits the current term structure observed in the market.

1.2.2 Spline Literature

As mentioned in Section 1.1.2, the mathematical spline was adapted from industrial practice by I. J. Schoenberg (Schoenberg, 1946) and developed further in the 1960s and 1970s. This section will describe a few of the developments at that time in terms of smoothing splines and generalized cross validation (GCV). The literature review presented here will be restricted to the part of spline research used in this dissertation, specifically the use of smoothing splines to estimate the zero curve.

In 1964, Schoenberg combined the work done thus far on spline interpolation and the problem of graduation of data discussed by E. T. Whittaker (Whittaker and Robinson, 1926). (Whittaker had introduced the problem of graduation, which was a precursor to splines and smoothing of data; many methods were proposed in Chapter 11 of

Chapter 1. Introduction

Whittaker and Robinson (1926).) Schoenberg used one of these methods in relation to spline functions; this led to the introduction of a criterion for evaluating the goodness of fit of a spline curve to data. This two part minimization criterion problem laid the ground work for the GCV work in subsequent research. Essentially, the first part of the minimization problem measures the distance of the spline curve from the data points and the second part measures the smoothness of the spline curve.

Reinsch (1967) modified the approach of Schoenberg to replace interpolation of the data with an algorithm for smoothing experimental data. In this article as well as Reinsch (1971), Reinsch cites a result from Schoenberg (1964) that states that the unique solution to the two part minimization problem described above is a natural spline of degree $2m - 1$ where m is less than or equal to the number of knot points.

The problem that arose from the work to this point, however, was the tradeoff between the two parts of the minimization problem. In the general problem, a single parameter (λ) coefficient on the second part of the problem measured the roughness of the spline curve. Since this parameter somehow had to be chosen, the choice of this parameter was the focus of much work. Reinsch (1967) presents an interval (dependent upon σ^2) within which to choose the roughness parameter. (Here, σ^2 is the variance of the error terms between the observed data points and the true but unknown curve.) Wahba (1975) suggests choosing the roughness parameter to be less than σ^2 by a “fudge factor”; this approach allows the author to determine an optimal λ value when σ^2 is known.

Since σ^2 is not generally a known value, Craven and Wahba (1979) developed an optimal method of smoothing that did not require knowledge of σ^2 . This method, known as generalized cross validation, is a generalization of ordinary cross validation

Chapter 1. Introduction

that allows for faster computation, an important consideration at a time when computational time was more critical than today. Further details of the GCV calculations are included in Chapter 2.

Fisher et al. (1994) utilize the GCV formulation for smoothing splines to estimate the yield curve from bond data. The entire formulation is summarized in Wahba (1990). Fisher et al. (1994) use the GCV method to choose an appropriate value of λ . They cite that the advantage to such an approach is that “the shape of the spline is controlled by a single parameter.” Chapter 2 also contains further details of the use of GCV in the estimation of the yield curve.

1.2.3 Bootstrap Literature

As mentioned in Section 1.1.3, the concept of the statistical bootstrap was introduced by Bradley Efron (Efron, 1979). Since then, the bootstrap methodology has been used in a variety of situations. This section will focus on the part of the literature dealing with regression problems. The basic bootstrap methodology deals with univariate independent and identically distributed samples, but in the case of regression problems, the data are no longer univariate and generally correlation exists between the independent variables. Therefore, the basic bootstrap has been extended to handle regression problems.

Extensions to the initial bootstrap methodology appeared soon after Efron’s original article. For instance, Bickel and Freedman (1981) provide asymptotic theory to show that the bootstrap introduced by Efron is valid in a large number of situations, but they also show results of the bootstrap failing for estimating θ from a uniform $(0, \theta)$

Chapter 1. Introduction

distribution. The first work to extend the bootstrap to regression problems was published two years after Efron's introduction of the bootstrap. Freedman (1981) examines a bootstrap approach for estimating the distribution of linear least squares estimates. The regression bootstrap resamples from the bootstrap centered residuals. The centered residuals are calculated by subtracting the mean of the residuals from each of the residuals, which are the differences of the observed value minus the predicted value of the dependent variable. This approach relies on the residuals obtained from the initial regression model to determine the bootstrap samples. The author notes that the bootstrap will generally fail if the residuals are not centered prior to performing the resampling procedure. This article also includes details for bootstrapping in the multiple linear regression case.

Efron and Gong (1983) provide a simple explanation of the bootstrap procedure and include an example of using the bootstrap to estimate the correlation coefficient in the simple linear regression setting. They claim that "it can be applied just as well to any statistic, simple or complicated, as to the correlation coefficient." Bickel and Freedman (1983) consider the bootstrap for a regression model where both the number of data points and the number of parameters are both large. Their results focus on the ratio of the number of parameters to the number of data points; if this ratio is small, the bootstrap approximation to the distribution of contrasts is valid. They also show that if this ratio does not tend to zero, the bootstrap approximation is not valid.

The first application of the bootstrap to generalized least squares appeared one year later (Freedman and Peters, 1984); the authors describe the bootstrap as a "technique for estimating standard errors," and apply the bootstrap in the context of an econometric linear model. The specific model that they present describes industry demand

Chapter 1. Introduction

for energy. They emphasize that the initial regression model must be valid in order for residual-based bootstrap techniques to work. This concern is also voiced by Weber (1985), where two different methods of regression bootstrapping are compared. The first method, the one discussed in the literature up to this point, relies heavily on the regression model by resampling from the residuals to reconstruct dependent variable values in the bootstrap samples. The author notes that this approach is susceptible to model misspecification since the bootstrap samples rely upon the model being correctly specified. The second approach, which Weber favors, is to resample with replacement directly from the original data points to create bootstrap samples. He states that this approach is robust against heterogeneity of the error variance, noting that despite Efron and Gong (1983) mentioning this observation in passing, it is not emphasized in the bootstrap literature. This second approach is also easier to implement than the residual bootstrap.

Shao (1988) shows that the bootstrap method implemented by resampling residuals, the first method described in the previous paragraph, is asymptotically correct only in the case of a homoskedastic error model. This article also finds that bootstrapping residuals in the heteroskedastic case results in the variance estimator having a larger mean square error than in the jackknife estimator.

Finally, Efron and Tibshirani (1993, Chapter 9) weigh in again to discuss the two different approaches to the regression bootstrap first compared by Weber (1985). Efron and Tibshirani call these approaches “bootstrapping residuals” and “bootstrapping pairs,” and they prefer bootstrapping pairs. However, they note that “bootstrapping is not a uniquely defined concept.” They also include examples of using the bootstrap in regression problems.

1.3 Hull-White Yield Curve Model

In this section, various aspects of the Hull-White model for short term interest rates are discussed. Section 1.3.1 presents the model and a brief description of the trinomial tree discretization of the model. (Appendix A contains a full description of the trinomial tree construction used in this research.) Section 1.3.2 presents an example of pricing an interest rate derivative using the trinomial tree.

Before describing the Hull-White model in detail, we take a moment to put the model in context of the history of interest rate models. Much of this history is covered in Section 1.2.1. The reader is reminded of four classes of interest rate models: equilibrium, no-arbitrage, market (or LIBOR), and infinite-dimensional. The first type of models developed were equilibrium models, but these models did not match the initial term structure of interest rates exactly and led to mispricing of bonds and more severe mispricing of options. (Hull (2003) notes that a 25% error in option pricing can result from as little as a 1% error in the pricing of the underlying bond.) For this reason, no-arbitrage models were developed; these models use the current term structure of interest rates as an input and are then able to match today's term structure exactly. The Hull-White model belongs to the no-arbitrage class of models.

More recent developments in interest rate modeling are related in the literature review. In general, the models have become more complex as computing power increases and as further extensions of previous models seem warranted.

1.3.1 Model Description and Discretization

The Hull-White model (Hull, 2003, page 546) is a model for the short rate that provides an exact fit to the initial term structure. The short rate (or short-term risk-free rate) is the interest rate applicable for a very short period of time. The term structure of interest rates is the relationship between interest rates and their maturities. The Hull-White model is specified as follows:

$$dr = [\theta(t) - ar]dt + \sigma dz \quad (1.1)$$

where r is the short rate, z is a Wiener process, and a and σ are constants. A Wiener process is a Brownian motion; it has mean change of zero, a variance rate of one (per year), and the values of its changes for any two different short time intervals are independent (Hull, 2003). The function $\theta(t)$ is then calculated from the initial term structure:

$$\theta(t) = F_t(0, t) + aF(0, t) + \frac{\sigma^2}{2a} (1 - e^{-2at}) \quad (1.2)$$

where $F(0, t)$ is the instantaneous forward rate for a maturity t at time zero. The subscript t denotes the partial derivative with respect to t . Hull notes that the third term of $\theta(t)$ is often small. Therefore, the majority of the drift of r is represented by $F_t(0, t) + a(F(0, t) - r)$, which is the slope of the instantaneous forward rate curve plus a term that causes r to revert back to that curve at a rate of a .

In the current research, a and σ are fixed constants. Further study is required to determine the effects of variability of estimates for a and σ on the eventual pricing of interest rate derivatives using the Hull-White model.

Chapter 1. Introduction

Hunt and Kennedy (2004, Chapter 17) praise the Hull-White model for its tractability, stating that this is important so that the model can be efficiently calibrated to market prices. The authors show that the solution to the Hull-White stochastic differential equation is a Gaussian process, which allows users to calibrate the model to bond and option prices. Also, the distribution of the short rate r_t (given r_s , where $t \geq s$) is Gaussian.

The Hull-White model is a continuous-time model. However, in the current analysis, a discrete-time approximation will be used. The implementation of the Hull-White model that will be used in this analysis is through the use of a recombining trinomial tree (Hull, 2003, page 552). A trinomial tree starts at time zero with one node and has three branches that extend to the right (one timestep into the future) to three nodes. From each of those three nodes, three more branches extend to nodes at the next timestep for a total of five nodes. In this work, we use recombining trees, meaning that branches recombine from one timestep to another. (For an example of a recombining trinomial tree, see Figure 1.1.) Hence, the total number of nodes at a given timestep is less than if this recombining did not take place. Indeed, a non-recombining trinomial tree has 3^n nodes at timestep n , whereas a recombining trinomial tree has only $2n + 1$ nodes at timestep n . The tree terminates on the horizontal (time) axis at a time called maturity. In this analysis, the timesteps are all the same length, and the amount of increase or decrease at each branch is initially the same. Hull (2003) notes that trinomial trees (rather than binomial trees) are preferred for interest rate trees since trinomial trees offer an extra degree of freedom which allows them to better match the mean reversion property of some interest rate models. See Appendix A for a more complete description of the trinomial tree construction.

A value for the zero curve is required at every timestep of the tree. For this reason, we use a spline approximation to the zero curve in Chapter 2. Also, in order to obtain values for the interest rate at each node at the final timestep of the tree, the value of the zero curve at one timestep past the final timestep of the tree is required.

1.3.2 Pricing Derivatives Using the Trinomial Tree

The trinomial trees constructed in this research can be used to price securities based on interest rates, including zero coupon bonds and interest rate derivatives, such as interest rate caplets. Caplets are the individual payments that make up the stream of payments comprising an interest rate cap. Interest rate caps are described in some detail in Section 1.1.1, but we will introduce them again here and go through a simple example of using a trinomial tree to price a caplet.

An interest rate cap is a financial derivative that uses a specified interest rate level as the underlying asset on which the payoff of the cap is determined. In this framework, a cap rate is compared to the floating rate at each of the specified payoff times. If the floating rate is above the cap rate, there is a payoff; otherwise, no payoff occurs at that time. These times are generally spaced quarterly throughout the year (once every three months), and this period of time is called the tenor of the cap agreement. Each payoff in the cap is called a caplet. If a payoff occurs, it is determined by multiplying the excess of the floating rate over the cap rate by the notional amount (specified as part of the cap agreement). An interest rate cap therefore provides the holder with insurance against rising interest rates, since the payoffs help offset other negative effects of rising rates.

Table 1.1: Zero Rates Used in Pricing Example

<u>Maturity (years)</u>	<u>Rate (%)</u>
1.0	3.824
2.0	4.512
3.0	5.086
4.0	5.500

We will now illustrate the use of the trinomial tree to price a caplet. The example tree is a modified version of the tree used in Hull (2003) as an illustration of constructing a trinomial tree. The zero rates used in this example are listed in Table 1.1. The caplet in this example will have a cap rate of 8%, a notional amount of \$10,000,000, and quarterly reset (payoff) dates. Thus, the tenor is three months. The caplet we are pricing is three years into the future. The trinomial tree will use the zero rates given in Table 1.1 and one timestep per year.

The resulting trinomial tree shown in Figure 1.1 is then used to price the caplet. Each node at the end of the tree is labeled with an interest rate (top line) and a Q value (bottom line). The interest rate value is the level of the interest rate at that particular node. The Q value for a particular node is the present value of a security that pays \$1 if that node is reached and pays nothing otherwise. (For more details on the construction of the trinomial tree and the calculation of the values in the tree, see Appendix A.)

Since the cap rate in this example is 8%, the only nodes that contribute to the pricing of the caplet are those that have interest rates greater than 8%. The interest

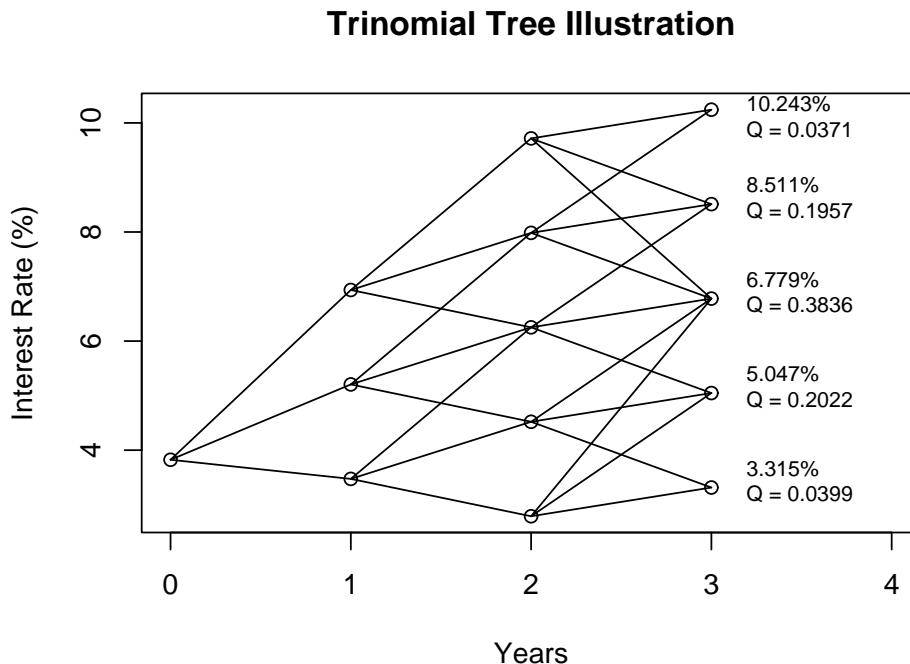


Figure 1.1: Trinomial Tree Illustration

rates throughout are expressed as annual rates, so for a quarterly reset date, the rates must be multiplied by 0.25. For each node above the cap rate, the excess of the node's interest rate over the cap rate is multiplied by the Q value of the node. Then, the price of the caplet is:

$$0.25 * [(0.10243 - 0.08) * 0.0371 + (0.08511 - 0.08) * 0.1957] * \$10,000,000 = \$4,580.45$$

This example shows how the trinomial tree described in Section 1.3.1 can be used to price an interest rate caplet. This pricing example required only the values of the

Chapter 1. Introduction

interest rate and Q for the final timestep. However, more complicated derivatives that are path-dependent may require calculations using values from nodes other than the nodes at the final timestep. In cases such as these, the tree construction is the same, but the values at nodes prior to the end of the tree must either be stored or used during the tree construction to determine the price of the derivative.

Another case mentioned at the beginning of this section is zero coupon bond pricing. The present value of a zero coupon bond maturing in three years in this example would be $De^{-3(.05086)} = 0.85849D$, where D is a notional amount for the bond; this result follows from the zero curve value at three years maturity being 5.086%. This value can also be derived from the Q values on the trinomial tree. Summing the Q values at the three year mark gives the value 0.8585, which is the same value as the present value of the three year zero coupon bond. This result holds because the Q values represent the present value of a security that pays \$1 if a particular node is reached. For a zero coupon bond, however, it does not matter which node is reached at maturity; the bond will pay off its face value. Therefore the present value of the bond is the sum of the present values of all the nodes at the bond's maturity.

Chapter 2

Modeling the Term Structure

The trinomial tree algorithm of Section 1.3 requires a value of the zero curve at every timestep. Therefore, to explore the effects on pricing of decreasing the intervals between timesteps, we need to model the term structure of the zero curve such that it can be evaluated at every timestep. For this purpose, we adapt and expand upon a spline method proposed for the term structure of interest rates by Fisher et al. (1994). The method involves the use of smoothing splines with a cubic B-spline basis. We instead use smoothing splines with a natural spline basis; this choice is explained further in Section 2.2. The smoothing spline is based on a regression spline and is then smoothed through the use of a smoothing parameter λ . This approach, recommended by Fisher et al., uses a large number of knot points and penalizes excess variability using the smoothing parameter. To choose λ , we minimize a generalized cross validation (GCV) value over a grid search for the optimal λ . We further explore extensions to the Fisher et al. (1994) method by comparing different timescales and adding a weighting scheme to the data to account for unequal amounts of variability at the short and long term maturities. We also extend their method to spline the zero curve directly; see Section 2.8.2 for a more detailed explanation of this extension.

The remainder of this chapter describes the process of modeling the term structure. Section 2.1 describes the data used to develop the term structure spline, Section 2.2

reports an analysis comparing the cubic B-spline basis and the natural spline basis, and Sections 2.3 through 2.5 describe the results of the various adaptations to the Fisher et al. (1994) method. Section 2.5 compares our spline results to those of the US Treasury Department and also summarizes the comparisons between the many variations of the spline process. In Section 2.6, we include a brief discussion of the treasury premium as this topic relates to our research. Section 2.7 describes a spline method to bypass the spline for the zero curve. Finally, Section 2.8 contains background calculations for the chapter.

2.1 Description of the Data and Notation

The treasury security price data we use to model the term structure of the interest rate was obtained from the 15 June 2005 *Wall Street Journal* (2005) and the 1 March 2006 *Wall Street Journal* (2006); we refer to these two data sets as the June 2005 data set and the February 2006 data set. All treasury bills and notes were used. Treasury bonds were also used, with the exception of inflation-indexed issues and callable bonds (denoted in USGAO (1999)). This follows the guidelines set forth by Fisher et al. (1994). The maturity dates were obtained from the Treasury department website¹. These dates were used to determine the amount of accrued interest to add to the quoted price for interest-bearing bonds and notes. (Bills are zero-coupon instruments, and therefore they have no accrued interest.)

The calculation of accrued interest was performed according to Fabozzi (1997); the semiannual coupon payment was multiplied by the accumulated time since the last

¹<http://www.publicdebt.treas.gov/of/ofaicqry.htm>

coupon payment (measured as the number of days since the last payment divided by the number of days in the current coupon payment period). This accrued interest amount was then added to the security's quoted price.

We use the notation $\delta(\tau)$ to refer to the discount function of a one unit zero coupon bond maturing at time τ . We define $l(\tau) = -\log(\delta(\tau))$, and the zero curve $z(\tau) = l(\tau)/\tau$. We also denote the natural spline basis functions with the notation ϕ and the coefficients on those basis functions β .

2.2 Comparison of Two Cubic Spline Bases

Fisher et al. (1994) advocate the use of a cubic B-spline basis to spline the zero curve. In our research, we have opted to use a natural cubic spline instead. This section reports the results of a comparison of the two methods.

For the i^{th} bond, the quoted market price is denoted P_i , and π_i is the predicted price. Degrees of freedom are determined by the number of interior knots used in the spline procedure. A cubic B-spline basis has degrees of freedom equal to the number of interior knots plus four; a natural cubic spline basis has degrees of freedom equal to the number of interior knots plus two². The error degrees of freedom is calculated as the difference between the number of data points (N) and the degrees of freedom for the spline fit. The mean squared error (MSE) is calculated by dividing the error sum of squares (SSE) by the error degrees of freedom. We conclude that there is no significant effect of using a natural cubic spline basis rather than a cubic B-spline basis. For the

²If no intercept term is used, then the extra degrees of freedom added to the number of interior knots is one less than stated here (three and one, respectively). We use an intercept term in our research.

Chapter 2. Modeling the Term Structure

Table 2.1: Spline Comparison (14 Jun 2005)

	Natural Spline	B-spline
N (# of bonds used in spline)	160	160
Mean $(P_i - \pi_i)$	0.00245	0.00245
StdDev $(P_i - \pi_i)$	0.08317	0.08296
Degrees of Freedom (fit)	53	55
Degrees of Freedom (error)	107	105
SSE = $\sum(P_i - \pi_i)^2$	1.10071	1.09532
MSE = $\frac{1}{df} \sum(P_i - \pi_i)^2$	0.01029	0.01043

Table 2.2: Spline Comparison (28 Feb 2006)

	Natural Spline	B-spline
N (# of bonds used in spline)	169	169
Mean $(P_i - \pi_i)$	0.00023	0.00026
StdDev $(P_i - \pi_i)$	0.20973	0.20524
Degrees of Freedom (fit)	56	58
Degrees of Freedom (error)	113	111
SSE = $\sum(P_i - \pi_i)^2$	7.39011	7.07684
MSE = $\frac{1}{df} \sum(P_i - \pi_i)^2$	0.06540	0.06376

remainder of the results, the spline basis used will be a natural cubic spline basis.

2.3 Effects of Changing the Timescale

We explore the effects of changing the timescale in the spline process for the interest rate term structure. Initially, the timescale of the bond data input into the spline procedure was on a standard linear timescale. In this section, the standard timescale approach is compared to a square root timescale approach.

Before presenting the results of this section, we provide some details of the computations involved in changing the timescale in the spline procedure. In the standard spline procedure proposed by Fisher et al., the spline is created for $l(\tau)$, where $l(\tau) = -\log(\delta(\tau)) = \tau z(\tau)$. (Here, $\delta(\tau)$ is the discount function with a current time of zero and $z(\tau)$ is the zero curve.) We spline $l(\tau)$ and then convert the spline to be a representation of the zero curve.

To illustrate changing the timescale that is input into the spline procedure, we will briefly discuss the process for the square root timescale. First, the maturities and times to payoff dates for the bond data are replaced by their respective square root values. Next, we modify the calculation of the penalty term in the spline procedure to use a natural spline basis calculated on a fine grid over maturities of 0 to $\sqrt{30}$ rather than 0 to 30 as in the standard timescale. In the standard timescale, $l(t) = \phi\beta$ and the zero curve is a plot of $z(t) = \phi\beta/t$ against \sqrt{t} . In the square root timescale, $l(t) = \phi\beta$ (calculated with the square root maturities and times to payoffs) and the zero curve is a plot of $z(t) = \phi\beta/t$ against \sqrt{t} . As shown in the remainder of this section, the square root approach allows for greater flexibility in the spline fit for short term maturities and a more rigid fit for long term maturities.

We first present the results of the interest rate spline for the standard and square

Chapter 2. Modeling the Term Structure

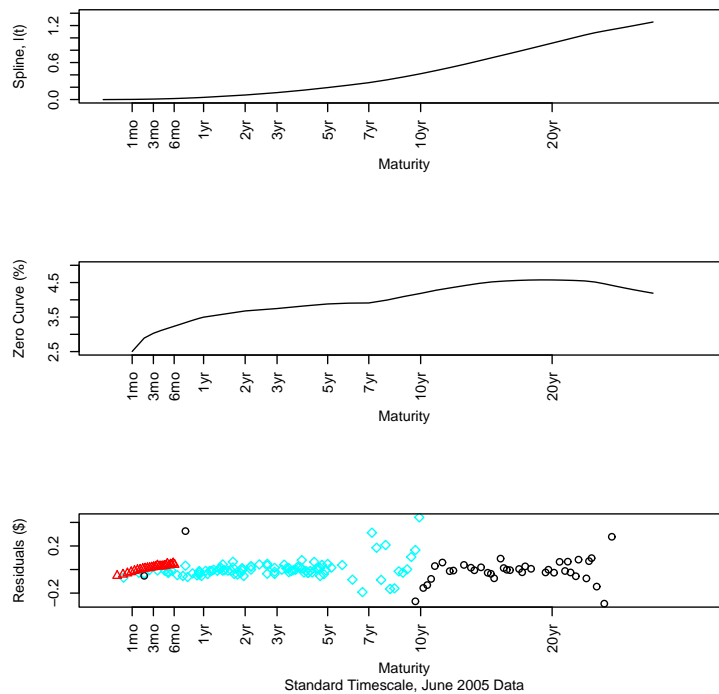


Figure 2.1: Standard Timescale (14 Jun 2005)

root timescales for the June 2005 bond data set, as well as a comparison of the residuals obtained using each timescale. Analogous results are also presented for the February 2006 data set.

June 2005 Data Set Results

In Figure 2.1, the top panel shows the spline for $l(\tau)$ using the bond data with a standard linear timescale. The second panel shows the zero curve $z(\tau)$, which is derived from $l(\tau)$ using the relationship $z(t) = l(t)/t$. The third panel displays the residuals, where the i^{th} residual is the difference between the market price of bond i and the

Chapter 2. Modeling the Term Structure

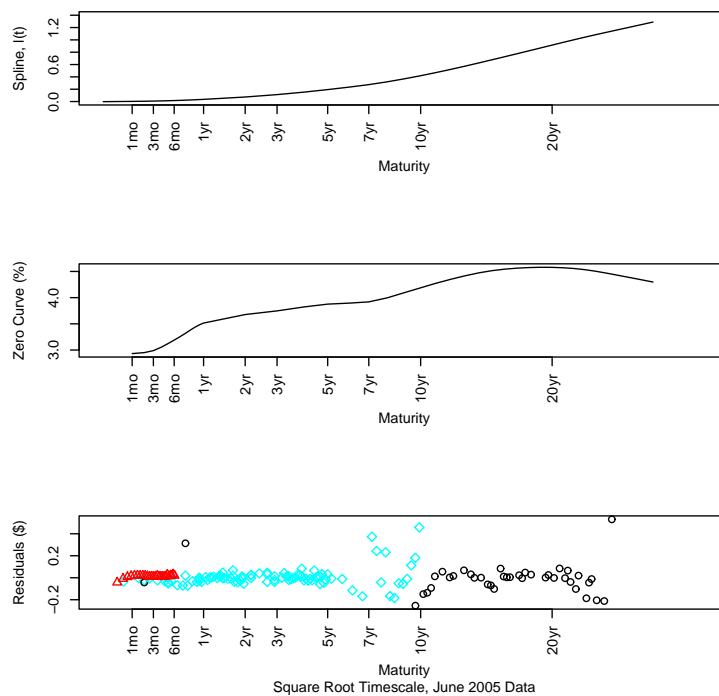


Figure 2.2: Square Root Timescale (14 Jun 2005)

predicted price of bond i based on the spline procedure. The colors of the residuals in the third panel of the plots in this chapter refer to the type of Treasury security the residual refers to: red refers to Treasury bills (maturity of one year or less), blue refers to Treasury notes (maturity of two to ten years), and black refers to Treasury bonds (maturity of ten years or longer). All three panels are plotted with a horizontal axis of time on a square root scale.

The results for the square root timescale are shown in Figure 2.2. The three panels correspond to the three panels from the figure for the standard timescale approach.

We now turn to comparing the standard and square root timescales in detail. First,

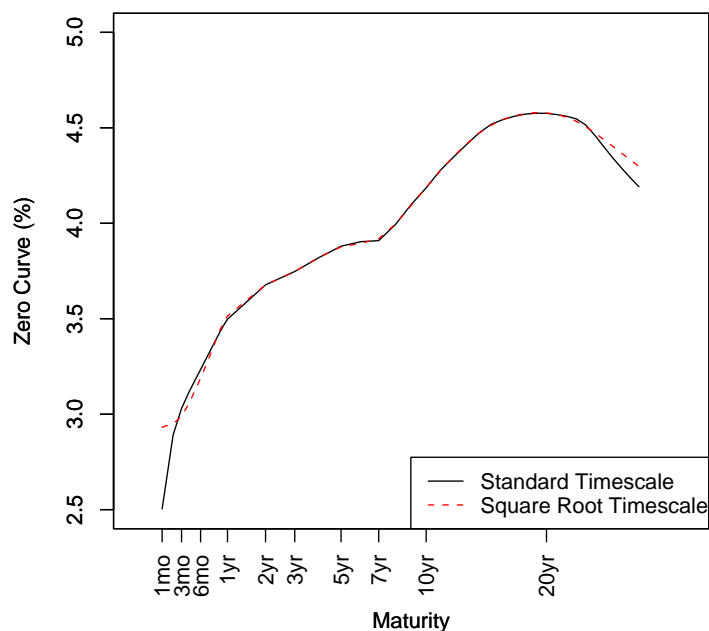


Figure 2.3: Zero Curve Comparison (14 Jun 2005)

the zero curves produced by each approach are shown in Figure 2.3. The solid line is the zero curve obtained from using a standard timescale, and the dashed line is the zero curve obtained from using a square root timescale. The square root timescale zero curve exhibits more flexibility at the short term maturities (specifically maturities less than 2 years), but the standard timescale zero curve exhibits more flexibility at the long term maturities (roughly 20 years or more). The curves presented in Figures 2.1 through 2.3 were plotted by calculating the zero curve value monthly from 0 to 12 months and then annually from 1 to 30 years.

To compare the residuals obtained from the two timescale approaches, we compare

Table 2.3: SSE Comparison (14 Jun 2005)

Bin	n	Standard timescale	Square root timescale
Overall	160	1.1007	1.3896
0-1 years	47	0.1620	0.1372
1-4 years	45	0.0371	0.0365
4-9 years	26	0.2978	0.3776
9-16 years	21	0.3792	0.4043
≥ 16 years	21	0.2247	0.4339

sums of squared residual terms both overall (for all maturities) and in five different bins based on maturities. (The sum of squared residuals (SSE) is $\sum(P_i - \pi_i)^2$, where the i^{th} quoted market price is denoted P_i and π_i is the i^{th} predicted price.) These comparisons are shown in Table 2.3. The SSE for the square root timescale is greater when computed for all 160 bonds in the 14 Jun 2005 data set; however, for the short term maturity bins ($[0, 1]$ and $(1, 4]$), the square root timescale approach has smaller sums of squared residuals. This again shows that the square root timescale is fitting the bond data set prices more closely (i.e. with more flexibility in the spline process) for the short maturities. Conversely, the SSE values for the standard timescale are smaller than for the square root timescale at long term maturity values (greater than 4 years). Table 2.4 reports the means and standard deviations of the residuals both overall and for each of the maturity range bins.

Table 2.4: Mean and Standard Deviation Comparison (14 Jun 2005)

Bin	n	Standard timescale	Square root timescale
Overall	160	0.0024 (0.0832)	0.0023 (0.0935)
0-1 years	47	0.0021 (0.0593)	-0.0001 (0.0546)
1-4 years	45	0.0022 (0.0290)	0.0048 (0.0284)
4-9 years	26	0.0001 (0.1091)	-0.0007 (0.1229)
9-16 years	21	0.0086 (0.1374)	0.0075 (0.1420)
≥ 16 years	21	0.0006 (0.1060)	0.0011 (0.1473)

February 2006 Data Set Results

Analogous results are now presented for the February 2006 bond data. Figures 2.4 and 2.5 display the results of the spline process for the standard and square root timescales, respectively. Figure 2.6 shows the comparison plot for the 28 February 2006 data. Once again, the square root timescale curve shows more flexibility in the first year of maturity and less flexibility at longer maturities, especially those over 16 years. Table 2.5 shows the SSE values for the overall curves as well as for the bins of maturities used in the previous tables of this section. Once again, the overall SSE for the square root timescale is greater than the overall SSE for the standard timescale. However, in the first two bins (up to 1 year and 1 to 4 years), the SSE for the square root timescale is less than the SSE for the standard timescale. In Table 2.6, the means and standard deviations of the residuals are reported for the overall residuals and for each of the bins.

Chapter 2. Modeling the Term Structure

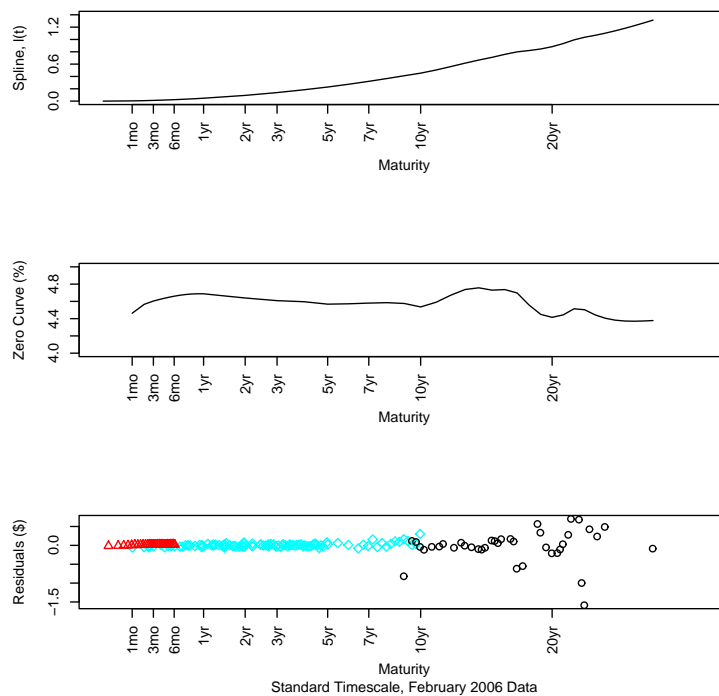


Figure 2.4: Standard Timescale (28 Feb 2006)

Table 2.5: SSE Comparison (28 Feb 2006)

Bin	n	Standard timescale	Square root timescale
Overall	169	7.3901	9.8405
0-1 years	49	0.0298	0.0280
1-4 years	51	0.0412	0.0390
4-9 years	27	0.7812	0.8264
9-16 years	22	0.2450	0.3702
≥ 16 years	20	6.2930	8.5769

Chapter 2. Modeling the Term Structure

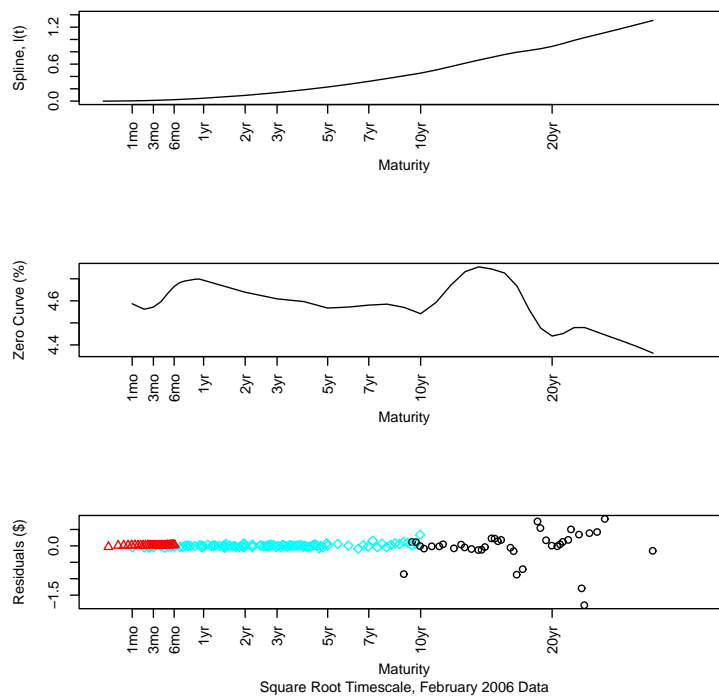


Figure 2.5: Square Root Timescale (28 Feb 2006)

Table 2.6: Mean and Standard Deviation Comparison (28 Feb 2006)

Bin	n	Standard timescale	Square root timescale
Overall	169	0.0002 (0.2097)	-0.0005 (0.2420)
0-1 years	49	-0.0009 (0.0249)	-0.0006 (0.0241)
1-4 years	51	0.0027 (0.0286)	0.0022 (0.0278)
4-9 years	27	-0.0094 (0.1731)	-0.0148 (0.1776)
9-16 years	22	0.0276 (0.1043)	0.0452 (0.1244)
≥ 16 years	20	-0.0204 (0.5751)	-0.0384 (0.6707)

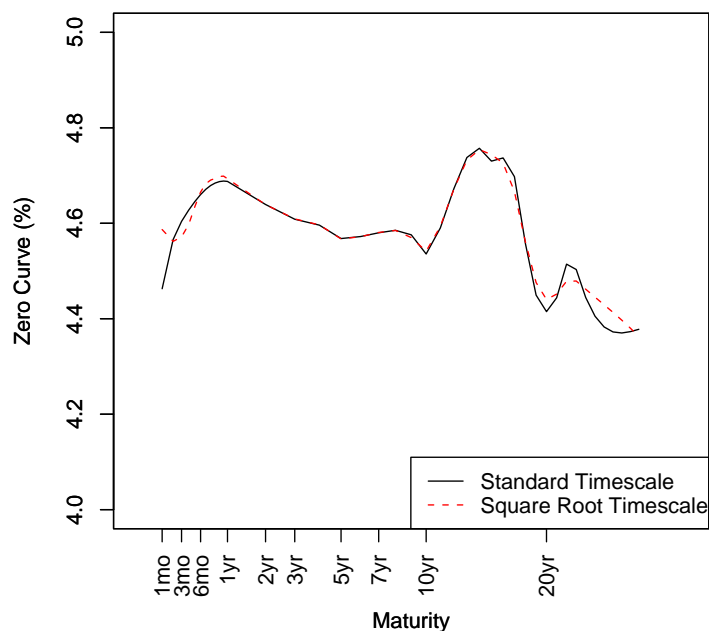


Figure 2.6: Zero Curve Comparison (28 Feb 2006)

2.4 Incorporating Weights into the Spline Fit

This section details the formulation of the criterion for determining the optimal smoothing parameter in the weighted case of the Fisher et al. (1994) (FNZ) zero curve spline analogous to the generalized cross-validation (GCV) criterion. The section will describe the reasoning behind adding weights to the market bond data and our adaptation for weighted data to the Craven and Wahba (1979) approach. The attempt to incorporate some type of weighting into the FNZ zero curve spline analysis was driven by the heteroskedasticity of the residuals in the plots of the residuals (market bond prices P minus

Chapter 2. Modeling the Term Structure

FNZ predicted bond prices π), as seen in the third panel of Figure 2.1. The residuals are small in the short term of maturities, but they increase at longer maturities.

Our solution is to incorporate a weight matrix W into the FNZ algorithm. The W matrix is an $n \times n$ diagonal matrix constructed element by element using a k -th nearest neighbor variance estimation algorithm, where n is the number of observations in the data set and k is the number of nearest neighbors included in the calculation. For example, the first element, corresponding to the first observation, finds the k nearest residuals on the maturity (horizontal) axis and averages the sum of those squared residuals. This is an estimate of the variance of the first observation's residual (s_i^2), defined in Equation 2.3 (page 42). The W matrix is constructed by inverting these variance estimates, such that $w_i = 1/s_i^2$ and $W = \text{diag}(w_1/\bar{w}, w_2/\bar{w}, \dots, w_n/\bar{w})$, where \bar{w} is the mean of the w_i values. Hence, W is a diagonal weight matrix standardized such that the diagonal elements have a mean of one.

Next, we refer back to the original GCV paper by Craven and Wahba (1979). The remainder of this section details the adaptation of their method of GCV calculation to find an analogous version of the GCV for the weighted data case. (The following description refers to Craven and Wahba (1979); we use the notation (CW 1.1) to refer to their Equation 1.1.) The authors start with a mathematical model:

$$y(t) = g(t) + \epsilon(t), \quad t \in [0, 1] \quad (\text{CW 1.1})$$

where $g(t)$ is a “smooth” curve, $\epsilon(t)$ is a white noise process, and $y(t)$ is the observed data. They simplify their notation for $y(t_j)$ to be y_j . They assume that $g \in W_2^{(m)}$,

Chapter 2. Modeling the Term Structure

where:

$$W_2^{(m)} = \{g : g^{(v)} \text{ abs. cont.}, v = 0, 1, \dots, m - 1, g^{(m)} \in \mathcal{L}_2[0, 1]\}.$$

The smoothing problem is then stated as: Find $f \in W_2^{(m)}$ to minimize:

$$\frac{1}{n} \sum_{j=1}^n (f(t_j) - y_u)^2 + \lambda \int_0^1 (f^{(m)}(u))^2 du. \quad (\text{CW 1.2})$$

This allows one parameter λ to control the smoothness of the solution. The authors denote the solution to (CW 1.2) by $g_{n,\lambda}$. The authors then define an $n \times n$ matrix $A(\lambda)$ as satisfying:

$$\begin{pmatrix} g_{n,\lambda}(t_1) \\ \vdots \\ g_{n,\lambda}(t_n) \end{pmatrix} = A(\lambda) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

This leads to a generalized cross validation (GCV) estimate, which is the minimizer of $V(\lambda)$ defined by:

$$V(\lambda) = \frac{1}{n} \|(I - A(\lambda))y\|^2 / \left[\frac{1}{n} \text{Tr}(I - A(\lambda)) \right]^2. \quad (\text{CW 3.3})$$

Fisher et al. (1994) use a similar GCV value derived for the nonlinear regression case they present in their spline methods for various interest rate curves. In order to do so, they derive an $A(\lambda)$ matrix based on a linearization of their original nonlinear least squares problem, detailed in Section 2.8.4; their $A(\lambda)$ also serves as a hat matrix in that $A(\lambda)Y(\beta^*(\lambda))$ provides the vector of fitted Y values. The FNZ $A(\lambda)$ is defined

Chapter 2. Modeling the Term Structure

as:

$$A(\lambda) = X(\beta^*(\lambda)) \left(X(\beta^*(\lambda))^T X(\beta^*(\lambda)) + \lambda H \right)^{-1} X(\beta^*(\lambda)). \quad (2.1)$$

We use Equation 2.1 to develop an analogous weighted criterion which we then use in place of the GCV criterion in the weighted case of our spline methodology. We incorporate the weights w_k described at the beginning of this section (where w_k are the diagonal elements of the weight matrix W) by modifying the numerator of (CW 3.3) to become:

$$\frac{1}{n} \sum_{k=1}^n w_k \left(\sum_{j=1}^n a_{kj} y_j - y_k \right)^2. \quad (2.2)$$

This analogous criterion is a generalization of the numerator of both (CW 3.3) and Equation 2.10 in Section 2.8.4; when the weight matrix W is equal to the identity matrix (implying all weights equal to one), the generalized GCV numerator, Equation 2.2, reduces to the numerator in (CW 3.3), which is also the numerator of the GCV criterion in Fisher et al. (1994).

Hence, the generalized GCV criterion we use is defined as:

$$\gamma_g(\lambda) = \frac{\frac{1}{n} \sum_{k=1}^n w_k \left(\sum_{j=1}^n a_{kj} y_j - y_k \right)^2}{\left(n - 2\text{tr}(A(\lambda)) \right)^2}.$$

In this fashion, the spline procedures are updated to include weights such as the ones described using nearest neighbor variance estimates as described in the beginning of this section. The next subsection details the results of this adjustment to the spline procedure and the choice of an optimal k , the number of nearest neighbors used in the weight calculation.

2.4.1 Optimal k Values

To determine the optimal value of k (the number of nearest neighbors used in the variance estimation technique which leads to the weights in the previous section), we compared a BIC analog measure over k values from 0 to 50. We did this for each of four situations for each data set: standard timescale for $l(\tau)$, square root timescale for $l(\tau)$, standard timescale for $z(\tau)$, and square root timescale for $z(\tau)$. (Recall that $l(\tau)$ and $z(\tau)$ refer to the zero curves produced using a spline for $l(\tau) = \tau z(\tau)$ and using a spline directly for $z(\tau)$, respectively.) Figures 2.7 and 2.8 display a measure (labeled qBIC) analogous to a BIC measure for each of the above four situations for $k = 0, \dots, 50$ for the June 2005 data set and the February 2006 data set, respectively. Our qBIC measure is defined as:

$$qBIC(k) = \sum_{i=1}^n \log(s_i^2) + \sum_{i=1}^n \frac{(P_i - \pi_i)^2}{s_i^2} + d \log(n),$$

where d is the number of effective parameters ($d = n/(k+1)$), P_i and π_i are the market price and predicted price, respectively, and:

$$s_i^2 = \frac{1}{k+1} \sum_{j \in K_i} (P_j - \pi_j)^2, \tag{2.3}$$

where K_i is the set of k nearest neighbors to the i^{th} observation as well as the i^{th} observation itself.

In the June 2005 data set, the minimum qBIC value occurs at $k = 15$ for the square root timescale for $z(\tau)$ spline, as shown in Figure 2.7. In the February 2006 data set, the minimum qBIC value occurs at $k = 23$ for the standard timescale for $l(\tau)$ spline,

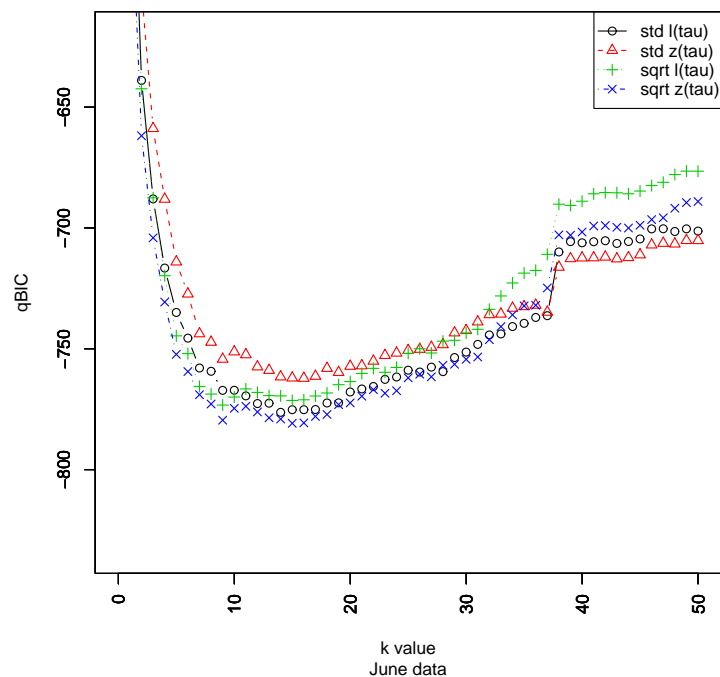


Figure 2.7: k Value Comparison (14 Jun 2005)

as shown in Figure 2.8. Hence, the optimal weighting situation varies depending on the data set. Overall, however, k values between 15 and 25 seem reasonable.

We also compare the λ values that produce the minimum GCV value for each value of k as well as compare these values to the value of λ in the unweighted case. Tables B.1 and B.2 (located in Appendix B) report these comparisons for the June 2005 data set and the February 2006 data set, respectively.

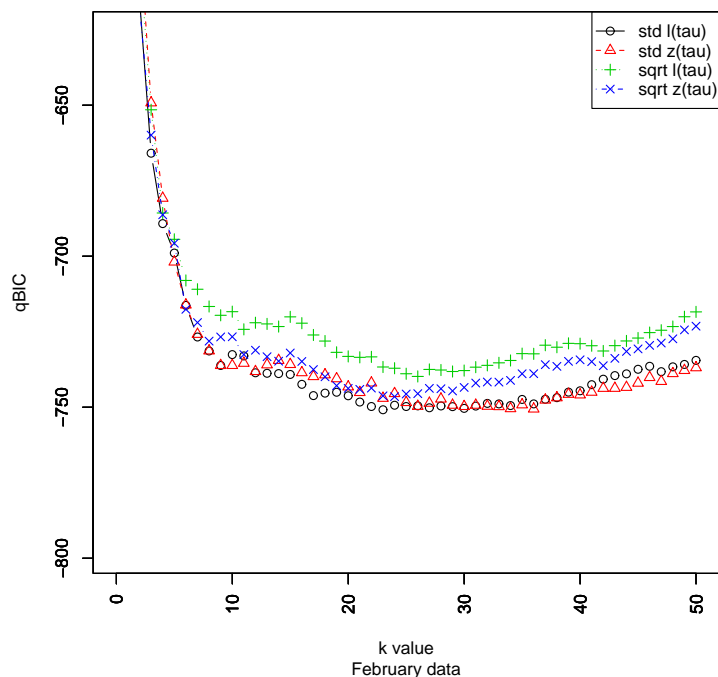


Figure 2.8: k Value Comparison (28 Feb 2006)

2.5 Comparing the Spline Methods

The US Treasury Department’s Office of Debt Management tabulates the daily treasury yield curve rates for maturities of 1, 3, and 6 months and 1, 2, 3, 5, 7, 10, and 20 years³. On a subsequent webpage⁴, the method used to determine this yield curve is described as a “quasi-cubic hermite spline function.” We use the curve values from the Treasury Department as a comparison for our yield curve estimation derived from our various

³<http://www.treas.gov/offices/domestic-finance/debt-management/interest-rate/yield.shtml>

⁴<http://www.treas.gov/offices/domestic-finance/debt-management/interest-rate/yieldmethod.html>

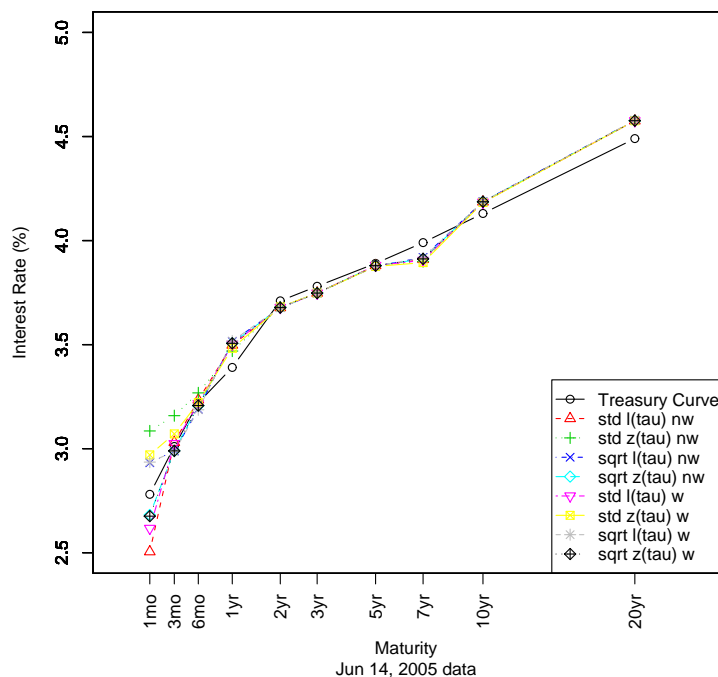


Figure 2.9: Yield Curve Estimation Comparison (14 Jun 2005)

adaptations of the spline methodology of Fisher et al. (1994). For each of the two bond data sets used in our analysis, we present two comparison plots, the first of which plots the zero curves of the eight different estimation methods as well as the treasury zero curve, and the second of which plots the difference between each of the eight methods and the treasury zero curve.

Figure 2.9 shows the comparison for the June 2005 data set, and the match is quite good. Figure 2.11 shows the February 2006 data set, where the match is not as good. An obvious result from comparing these two plots is that the yield curve changed shape dramatically from 14 June 2005 to 28 February 2006. Figures 2.10 and 2.12 show the

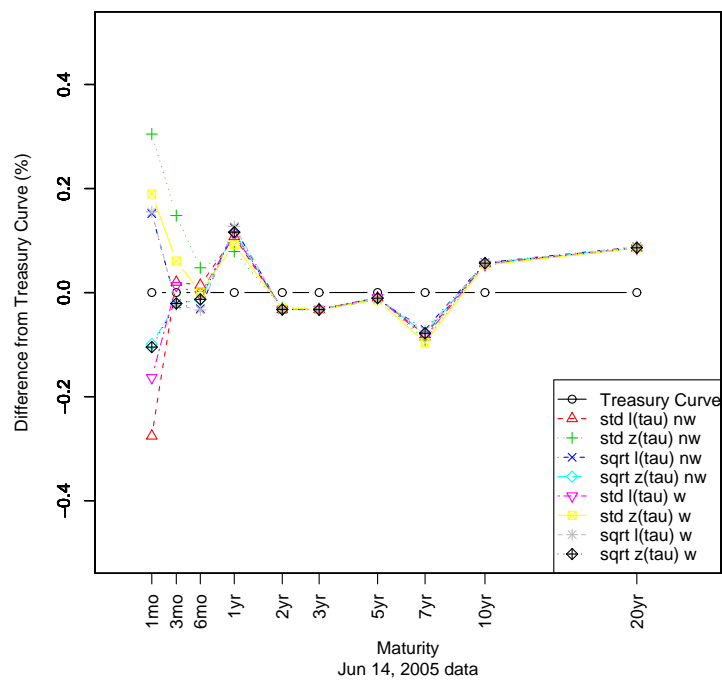


Figure 2.10: Yield Curve Estimation Difference Plot (14 Jun 2005)

difference between each spline curve and the Treasury curve for the June and February data sets, respectively.

For the June 2005 data set, the optimal spline method is the square root timescale for $z(\tau)$ with weights using $k = 15$. For the February 2006 data set, the optimal spline method is the standard timescale for $l(\tau)$ with weights using $k = 23$.

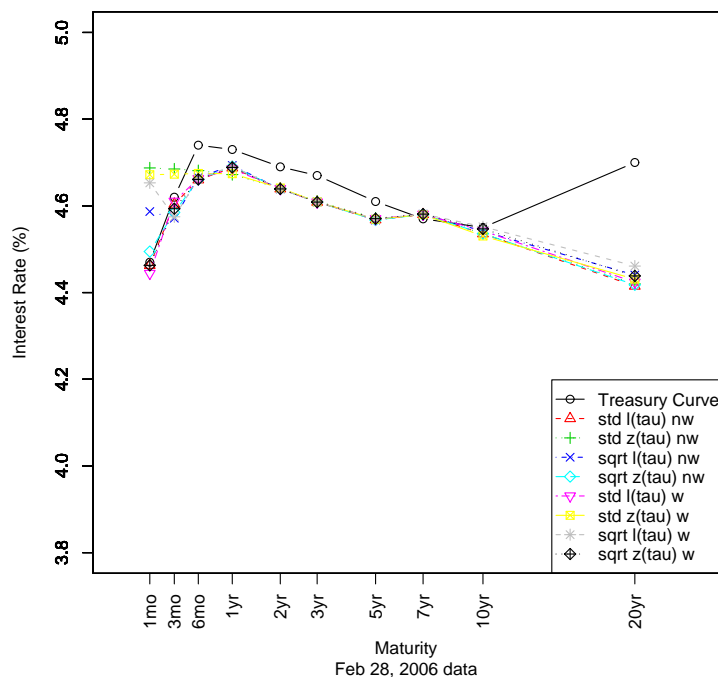


Figure 2.11: Yield Curve Estimation Comparison (28 Feb 2006)

2.6 Discussion of the Treasury Premium

The analysis in this chapter uses Treasury security price data to construct a spline for the riskless interest rate term structure. However, this may not be the most accurate estimate of the riskless rate. In Feldhütter and Lando (2005), the authors break down the swap spread (over the 6 year period from 1997 through the end of 2002) into three components: a factor specific to the swap market, a credit risk element from the LIBOR market, and a convenience yield from holding Treasury securities. In this analysis, we are interested in the convenience yield component since this is the difference between

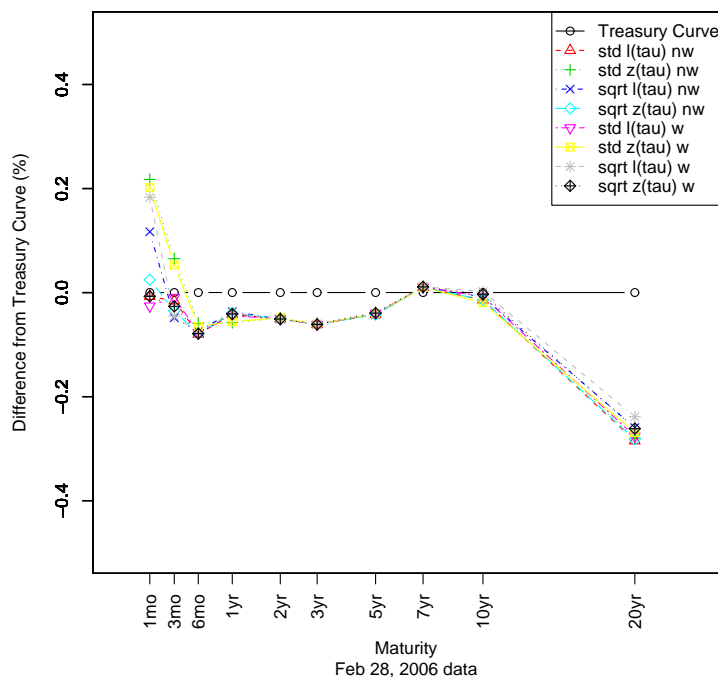


Figure 2.12: Yield Curve Estimation Difference Plot (28 Feb 2006)

the Treasury rate and the riskless rate. They show that the size of this component likely fluctuates over time. In fact, for the time period they investigated, it was negative and tending toward zero.

This result implies that there exists a convenience premium for Treasury instruments due to their liquidity (among other factors). The implication for the current research is that the interest rate being used in the Hull-White trinomial tree to discount future payments might not be as close to the true riskless rate as one would desire. An extension to this work could involve an adjustment to the discount rates in the tree so that they are closer to the riskless rate, although this would likely require

more data than is used in this research.

2.7 Bypassing the Zero Curve Spline

To bypass the process of estimating the zero curve with a spline approach, we spline the function $\alpha(\tau)$ directly (using calculations derived in Section 2.8.3) and use the spline estimation of $\alpha(\tau)$, denoted $\alpha_s(\tau)$, to construct a Hull-White trinomial tree without estimating the zero curve. Given values for a , σ , and $\alpha(t)$ in the Hull-White framework (where a and σ are constants and $\alpha(t)$ is a function of time used in the Hull-White tree construction), a trinomial tree can be constructed. The iterative nature of the construction of the trinomial tree now depends on $\alpha(\tau)$ as an input rather than $z(\tau)$. Instead of using the zero curve values to determine Q and α values at each timestep, α values are used to determine Q and zero curve values at each timestep. In this way, a given α value yields a trinomial tree for the zero rate.

In Section 3.3, we present results of a bootstrap analysis using this method of tree construction to price interest rate caplets.

2.8 Background Calculations

Calculations too lengthy to include earlier in the chapter are included in this section. The following two subsections contain background calculations pertaining to the splines for $l(\tau)$ and $z(\tau)$. This section also includes details of the spline for $\alpha(\tau)$ and the derivation of the FNZ GCV criterion.

2.8.1 Calculations for $l(\tau)$

This section contains the calculations necessary to obtain $\pi_i(\beta)$ and $X(\beta)$ in the spline method for $l(\tau)$, described in Fisher et al. (1994). Recall that $l(\tau) = -\log(\delta(\tau))$, where $\delta(\tau)$ is the discount function of a one unit zero coupon bond maturing at time τ . For each bond i (n bonds are denoted with subscripts $i = 1, \dots, n$), there are m_i coupon payments. The amounts and times until these payments are denoted with the vectors c_i and τ_i , respectively. Both c_i and τ_i are $m_i \times 1$ vectors. Also, each bond has a natural spline basis denoted $\tilde{\phi}(\tau_i)$, which is an $m_i \times \kappa$ matrix, where κ is the number of knots in the natural spline basis as well as the length of the β vector. The natural spline basis is defined as:

$$\tilde{\phi}(\tau_i) = \left(\tilde{\phi}_1(\tau_i), \dots, \tilde{\phi}_\kappa(\tau_i) \right),$$

where $\tilde{\phi}_k(\tau_i) = (\tilde{\phi}_k(\tau_{i,1}), \dots, \tilde{\phi}_k(\tau_{i,m_i}))^T$.

The natural spline basis is provided by the `ns()` function in the statistical package *R*. We will simplify the Fisher et al. notation such that $\tilde{\phi}_k(\tau_{i,l}) \equiv \phi(i, l, k)$; hence, i denotes the bond, and l and k represent the row and column of $\tilde{\phi}(\tau_i)$, respectively.

Next, following Section 2.2 of Fisher et al. (1994), we define $g(h(\cdot), \tau) \equiv \delta(\tau)$, where $h(\tau)$ is an arbitrary function of the term structure, and we parameterize $h(\tau)$ as a natural spline. Then, $h_s(\tau, \beta) = \phi(\tau)\beta$, and $\delta_s(\tau, \beta) = g(h_s(\cdot, \beta), \tau)$. Since $l(\tau) = -\log(\delta(\tau))$ and $l(\tau)$ is our term structure function of interest, we let $h(\tau) = l(\tau)$. Then, $\delta(\tau) = -e^{h(\tau)} \equiv g(h(\cdot), \tau)$. We then define the present value of the i^{th} bond to be:

$$\pi_i(\beta) = c_i^T \delta_s(\tau_i, \beta) = c_i^T \tilde{g}(h_s(\cdot, \beta), \tau_i) = c_i^T \tilde{g}(\phi(\cdot)\beta, \tau_i), \quad (2.4)$$

Chapter 2. Modeling the Term Structure

where $\tilde{g}(\phi(\cdot)\beta, \tau_i) \equiv (g(\phi(\cdot)\beta, \tau_{i,1}), \dots, g(\phi(\cdot)\beta, \tau_{i,m_i}))$, an $m_i \times 1$ vector. From here, we need to find expressions for both $\pi_i(\beta)$ and $X(\beta)$, where $X(\beta)$ is defined as the partial derivative of $\Pi(\beta)$ with respect to β^T and $\Pi(\beta)$ is the vector of bond present values, $\pi(\beta)$.

Using the above relationship between $\delta_s(\tau_i, \beta)$ and $\tilde{g}(h_s(\cdot, \beta), \tau_i)$, it can be shown that $\pi_i(\beta) = c_i^T \exp\{-\phi(\tau_i)\beta\}$.

Next, we turn to the calculation of:

$$X(\beta) = \frac{\partial \Pi(\beta)}{\partial \beta^T}.$$

The function $X(\beta)$ can be thought of as n rows of an overall matrix, with each row corresponding to an individual bond in the bond data set. Each row contains κ columns, with each column corresponding to each element of β . The function $X(\beta)$ is as follows:

$$X(\beta) = \begin{pmatrix} \frac{\partial \pi_1(\beta)}{\partial \beta_1} & \dots & \frac{\partial \pi_1(\beta)}{\partial \beta_\kappa} \\ \vdots & \ddots & \vdots \\ \frac{\partial \pi_n(\beta)}{\partial \beta_1} & \dots & \frac{\partial \pi_n(\beta)}{\partial \beta_\kappa} \end{pmatrix}.$$

Chapter 2. Modeling the Term Structure

We now examine the j^{th} element of $X(\beta)$ for bond i :

$$\begin{aligned}
\frac{\partial \pi_i(\beta)}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} (c_i^T \exp\{-\phi(\tau_i)\beta\}) \\
&= \frac{\partial}{\partial \beta_j} \sum_{l=1}^{m_i} c_{i,l} \exp\{-\phi(i, l, \cdot)\beta\} \\
&= \sum_{l=1}^{m_i} c_{i,l} \frac{\partial}{\partial \beta_j} \exp\{-\phi(i, l, \cdot)\beta\} \\
&= \sum_{l=1}^{m_i} c_{i,l} \frac{\partial}{\partial \beta_j} \exp\left\{-\sum_{k=1}^{\kappa} \phi(i, l, k)\beta_k\right\} \\
&= -\sum_{l=1}^{m_i} c_{i,l} \exp\left\{-\sum_{k=1}^{\kappa} \phi(i, l, k)\beta_k\right\} \phi(i, l, j) \\
&= -\sum_{l=1}^{m_i} c_{i,l} \exp\{-\phi(\tau_i)\beta\} \phi(i, l, j) \\
\frac{\partial \pi_i(\beta)}{\partial \beta_j} &= -c_i^T (\exp\{-\phi(\tau_i)\beta\} \bullet \phi(i, \cdot, j)),
\end{aligned}$$

where \bullet represents the Hadamard product⁵. Using this formulation, all the elements of $X(\beta)$ can be computed to form the entire matrix. This matrix is then used in the remainder of the method proposed by Fisher et al. (1994). Note that the formula given by Fisher et al. (1994) for $\frac{\partial \pi_i(\beta)}{\partial \beta^T}$ in the $l(\tau)$ case is incorrect. Table 2.7 shows the updated entry for the $l(\tau)$ case in Fisher et al. (1994). Note that we include here only the j^{th} term of this expression in this table.

It should also be noted that we use a natural spline basis, whereas Fisher et al. use a cubic B-spline basis. This choice is discussed in Section 2.2.

⁵For more information on the Hadamard product, or component-wise matrix multiplication, see Horn and Johnson (1991, Chapter 5). Briefly, the ‘‘Hadamard product of $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{ij}] \in M_{m,n}$ is defined by $A \bullet B = [a_{ij}b_{ij}] \in M_{m,n}$.’’

2.8.2 Calculations for $z(\tau)$

This section contains the calculations necessary to obtain $\pi_i(\beta)$ and $X(\beta)$ in the spline method for $z(\tau)$, which is extended from the approach by Fisher et al. (1994). Instead of creating a spline for $l(\tau)$, we use the methods of Fisher et al. (1994) (described in the previous section for the $l(\tau)$ case) to develop a spline for $z(\tau)$. In this section, we only describe the differences encountered when developing the spline for $z(\tau)$ rather than $l(\tau)$.

In the $z(\tau)$ case, the arbitrary function of the term structure, $h(\tau)$, becomes:

$$z(\tau) = \frac{-\log(\delta(\tau))}{\tau}.$$

Then, $\delta(\tau) = \exp\{-\tau h(\tau)\} \equiv g(h(\cdot), \tau)$. This also implies that $\delta_s(\tau, \beta) = \exp\{-\tau \phi(\tau)\beta\}$.

The expression for $\pi_i(\beta)$ follows:

$$\pi_i(\beta) = c_i^T \delta_s(\tau_i, \beta) = c_i^T \exp\{-\tau_i \bullet \phi(\tau_i)\beta\}.$$

The derivation for the j^{th} element of $X(\beta)$ for bond i is:

$$\begin{aligned} \frac{\partial \pi_i(\beta)}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} (c_i^T \exp\{-\tau_i \bullet \phi(\tau_i)\beta\}) \\ &= \sum_{l=1}^{m_i} c_{i,l} \frac{\partial}{\partial \beta_j} \exp\{-\tau_{i,l} \phi(i, l, \cdot)\beta\} \\ &= -\sum_{l=1}^{m_i} c_{i,l} \exp\{-\tau_{i,l} \phi(\tau_i)\beta\} \tau_{i,l} \phi(i, l, j) \\ \frac{\partial \pi_i(\beta)}{\partial \beta_j} &= -c_i^T (\exp\{-\tau_i \bullet \phi(\tau_i)\beta\} \bullet (\tau_i \bullet \phi(i, \cdot, j))). \end{aligned}$$

Table 2.7: Updated Entries of Fisher et al. (1994) Table

$h(\tau)$	$\pi_i(\beta)$	$\partial\pi_i(\beta)/\partial\beta_j$
$l(\tau)$	$c_i^T \exp\{-\phi(\tau_i)\beta\}$	$-c_i^T (\exp\{-\phi(\tau_i)\beta\} \bullet \phi(i, \cdot, j))$
$z(\tau)$	$c_i^T \exp\{-\tau_i \bullet \phi(\tau_i)\beta\}$	$-c_i^T (\exp\{-\tau_i \bullet \phi(\tau_i)\beta\} \bullet (\tau_i \bullet \phi(i, \cdot, j)))$
$\alpha(\tau)$	$c_i^T \exp\{-\psi(\tau_i)\beta + \int_0^T m(s)ds\}$	$-c_i^T \left(\exp\{-\psi(\tau_i)\beta + \int_0^T m(s)ds\} \bullet \psi(i, \cdot, j) \right)$

The results for $\pi_i(\beta)$ and $\partial\pi_i(\beta)/\partial\beta_j$ in the $z(\tau)$ case are tabulated in Table 2.7.

2.8.3 Calculations for $\alpha(\tau)$

This section contains the calculations necessary to obtain $\pi_i(\beta)$ and $X(\beta)$ in the spline method for $\alpha(\tau)$, which is extended from the approach by Fisher et al. (1994). Instead of creating a spline for $l(\tau)$, we use the methods of Fisher et al. (1994) (described in the Section 2.8.1 for the $l(\tau)$ case) to develop a spline for $\alpha(\tau)$. In this section, we only describe the differences encountered when developing the spline for $\alpha(\tau)$ rather than $l(\tau)$.

Standard Notation

In the various sources for the Hull-White model and trinomial tree framework, there are many different notations. We will adopt the notation used in the fifth edition of Hull (2003). In this section, we briefly present the standard notation to ease confusion in the following section.

First, we present the notation used in the Hull-White model. The Hull-White model

Chapter 2. Modeling the Term Structure

for the instantaneous short rate r is specified as follows:

$$dr = [\theta(t) - ar]dt + \sigma dz, \quad (2.5)$$

where r is the short rate, z is a Wiener process, and a and σ are constants. The function $\theta(t)$ is defined as:

$$\theta(t) = \frac{\partial F(0, t)}{\partial t} + aF(0, t) + \frac{\sigma^2}{2a}(1 - e^{-2at}), \quad (2.6)$$

where $F(0, t)$ is the instantaneous forward rate for a maturity t at time zero.

When switching over to the tree construction process, Hull assumes that the δt rate, R , follows the same process as r , stating that this is reasonable as δt tends to zero. Then, he constructs a tree (symmetric around zero) for a variable R^* that is initially zero and follows the process:

$$dR^* = -aR^*dt + \sigma dz.$$

In the second stage of the tree construction, the R^* tree is converted to a tree for R by utilizing $\alpha(t)$ which is defined as the difference $R(t) - R^*(t)$. Analytically, it follows that:

$$\alpha(t) = F(0, t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2 \quad (2.7)$$

is a solution to the differential equation $d\alpha = [\theta(t) - a\alpha(t)]dt$. However, this representation of $\alpha(t)$ is not used in the iterative tree construction process because the tree produced with Equation 2.7 is not entirely consistent with the initial term structure.

Calculations

We start with Equation 2.7. Even though this is not the representation of $\alpha(\tau)$ used in the iterative tree construction process, we use it here since it is analytically tractable.

In the $\alpha(\tau)$ case, the arbitrary function of the term structure, $h(\tau)$, becomes:

$$\alpha(\tau) = \frac{-d}{d\tau} \log(\delta(\tau)) + \frac{\sigma^2}{2a^2}(1 - e^{-a\tau})^2.$$

Then, $\delta(\tau) = \exp\{\int(-h(\tau) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2)d\tau\} \equiv g(h(\cdot), \tau)$. This also implies that $\delta_s(\tau, \beta) = \exp\{\int(-\phi(\tau)\beta + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2)d\tau\}$. The expression for $\pi_i(\beta)$ follows:

$$\pi_i(\beta) = c_i^T \delta_s(\tau_i, \beta) = c_i^T \exp\left\{\int_0^T (-\phi(s_i)\beta + m(s))ds\right\},$$

where $m(t) = \frac{\sigma^2}{2a^2}(1 - e^{-at})^2$. We then define:

$$\psi(\tau) = \int_0^T \phi(s)ds.$$

The derivation for the j^{th} element of $X(\beta)$ for bond i is:

$$\begin{aligned} \frac{\partial \pi_i(\beta)}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} \left(c_i^T \exp\left\{-\psi(\tau_i)\beta + \int_0^T m(s)ds\right\} \right) \\ &= \sum_{l=1}^{m_i} c_{i,l} \frac{\partial}{\partial \beta_j} \exp\left\{-\psi(i, \cdot, l)\beta + \int_0^T m(s)ds\right\} \\ &= -\sum_{l=1}^{m_i} c_{i,l} \exp\left\{-\psi(i, \cdot, l)\beta + \int_0^T m(s)ds\right\} \psi(i, l, j) \\ \frac{\partial \pi_i(\beta)}{\partial \beta_j} &= -c_i^T \left(\exp\left\{-\psi(\tau_i)\beta + \int_0^T m(s)ds\right\} \bullet \psi(i, \cdot, j) \right). \end{aligned}$$

The results for $\pi_i(\beta)$ and $\partial\pi_i(\beta)/\partial\beta_j$ in the $\alpha(\tau)$ case are tabulated in Table 2.7.

2.8.4 Derivation of FNZ GCV Criterion

This section describes the derivation of the Fisher et al. (1994) (FNZ) generalized cross-validation (GCV) criterion. The setup of the previous subsections applies here as well. Define P as the vector of market prices for the n bonds in the data set, and define $\Pi(\beta)$ as the present values of the n bonds defined by $\pi_i(\beta)$ in Equation 2.4 in Section 2.8.1. The authors state that $h_s(\tau, \beta^*)$ is a regression spline, where β^* solves:

$$\min_{\beta} [(P - \Pi(\beta))^T (P - \Pi(\beta))]. \quad (2.8)$$

They solve Equation 2.8 as a nonlinear least squares problem, linearizing $\Pi(\beta)$ around β^0 , an initial guess for β . The linearization is:

$$\Pi(\beta) \approx \Pi(\beta^0) + (\beta - \beta^0)X(\beta^0),$$

where $X(\beta^0) = \frac{\partial\Pi(\beta)}{\partial\beta^T}$ evaluated at $\beta = \beta^0$. They also define $Y(\beta^0) = P - \Pi(\beta^0) + \beta^0 X(\beta^0)$. Equation 2.8 is then rearranged to be:

$$\min_{\beta} \left[\left(Y(\beta^0) - X(\beta^0)\beta \right)^T \left(Y(\beta^0) - X(\beta^0)\beta \right) \right]. \quad (2.9)$$

An iterative approach is used to solve Equation 2.9. The first iteration β^1 is:

$$\beta^1 = \left(X(\beta^0)^T X(\beta^0) \right)^T X(\beta^0)^T Y(\beta^0).$$

Chapter 2. Modeling the Term Structure

In their spline method, Fisher et al. (1994) use a smoothing penalty defined as:

$$\lambda \int_0^T h''(\tau)^2 d\tau,$$

where $h(\tau)$ is the function being approximated by the spline method. Replacing $h(\tau)$ with the spline representation $h_s(\tau, \beta)$, the smoothing penalty can be written as $\lambda\beta^T H\beta$, where H is determined by the natural spline basis functions and the knot points. Then, the minimization problem, for a general λ , is:

$$\min_{\beta(\lambda)} \left[\left(P - \Pi(\beta(\lambda)) \right)^T \left(P - \Pi(\beta(\lambda)) \right) + \lambda\beta(\lambda)^T H\beta(\lambda) \right].$$

They go on to define $A(\lambda)$, where $A(\lambda)Y(\beta^*(\lambda))$ is the vector of fitted Y values:

$$A(\lambda) = X(\beta^*(\lambda)) \left(X(\beta^*(\lambda))^T X(\beta^*(\lambda)) + \lambda H \right)^{-1} X(\beta^*(\lambda)).$$

Finally, they choose the value of λ that minimizes their GCV criterion:

$$\gamma(\lambda) = \frac{\left((I - A(\lambda))Y(\beta^*(\lambda)) \right)^T \left((I - A(\lambda))Y(\beta^*(\lambda)) \right)}{\left(n - 2\text{tr}(A(\lambda)) \right)^2}. \quad (2.10)$$

Chapter 3

Bootstrap Analysis

This chapter begins with an introduction of the use of the statistical bootstrap in the simple regression setting using an example described by Efron and Gong (1983). Following the introduction, there will be a description of the bootstrap analysis performed (using the bond price data described in Section 2.1) followed by the results of the analysis.

Efron and Gong (1983, Section 2) use bootstrap methodology in a regression setting to estimate the sampling distribution of the correlation coefficient for bivariate data. The data used are from 15 American law schools in 1973; the bivariate data, $x_i = (y_i, z_i)$, consists of the average LSAT score of students entering school i , (y_i) , and the average undergraduate GPA of students entering school i , (z_i) . They observe the correlation coefficient to be 0.776, but to determine a standard error, they elect to use the nonparametric bootstrap. First, they construct the empirical distribution function \hat{F} as an estimate of the bivariate distribution F , from which the data points x_1, x_2, \dots, x_{15} were observed. The construction of \hat{F} is as follows:

$$\hat{F} : \text{mass } \frac{1}{n} \text{ on each observed data point } x_i, \quad i = 1, 2, \dots, n.$$

The second step is to obtain a random sample of 15 observations from the original set of observations, called a *bootstrap sample*; this sampling is done *with replacement*.

Chapter 3. Bootstrap Analysis

The bootstrap sample is then used to calculate a *bootstrap replication* of the statistic of interest; in this case, the statistic of interest is the correlation coefficient. This is repeated many times (the number of times is usually referred to as B), each time recording the estimate of the statistic of interest. Finally, the bootstrap estimate of the standard error of the correlation coefficient is:

$$\hat{\sigma}_B = \left[\left(\sum_{b=1}^B (\hat{\rho}^{*b} - \hat{\rho}^*)^2 \right) / (B - 1) \right]^{1/2}, \quad \hat{\rho}^* = \frac{\sum \hat{\rho}^{*b}}{B},$$

where $\hat{\rho}^{*b}$ is the b^{th} bootstrap replication of the correlation coefficient. In this example, Efron and Gong use $B = 1000$ bootstrap replications to estimate the standard error of the correlation coefficient, obtaining $\hat{\sigma}_B = 0.127$.

The choice of B is discussed further in Efron and Tibshirani (1993, Section 6.4), where two rules of thumb are presented. The first is that for estimating standard errors, as few as $B = 25$ replications can be informative and $B = 50$ is often enough. The second rule of thumb is that $B > 200$ replications are rarely needed to estimate a standard error, but that larger values of B are needed for estimating confidence intervals. The analysis presented in the remainder of this chapter uses $B = 1000$ replications.

3.1 Description of the Method

The bootstrap analysis was performed using both of the *Wall Street Journal* (2005, 2006) bond price data sets described in Section 2.1. Following the results of Section 2.5, the spline fit was for the zero curve directly using the weighted spline method described

Chapter 3. Bootstrap Analysis

in Section 2.4. The knots for the spline were determined using quantiles to obtain roughly one third as many knots as there were prices being used for the spline. (This number of knots is suggested by Fisher et al. (1994).) Each of the 1000 bootstrap samples was the same size as the original data set, using sampling with replacement. For each iteration, the bootstrapped bond price data were used to determine a spline fit to the discount curve. This curve was then used as the input to the Hull-White trinomial tree algorithm described in Section 1.3.1. The output from the trinomial tree was used to determine prices for caplets at five different strike levels. The results of this analysis are presented in Section 3.2.2.

Two versions of each analysis were performed. First, the knots for the spline were determined outside of the bootstrap loop (referred to as the “outside” knots version of the analysis). Second, the knots for the spline were determined inside of the bootstrap loop, making the knot calculation part of the bootstrap process itself (referred to as “inside” knots). One reason to calculate the knots outside the bootstrap loop is to avoid non-unique knot points being used in the calculation of the natural spline. This approach is philosophically different from including the calculation of knots inside the bootstrap loop. Calculating the knots inside the loop would bootstrap the entire process of fitting a spline to a subsampled portion of the data set. To bypass the duplicate knots issue, knots were calculated using only the unique values of maturities in each bootstrap sample of bonds. Results of the two approaches are compared in Section 3.2.2.

Finally, Section 3.3 presents results of a bootstrap analysis for the $\alpha(\tau)$ spline. Because of computational issues, this analysis was only performed for the case of outside knots.

3.2 Bootstrap Results

Two bootstrap analyses were performed. The first involved only the discount (or zero) curve, and the second involved the entire process of splining the discount curve and pricing a caplet. The following sections describe the results of these two analyses.

3.2.1 Discount Curve Spline Bootstrap Results

The discount curve was bootstrapped first using knots calculated outside the bootstrap loop and then using knots calculated inside the bootstrap loop. The 1000 bootstrap curves were calculated at one month intervals for the first year and annually from one to thirty years. Figure 3.1 shows both discount curves calculated for the June 2005 data set. (The solid black line represents the curve produced using the outside knots and the dashed red line represents the curve produced using inside knots.)

The splined discount curves are very similar, and the standard deviations fluctuate over maturity in similar patterns. The standard deviations increase near the start and end of the thirty year time period as well as in the six to twelve month area.

Next, we turn to the February 2006 data set. Figure 3.2 (analogous to Figure 3.1) shows a very flat bootstrapped zero curve at about 4.5% for maturities out to 10 years. The behavior after 10 years is less flat, but both the outside (solid line) and inside (dashed line) knot bootstrap curves move together until they diverge after 25 years. At this same point in maturities, the data become much more sparse and the standard deviations of the estimates of the zero curve increase dramatically. The standard deviation plot does not show a similar pattern around six to twelve months as the June 2005 data set analysis does; however, the behavior at the very short and very long

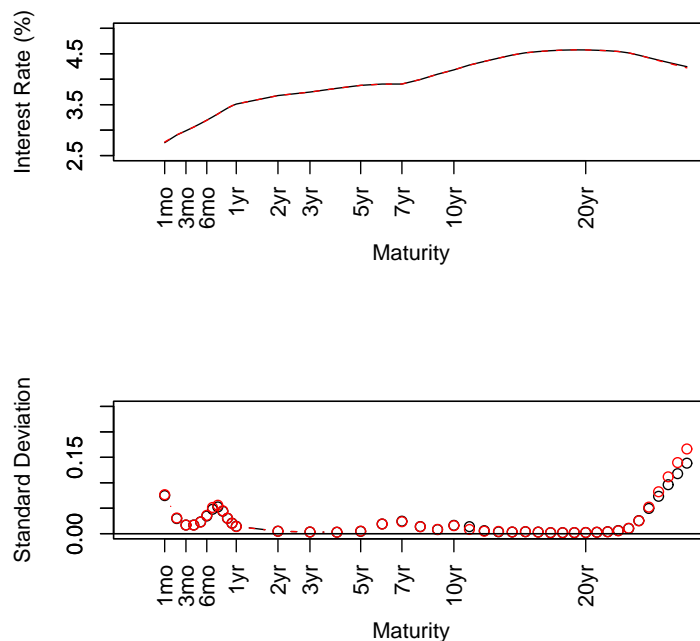


Figure 3.1: Discount Curve Bootstrap Results (14 Jun 2005)

maturities is similar.

3.2.2 Caplet Pricing Bootstrap Results

June 2005 Data Set Results

For the bootstrap analysis of caplet pricing, we again performed two different versions of the analysis. An interest rate cap is a derivative that pays a notional amount multiplied by the excess of a specified interest rate over the agreed-upon strike rate (also called the *cap rate*). The payment is structured as a stream of periodic payments,

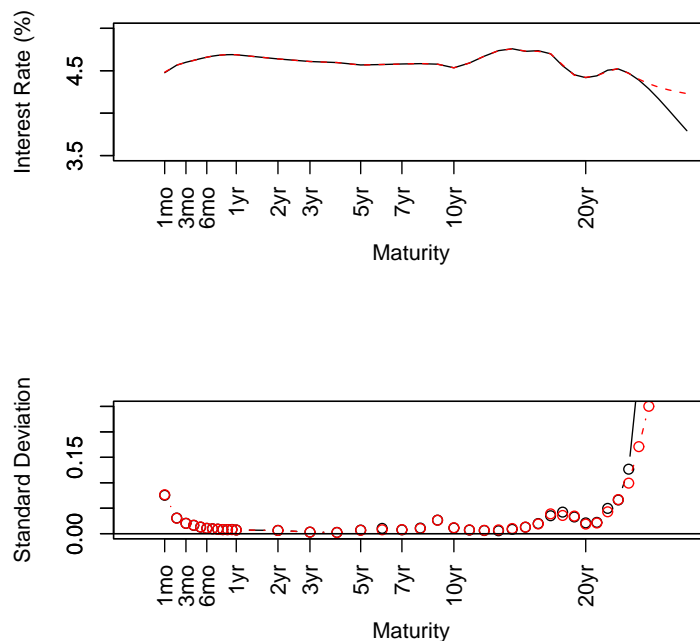


Figure 3.2: Discount Curve Bootstrap Results (28 Feb 2006)

each of which is called a caplet. The length of time between payments is called the *tenor* of the agreement; generally, this length is three months (or a quarter of a year). The analysis that follows prices caplets that mature 5 years from the present with a notional amount of \$400 and a tenor of three months.

In Figure 3.3, the five lines refer to the five strike levels of caplets used in the bootstrap analysis: 0%, 2.5%, 3.85%, 5%, 10% (numbered 1 through 5, respectively). Each of the five lines is the difference (in \$ amounts) between the bootstrap mean price and the price obtained using the full data set. The horizontal axis refers to the number of timesteps per year ($\frac{1}{\delta t}$, where δt is the length of one timestep). The first four strike

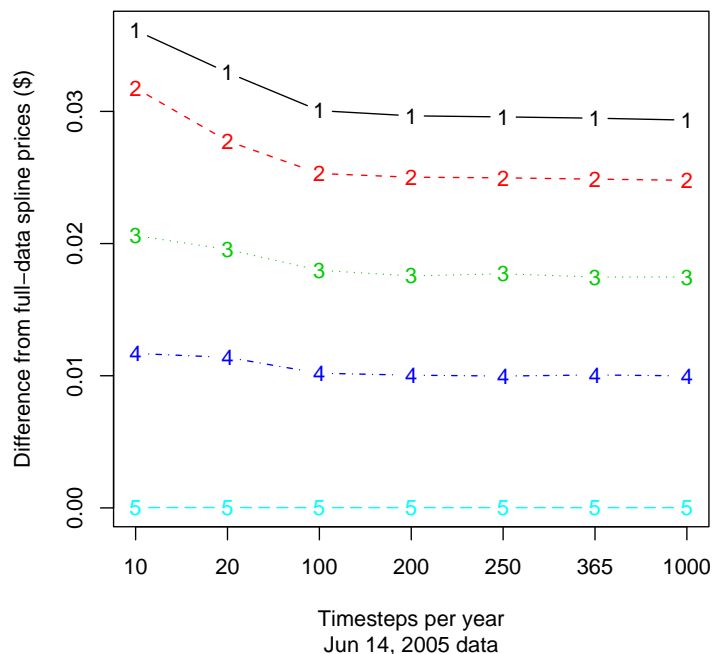


Figure 3.3: Caplet Pricing Bootstrap Means (outside knots)

levels all clearly exhibit a similar pattern over the various numbers of timesteps per year. However, while the largest strike level (10%) also exhibits this pattern, it is on a much smaller scale and is not perceptible in Figure 3.3. Also note that all of these prices exhibit a positive bias from the prices generated using the full data set of bonds.

Figure 3.4 shows the coefficients of variation of the bootstrap prices. Once again, the first four strike levels show a similar pattern, and the fifth (highest) strike level exhibits a somewhat different pattern. From this plot, it appears that further increase of the number of timesteps above 100 per year does not reduce the coefficient of variation any further. This suggests that at this point, the variability resulting from the bootstrap

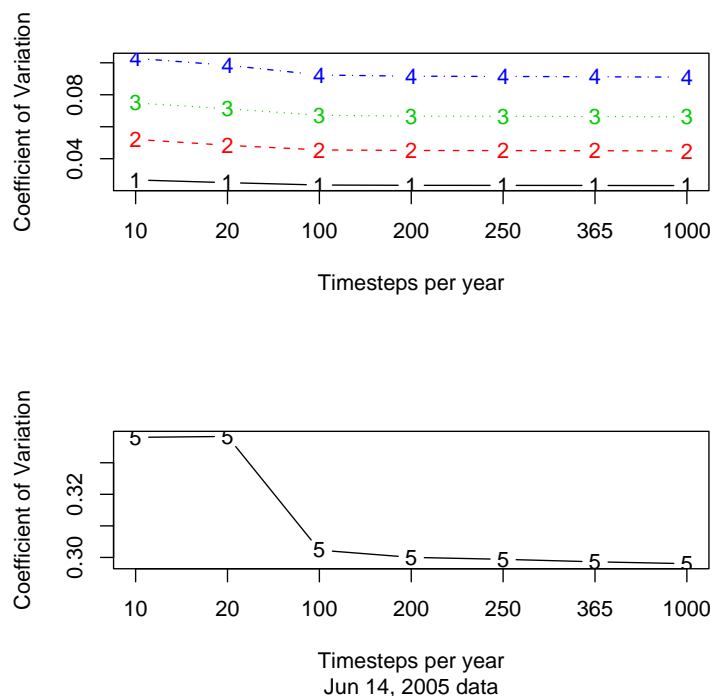


Figure 3.4: Caplet Pricing Bootstrap Coefficients of Variation (outside knots)

procedure is greater than the variability resulting from the tree procedure.

Hull (2003) does not specifically discuss issues arising from δt getting too small (i.e. the number of timesteps per year getting too large). However, an example in Hull (2003) shows the convergence of the tree procedure to analytic methods for pricing bonds where Hull only increases the number of timesteps to 500 over 3 years ($\delta t = \frac{3}{500} = .0167$).

Figure 3.5 shows the difference between the bootstrap mean prices and the full data set prices when the inside knots method was used. In this version of the bootstrap, the knots for the natural spline were obtained inside each bootstrap loop. Therefore,

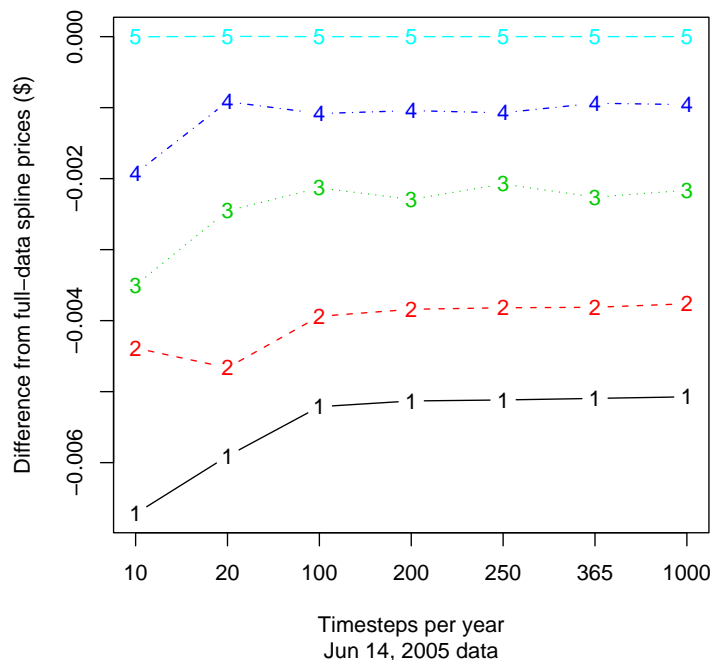


Figure 3.5: Caplet Pricing Bootstrap Means (inside knots)

the entire process of splining the interest rate zero curve and pricing caplets using the trinomial tree discretization of the Hull-White model occurred inside of each bootstrap loop. Note that when the knots are calculated inside the bootstrap loop, a negative bias arises between the bootstrap prices and the prices generated using the full data set of bonds. In both Figures 3.3 and 3.5, the absolute bias decreases as the strike level increases.

Figure 3.6 is the analogous plot to Figure 3.4 for knots calculated inside the bootstrap loop. The pattern exhibited by the coefficients of variation is somewhat different. There is no longer a noticeable decreasing trend in the variation as the number of

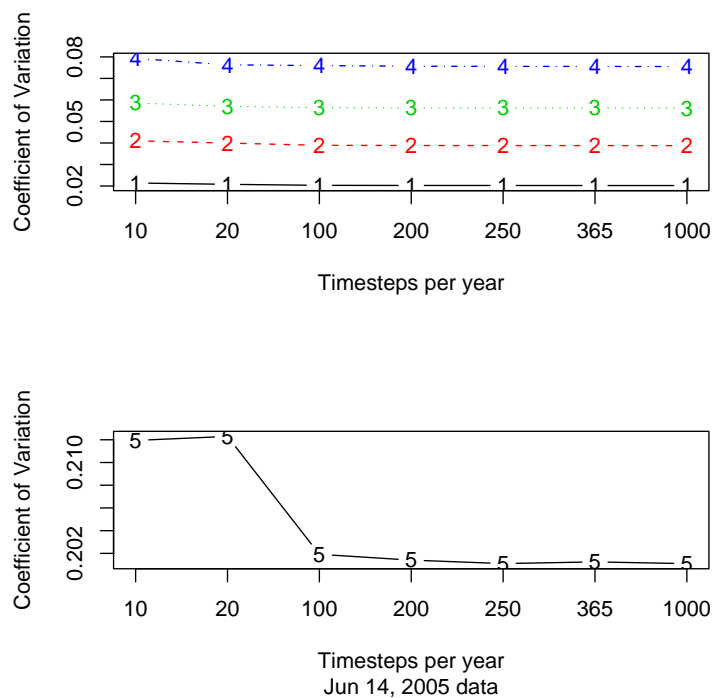


Figure 3.6: Caplet Pricing Bootstrap Coefficients of Variation (inside knots)

timesteps per year increases. However, the magnitude of the coefficients of variation is less in Figure 3.6. Considerably less evidence exists here (than in the outside knots case) to suggest that the coefficient of variation does not decrease once the number of timesteps exceeds 100 per year.

February 2006 Data Set Results

This section reports results analogous to the previous section for bond data from the 1 March 2006 issue of the *Wall Street Journal* (2006). (The data is 28 February 2006 market data, published on 1 March 2006.) This analysis proceeds in the same manner

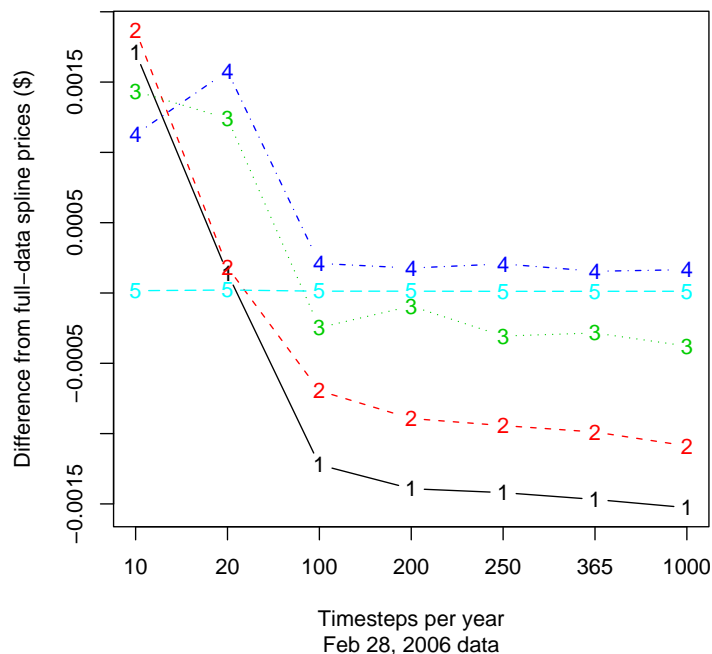


Figure 3.7: Caplet Pricing Bootstrap Means (outside knots)

as the previous analysis on the June 2005 data set.

The bootstrap analysis for caplet pricing compares the bootstrap prices of the caplets at the five strike levels from the previous section to the prices obtained using a zero curve obtained using a spline of the full data set. The next four figures are analogous to Figure 3.3 through Figure 3.6. Figure 3.7 is the only figure that shows a sign change in the bias from the full-spline prices. Furthermore, this bias does not decrease in absolute value for all the strike levels. Figure 3.8 shows a similar pattern for the coefficient of variation as in the June 2005 data set analysis.

We now turn to the inside knots analysis for the February 2006 bond data. In

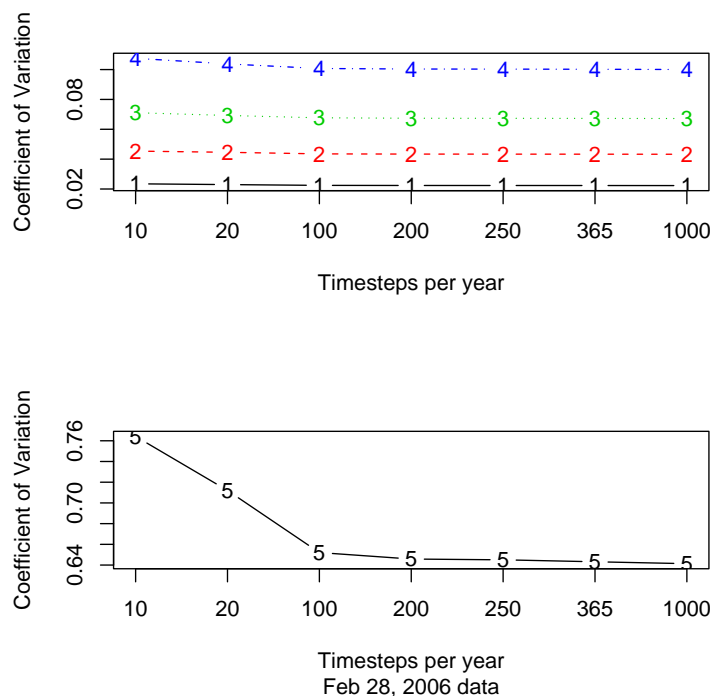


Figure 3.8: Caplet Pricing Bootstrap Coefficients of Variation (outside knots)

Figure 3.9, we see a small change from the June 2005 data set analysis. Instead of the differences getting smaller in absolute value as the number of timesteps per year increases, the differences get larger.

The coefficients of variation shown in Figure 3.10 are similar in pattern and magnitude to those shown in Figure 3.6. However, only the highest strike level shows a substantial decrease as the number of timesteps per year increases. Also, there is little to no further decrease after increasing the number of timesteps per year to 100.

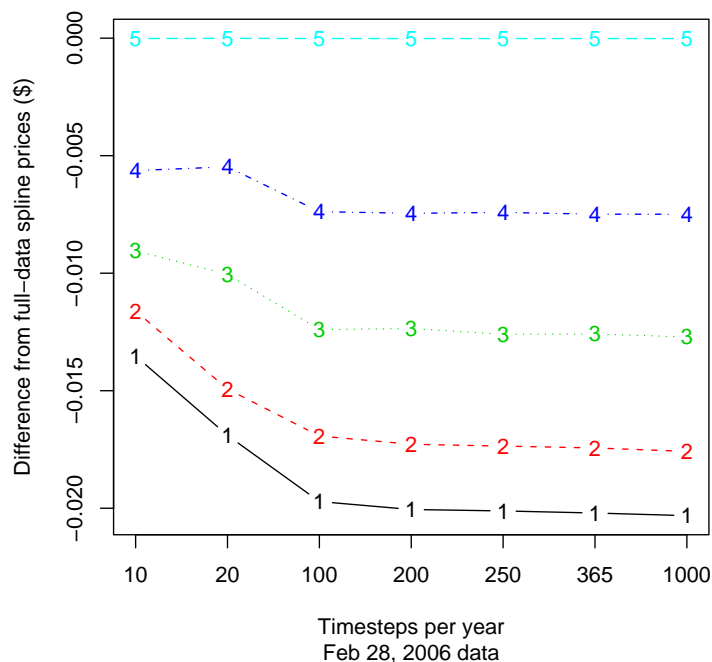


Figure 3.9: Caplet Pricing Bootstrap Means (inside knots)

3.3 Bootstrap Analysis of the $\alpha(\tau)$ Spline

We now present the results of a bootstrap analysis of the $\alpha(\tau)$ spline described in Section 2.7. This analysis consists of only outside knots since calculation of inside knots would be too computationally intensive. We present results for both the June and February data sets.

Figure 3.11 shows the differences between the full-spline caplet prices and the mean bootstrap prices of caplets at the same five strike levels as in the previous sections of this chapter for the June 2005 data set. The differences are negative, and they do not

Chapter 3. Bootstrap Analysis

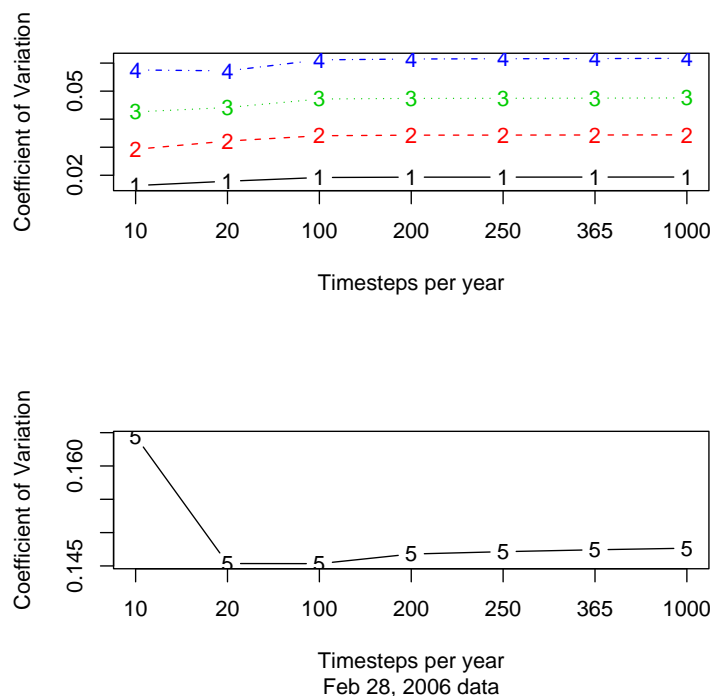


Figure 3.10: Caplet Pricing Bootstrap Coefficients of Variation (inside knots)

decrease in absolute value as the number of timesteps per year increases.

Figure 3.12 shows the coefficients of variation for the June 2005 data set bootstrap analysis. This plot is similar to the coefficient of variation plots for the bootstrap analysis using the zero curve as the input to the Hull-White trinomial tree. Only the highest strike level shows evidence of the variation no longer decreasing after the number of timesteps per year increases to at least 100. The other four strike levels show little to no decrease in variation as the number of timesteps per year increases.

Figure 3.13 is the analogous plot to Figure 3.11 for the February 2006 data set. It shows a similar pattern to the June 2005 data set analysis as well. The differences

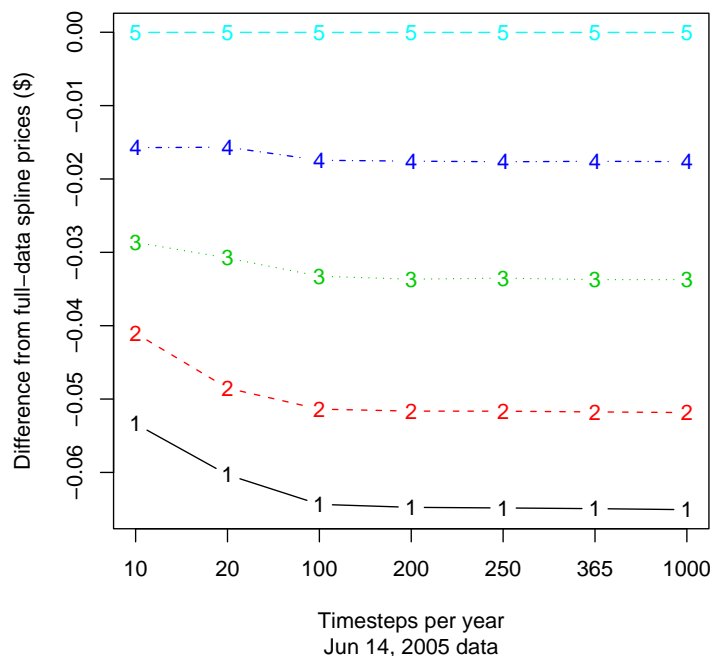


Figure 3.11: Caplet Pricing Bootstrap Means (outside knots)

between the bootstrap mean prices and the full-spline prices do not decrease in absolute value.

Figure 3.14 is the analogous plot to Figure 3.12 for the February 2006 data set. It differs from the June 2005 data set analysis in that it shows a decrease in variation as timesteps per year increases to 100, and there is little to no very decrease in variation after 100 timesteps per year.

Overall, there is no evidence to suggest any improvement in the variability of caplet pricing using the $\alpha(\tau)$ spline versus the $z(\tau)$ spline.

Chapter 3. Bootstrap Analysis

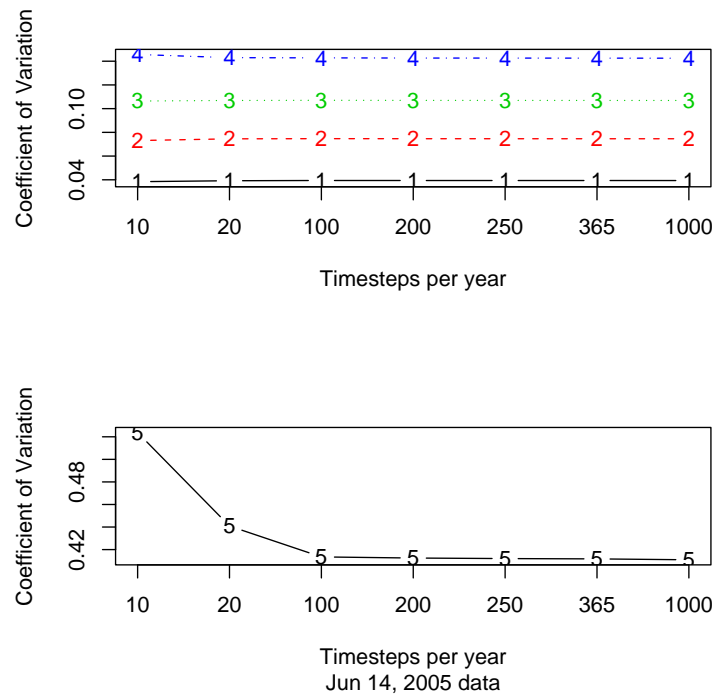


Figure 3.12: Caplet Pricing Bootstrap Coefficients of Variation (outside knots)

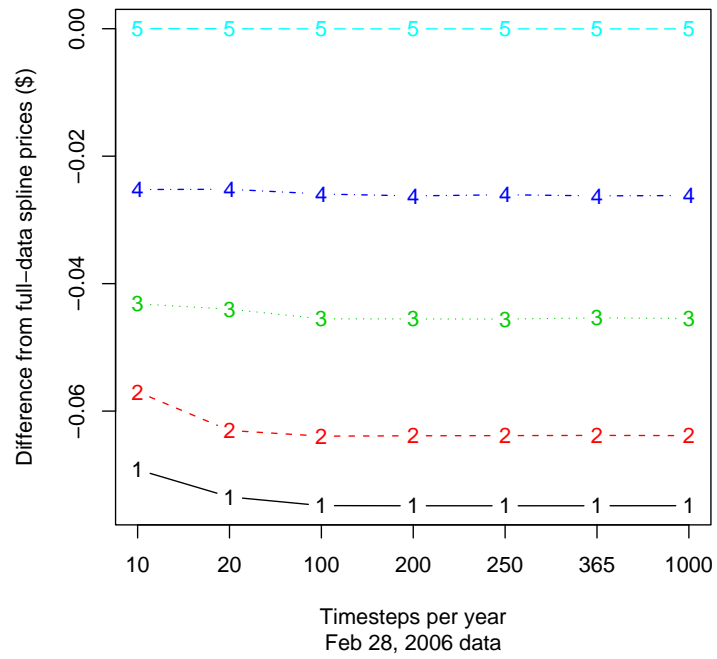


Figure 3.13: Caplet Pricing Bootstrap Means (outside knots)

Chapter 3. Bootstrap Analysis

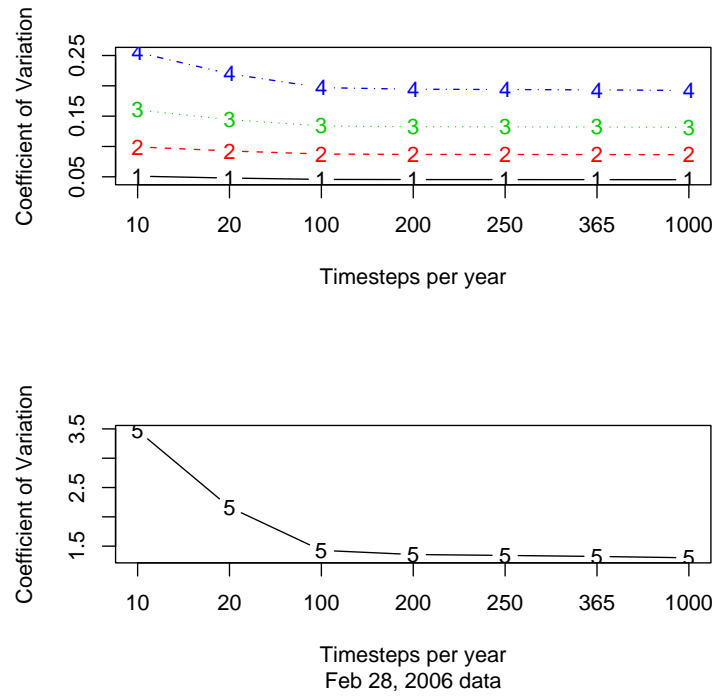


Figure 3.14: Caplet Pricing Bootstrap Coefficients of Variation (outside knots)

Bibliography

- Treasury Bonds, Notes and Bills. *Wall Street Journal*, June 15, 2005.
- Treasury Bonds, Notes and Bills. *Wall Street Journal*, March 1, 2006.
- J. Ahlberg, E. Nilson, and J. Walsh. *The Theory of Splines and Their Applications*. Academic Press, New York, 1967.
- P. Bickel and D. Freedman. Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics*, 9(6):1196–1217, November 1981.
- P. Bickel and D. Freedman. *A Festschrift for Erich L. Lehman*, chapter Bootstrapping Regression Models with Many Parameters, pages 28–48. Wadsworth, Inc., 1983.
- F. Black. Interest Rates as Options. *The Journal of Finance*, 50(5):1371–1376, December 1995.
- F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3):637–654, May–June 1973.
- A. Brace, D. Gątarek, and M. Musiela. The Market Model of Interest Rate Dynamics. *Mathematical Finance*, 7(2):127–147, April 1997.
- P. Craven and G. Wahba. Smoothing Noisy Data with Spline Functions. *Numerische Mathematik*, 31:377–403, 1979.
- D. Duffie and R. Kan. Multi-Factor Term Structure Models. *Philosophical Transactions: Physical Sciences and Engineering*, 347(1684):577–586, June 1994.
- B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, January 1979.
- B. Efron and G. Gong. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, 37(1):36–48, February 1983.
- B. Efron and R. Tibshirani. *Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.
- F.J. Fabozzi, editor. *The Handbook of Fixed Income Securities*. McGraw-Hill, New York, 5th edition, 1997.
- P. Feldhütter and D. Lando. Decomposing Swap Spreads. September 2005.

BIBLIOGRAPHY

- M. Fisher, D. Nychka, and D. Zervos. Fitting the term structure of interest rates with smoothing splines. *Federal Reserve Working Paper*, 1994.
- D. Freedman. Bootstrapping Regression Models. *The Annals of Statistics*, 9(6):1218–1228, November 1981.
- D. Freedman and S. Peters. Bootstrapping a Regression Equation: Some Empirical Results. *Journal of the American Statistical Association*, 79(385):97–106, March 1984.
- R. Goldstein. The Term Structure of Interest Rates as a Random Field. *The Review of Financial Studies*, 13(2):365–384, Summer 2000.
- D. Heath, R. Jarrow, and A. Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60(1):77–105, January 1992.
- T. Ho and S. Lee. Term Structure Movements and Pricing Interest Rate Contingent Claims. *The Journal of Finance*, 41(5):1011–1029, December 1986.
- R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- J. Hull. *Options, Futures, & Other Derivatives*. Prentice Hall, New Jersey, 5th edition, 2003.
- J. Hull and A. White. Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies*, 3(4):573–592, 1990.
- J. Hull and A. White. Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models. *Journal of Derivatives*, 2(2):37–48, Winter 1994.
- P. Hunt and J. Kennedy. *Financial Derivatives in Theory and Practice*. Wiley, 2004.
- L. Izzi and B. Racheva. The term structure of interest rates in the economic and monetary union. *Mathematical Methods of Operations Research*, 55:187–224, 2002.
- F. Jamshidian. LIBOR and swap market models and measures. *Finance and Stochastics*, 1:293–330, 1997.
- K. Miltersen, K. Sandmann, and D. Sondermann. Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates. *The Journal of Finance*, 52(1):409–430, March 1997.

BIBLIOGRAPHY

- S. Neftci. *An Introduction to the Mathematics of Financial Derivatives*. Academic Press, Amsterdam, 2nd edition, 2000.
- C. Reinsch. Smoothing by Spline Functions. *Numerische Mathematik*, 10:177–183, 1967.
- C. Reinsch. Smoothing by Spline Functions II. *Numerische Mathematik*, 16:451–454, 1971.
- R. Rendleman and B. Bartter. The Pricing of Options on Debt Securities. *Journal of Financial and Quantitative Analysis*, 15:11–24, March 1980.
- I. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. *Quarterly of Applied Mathematics*, 4:45–99, 112–141, 1946.
- I. Schoenberg. Spline Functions and the Problem of Graduation. *Proceedings of the National Academy of Sciences*, 52:947–950, 1964.
- J. Shao. On Resampling Methods for Variance and Bias Estimation in Linear Models. *The Annals of Statistics*, 16(3):986–1008, September 1988.
- D. Sornette. “String” formulation of the dynamics of the forward interest rate curve. *The European Physical Journal B*, 3:125–137, 1998.
- USGAO. Federal Debt: Debt Management in a Period of Budget Surplus. Technical report, United States General Accounting Office, September 1999.
- G. Wahba. Smoothing Noisy Data with Spline Functions. *Numerische Mathematik*, 24:383–393, 1975.
- G. Wahba. *Spline Models for Observational Data*. Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1990.
- N. Weber. On Resampling Techniques for Regression Models. *Pacific Statistical Congress; proceedings of the Pacific Statistical Congress – 1985, Auckland, New Zealand, 20–24 May, 1985*, pages 51–55, 1985.
- E. Whittaker and G. Robinson. *The Calculus of Observations*. Blackie & Son Limited, London, 2nd edition, 1926.
- H. Yan. Dynamic Models of the Term Structure. *Financial Analysts Journal*, 57(4): 60–74, July/August 2001.

Appendix

Appendix A

Trinomial Tree Construction

The trinomial tree used in this research is constructed in two stages; the construction process is presented in Hull (2003, Chapter 23). In the first stage, the initial node is set at an interest rate of zero. Each node on the tree is denoted by its horizontal and vertical positions (i, j) , where i is the number of the timestep ($i = 0, 1, 2, \dots$) and j is the number of levels above or below the vertical position of the starting node ($j = \dots, -2, -1, 0, 1, 2, \dots$). We define δt as the length of one timestep along the tree and δR as the size of one vertical increment on the tree. The branching continues through the final timestep.

There are three types of branching, illustrated in Figure A.1; these correspond to the standard branching (a) and the alternative branching used at the top (b) and bottom (c) of the tree. The standard branching has a middle branch that is “flat” and upper and lower branches that go up and down one “level”, respectively. By “flat”, we mean that the vertical position of the node does not change, and by “level”, we mean the amount of vertical change in one timestep. This amount of vertical change is predetermined by a formula given later in this section. As suggested by Hull (2003), we use $\delta R = \sigma\sqrt{3\delta t}$. (Hull states that this proves to be a good choice of δR for error minimization.) There is also a maximum and minimum vertical position on the tree, determined by j_{max} and j_{min} . Once the branching reaches the node denoted j_{max} , the

Appendix A. Trinomial Tree Construction

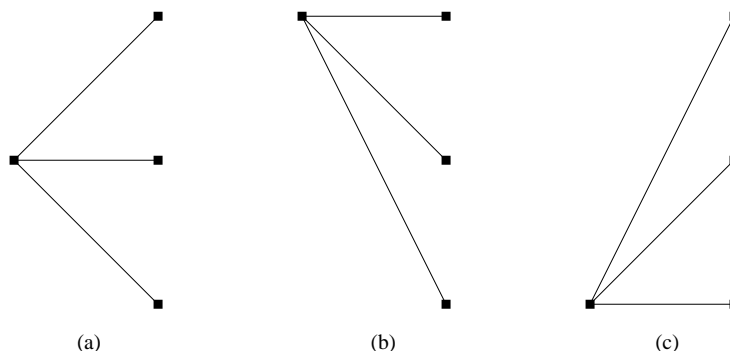


Figure A.1: Trinomial Tree Branching

branching changes to “flat”, “one level down”, “two levels down” (see Figure A.1(b)). Likewise, at j_{min} , the branching changes to “two levels up”, “one level up”, “flat” (see Figure A.1(c)). Each vertical position on the tree has three probabilities associated with it (denoted p_u , p_m , p_d for the probability of going forward on each of the three branches at a given node). These probabilities change at the nodes located at j_{max} and j_{min} on the tree. The calculation of j_{max} and j_{min} is determined by the need to keep these probabilities non-negative and summing to one. (For a tree that is symmetric around $j = 0$, $j_{min} = -j_{max}$.)

We now define R^* as a variable that is initially zero and follows the process:

$$dR^* = -aR^*dt + \sigma dz. \quad (\text{A.1})$$

Then, since $R^* = 0$ at time zero, the value of R^* at node (i, j) is $j\delta R$. The first stage of the tree building procedure concerns the above differential equation where R^* is a variable that initially starts at zero and gets adjusted in the second stage to reflect the initial term structure. The mean and variance of the change in R^* over a short time

Appendix A. Trinomial Tree Construction

(one timestep, δt) are $-aR^*\delta t$ and $\sigma^2\delta t$, respectively.

The determination of the three probabilities uses a system of three equations in the three probabilities. (A separate set of equations arises for each of the three branching configurations.) To illustrate this, we go through the process of determining the probabilities for the standard (middle of the tree) branching situation. The system of equations arises in a “method of moments” fashion by equating the mean and variance of the drift at one timestep of the dR^* process to the first and second sample moments. The third equation results from requiring the three probabilities to sum to one. Hence, the following system of equations relates to the standard branching which occurs in the middle of the tree:

$$(1)\delta R p_u + (0)\delta R p_m + (-1)\delta R p_d = -aR^*\delta t \quad (\text{A.2})$$

$$(1\delta R)^2 p_u + (0\delta R)^2 p_m + (-1\delta R)^2 p_d = \sigma^2\delta t + (-aR^*\delta t)^2 \quad (\text{A.3})$$

$$p_u + p_m + p_d = 1. \quad (\text{A.4})$$

This system results in the following probabilities for moving along each branch of the tree in standard branching:

$$p_u = \frac{(-\delta R^2 a j + \sigma^2 + (a j \delta R)^2 \delta t) \delta t}{2\delta R^2} \quad (\text{A.5})$$

$$p_m = \frac{-(-\delta R^2 + \sigma^2 \delta t + (a j \delta R)^2 \delta t^2)}{\delta R^2} \quad (\text{A.6})$$

$$p_d = \frac{(\delta R^2 a j + \sigma^2 + (a j \delta R)^2 \delta t) \delta t}{2\delta R^2}. \quad (\text{A.7})$$

Similar sets of probabilities exist for the alternative branching situations, which occur at the top and bottom of the tree. The quantity j_{max} must be calculated to

Appendix A. Trinomial Tree Construction

determine the top and bottom of the tree; the range of possible values of j_{max} is derived from the necessity of keeping all the probabilities non-negative. (Since the probabilities are all derived using systems of equations that include the three of them summing to one, valid probabilities arise simply from ensuring that they are each non-negative.) In standard branching, $p_u > 0$ and $p_d > 0$ for all values of j , and $p_m > 0$ when $j \in \left(\frac{-\sqrt{2/3}}{a\delta t}, \frac{\sqrt{2/3}}{a\delta t} \right)$. Therefore, branching must change from the standard branching when j is outside this range; thus, the largest value that j_{max} can take is the largest integer less than or equal to $\frac{\sqrt{2/3}}{a\delta t} \approx \frac{0.816}{a\delta t}$. (The quantity j_{max} must be integer-valued, as it is the node where the branching changes.) In the alternative branching at the top of the tree, $p_u > 0$ and $p_d > 0$ for all values of j , and $p_m > 0$ when $j \in \left(\frac{1-\sqrt{2/3}}{a\delta t}, \frac{1+\sqrt{2/3}}{a\delta t} \right)$. Hence, top branching cannot occur if $j < \frac{1-\sqrt{2/3}}{a\delta t} \approx \frac{0.184}{a\delta t}$; then, the smallest value that j_{max} can take is the smallest integer greater than or equal to $\frac{1-\sqrt{2/3}}{a\delta t} \approx \frac{0.184}{a\delta t}$. The alternative branching at the bottom of the tree has symmetric properties. (Because this is a symmetric tree around $j = 0$, the value of j_{min} is taken as $-j_{max}$.) Then, the range for j_{max} is $\left(\frac{0.184}{a\delta t}, \frac{0.816}{a\delta t} \right)$ and the range for j_{min} is $\left(\frac{-0.816}{a\delta t}, \frac{-0.184}{a\delta t} \right)$. Hull (2003) states that changing at the first possible node (i.e. the smallest integer greater than $\frac{0.184}{a\delta t}$) is most efficient computationally.

This concludes the first stage of constructing the trinomial tree. At this stage, it still emanates from zero at time zero and is symmetric around zero. Furthermore, the probabilities at each node only depend on their vertical position on the tree, j . The second stage of the tree construction involves the use of the initial term structure so that the tree matches that initial term structure exactly.

The second stage of constructing the trinomial tree involves converting the tree for the R^* process to a tree for the R process, where R is the δt rate and follows the same

Appendix A. Trinomial Tree Construction

process as the instantaneous short rate r . First, we define α_i as the value of R at time $i\delta t$ on the R -tree minus the corresponding value of R^* at the same time on the R^* -tree. Hence, $\alpha(t) = R(t) - R^*(t)$.

We also define $Q_{i,j}$ as the present value of a security that pays off \$1 if node (i, j) is reached and nothing otherwise. The values for α_i and $Q_{i,j}$ are calculated so that the initial term structure is matched exactly. Hence, it is necessary to obtain values at each timestep of the tree for the zero rate (or zero-coupon interest rate), which is the interest rate that would be earned on a bond that provides no coupon payments. Hull (2003) goes through a detailed illustration of constructing a tree for R using zero rates at semiannual maturities over three years. Here, we include the given formulae for the α_i and $Q_{i,j}$ values. It is a forward induction process starting at node $(0,0)$ where α_0 is set equal to the initial δt -period interest rate and $Q_{0,0}$ is 1.0. Assuming that the $Q_{i,j}$'s have been determined for $i \leq m$ ($m \geq 0$), we need to determine α_m such that the tree correctly prices a zero-coupon bond maturing at the next timestep. The interest rate at node (m, j) is $\alpha_m + j\delta R$, so the price of the zero-coupon bond maturing at the next timestep is:

$$P_{m+1} = \sum_{j=-n_m}^{n_m} Q_{m,j} e^{-(\alpha_m + j\delta R)\delta t},$$

where n_m is the number of nodes on either side of $j = 0$ at time $m\delta t$. The solution to the above equation is:

$$\alpha_m = \frac{\log \left(\sum_{j=-n_m}^{n_m} Q_{m,j} e^{-j\delta R\delta t} \right) - \log (P_{m+1})}{\delta t}.$$

Using the α_m derived from the above equation, we can calculate the $Q_{i,j}$'s for the

Appendix A. Trinomial Tree Construction

next timestep using:

$$Q_{m+1,j} = \sum_k Q_{m,k} q(k,j) e^{-(\alpha_m + k\delta R)\delta t},$$

where $q(k, j)$ is the probability of moving from node (m, k) to node $(m + 1, j)$ and the summation is taken over all values of k for which this is nonzero (Hull, 2003).

Appendix B

λ Comparison Tables

Table B.1: λ Values in Weighted Case (14 Jun 2005)

	Standard Timescale		Square root timescale	
k	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
(unweighted)	2.95	4.00	1.34	1.76
0	2.94	3.95	1.32	1.75
1	2.89	3.90	1.34	1.73
2	2.78	3.43	1.30	1.75
3	2.81	3.43	1.27	1.77
4	2.76	3.41	1.28	1.73
5	2.78	3.36	1.24	1.77
6	2.79	3.38	1.21	1.75
7	2.75	3.33	1.16	1.73
8	2.75	3.35	1.16	1.73
9	2.74	3.34	1.19	1.70
10	2.69	3.33	1.27	1.80
11	2.69	3.35	1.31	1.79
12	2.70	3.35	1.32	1.78
13	2.69	3.34	1.34	1.78
14	2.69	3.33	1.33	1.78
15	2.66	3.33	1.32	1.78

Continued on next page

Appendix B. λ Comparison Tables

Table B.1 – continued from previous page

k	Standard Timescale		Square root timescale	
	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
16	2.67	3.33	1.31	1.78
17	2.69	3.34	1.30	1.79
18	2.69	3.37	1.30	1.80
19	2.71	3.39	1.30	1.81
20	2.73	3.41	1.30	1.79
21	2.73	3.41	1.30	1.79
22	2.74	3.41	1.28	1.76
23	2.74	3.43	1.28	1.78
24	2.75	3.43	1.26	1.77
25	2.77	3.43	1.29	1.79
26	2.77	3.43	1.29	1.78
27	2.79	3.43	1.28	1.79
28	2.73	3.42	1.35	1.83
29	2.73	3.41	1.33	1.82
30	2.73	3.39	1.31	1.79
31	2.73	3.40	1.31	1.78
32	2.74	3.42	1.32	1.77
33	2.70	3.38	1.31	1.76
34	2.72	3.41	1.31	1.76
35	2.73	3.41	1.31	1.77
36	2.74	3.39	1.34	1.78
37	2.75	3.32	1.40	1.85
38	2.97	3.49	1.58	2.03
39	2.97	3.50	1.55	2.00

Continued on next page

Appendix B. λ Comparison Tables

Table B.1 – continued from previous page

k	Standard Timescale		Square root timescale	
	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
40	2.97	3.50	1.56	2.00
41	2.97	3.51	1.56	2.00
42	2.97	3.52	1.56	2.00
43	2.96	3.52	1.56	1.99
44	2.97	3.53	1.57	2.00
45	2.97	3.55	1.58	2.01
46	2.98	3.59	1.58	2.02
47	2.98	3.60	1.59	2.02
48	2.97	3.59	1.59	2.03
49	2.97	3.59	1.60	2.05
50	2.97	3.59	1.60	2.05

Table B.2: λ Values in Weighted Case (28 Feb 2006)

k	Standard Timescale		Square root timescale	
	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
(unweighted)	2.76	4.78	0.72	1.38
0	2.77	4.73	0.71	1.39
1	2.92	4.67	0.71	1.43
2	2.61	4.48	0.80	1.46
3	2.59	4.41	0.83	1.62
4	2.77	4.42	0.83	1.71

Continued on next page

Appendix B. λ Comparison Tables

Table B.2 – continued from previous page

k	Standard Timescale		Square root timescale	
	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
5	2.91	4.31	0.96	1.73
6	2.85	4.33	1.00	1.65
7	2.92	4.33	1.06	1.69
8	2.89	4.28	1.09	1.76
9	2.97	4.28	1.11	1.89
10	3.10	4.28	1.16	1.93
11	3.08	4.29	1.12	1.89
12	3.10	4.35	1.14	1.93
13	3.07	4.38	1.17	1.91
14	3.08	4.44	1.18	1.90
15	3.09	4.39	1.18	1.90
16	3.08	4.37	1.18	1.90
17	3.08	4.40	1.18	1.91
18	3.08	4.40	1.19	1.91
19	3.10	4.41	1.18	1.93
20	3.12	4.38	1.19	1.95
21	3.11	4.38	1.18	1.98
22	3.14	4.53	1.17	1.98
23	3.15	4.44	1.16	1.99
24	3.19	4.50	1.14	1.97
25	3.19	4.44	1.16	2.03
26	3.22	4.40	1.17	2.06
27	3.20	4.42	1.20	2.06
28	3.22	4.44	1.19	2.04

Continued on next page

Appendix B. λ Comparison Tables

Table B.2 – continued from previous page

k	Standard Timescale		Square root timescale	
	$l(\tau)$	$z(\tau)$	$l(\tau)$	$z(\tau)$
29	3.21	4.39	1.18	2.02
30	3.20	4.38	1.18	2.03
31	3.20	4.38	1.19	2.06
32	3.19	4.36	1.19	2.05
33	3.20	4.35	1.19	2.03
34	3.18	4.32	1.19	2.03
35	3.18	4.31	1.20	2.04
36	3.17	4.31	1.18	2.02
37	3.20	4.37	1.18	2.02
38	3.19	4.38	1.18	2.01
39	3.19	4.38	1.18	2.00
40	3.20	4.38	1.18	2.02
41	3.20	4.36	1.17	2.01
42	3.20	4.34	1.17	2.01
43	3.21	4.34	1.18	2.04
44	3.19	4.32	1.18	2.06
45	3.20	4.33	1.18	2.05
46	3.20	4.36	1.19	2.04
47	3.18	4.33	1.19	2.04
48	3.18	4.35	1.19	2.04
49	3.18	4.36	1.20	2.05
50	3.18	4.36	1.20	2.05