Short Rate Models with Nonlinear Drift and Jumps

by

Amir Memartoluie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Many financial contracts can be regarded as derivative securities where the underlying state variable is one or more rates of interest. A partial list of such contracts would include zero-coupon bonds, coupon paying bonds, callable bonds, convertible bonds, retractable/extendable bonds, etc., along with a number of popular interest rate derivatives such as swaps, swaptions, caps, and floors. A commonly used strategy for valuing these contracts is to base a continuous time model for the stochastic behaviour of the short term rate of interest. Three key features of most of the models currently in use are (i) the drift, or expected change over a short time period in the level of the short term interest rate, is a linear function; (ii) the conditional variance of changes in short term interest rates is not strongly related to the level of interest rates; and (iii) the short term interest rate is assumed to follow a diffusion process, which effectively means that it cannot change too rapidly over short periods of time. Each of these assumptions appears to be made primarily for modelling convenience, as they make it possible in some cases to derive analytical expressions for the values of bonds and European-style bond options. If such solutions are not available, then numerical techniques such as Monte Carlo simulation or the numerical solution of partial differential equations are needed.

However, available econometric evidence indicates that all of the assumptions noted above are questionable: changes in short term interest rates may be characterized by drift which is nonlinear and by conditional variance that depends more heavily on the level of interest rates than is assumed in models with analytic solutions. Moreover, they may be better approximated by a jump-diffusion process which allows for sudden discontinuous changes. Consequently, it is of interest to develop numerical techniques to value interest rate derivative securities for cases where the short term interest rate follows a jump-diffusion process featuring non-linear drift. This thesis describes and illustrates the use of such techniques.

Acknowledgements

This thesis would not have been possible without the support and guidance of my supervisor, Professor Ken Vetzal. From the formative stages of this thesis, to the final draft, I owe an immense debt of gratitude to him. He was abundantly helpful and offered invaluable assistance in every step of the way.

I would also like to express my gratefulness to Professor Peter Forsyth for providing me unflinching encouragement and support in various ways over the course of last year. Last but not least, I would like to offer my special thanks to Professor Ranjini Jha and Tony Wirjanto for taking out time to read my thesis on such a short notice.

Dedication

I dedicate this thesis to my parents.

Contents

List of Tables **x**

List of Figures **xi**

1 Introduction **1**

2 Review of Relevant Prior Literature **4**

2.1 One-Factor Models of the Short Rate 4

2.1.1 Vasicek Model 5

2.1.2 Cox-Ingersoll-Ross (CIR) Model 6

2.1.3 A Generalized Approach 7

2.2 Empirical Evidence 8

2.2.1 Conditional Heteroskedasticity 9

2.2.2 Linearity of Drift 9

2.2.3 Sample Path Continuity 10

3 Underlying Structure of Model **11**

3.1 The Market Price of Risk 13

3.2 The Term Structure Equation 15

3.3 Bond Price Under Classic Models 16

4 Short Rate Models With Jumps **19**

4.1 The Pricing Equation of The Short Rate Model With Jumps 20

5 Numerical Methods **24**

5.1 Numerical Solution Without Jumps 24

5.1.1 Explicit Scheme, Central Difference 26

5.1.2 Implicit Scheme, Central Difference 26

5.1.3 Explicit Scheme, Forward Difference 27

5.1.4 Implicit Scheme, Forward Difference 28

5.1.5 Explicit Scheme, Backward Difference 28

5.1.6 Implicit Scheme, Backward Difference 29

5.1.7 Upstream Weighting 30

5.1.8 Summary of Implicit and Explicit Schemes 31

5.1.9 M-Matrices 32

5.1.10 Boundary Conditions 32

5.1.11 Stability and Monotonicity Considerations 40

5.2 Numerical Solution in the Presence of Jumps 41

6 Numerical Tests **46**

6.1 Bond Prices Without Jumps 47

6.2 Bond Option Prices Without Jumps 58

6.3 Bond Prices Under Jump-Diffusion 62

6.4 Bond Option Prices Under Jump-Diffusion 69

7 Sensitivity Analysis **71**

7.1 Sensitivity of the Yield Curve Under the QTS-Jump Model 75

7.2 Sensitivity of Bond Option Prices Under the QTS-Jump Model 78

8 Conclusions and Future Research **81**

APPENDICES **83**

A Proof of Risk-Neutral Expectation Formulation for Bond Prices **84**

B Pricing Bond Options Consistent with the Initial Yield Curve **86**

References **89**

List of Tables

2.1 Alternative short rate models with linear drift. 8

6.1 Short rate evolution and input parameters for the various models. 47
6.2 Analytic bond prices for the CIR model. 48
6.3 Monte Carlo simulation of bond prices for the CIR model. 48
6.4 CIR bond prices for PDE methods. 50
6.5 CKLS bond prices for PDE methods. 54
6.6 Monte Carlo simulation of bond prices for the CKLS model 55
6.7 QTS bond prices for PDE methods. 56
6.8 Monte Carlo simulation of bond prices for the QTS model. 56
6.9 Analytic bond option prices for the CIR model. 59
6.10 CIR bond option prices for PDE methods. 60
6.11 CKLS bond option prices for PDE methods. 61
6.12 QTS bond option prices for PDE methods. 61
6.13 Monte Carlo simulation of bond prices for the CIR model with jumps. . . . 63
6.14 Monte Carlo simulation of bond prices for the CKLS model with jumps. . . 63
6.15 Monte Carlo simulation of bond prices for the QTS model with jumps. . . 63
6.16 CIR-jump bond prices for PIDE methods. 64
6.17 CKLS-jump bond prices for PIDE methods. 65
6.18 QTS-jump bond prices for PIDE methods. 66
6.19 CIR-jump bond option prices for PIDE methods. 69

6.20 CKLS-jump bond option prices for PIDE methods. 70
6.21 QTS-jump bond option prices for PIDE methods. 70

7.1 Sensitivity of bond option prices under the QTS-jump model. 80

List of Figures

6.1 Bond price surface for the CIR model. 51
6.2 Bond price surface for the CKLS model. 52
6.3 Bond price surface for the CKLS model, modified volatility. 53
6.4 Bond price surface for the QTS model. 57
6.5 Option price surface for the CIR model. 59
6.6 Bond price surface for the CIR-jump model. 67
6.7 Bond price surface for the CKLS-jump model. 67
6.8 Bond price surface for the QTS-jump model. 68

7.1 Yield curves for the CIR, CKLS, and QTS models. 73
7.2 Yield curves for the CIR-jump, CKLS-jump, and QTS-jump models. 74
7.3 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_{-1}. 75
7.4 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_0. 76
7.5 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_1. 76
7.6 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_2. 76
7.7 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to σ. . 77
7.8 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to λ_J. 77
7.9 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to μ_J. 78
7.10 Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to γ_J. 78

Chapter 1

Introduction

Fixed income markets are some of the largest financial markets in the world. According to the Bank for International Settlements [18], the global amount of international debt securities outstanding at the end of 2008 was about $23.9 trillion (all amounts in USD). At the same time, the global amount of domestic debt securities outstanding was around $60 trillion ($24.6 trillion of which was U.S. domestic debt). As a point of comparison, the world's largest stock exchange group is NYSE Euronext. Trading takes place in almost 5,000 firms from more than 50 different countries, with slightly over 2/3 of the world's 100 largest firms included. As of June 2008, the total market capitalization of the firms listed on the NYSE Euronext group was about $27 trillion [17]. These figures indicate that the overall size of fixed income markets is substantially larger than that of equity markets.

Even this understates the case, however, as it overlooks the size of derivatives markets. According to [18], the notional principal of exchange-traded futures contracts outstanding world wide as of December 2008 was $19.4 trillion, of which $18.7 trillion were interest rate futures contracts. The notional principal of outstanding exchange-traded options contracts was $37.2 trillion, with interest rate contracts accounting for $33 trillion. Moreover, in the "over-the-counter" (OTC) markets, the global notional amounts outstanding of derivatives in December 2008 was almost $600 trillion, of which some $420 trillion were interest rate contracts.

Given the enormous size of these markets, it is important for issuers of interest rate dependent contracts (bonds and their derivatives) to have good models for valuing the contracts and assessing risk exposures. A wide variety of different models have been proposed in the literature. For general reviews, see for example [7] or [22]. In broad terms, following [21], these can be classified into *equilibrium models* and *no-arbitrage* models. Equilibrium

models are based on developing a model for the process for the evolution of the short term rate of interest r and exploring its implications for pricing bonds and derivatives such as bond options. A potential downside to this is that it is not possible for such models to exactly fit the prices of all observed bonds. This can lead to very inaccurate values for bond options. By contrast, no-arbitrage models are designed to exactly fit observed prices of bonds, typically by modelling changes in the entire term structure from its given current state. In general, such models exhibit path-dependency and can be relatively complicated to implement, especially for derivative contracts with early exercise features. As noted by [16], it is possible to make *ad hoc* adjustments to equilibrium models by incorporating time-dependent parameters so as to ensure that they do capture the current term structure. While this is not entirely satisfactory from a modelling standpoint because it can lead to overfitting, it is relatively easy to implement compared to most no-arbitrage models.

In addition to ease of use, another potential advantage offered by equilibrium models is bond valuation. The equilibrium approach allows one to compare observed bond prices and model prices. In principle, this might lead to profitable trading strategies among bonds. Of course, this is most likely to occur in markets where bonds are relatively thinly traded. On the other hand, no-arbitrage models simply assume that all existing bonds are correctly priced and proceed from there to derivative valuation, without considering the possibility of the bonds themselves being mis-priced (and, if such mis-pricing is present, impounding these errors into the prices of derivatives).

Based on the considerations outlined above, this thesis will focus exclusively on the equilibrium approach. Relative to the existing literature, the primary focus is on the implications of discontinuous jumps and non-linear drift. The motivation for this is primarily based on existing empirical evidence about changes in short term interest rates. Many current pricing models are based on modelling these changes as continuous time diffusion processes. In general terms, this means that over short time intervals, changes in interest rates arise from two factors: an expected change or drift, and a random noise term. However, the drift in the pricing models is typically assumed to be a mean-reverting linear function of the level of interest rates. There is a tendency for interest rates to return to some long run average level, but the speed at which that happens does not depend on how far away interest rates are from their long run average. The advantage of a nonlinear drift specification is that it allows for much faster reversion when interest rates are extremely high or low. For example, suppose the long run average level is 6%. Nonlinear drift means that interest rates will be moving back towards this average more quickly if they are currently at 20% than if they are at 8%. This makes sense intuitively since it would be difficult for the economy to sustain very high interest rates for a long time. The second feature of a diffusion process is that the noise term cannot be very large because the sample paths

of such a process are continuous. This means that sudden large changes in interest rates cannot be captured by such a model. However, interest rates often move quickly and by large amounts, perhaps due to factors such as macroeconomic news announcements. This can only be modelled by allowing for discontinous jumps in the path of interest rates. Overall, there seems to be evidence that changes in interest rates should be modelled using more flexible specifications than linear drift diffusion processes.

The balance of the thesis is organized as follows. Chapter 2 reviews the the relevant literature in general terms, highlighting the main issues to be addressed in this thesis. Chapter 3 reviews the established theory in the pure diffusion case. Chapter 4 extends this to cases where the short term interest rate r can exhibit discontinuous jumps. Chapter 5 describes the proposed numerical algorithms. Chapter 6 then provides a series of tests to evaluate the convergence of the algorithms. Chapter 7 provides a sensitivity analysis, showing the effects of changes in the values of the various parameters on bond and bond option prices. Chapter 8 summarizes and concludes.

Chapter 2

Review of Relevant Prior Literature

As noted above, the prior literature is extensive, but we will focus only on equilibrium models. We will concentrate on single factor models. This is primarily for simplicity, but it should also be noted that studies such as [25, 26] have reported that on the order of 90% of the variation in bond returns can be explained by a single factor.

2.1 One-Factor Models of the Short Rate

This material is standard in the literature, and may be found in sources such as [6]. It is provided here for convenience. Let $B(t)$ be the value of a *bank account* (or *money account*) at time $t \geq 0$. The dynamics of the bank account happen according to:

$$dB(t) = r(t)B(t)dt, \tag{2.1}$$

where $r(t)$ is a function of time. Assuming $B(0) = 1$,

$$B(t) = \exp\left\{\int_0^t r(u)du\right\}. \tag{2.2}$$

Since investing a unit amount at time 0 yields the value $B(t)$ at time t, $r(t)$ is called the *instantaneous rate* at which the bank account accrues. We refer to this rate as the *instantaneous spot rate*, or simply as the *short rate*. One can easily see that the *discount process* at any point of time becomes

$$\frac{1}{B(t)} = \exp\left\{-\int_0^t r(u)du\right\}. \tag{2.3}$$

The simplest model for the short rate is represented by the following stochastic differential equation

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t) \tag{2.4}$$

in which $\mu(t, r(t))$ and $\sigma(t, r(t))$ denote the instantaneous drift and diffusion of the process at any point of time and $W(t)$ is a Brownian motion under the real-world probability measure P. Since the above interest rate is defined by one stochastic differential equation, we call the model a *single factor* model. Although these types of models are too simplistic in some contexts because they imply that changes in yields of all maturities are (locally) perfectly correlated, they enable us to lay a foundation for more elaborate models.

We will now examine two classic models, those of Vasicek [30], and Cox, Ingersoll, and Ross [11]. We will compare these models and discuss some of the advantages and shortcomings of each of them. Our discussion of their drawbacks will serve as a basis for the types of models that we will be implementing in this work.

2.1.1 Vasicek Model

The Vasicek model [30] assumes that the instantaneous short rate under the actual or real world probability measure P evolves according to an Ornstein-Uhlenbeck process:

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma dW(t), \qquad r(0) = r_0,$$

where θ, κ, and σ are positive constants and $dW(t)$ is the increment of a Wiener process. By integrating the above equation over the time interval $[s, t]$ where $0 \leq s \leq t$, we obtain

$$r(t) = r(s)e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) + \sigma \int_s^t e^{-\kappa(t-u)} dW(u)$$

This implies that $r(t)$ is normally distributed with mean and variance given respectively by

$$E(r(t)|r(s)) = r(s)e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)$$
$$Var(r(t)|r(s)) = \frac{\sigma^2}{2\kappa}\left[1 - e^{-2\kappa(t-s)}\right]$$

The above equations imply that as t goes to infinity, the expected level of this process approaches θ; The constant κ is the speed of mean reversion. Note that the drift of the process is positive for $r(t) < \theta$ and negative for $r(t) > \theta$. Therefore the process tends to drift back to θ.

One of the positive aspects of this model is having a closed form solution for $r(t)$. The fact that the process is mean-reverting is another desirable property, since it is consistent with the empirical observation that interest rates tend to remain within certain ranges, as opposed to growing very large or small over the long run.

These properties make the model analytically tractable, i.e. both analytic expressions and distributions of several useful interest rate derivative claims such zero-coupon bonds, interest rate caps, floors and swaptions are obtained easily under this model. (This applies only to European-style derivatives, where the contract holder does not have the right to determine when to exercise any embedded optionality.) A major deficiency of the Vasicek model is that there is a positive probability that the process $r(t)$ assumes a negative value. This would represent an arbitrage opportunity since one could borrow a quantity of funds, hold it as cash and pay back less than the amount borrowed. This makes the model unrealistic at times, producing inaccurate values for derivative claims such as options.

2.1.2 Cox-Ingersoll-Ross (CIR) Model

Cox, Ingersoll and Ross (CIR) [11] introduced their model in 1985. They incorporated a "square-root" term in the diffusion coefficient of the spot rate dynamics proposed by Vasicek. This model has provided a solid foundation for many years due to its analytic tractability and the fact that a negative spot rate cannot happen under this model. The process under the real-world probability measure P is defined by:

$$dr(t) = \kappa[\theta - r(t)]dt + \sigma\sqrt{r(t)}dW(t), \qquad r(0) = r_0,$$

where κ, θ and σ are positive constants. Note that the drift of the process is the same as the Vasicek model while the new $\sqrt{r(t)}$ term is introduced in the diffusion.

The origin remains inaccessible to the process (i.e. the process remains strictly positive) as long as the condition $2\kappa\theta > \sigma^2$ is met. The fact that negative interest rates cannot happen can be regarded as one of the advantages of the CIR model over the Vasicek model. If this condition is not satisfied and $r(t)$ reaches zero, the square-root term multiplying $dW(t)$ will disappear and the drift term will reduce to $\kappa\theta$. Since $\kappa\theta > 0$ the process will be reflected back above zero.

Under the real world measure P, the process $r(t)$ has a noncentral chi-squared distri-

bution. Let p_Y denote the density function of the random variable Y,

$$\begin{aligned} p_{r(t)}(x) &= p_{\chi^2(\nu,\lambda_t)/c_t}(x) = c_t p_{\chi^2(\nu,\lambda_t)}(c_t x) \\ c_t &= \frac{4\kappa}{\sigma^2(1-\exp(\kappa t))} \\ \nu &= 4\kappa\theta/\sigma^2 \\ \lambda_t &= c_t r_0 \exp(-\kappa t) \end{aligned}$$

where the noncentral chi-squared distribution function $\chi^2(\nu, \lambda)$ with ν degrees of freedom and non-centrality parameter λ has density

$$\begin{aligned} p_{\chi^2(\nu,\lambda)}(z) &= \sum_{i=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)^i}{i!} p_{\Gamma(i+\nu/2,1/2)}(z), \\ p_{\Gamma(i+\nu/2,1/2)}(z) &= \frac{(1/2)^{i+\nu/2}}{\Gamma(i+\nu/2)} z^{i-1+\nu/2} e^{-z/2} = p_{\chi^2(\nu+2i)}(z) \end{aligned}$$

The CIR model retains the analytic tractability of the Vasicek model, while offering the advantage of avoiding negative interest rates. It is also worth noting that both the CIR and Vasicek models are special cases of the general affine yield framework developed in [15]. This class of models is popular because it is analytically tractable and can incorporate multiple factors.

2.1.3 A Generalized Approach

As noted by Longstaff et al. [8], both the Vasicek and CIR short-rate models can be nested in the following stochastic differential equation:

$$dr(r) = [\alpha + \beta r(t)]dt + \sigma r^{\gamma} dW_t$$

The Vasicek model is derived by setting γ to zero and the CIR model by setting γ to 0.5. In general, other values of γ do not lead to analytically tractable models. In addition to these two models, a variety of others in the literature such as Merton, Dothan, GBM, Brennan-Schwartz, CIR VR and CEV fall under this category as well. More detailed descriptions of these models can be found in [8] or [7]; here we simply note the differences in the model specifications in Table 2.1:

Model	Short rate evolution
Merton	$dr = \alpha dt + \sigma dW$
Vasicek	$dr = (\alpha + \beta r)dt + \sigma dW$
CIR SR	$dr = (\alpha + \beta r)dt + \sigma r^{1/2} dW$
Dothan	$dr = \sigma r dW$
GBM	$dr = \beta r dt + \sigma r dW$
Brennan-Schwartz	$dr = (\alpha + \beta r)dt + \sigma r dW$
CIR VR	$dr = \sigma r^{3/2} dW$
CEV	$dr = \beta r dt + \sigma r^{\gamma} dW$

Table 2.1: Alternative short rate models with linear drift.

2.2 Empirical Evidence

As indicated in Table 2.1, a wide variety of specifications are nested in the linear drift single factor diffusion

$$dr(t) = [\alpha + \beta r(t)]dt + \sigma r^{\gamma} dW(t). \tag{2.5}$$

Recall that the drift in the Vasicek and CIR models was written in the form $\kappa(\theta - r)$. Matching coefficents, we see that $\kappa = -\beta$ and $\theta = -\alpha/\beta$. Therefore, in equation (2.5), the desirable feature of mean-reversion requires that $\beta < 0$, and the implied reversion level of the short rate is $-\alpha/\beta$. While analytic tractability is one criterion by which to judge the usefulness of these models, it should also be noted that the availability of reliable and robust numerical methods implies that, at a more fundamental level, the models should be judged on their empirical performance.

A variety of empirical studies have attempted to evaluate how well the various models nested in equation (2.5) perform in terms of capturing the statistical behaviour of changes in short term interest rates. In general, it can be difficult to draw firm conclusions because different authors use different proxies for the short rate r (e.g. 1-month Treasury bill yields, 3-month Treasury bill yields, the U.S. Federal funds rate, short-term Eurodollar deposit rates, etc.) and because the studies sometimes use quite different sample periods (e.g. studies which exclude the early 1980s in which very high interest rates were observed reach somewhat different conclusions than studies which incorporate this period). However, three main themes emerge from this empirical literature. These are summarized in the following (sub)sections.

2.2.1 Conditional Heteroskedasticity

Most studies which attempt to estimate the parameter γ in equation (2.5) find a value substantially above zero. For example, γ is estimated to be approximately 1.5 in [8] and about 1.3 in [29]. This implies that changes in interest rates are conditionally heteroskedastic, in that that the conditional variance of interest rate changes strongly increases with the level of interest rates. Models such as Vasicek [30] assume that this volatility is constant, and appear to be seriously mis-specified from this perspective. While the CIR [11] model does incorporate conditional heteroskedasticity, the value of $\gamma = 0.5$ assumed in that model is far lower than what appears to be needed to capture this aspect of the data. Note that the inability of standard models to match the conditional variance of changes in interest rates has led some authors (e.g. [4, 31]) to investigate stochastic volatility models, which are basically extensions of equation (2.5) to cases where σ follows a second diffusion process.

2.2.2 Linearity of Drift

Linear drift is a feature of specification (2.5). In other words, the drift term is a linear function of the level of the short rate. As noted above, if $\alpha > 0$ and $\beta < 0$, the short rate reverts to the level $-\alpha/\beta$. However, the speed of mean reversion (measured by β) is constant, independent of the level of interest rates. This means that interest rates do not revert faster (or slower) if they are far from their long run reversion levels. This aspect of these models has been intensively debated in the literature (see [29] for a more detailed review). Studies such as [2, 28, 3, 24, 29] all report evidence that is inconsistent with the hypothesis of linear drift. It is worth noting that the statistical methods used in some of these studies have been criticized in [9], which shows using Monte Carlo experiments that the estimation techniques are unreliable, finding evidence of nonlinear drift even when the model is known (in the experiment) to be linear. The authors of [9] conclude that there is "no definitive answer" (p. 387) to the question of whether the short rate drift is nonlinear, because "time series methods alone are not capable of producing [reliable] evidence of nonlinearity" (p. 387). However, in [29] both time series and cross-sectional methods are used, and the short rate drift is found to be nonlinear.

In general, there are two ways that non-linear drift has been modelled. First, a non-parametric function $\mu(r,t)$ has been estimated by some authors. Second, a parametric functional form has been specified which incorporates nonlinearity. The typical representation of this form is

$$dr(t) = \left[\alpha_{-1}/r(t) + \alpha_0 + \alpha_1 r(t) + \alpha_2 r(t)^2\right] dt + \sigma r(t)^{\gamma} dW(t) \tag{2.6}$$

where α_{-1}, α_0, α_1, α_2, and $\gamma > 0$ are constant parameters. Due to the presence of the quadratic term r^2 in the drift, we will refer to this specification as a quadratic term structure (QTS) model.

2.2.3 Sample Path Continuity

All of the models discussed thus far are continuous-time diffusion models. They imply that the sample paths of the short rate are continuous functions. Several studies have investigated the suitability of this modelling assumption compared to the alternative of a jump-diffusion process which allows for continuity most of the time but also for discrete discontinuous jumps of random size at random points in time. Evidence in favour of such models is reported in [12, 24, 23], among others. In particular, it is reported in [24] that "jumps are both economically and statistically important ... and generate more than half the conditional variance of interest rate changes" (p. 255). One reason for this is the impact of macroeconomic news announcments on bond markets ([5, 20]).

To briefly summarize the three main empirical findings reported here:

- Changes in interest rates consistently exhibit conditional heteroskedasticity across many empirical studies;
- Changes in interest rates are found by some authors to be characterized by nonlinear drift (though the evidence for this is perhaps less solid than for heteroskedasticity);
- Several relatively recent studies have reported that changes in interest rates can be usefully modelled by a jump-diffusion process.

Note that each of these on its own casts significant doubt on the empirical suitability of the classic models of Vasicek [30] and CIR [11]. Taken together, the evidence against these models is convincing. Consequently, our main focus from here on will be on numerical methods which can be used in more general short rate models.

Chapter 3

Underlying Structure of Model

This chapter reviews the pricing theory for single factor short rate models. Although this material is standard, and can be found in many sources in the literature, we include it here for completeness. The treatment here closely follows that of [6].

Let $P(t,T)$ denote the price at time t of a pure discount bond paying \$1 at maturity time T. Since the dynamics of the short rate r play an important role in determining $P(t,T)$ over the interval $[t,T]$, a reasonable starting point to specify the behaviour of r. Suppose that the dynamics of r under the objective probability measure P is given by equation (2.4). Assume further that the dynamics of the money account $B(t)$ follow equation (2.1), and that bonds are available for every maturity. We can express the above assumptions more formally:

Assumption 1: *We assume the existence of one externally given (locally risk-free) asset. The price dynamics of this asset is given by equation (2.2), for which the dynamics of the short rate of interest, under the objective probability measure P, is given by equation (2.4).*

Assumption 2: *We assume that there exists a market for zero coupon T-maturity bonds for every value of T.*

Thus our market contains an infinite number of bonds plus a risk-free asset. In other words, in our market, the risk-free asset is regarded as the underlying asset, while all bonds are derivatives of the externally given short rate of interest. We will proceed to investigate the relationship between the price processes of bonds with different maturities in the absence of arbitrage.

As an important observation, one shall keep in mind that while bonds are viewed as interest rate derivatives, they are not uniquely determined by P-dynamics of the short rate. This can be shown in light of the following theorem (for proof please see [6]):

Theorem 1. *Let M denote the number of underlying traded assets in the model not counting the risk-free asset, and let R denote the number of random factors. Generically we have the following relations:*

- *The model is arbitrage free if and only if $M \leq R$.*
- *The model is complete if and only if $M \geq R$.*
- *The model is complete and arbitrage free if and only if $M = R$.*

Proof. See [6]. □

In our case the number of externally given traded assets excluding the risk-free asset $M = 0$. Also the fact that we have one driving Wiener process in equation (2.4) gives the number of random factors $R = 1$. Therefore, according to the above theorem we have an arbitrage free incomplete market. Intuitively, an investor has only the option of investing money in the bank account and waiting for it to grow according to the dynamics of equation (2.2); thus it is impossible to replicate any interesting derivative, even such a simple one as T-maturity pure discount bond.

So, we have the following results:

- We cannot completely determine the price of a particular bond using the dynamics of the short rate of interest (2.4) in the absence of arbitrage.
- This is a direct consequence of having insufficient number of underlying assets in our market, making it impossible for us to price a derivative in terms of the price of some underlying assets using arbitrage pricing techniques.

Therefore, we will not be able to derive a unique price for a particular bond. However, this does not imply that bond prices can take any arbitrary form. Instead:

- In order for bond market to be arbitrage free, the prices of bonds with different maturities should satisfy a certain internal consistency relationship.
- We should be able to uniquely determine the prices of all bonds in terms of price of a given particular "benchmark" bond and the assumed dynamics of r.

The new benchmark bond will increase the number of externally given assets to $M = 1$; as a result of that we achieve completeness in the market, since we have $R = M = 1$.

3.1 The Market Price of Risk

A concept of fundamental importance in our context is the *market price of risk*. This will be used to link the prices of bonds of all maturities. Although we are concentrating exclusively on the single factor context, the ideas can be easily extended to multi-factor models.

Recall that we are assuming that the prices of pure discount bonds $P(t,T)$ depend on the single stochastic variable r and that such bonds are available for every maturity T. Denote the price of a T-maturity bond by

$$P(t,T) = F(t, r(t); T),$$

where F is a smooth function of three variables. To further simplify notation, since we can regard T as a parameter, we will let $F(t, r(t); T = F^T(t, r(t))$. Since at T the bond is worth \$1, we have the following simple boundary condition:

$$F^T(T, r(T)) = 1, \quad \forall r. \tag{3.1}$$

Suppose we form a portfolio consisting of two bonds with different times to maturitiy S and T. We are interested in the evolution of the value dynamics of our portfolio. To start, we need to apply Itô's formula to F^T. Recall that:

Itô's formula: *Assume that the process $X(t)$ has a stochastic differential given by*

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

and let f be a twice differentiable function. Define the process Z by $Z(t) = f(t, X(t))$. Then Z has the stochastic differential given by

$$df(t, X(t)) = \left\{ \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} dt + \sigma \frac{\partial f}{\partial x} dW(t)$$

Using this, we can derive the price dynamics of the T- maturity bond (with a corresponding equation for the S-maturity bond) when the short rate of interest evolves according to equation (2.4):

$$dF^T = F^T \left\{ \alpha_T dt + \sigma_T dW(t) \right\} \tag{3.2}$$

where

$$\alpha_T = \frac{1}{F^T}(F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T), \tag{3.3}$$

$$\sigma_T = \frac{1}{F^T}(\sigma F_r^T). \tag{3.4}$$

Let $\pi = (w_T, w_S)$ denote the portfolio weights with $w_S + w_T = 1$, and let V be the value of this portfolio. Then the value dynamics of our portfolio are:

$$dV = V\left\{w_T\frac{dF^T}{F^T} + w_S\frac{dF^S}{F^S}\right\} \tag{3.5}$$

Substituting the appropriate differentials for T and S from (3.2) gives:

$$dV = V\left\{(w_T\alpha_T + w_S\alpha_S)dt\right\} + V\left\{(w_T\sigma_T + w_S\sigma_S)dW\right\} \tag{3.6}$$

If the portfolio V satisfies the following system of equations

$$w_S + w_T = 1$$
$$w_T\sigma_T + w_S\sigma_S = 0,$$

we have a locally risk-free portfolio, since the dW term in equation (3.5) will disappear. Therefore the value dynamics are reduced to

$$dV = V\left\{w_T\alpha_T + w_S\alpha_S\right\}dt \tag{3.7}$$

By solving the above system of equations we have:

$$w_T = -\frac{\sigma_S}{\sigma_T - \sigma_S} \tag{3.8}$$

$$w_S = \frac{\sigma_T}{\sigma_T - \sigma_S} \tag{3.9}$$

By substituting these coefficients in the equation (3.7), we will have

$$dV = V\left\{\frac{\alpha_S\sigma_T - \alpha_T\sigma_S}{\sigma_T - \sigma_S}\right\}dt \tag{3.10}$$

Since this portfolio is risk-free, the no-arbitrage principle implies that it must have a rate of return equal to the risk-free rate:

$$\frac{\alpha_S\sigma_T - \alpha_T\sigma_S}{\sigma_T - \sigma_S} = r(t) \quad \forall t. \tag{3.11}$$

By reshuffling the terms we get

$$\frac{\alpha_S(t) - r(t)}{\sigma_S(t)} = \frac{\alpha_T(t) - r(t)}{\sigma_T(t)} \tag{3.12}$$

A close examination of equation (3.12) shows that the left hand is independent of T and the right hand side is independent of S; this leads to the following crucial result:

Proposition 2. *Assume that we have an arbitrage free bond market. Then there exists a process λ such that the relation*

$$\frac{\alpha_T(t) - r(t)}{\sigma_T(t)} = \lambda(t) \tag{3.13}$$

holds for all t and for every choice of maturity time T.

An important characteristic of the process $\lambda(t)$ is that it holds regardless of the choice of time to maturity T. In the numerator of (3.13) we have the term $\alpha_T(t) - r(t)$. According to equation (3.2), α_T is the local rate of return on the T-maturity bond, whereas r is the rate of return of the risk-free asset. Therefore the difference $\alpha_T(t) - r(t)$ is the risk premium of the T-maturity bond. In the denominator, we have $\sigma_T(t)$, the local volatility of the T-maturity bond. Thus we can think of λ as representing the risk premium per unit volatility. The process λ is known as the *market price of risk*. Using this terminology, we can rephrase Proposition 2 as follows: *In an arbitrage free market, all bonds will have the same market price of risk, regardless of their maturity times.*

3.2 The Term Structure Equation

One of the most fundamental results in the theory of interest rates can be obtained by substituting α_T and σ_T from (3.3) and (3.4) into equation (3.13) for the market price of risk. The result is presented in the following proposition:

Proposition 3. *In an arbitrage free bond market, F^T will satisfy the term structure equation:*

$$F_t^T + \{\mu - \lambda\sigma\}F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T - rF^T = 0, \tag{3.14}$$

subject to the boundary condition (3.1).

Although the term structure equation above is a standard partial differential equation (PDE), it is important to note that λ is not determined inside the model. In particular, it must be specified exogenously, as with the drift μ and volatility σ.

It is also important to note that we can express the solution to the term structure equation (3.14) in the form of an expectation under an alternative "risk-neutral" probability measure Q:

$$F^T(t, r(t)) = E_t^Q[e^{-\int_t^T r(s)ds}]. \tag{3.15}$$

In (3.15, E^Q denotes expectation under the measure Q, for which the dynamics of the short rate is governed not by (2.4), but instead by:

$$dr(t) = \{\mu(r,t) - \lambda(t)\sigma(r,t)\}dt + \sigma(r,t)dW(t) \tag{3.16}$$

Proofs of this result rely on the Feynman-Kac formula and can be found in sources such as [6]. For completeness, we provide a proof in Appendix A. This important result implies that Monte Carlo methods can be used to find numerical solutions for bond prices by repeatedly simulating the process for r under the Q measure and calculating the expectation in (3.15). More generally, a European-style derivative with payoff function $H(T, r(T))$ at maturity T solves the term structure equation (3.14) subject to the boundary condition $F^T(T, r(T)) = H(T, r(T))$. Moreover, the value of such a contract can be found by calculating

$$E_t^Q\left[e^{-\int_t^T r(s)ds} H(T, r(T))\right].$$

The bond pricing result (3.15) is a simple special case where $H(T, r(T)) = 1$.

3.3 Bond Price Under Classic Models

As mentioned above in Sections 2.1.1 and 2.1.2, analytic solutions are available for bond prices for the Vasicek [30] and CIR [11] models. However, an implication of the results above is that this will require us to specify the market price of risk $\lambda(t)$. In fact, the analytic solutions rely on particular forms of the market price of risk function which preserve the analytic tractability of the model when we switch to the Q measure from the P measure.

Consider first the Vasicek model, and assume that:

$$\lambda(t) = \lambda_0 + \lambda_1 r(t)/\sigma,$$

where λ_0 and λ_1 are constants. Under the Q-measure, we have

$$\begin{aligned} dr(t) &= [\kappa(\theta - r(t)) - \lambda_0\sigma_0 - \lambda_1 r(t)]dt + \sigma dW(t) \\ &= [(\kappa\theta - \lambda_0\sigma) - (\kappa + \lambda_1)r(t)]dt + \sigma dW(t), \end{aligned}$$

which can be expressed in the form:

$$dr(t) = [a(b - r(t))]dt + \sigma dW(t),$$

where a and b are constants. As shown in [30], we can then analytically calculate the price of a zero-coupon bond at time t for a payoff of \$1 at the maturity T as:

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)} \tag{3.17}$$

where

$$\begin{aligned} B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \\ A(t,T) &= \exp\left\{ \frac{[B(t,T) - T + t][a^b - \sigma^2/2]}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a} \right\}. \end{aligned}$$

Similarly, for the CIR model, if we assume that $\lambda(t) = \lambda_1\sqrt{r(t)}/\sigma$, we find that under measure Q:

$$dr(t) = [\kappa\theta - (\kappa + \lambda_1)r(t)]dt + \sigma\sqrt{r(t)}dW(t),$$

which can be expressed in the form:

$$dr(t) = [a(b - r(t))]dt + \sigma\sqrt{r(t)}dW(t),$$

with a and b being constants. From [11], we have the analytic solution:

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)} \tag{3.18}$$

where

$$\begin{aligned} B(t,T) &= \frac{2[e^{c(T-t} - 1]}{[c + a][e^{c(T-t)} - 1] + 2c} \\ A(t,T) &= \left(\frac{2ce^{(a+c)(T-t)/2}}{[c + a][e^{c(T-t} - 1] + 2c} \right)^{2ab/\sigma^2} \\ c &= \sqrt{a^2 + 2\sigma^2}. \end{aligned}$$

It is worth emphasizing that the Vasicek and CIR models are two of the only cases for short rate models where analytic solutions are known. In almost every other model (including variations of the Vasicek and CIR specifications with different forms of the market price of risk), numerical methods must be relied on. Moreover, such methods are necessary in every non-trivial case for American-style interest rate derivatives.

Chapter 4

Short Rate Models With Jumps

We now proceed to present an extension of the term structure equation (3.14) which incorporates the assumption that changes in the short rate are modelled by a jump-diffusion process. We begin by briefly outlining the stochastic setting of the models.

The general single factor model for r is assumed to be replaced by

$$dr(t) = \mu(t, r(t))dt + \sigma(r, r(t))dW(t) + r(t)\left\{j(t) - 1\right\}dN(t). \tag{4.1}$$

As in (3.14), the diffusive drift μ and volatility σ are assumed to be functions of time and the short rate r. However, we now have the additional feature of a Poisson process $N(t)$ which has an assumed intensity $\hat{\lambda}$. This implies that during a small time interval dt:

$$\begin{aligned}
&\text{Prob(short rate jumps once)} = \text{Prob}(dN(t) = 1) \simeq \hat{\lambda}dt,\\
&\text{Prob(short rate jumps more than once)} = \text{Prob}(dN(t) \geq 2) \simeq 0,\\
&\text{Prob(short rate does not jump)} = \text{Prob}(dN(t) = 0) \simeq 1 - \hat{\lambda}dt.
\end{aligned}$$

Assume that during dt a jump occurs, taking the short rate from its level immediately before the jump of $r(t^-)$ to $r(t) = j(t)r(t^-)$. The percentage change in the short rate due to the jump is given by

$$\frac{dr(t)}{r(t)} = \frac{j(t)r(t^-) - r(t^-)}{r(t^-)} = j(t) - 1.$$

We have two independent sources of randomness creating jumps. The first is the Poisson process $N(t)$, which accounts for the timing of jumps. Conditional on a jump occurring, the size of the jump is independently determined by $j(t)$, with the relative jump size being $j(t) - 1$.

4.1 The Pricing Equation of The Short Rate Model With Jumps

While it is possible to use general equilibrium arguments to derive the pricing equation for a short rate model with jumps (see [1] for example), we will use much simpler hedging arguments along the lines presented in [19]. We begin by noting Itô's formula for a jump-diffusion process:

Itô's formula for a jump-diffusion process: *Assume that the process $X(t)$ is a jump-diffusion:*

$$X(t) = X(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)ds + \sum_{i=1}^{N(t)} \Delta X(i),$$

where $N(t)$ is a compound Poisson process. Then for a twice differentiable function $f(t,(X(t))$ we have:

$$df(t,X(t)) = \left\{ \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\} dt + \sigma \frac{\partial f}{\partial x} dW(t) + f(X(t^-) + \Delta X(t)) - f(X(t^-)).$$

As above, let $F^T(r,T)$ denote the value of a T-maturity pure discount bond. Applying Itô's formula and simplifying notation, we have:

$$dF^T = \alpha_T dt + \sigma_T dW(t) + J_T dN(t) \tag{4.2}$$

where

$$\alpha_T = F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T, \tag{4.3}$$

$$\sigma_T = \sigma F_r^T, \tag{4.4}$$

$$J_T = F^T(j(t)r,t) - F^T(r,t). \tag{4.5}$$

We will assume for the moment that there is only one possible jump size J, i.e. following a jump $r \to Jr$, with J being a known constant. In order to hedge the two sources of randomness caused by the diffusion dW and the jump dN, we need to form a portfolio V containing three bonds. Let the maturities of these bonds be denoted by S, T, and U. Our portfolio weights will be assumed to satisfy $w_S + w_T + w_U = 1$. The value of our portfolio will evolve according to:

$$dV = \left\{ w_S dF^S + w_T dF^T + w_U dF^U \right\}. \tag{4.6}$$

By substituting (4.2) and its counterparts for S and U in (4.6) we have:

$$\begin{aligned} dV = &(w_S\alpha_S + w_T\alpha_T + w_U\alpha_U)dt \\ &+ (w_S\sigma_S + w_T\sigma_T + w_U\sigma_U)dW \\ &+ (w_S J_S + w_T J_T + w_U J_U)dN(t) \end{aligned}$$

By setting the coefficients of $dW(t)$ and $dN(t)$ to zero, we can hedge the randomness caused by both the diffusion and jump terms. Therefore if we define the portfolio V by the following system of equations (in addition to the constraint that the portfolio weights sum to one):

$$\begin{aligned} w_S\sigma_S + w_T\sigma_T + w_U\sigma_U &= 0 \\ w_S J_S + w_T J_T + w_U J_U &= 0, \end{aligned}$$

we have a portfolio which is locally risk-free. The value of the porfolio should then grow at the risk-free rate:

$$dV = rV dt = r(w_S F^S + w_T F^T + w_U F^U)dt.$$

Consequently:

$$(w_S\alpha_S + w_T\alpha_T + w_U\alpha_U) = r(w_S F^S + w_T F^T + w_U F^U),$$

or equivalently,

$$w_S(\alpha_S - rF^S) + w_T(\alpha_T - rF^T) + w_U(\alpha_U - rF^U) = 0.$$

This implies that we have the system:

$$\begin{aligned} w_S\sigma_S + w_T\sigma_T + w_U\sigma_U &= 0 \\ w_S J_S + w_T J_T + w_U J_U &= 0 \\ w_S(\alpha_S - rF^S) + w_T(\alpha_T - rF^T) + w_U(\alpha_U - rF^U) &= 0. \end{aligned}$$

In order for this system to have a nontrivial solution, the equations must be linearly dependent. In other words, we need to be able to write the coefficients of the last equation as a linear combination of the previous equations:

$$\begin{aligned} \alpha_S - rF^S &= \lambda_W\sigma_S + \lambda_J J_S \\ \alpha_T - rF^T &= \lambda_W\sigma_T + \lambda_J J_T \\ \alpha_U - rF^U &= \lambda_W\sigma_U + \lambda_J J_U. \end{aligned}$$

After dropping the maturity indices, we have:

$$\alpha - rF = \lambda_W \sigma + \lambda_J J.$$

By substituting the definitions of α, σ and J from (4.3), (4.4), and (4.5) and rearranging the terms we derive the following:

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - rF^T + \lambda_J(F^T(Jr,t) - F^T(r,t)) = 0. \tag{4.7}$$

Now suppose that there is a finite number n of possible jump states which occur with probability λ_J^i $(i = 1, \ldots, n)$, where after each jump the short rate jumps to $j(t_i)r(t_i)$, i.e.

$$r(t_i) \to j(t_i)r(t_i), \quad i = 1, \ldots, n.$$

We can form a hedging portfolio using $(n+2)$ bonds with different maturities T_i $(i = 1, \ldots, n+2)$:

$$V = \sum_{i=1}^{n+2} w_i F^{T_i} \qquad \text{where } T_i \neq T_k \quad \text{for } i \neq k.$$

If we follow a similar approach to hedge diffusion and jumps, we obtain the following:

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - rF^T + \sum_{i=1}^{n} \lambda_J^i (F^T(j(t_i)r(t_i), t_i) - F^T(r(t_i), t_i)) = 0. \tag{4.8}$$

Define

$$p(j_i) = \frac{\lambda_J^i}{\sum_{i=1}^n \lambda_J^i}$$
$$\lambda^* = \sum_{i=1}^{n} \lambda_J^i.$$

By using the above identities, we can rewrite equation (4.8) as:

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - rF^T + \lambda^* \sum_{i=1}^{n} p(j_i)(F^T(j(t_i)r(t_i), t) - F^T(r(t_i), t_i)) = 0. \tag{4.9}$$

Observe that since $\lambda_j^i \geq 0$, $p(j_i) \geq 0$ and thus $\lambda^* \geq 0$. If we take the limit as the number of jump states tends to infinity, $p(j_i)$ will tend to a continuous distribution, and equation (4.9) becomes the following partial integro differential equation (PIDE):

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - rF^T + \lambda^* \int_0^\infty p(j(t))[F^T(j(t)r, t) - F^T(r,t)]dj(t) = 0. \tag{4.10}$$

By breaking the integral into two parts and simplifying, we can rewrite (4.10) in the form:

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - (r + \lambda^*)F^T + \lambda^* \int_0^\infty p(j(t))F^T(j(t)r, t)dj(t) = 0. \quad (4.11)$$

Note that the jump intensity λ^* in equation (4.11) will not generally be the same as the actual jump intensity $\hat{\lambda}$. This should not come as a surprise. We are effectively using a risk-adjusted jump intensity, much like the market price of risk changes the expected drift of the diffusion term. Similarly, the parameters of the jump size distribution can be different under the risk-adjusted pricing measure as compared to the actual physical measure. Of course, if we estimate parameters by calibrating to observed prices of financial instruments, we will be finding the risk-adjusted jump parameters.

Equation (4.11) is the general form of the pricing PIDE that will be used in our numerical tests. It is worth concluding this section by noting that this equation is somewhat different from the standard pricing PIDE for equity options (see, e.g. [13]). The reason is that in this interest rate derivative context, the basic underlying state variable r is not a traded asset. If we imagine that r is a traded asset, then it must satisfy equation (4.10). Suppose that we rewrite this equation, using some other state variable x in place of the short rate r (and dropping the T-superscript for simplicity):

$$F_t + \frac{1}{2}\sigma^2 F_{xx} + (\mu - \lambda_W \sigma)F_x - rF + \lambda^* \int_0^\infty p(j(t))(F(j(t)x) - F(x)) = 0.$$

If x is a traded asset, then $F = x$ must satisfy the above equation. This implies that $F_{xx} = F_t = 0$ and $F_x = 1$, and in turn:

$$(\mu - \lambda_W \sigma) - rx + \lambda^* \int_0^\infty p(j(t))x[j(t) - 1]dj(t) = 0$$
$$\Rightarrow (\mu - \lambda_W \sigma) = [r - \lambda^* E(j(t) - 1)]x.$$

This implies that we would replace the term in the pricing PIDE involving the risk-adjusted drift $(\mu - \lambda_W \sigma)$ by the risk-free rate minus the expected relative jump size. The standard PIDE for equity options reflects this change because the underlying stock price is a traded asset.

Chapter 5

Numerical Methods

This chapter describes the numerical techniques that will be used to solve the various pricing equations. We start by considering the case where there are no jumps. The methods used in this case are somewhat similar to those described in [32].

5.1 Numerical Solution Without Jumps

In the case where there is diffusion but no jumps, interest rate derivatives can be valued by solving equation (3.14). We can rewrite equation (3.14) in the form:

$$F_\tau = \{\mu - \lambda\sigma\}F_r + \frac{1}{2}\sigma^2 F_{rr} + rF, \tag{5.1}$$

where $\tau = T - t$ denotes time running in the backwards direction from the expiration date T to the current time t. With this change, the terminal payoff condition for the contract of interest ($F(r,T) = 1 \ \forall \ r$ in the case of a pure discount bond having a par value of \$1) becomes an initial condition for our numerical methods.

In this section, we will describe the general numerical methods that we use for solving this class of PDEs numerically. Equation (5.1) is of the general form:

$$u_\tau = a(x)u_{xx} + b(x)u_x + c(x)u \tag{5.2}$$

where u_τ, u_x and u_{xx} represent first and second derivatives with respect to time and state variable. In order to introduce the different numerical schemes which we use for solving

this type of equation, we start by defining general approximation methods of the terms present in equation (5.2). *Central*, *forward*, and *backward* derivative approximations for u_x on a discrete grid $x_i,\ i = 1, \ldots, N$ are defined as:

$$\begin{aligned} u_x &= \frac{u_{i+1} - u_{i-1}}{\Delta x_{i+1} + \Delta x_i} \quad \text{(central)} \\ &= \frac{u_{i+1} - u_i}{\Delta x_{i+1}} \qquad \text{(forward)} \\ &= \frac{u_i - u_{i-1}}{\Delta x_i} \qquad \text{(backward)} \end{aligned}$$

where the grid spacing is given by:

$$\begin{aligned} \Delta x_{i+1} &= x_{i+1} - x_i \\ \Delta x_i &= x_i - x_{i-1} \end{aligned}$$

Similarly, u_{xx} and u_τ are approximated by:

$$\begin{aligned} u_\tau &= \frac{u_i^{n+1} - u_i^n}{\Delta \tau} \\ u_{xx} &= \frac{(u_{i+1} - u_i)/\Delta x_{i+1} + (u_{i-1} - u_i)/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} \end{aligned}$$

where the time step size is $\Delta\tau$. Note that we are using an unevenly-spaced grid for x. This offers considerable advantages because we can use a fine spacing in regions where we want the computed solution to be highly accurate and a coarse spacing for other regions. It is also possible to use an unevenly spaced grid in the time direction, but as our descriptions to follow just depend on the solution values at two different time points, we will just let the time step size be denoted by $\Delta\tau$. As we will see, it is advantageous to use central differences for some parts of the grid and forward or backward differences for other parts of it. The spatial derivative approximations u_x and u_{xx} can be evaluated either at time level n or at time level $n + 1$. If we use the time level n, we will have an *explicit* scheme. If we use the time level $n + 1$, then our scheme is *implicit*. We can also combine them, averaging the values at the two time levels. This is known as a *Crank-Nicolson* scheme. In the following (sub)sections, we will describe the resulting discrete equations for these various cases.

5.1.1 Explicit Scheme, Central Difference

In this method, we use $(i-1)$-th, i-th and $(i+1)$-th grid nodes at the n-th time step for calculating u_i^{n+1}:

$$\frac{u_i^{n+1}-u_i^n}{\Delta\tau} = a_i\left[\frac{(u_{i+1}^n-u_i^n)/\Delta x_{i+1}+(u_{i-1}^n-u_i^n)/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right]+b_i\left[\frac{u_{i+1}^n-u_{i-1}^n}{\Delta x_{i+1}+\Delta x_i}\right]+c_iu_i^n.$$

In the above, a_i, b_i, and c_i denote the values of $a(x)$, $b(x)$, and $c(x)$ at the i-th grid node. For future reference, note that in our applications $a_i \geq 0$and $c_i = -r_i \leq 0$, while b_i can be of either sign. The expression above can be written in the form:

$$\begin{aligned}
u_i^{n+1} &= u_{i-1}^n\left[\frac{a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}-\frac{b_i\Delta\tau}{\Delta x_{i+1}+\Delta x_i}\right]\\
&\quad+u_i^n\left[1+c_i\Delta\tau-\frac{a_i\Delta t/\Delta x_{i+1}+a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right]\\
&\quad+u_{i+1}^n\left[\frac{a_i\Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1}+\Delta x_i)/2}+\frac{b_i\Delta\tau}{\Delta x_{i+1}+\Delta x_i}\right]\\
\Rightarrow u_i^{n+1} &= \alpha_i^c\Delta\tau u_{i-1}^n+\left[1-(\alpha_i^c+\beta_i^c-c_i)\Delta\tau\right]u_i^n+\beta_i^c\Delta\tau u_{i+1}^n,
\end{aligned}$$

where

$$\begin{aligned}
\alpha_i^c &= a_i/\left[\Delta x_i(\Delta x_{i+1}+\Delta x_i)/2\right]-b_i/\left[\Delta x_{i+1}+\Delta x_i\right]\\
\beta_i^c &= a_i/\left[\Delta x_{i+1}(\Delta x_{i+1}+\Delta x_i)/2\right]+b_i/\left[\Delta x_{i+1}+\Delta x_i\right].
\end{aligned}$$

5.1.2 Implicit Scheme, Central Difference

In this method, we use $(i-1)$-th, i-th and $(i+1)$-th grid nodes at the $(n+1)$-th time step for calculating u_i^n:

$$\frac{u_i^{n+1}-u_i^n}{\Delta\tau} = a_i\left[\frac{(u_{i+1}^{n+1}-u_i^{n+1})/\Delta x_{i+1}+(u_{i-1}^{n+1}-u_i^{n+1})/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right]+b_i\left[\frac{u_{i+1}^{n+1}-u_{i-1}^{n+1}}{\Delta x_{i+1}+\Delta x_i}\right]+c_iu_i^{n+1}.$$

This can be written in the form:

$$\begin{aligned} u_i^n &= u_{i-1}^{n+1}\left[-\frac{a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}+\frac{b_i\Delta\tau}{\Delta x_{i+1}+\Delta x_i}\right] \\ &\quad + u_i^{n+1}\left[1-c_i\Delta\tau+\frac{a_i\Delta t/\Delta x_{i+1}+a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right] \\ &\quad + u_{i+1}^{n+1}\left[-\frac{a_i\Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1}+\Delta x_i)/2}-\frac{b_i\Delta t}{\Delta x_{i+1}+\Delta x_i}\right] \\ \Rightarrow u_i^n &= -\alpha_i^c\Delta\tau u_{i-1}^{n+1}+\left[1+(\alpha_i^c+\beta_i^c-c_i)\Delta\tau\right]u_i^{n+1}-\beta_i^c\Delta t u_{i+1}^{n+1}, \end{aligned}$$

where, as for the explicit case:

$$\begin{aligned} \alpha_i^c &= a_i/\left[\Delta x_i(\Delta x_{i+1}+\Delta x_i)/2\right]-b_i/\left[\Delta x_{i+1}+\Delta x_i\right] \\ \beta_i^c &= a_i/\left[\Delta x_{i+1}(\Delta x_{i+1}+\Delta x_i)/2\right]+b_i/\left[\Delta x_{i+1}+\Delta x_i\right]. \end{aligned}$$

5.1.3 Explicit Scheme, Forward Difference

When a scheme with central difference approximations is used,

$$\alpha_i^c = a_i/\left[\Delta x_i(\Delta x_{i+1}+\Delta x_i)/2\right]-b_i/\left[\Delta x_{i+1}+\Delta x_i\right]$$

can become negative. This is due to presence of the term $b_i/\left[\Delta x_{i+1}+\Delta x_i\right]$ in α_i^c (assuming $b_i > 0$). As discussed below, this can lead to undesirable effects in our numerical solution. As an alternative, we can use a forward difference approximation for u_x. For the explicit case:

$$\frac{u_i^{n+1}-u_i^n}{\Delta\tau} = a_i\left[\frac{(u_{i+1}^n-u_i^n)/\Delta x_{i+1}+(u_{i-1}^n-u_i^n)/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right]+b_i\left[\frac{u_{i+1}^n-u_i^n}{\Delta x_{i+1}}\right]+c_i u_i^n,$$

so

$$\begin{aligned} u_i^{n+1} &= u_{i-1}^n\left[\frac{a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}\right] \\ &\quad + u_i^n\left[1+c_i\Delta\tau-\frac{a_i\Delta t/\Delta x_{i+1}+a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1}+\Delta x_i)/2}-\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ &\quad + u_{i+1}^n\left[\frac{a_i\Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1}+\Delta x_i)/2}+\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ \Rightarrow u_i^{n+1} &= \alpha_i^f\Delta\tau u_{i-1}^n+\left[1-(\alpha_i^f+\beta_i^f-c_i)\Delta\tau\right]u_i^n+\beta_i^f\Delta\tau u_{i+1}^n, \end{aligned}$$

where

$$\alpha_i^f = a_i / \left[\Delta x_i(\Delta x_{i+1} + \Delta x_i)/2\right]$$
$$\beta_i^f = a_i / \left[\Delta x_{i+1}(\Delta x_{i+1} + \Delta x_i)/2\right] + b_i / \left[\Delta x_{i+1}\right].$$

5.1.4 Implicit Scheme, Forward Difference

As above, it is possible for the central difference approximation to lead to a negative coefficient. If this happens when $b_i > 0$, then we can use a forward difference instead:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = a_i \left[\frac{(u_{i+1}^{n+1} - u_i^{n+1})/\Delta x_{i+1} + (u_{i-1}^{n+1} - u_i^{n+1})/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2}\right] + b_i \left[\frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x_{i+1}}\right] + c_i u_i^{n+1},$$

so

$$\begin{aligned}
u_i^n &= u_{i-1}^{n+1}\left[-\frac{a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2}\right] \\
&\quad + u_i^{n+1}\left[1 - c_i\Delta\tau + \frac{a_i\Delta\tau/\Delta x_{i+1} + a_i\Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} + \frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\
&\quad + u_{i+1}^{n+1}\left[-\frac{a_i\Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1} + \Delta x_i)/2} - \frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\
\Rightarrow u_i^n &= -\alpha_i^f\Delta\tau u_{i-1}^{n+1} + \left[1 + (\alpha_i^f + \beta_i^f - c_i)\Delta\tau\right] u_i^{n+1} - \beta_i^f\Delta\tau u_{i+1}^{n+1},
\end{aligned}$$

where, as for the explicit case:

$$\alpha_i^f = a_i / \left[\Delta x_i(\Delta x_{i+1} + \Delta x_i)/2\right]$$
$$\beta_i^f = a_i / \left[\Delta x_{i+1}(\Delta x_{i+1} + \Delta x_i)/2\right] + b_i / \left[\Delta x_{i+1}\right].$$

5.1.5 Explicit Scheme, Backward Difference

Recall that in a scheme with central difference approximations, we have:

$$\beta_i^c = a_i / \left[\Delta x_{i+1}(\Delta x_{i+1} + \Delta x_i)/2\right] + b_i / \left[\Delta x_{i+1} + \Delta x_i\right].$$

This can potentially become negative if $b_i < 0$. In this case, we can use a backward difference approximation for u_x. This will remove the term $b_i / \left[\Delta x_{i+1} + \Delta x_i\right]$ from the above expression. The scheme is:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = a_i \left[\frac{(u_{i+1}^n - u_i^n)/\Delta x_{i+1} + (u_{i-1}^n - u_i^n)/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2}\right] + b_i \left[\frac{u_i^n - u_{i-1}^n}{\Delta x_i}\right] + c_i u_i^n,$$

so

$$\begin{aligned} u_i^{n+1} &= u_{i-1}^n \left[\frac{a_i \Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} - \frac{b_i \Delta\tau}{\Delta x_i} \right] \\ &\quad + u_i^n \left[1 + c_i \Delta\tau - \frac{a_i \Delta\tau/\Delta x_{i+1} + a_i \Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} + \frac{b_i \Delta\tau}{\Delta x_i} \right] \\ &\quad + u_{i+1}^n \left[\frac{a_i \Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1} + \Delta x_i)/2} \right] \\ \Rightarrow u_i^{n+1} &= \alpha_i^b \Delta\tau u_{i-1}^n + \left[1 - (\alpha_i^b + \beta_i^b - c_i)\Delta\tau \right] u_i^n + \beta_i^b \Delta\tau u_{i+1}^n, \end{aligned}$$

where

$$\begin{aligned} \alpha_i^b &= a_i / \left[\Delta x_i (\Delta x_{i+1} + \Delta x_i)/2 \right] - b_i / \left[\Delta x_i \right] \\ \beta_i^b &= a_i / \left[\Delta x_{i+1} (\Delta x_{i+1} + \Delta x_i)/2 \right]. \end{aligned}$$

5.1.6 Implicit Scheme, Backward Difference

As for the explicit case, we may need to turn to a backward difference approximation of u_x if the central difference leads to a negative coefficient because $b_i < 0$:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = a_i \left[\frac{(u_{i+1}^{n+1} - u_i^{n+1})/\Delta x_{i+1} + (u_{i-1}^{n+1} - u_i^{n+1})/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} \right] + b_i \left[\frac{u_i^{n+1} - u_{i-1}^{n+1}}{\Delta x_i} \right] + c_i u_i^{n+1},$$

so

$$\begin{aligned} u_i^n &= u_{i-1}^{n+1} \left[-\frac{a_i \Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} + \frac{b_i \Delta\tau}{\Delta x_i} \right] \\ &\quad + u_i^{n+1} \left[1 - c_i \Delta\tau + \frac{a_i \Delta\tau/\Delta x_{i+1} + a_i \Delta\tau/\Delta x_i}{(\Delta x_{i+1} + \Delta x_i)/2} - \frac{b_i \Delta\tau}{\Delta x_i} \right] \\ &\quad + u_{i+1}^{n+1} \left[-\frac{a_i \Delta\tau/\Delta x_{i+1}}{(\Delta x_{i+1} + \Delta x_i)/2} \right] \\ \Rightarrow u_i^n &= -\alpha_i^b \Delta\tau u_{i-1}^{n+1} + \left[1 + (\alpha_i^b + \beta_i^b - c_i)\Delta\tau \right] u_i^{n+1} - \beta_i^b \Delta\tau u_{i+1}^{n+1}, \end{aligned}$$

where, as for the explicit case:

$$\begin{aligned} \alpha_i^b &= a_i / \left[\Delta x_i (\Delta x_{i+1} + \Delta x_i)/2 \right] - b_i / \left[\Delta x_i \right] \\ \beta_i^b &= a_i / \left[\Delta x_{i+1} (\Delta x_{i+1} + \Delta x_i)/2 \right]. \end{aligned}$$

5.1.7 Upstream Weighting

As pointed out above, when we use central difference approximations, we can get negative coefficients. This happens when the magnitude of b gets relatively large—this will occur for extreme values on our grid. As the interest rate gets very high, the drift will become more and more negative, and as the interest rate gets very low, the drift becomes more positive. We can avoid negative coefficients by switching to backward or forward differences. This "upstream weighting" is described in the following algorithm:

```
/* Definition of upstream weighting */
int i;
for (i = 0; i < n; i++)
{
  if (alpha_forward[i] >= 0 && beta_forward[i] >= 0)
  {
    alpha_upstream[i] = alpha_forward[i];
    beta_upstream[i] = beta_forward[i];
  }
  else
  {
    alpha_upstream[i] = alpha_backward[i];
    beta_upstream[i] = beta_backward[i];
  }
}
/* Choice of central or upstream */
for (i=0; i<n; i++)
{
  if (alpha_central[i] >= 0 && beta_central[i] >= 0)
  {
    alpha[i] = alpha_central[i];
    beta[i] = beta_central[i];
  }
  else
  {
    alpha[i] = alpha_upstream[i];
    beta[i] = beta_upstream[i];
  }
}
```

In other words, we start with central weighting at all nodes, and switch to upstream weighting at a node if a discrete coefficient at that node is negative.

5.1.8 Summary of Implicit and Explicit Schemes

From now on, we will use the coefficients α_i and β_i with the understanding the central/upstream weighting has been selected so that these coefficients are non-negative. With this in mind, we can write an explicit method in the form:

$$u_i^{n+1} = \alpha_i \Delta\tau u_{i-1}^n + \left[1 - (\alpha_i + \beta_i - c_i)\Delta\tau\right] u_i^n + \beta_i \Delta\tau u_{i+1}^n.$$

Similarly, an implicit scheme is defined as:

$$u_i^n = -\alpha_i \Delta\tau u_{i-1}^{n+1} + \left[1 + (\alpha_i + \beta_i - c_i)\Delta\tau\right] u_i^{n+1} - \beta_i \Delta\tau u_{i+1}^{n+1}.$$

We can take a weighted average of these two methods:

$$\begin{aligned}(1-\theta)[u_i^{n+1} - u_i^n] &= (1-\theta)\left\{\alpha_i \Delta\tau u_{i-1}^n + [-(\alpha_i + \beta_i - c_i)\Delta\tau] u_i^n + \beta_i \Delta\tau u_{i+1}^n\right\} \\ +\theta[u_i^{n+1} - u_i^n] \quad &+ \theta\left\{\alpha_i \Delta\tau u_{i-1}^{n+1} - [(\alpha_i + \beta_i - c_i)\Delta\tau] u_i^{n+1} \beta_i \Delta\tau_{i+1}^{n+1}\right\}.\end{aligned}$$

Choosing $\theta = 0$ gives the explicit scheme and setting $\theta = 1$ gives us the implicit scheme. A Crank-Nicolson method results from the choice of $\theta = 1/2$.

We can rewrite the above scheme in the form:

$$\begin{aligned}-\alpha_i \theta \Delta\tau u_{i-1}^{n+1} + [1 + (\alpha_i + \beta_i - c_i)\theta\Delta\tau] u_i^{n+1} - \beta_i \theta \Delta\tau u_{i+1}^{n+1} = \\ \alpha_i (1-\theta)\Delta\tau u_{i-1}^n + [1 - (\alpha_i + \beta_i - c_i)(1-\theta)\Delta\tau] u_i^n + \beta_i (1-\theta)\Delta\tau u_{i+1}^n.\end{aligned}$$

This can be written in matrix form as:

$$\left[I + \theta A \Delta\tau\right] u^{n+1} = \left[I - (1-\theta) A \Delta\tau\right] u^n, \tag{5.3}$$

where the i-th row of the tridiagonal matrix A is given by

$$[A]_i = [\ldots, -\alpha_i,\ (\alpha_i + \beta_i - c_i),\ -\beta_i, \ldots].$$

5.1.9 M-Matrices

It will turn out that a monotonicity property will be a desirable feature for our numerical scheme. In matrix form, this property is related to the concept of an M-matrix. A matrix B is an M-matrix if

- $\text{diag}(B) > 0$;
- $\text{offdiag}(B) \leq 0$;
- $\text{rowsum}(B) > 0$.

If B is an M-matrix, then

- $B^{-1} \geq 0$;
- $\text{diag}(B^{-1}) > 0$.

In other words, B^{-1} is a matrix which has nonnegative entries. Moreover the diagonal entries of B^{-1} are strictly positive.

Consider the fully implicit case where $\theta = 1$, so that

$$[I + A\Delta\tau]u^{n+1} = u^n \Rightarrow u^{n+1} = [I + A\Delta\tau]^{-1}u^n.$$

In this case, we would like $[I + A\Delta\tau]$ to be an M-matrix. The use of upstream weighting guarantees that α_i and β_i are nonnegative, and c_i is just minus one times the rate of interest (which is assumed to be nonnegative). Therefore, $[I + A\Delta t]$ is an M-matrix.

5.1.10 Boundary Conditions

The discussion thus far has focussed just on interior nodes, ignoring boundary conditions. Although the payoff will generate an initial condition u_i^0, $i = 1, \ldots, N$, we also need to consider the specification of boundary behaviour at the extreme edges of our computational domain, i.e. u_1^n and u_N^n for all time levels n. In order to incorporate the effects of boundary conditions, we should consider several cases:

Case 1: Suppose that we are given Dirichlet conditions at $x_{\min} = x_1$ and $x_{\max} = x_N$. This just involves specifying $u_1 = u_1^*$ and $u_N = u_N^*$, where u_1^* and u_N^* are given values. This means that the first and last rows of A are identically zero, so $I + A\Delta t$ will be an M-matrix.

Case 2: If $a(x) = b(x) = 0$ at x_1 then the PDE will reduce to $c(x)u = u_\tau$ at x_1. An example of this is the Black-Scholes equation for equity options. When we use an explicit scheme (for $i = 1$) we will have:

$$\begin{aligned}\frac{u_i^{n+1} - u_i^n}{\Delta\tau} &= c_i u_i^n \\ u_i^{n+1} &= u_i^n[1 + c_i\Delta\tau].\end{aligned}$$

Simlarly, for an implicit scheme (at $i = 1$) we get:

$$\begin{aligned}\frac{u_i^{n+1} - u_i^n}{\Delta\tau} &= c_i u_i^{n+1} \\ u_i^{n+1}[1 - c_i\Delta\tau] &= u_i^n\end{aligned}$$

The first row of A becomes $-c_1 \geq 0$ on the diagonal, and so $[I + A\Delta t]$ is an M-matrix.

Case 3: Assume $a(x) = c(x) = 0$, $b(x) > 0$ at x_1, so that our PDE becomes $b(x)u_x = u_\tau$ at x_1. An example of this is the CIR model. (Note that this does ignore some subtle issues about whether this boundary is attainable, see [14] for further discussion). We will use a one-sided (forward) difference for the u_x term.

With an explicit scheme (for $i = 1$) we will have:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = b_i\left[\frac{u_{i+1}^n - u_i^n}{\Delta x_{i+1}}\right],$$

so

$$\begin{aligned}u_i^{n+1} &= u_i^n\left[1 - \frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ &\quad + u_{i+1}^n\left[\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ \Rightarrow u_i^{n+1} &= [1 - \gamma_i\Delta\tau]u_i^n + \gamma_i\Delta\tau u_{i+1}^n,\end{aligned}$$

where

$$\gamma_i = \frac{b_i\Delta\tau}{\Delta x_{i+1}}.$$

Similarly, for an implicit scheme (at $i = 1$) we get:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = b_i\left[\frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x_{i+1}}\right],$$

so

$$\begin{aligned} u_i^n &= u_i^{n+1}\left[1+\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ &\quad - u_{i+1}^{n+1}\left[\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\ \Rightarrow u_i^n &= [1+\gamma_i\Delta\tau]u_i^{n+1} - \gamma_i\Delta\tau u_{i+1}^{n+1}, \end{aligned}$$

where, as for the explicit case:

$$\gamma_i = \frac{b_i\Delta\tau}{\Delta x_{i+1}}$$

The first row of A has $\gamma_1 > 0$ on the diagonal and $-\gamma_1$ on the upper off-diagonal, and so $[I + A\Delta t]$ is an M-matrix.

Case 4: Assume $a(x),\ b(x),\ c(x) \neq 0$ at x_1. The PDE remains $a(x)u_{xx} + b(x)u_x + c(x)u = u_\tau$. An example of this is the case of quadratic term structure models, where the drift $b(x_1) > 0$ and $b(x_1) \gg a(x_1)$ (recall that the drift term in these models contains a term involving $1/x$, so we cannot have $x_1 = 0$, but the first order term will get larger and larger as $x_1 \to 0$). We can use one-sided difference approximations, as follows. The nodes to be used are $u_i,\ u_{i+1},\ u_{i+2}$ (with $i = 1$), and the grid spacing is $\Delta x_{i+1} = x_{i+1} - x_i$, $\Delta x_{i+2} = x_{i+2} - x_{i+1}$. We can use Taylor's series expansions to write:

$$\begin{aligned} u_{i+1} &= u_i + (u_x)_i\Delta x_{i+1} + (u_{xx})_i\frac{\Delta x_{i+1}^2}{2} + \cdots \\ u_{i+2} &= u_i + (u_x)_i\left[\Delta x_{i+1} + \Delta x_{i+2}\right] + (u_{xx})_i\frac{\left[\Delta x_{i+1} + \Delta x_{i+2}\right]^2}{2} + \cdots . \end{aligned}$$

We can just use a forward difference to approximate the first derivative:

$$(u_x)_i = \frac{u_{i+1} - u_i}{\Delta x_{i+1}}.$$

To approximate the second derivative, rewrite the equations above (ignoring higher order terms) as:

$$\begin{aligned} \frac{u_{i+1} - u_i}{\Delta x_{i+1}} &= (u_x)_i + (u_{xx})_i\frac{\Delta x_{i+1}}{2} \\ \frac{u_{i+2} - u_i}{\Delta x_{i+1} + \Delta x_{i+2}} &= (u_x)_i + (u_{xx})_i\frac{\left[\Delta x_{i+1} + \Delta x_{i+2}\right]}{2}. \end{aligned}$$

Subtract the first equation from the second to get:

$$\begin{aligned}\frac{u_{i+2}-u_i}{\Delta x_{i+2}+\Delta x_{i+1}}-\frac{u_{i+1}-u_i}{\Delta x_{i+1}} &= (u_{xx})_i\left[\frac{\Delta x_{i+1}+\Delta x_{i+2}}{2}-\frac{\Delta x_{i+1}}{2}\right]\\ &= (u_{xx})_i\left[\frac{\Delta x_{i+2}}{2}\right].\end{aligned}$$

This gives the second derivative as:

$$(u_{xx})_i=\frac{u_{i+2}-u_i}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}-\frac{u_{i+1}-u_i}{\Delta x_{i+2}\Delta x_{i+1}/2}.$$

For an explicit scheme we have:

$$\frac{u_i^{n+1}-u_i^n}{\Delta\tau}=a_i\left[\frac{u_{i+2}^n-u_i^n}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}-\frac{u_{i+1}^n-u_i^n}{\Delta x_{i+2}\Delta x_{i+1}/2}\right]+b_i\left[\frac{u_{i+1}^n-u_i^n}{\Delta x_{i+1}}\right]+c_iu_i^n,$$

so

$$\begin{aligned}u_i^{n+1} &= u_i^n\left[1+c_i\Delta\tau-\frac{a_i\Delta\tau}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}+\frac{a_i\Delta\tau}{\Delta x_{i+2}\Delta x_{i+1}/2}-\frac{b_i\Delta\tau}{\Delta x_{i+1}}\right]\\ &\quad+u_{i+1}^n\left[\frac{b_i\Delta\tau}{\Delta x_{i+1}}-\frac{a_i\Delta\tau}{\Delta x_{i+2}\Delta x_{i+1}/2}\right]\\ &\quad+u_{i+2}^n\left[\frac{a_i\Delta\tau}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}\right]\\ \Rightarrow u_i^{n+1} &= u_i^n\left[1-(\gamma_i+\delta_i-c_i)\Delta\tau\right]+\gamma_i\Delta\tau u_{i+1}^n+\delta_i\Delta\tau u_{i+2}^n,\end{aligned}$$

where

$$\begin{aligned}\gamma_i &= \frac{b_i}{\Delta x_{i+1}}-\frac{a_i}{\Delta x_{i+2}\Delta x_{i+1}/2}\\ \delta_i &= \frac{a_i}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}.\end{aligned}$$

Similarly, for an implicit scheme:

$$\frac{u_i^{n+1}-u_i^n}{\Delta\tau}=a_i\left[\frac{u_{i+2}^{n+1}-u_i^{n+1}}{\Delta x_{i+2}\left[\Delta x_{i+1}+\Delta x_{i+2}\right]/2}-\frac{u_{i+1}^{n+1}-u_i^{n+1}}{\Delta x_{i+2}\Delta x_{i+1}/2}\right]+b_i\left[\frac{u_{i+1}^{n+1}-u_i^{n+1}}{\Delta x_{i+1}}\right]+c_iu_i^{n+1},$$

so

$$
\begin{aligned}
u_i^n &= u_i^{n+1}\left[1 - c_i\Delta\tau + \frac{a_i\Delta t}{\Delta x_{i+2}\left[\Delta x_{i+1} + \Delta x_{i+2}\right]/2} - \frac{a_i\Delta\tau}{\Delta x_{i+2}\Delta x_{i+1}/2} + \frac{b_i\Delta\tau}{\Delta x_{i+1}}\right] \\
&\quad + u_{i+1}^{n+1}\left[-\frac{b_i\Delta\tau}{\Delta x_{i+1}} + \frac{a_i\Delta\tau}{\Delta x_{i+2}\Delta x_{i+1}/2}\right] \\
&\quad + u_{i+2}^{n+2}\left[-\frac{a_i\Delta\tau}{\Delta x_{i+2}\left[\Delta x_{i+1} + \Delta x_{i+2}\right]/2}\right] \\
\Rightarrow u_i^n &= u_i^{n+1}\left[1 + (\gamma_i + \delta_i - c_i)\Delta\tau\right] - \gamma_i\Delta\tau u_{i+1}^n - \delta_i\Delta\tau u_{i+2}^n
\end{aligned}
$$

where, as for the explicit case:

$$
\begin{aligned}
\gamma_i &= \frac{b_i}{\Delta x_{i+1}} - \frac{a_i}{\Delta x_{i+2}\Delta x_{i+1}/2} \\
\delta_i &= \frac{a_i}{\Delta x_{i+2}\left[\Delta x_{i+1} + \Delta x_{i+2}\right]/2}.
\end{aligned}
$$

The first row of A has $\gamma_1 + \delta_1 - c_1$ on the main diagonal, and all of these terms are positive since $b_1 \gg a_1,\ \gamma_1 > 0$.[1] The first term to the right of the diagonal is $-\gamma_1$ and the second one is $-\delta_1$, so $[I + A\Delta\tau]$ is an M-matrix. However, it is no longer tridiagonal.

If desired, tridiagonality can be restored as follows. The first two equations of the linear system for the general Crank-Nicolson scheme are:

$$
\begin{aligned}
\left[1 + (\gamma_1 + \delta_1 - c_1)\theta\Delta\tau\right] u_1^{n+1} - \gamma_1\theta\Delta\tau u_2^{n+1} - \delta_1\theta\Delta\tau u_3^{n+1} &= y_1 \\
-\alpha_2\theta\Delta\tau u_1^{n+1} + \left[1 + (\alpha_2 + \beta_2 - c_2)\theta\Delta\tau\right] u_2^{n+1} - \beta_2\theta\Delta\tau u_3^{n+1} &= y_2,
\end{aligned}
$$

where

$$
\begin{aligned}
y_1 &= \left[1 - (\gamma_1 + \delta_1 - c_1)(1-\theta)\Delta\tau\right] u_1^n + \gamma_1(1-\theta)\Delta\tau u_2^n + \delta_1(1-\theta)\Delta\tau u_3^n \\
y_2 &= \alpha_2(1-\theta)\Delta\tau u_1^n + \left[1 - (\alpha_2 + \beta_2 - c_2)(1-\theta)\Delta\tau\right] u_2^n + \beta_2(1-\theta)\Delta\tau u_3^n.
\end{aligned}
$$

[1] More accurately, the terms might not all be positive, because γ_i depends on the grid spacing. However, if necessary we can change the grid spacing so that the terms are positive.

Take the first equation and subtract δ_1/β_2 times the second equation to get:

$$\left[1+(\gamma_1+\delta_1-c_1)\theta\Delta\tau+\frac{\delta_1}{\beta_2}\alpha_2\theta\Delta\tau\right]u_1^{n+1} + \left[-\gamma_1\theta\Delta t-\frac{\delta_1}{\beta_2}\left(1+(\alpha_2+\beta_2-c_2)\theta\Delta t\right)\right]u_2^{n+1} + \underbrace{\left[-\delta_1\theta\Delta t-\frac{\delta_1}{\beta_2}\beta_2\theta\Delta t\right]}_{0}u_3^{n+1}=y_1-\frac{\delta_1}{\beta_2}y_2.$$

Case 5: $a(x),\ b(x),\ c(x)\neq 0$ at $x_N=x_{\max}$. This can be used in mean-reverting cases (CIR, quadratic term structure) when $b(x_N)<0$ and $-b(x_N)\gg a(x_N)$ (so that the PDE is essentially hyperbolic). We can use one-sided difference approximations, as follows. The nodes to be used are u_i, u_{i-1}, u_{i-2} (with $i=N$), and the grid spacing is $\Delta x_i=x_i-x_{i-1}$, $\Delta x_{i-1}=x_{i-1}-x_{i-2}$. Taylor's series expansions give:

$$u_{i-1}=u_i-(u_x)_i\Delta x_i+(u_{xx})_i\frac{\Delta x_i^2}{2}+\cdots$$
$$u_{i-2}=u_i-(u_x)_i\left[\Delta x_i+\Delta x_{i-1}\right]+(u_{xx})_i\frac{\left[\Delta x_i+\Delta x_{i-1}\right]^2}{2}+\cdots.$$

To approximate the first derivative, use a backward difference:

$$(u_x)_i=\frac{u_i-u_{i-1}}{\Delta x_i}.$$

For the second derivative, rewrite the equations above (ignoring higher order terms) as:

$$\frac{u_i-u_{i-1}}{\Delta x_i}=(u_x)_i-(u_{xx})_i\frac{\Delta x_i}{2}$$
$$\frac{u_i-u_{i-2}}{\Delta x_i+\Delta x_{i-1}}=(u_x)_i-(u_{xx})_i\frac{\left[\Delta x_i+\Delta x_{i-1}\right]}{2}$$

Subtract the first equation from the second:

$$\frac{u_i-u_{i-2}}{\Delta x_i+\Delta x_{i-1}}-\frac{u_i-u_{i-1}}{\Delta x_i}=-(u_{xx})_i\left[\frac{\Delta x_i+\Delta x_{i-1}}{2}-\frac{\Delta x_i}{2}\right]$$
$$=-(u_{xx})_i\left[\frac{\Delta x_{i-1}}{2}\right].$$

This gives an approximation to the second derivative as:

$$(u_{xx})_i = \frac{u_{i-2} - u_i}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} - \frac{u_{i-1} - u_i}{\Delta x_{i-1}\Delta x_i/2}.$$

For an explicit scheme we will have:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = a_i\left[\frac{u_{i-2}^n - u_i^n}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} - \frac{u_{i-1}^n - u_i^n}{\Delta x_{i-1}\Delta x_i/2}\right] + b_i\left[\frac{u_i^n - u_{i-1}^n}{\Delta x_i}\right] + c_i u_i^n,$$

so

$$\begin{aligned} u_i^{n+1} &= u_{i-2}^n\left[\frac{a_i\Delta\tau}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2}\right] \\ &\quad + u_{i-1}^n\left[\frac{-b_i\Delta\tau}{\Delta x_i} - \frac{a_i\Delta\tau}{\Delta x_{i-1}\Delta x_i/2}\right] \\ &\quad + u_i^n\left[1 + c_i\Delta\tau - \frac{a_i\Delta\tau}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} + \frac{a_i\Delta\tau}{\Delta x_{i-1}\Delta x_i/2} + \frac{b_i\Delta t}{\Delta x_i}\right] \\ \Rightarrow u_i^{n+1} &= \delta_i\Delta\tau u_{i-2}^n + \gamma_i\Delta\tau u_{i-1}^n + u_i^n\left[1 - (\gamma_i + \delta_i - c_i)\Delta\tau\right], \end{aligned}$$

where

$$\begin{aligned} \delta_i &= \frac{a_i\Delta\tau}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} \\ \gamma_i &= \frac{-b_i\Delta t}{\Delta x_i} - \frac{a_i\Delta\tau}{\Delta x_{i-1}\Delta x_i/2}. \end{aligned}$$

Similarly, for an implicit scheme we get:

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = a_i\left[\frac{u_{i-2}^{n+1} - u_i^{n+1}}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} - \frac{u_{i-1}^{n+1} - u_i^{n+1}}{\Delta x_{i-1}\Delta x_i/2}\right] + b_i\left[\frac{u_i^{n+1} - u_{i-1}^{n+1}}{\Delta x_i}\right] + c_i u_i^{n+1},$$

so

$$\begin{aligned} u_i^n &= u_{i-2}^{n+1}\left[-\frac{a_i\Delta\tau}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2}\right] \\ &\quad + u_{i-1}^{n+1}\left[\frac{b_i\Delta\tau}{\Delta x_i} + \frac{a_i\Delta\tau}{\Delta x_{i-1}\Delta x_i/2}\right] \\ &\quad + u_i^{n+1}\left[1 - c_i\Delta\tau + \frac{a_i\Delta\tau}{\Delta x_{i-1}\left[\Delta x_i + \Delta x_{i-1}\right]/2} - \frac{a_i\Delta\tau}{\Delta x_{i-1}\Delta x_i/2} - \frac{b_i\Delta\tau}{\Delta x_i}\right] \\ \Rightarrow u_i^n &= -\delta_i\Delta\tau u_{i-2}^{n+1} - \gamma_i\Delta\tau u_{i-1}^{n+1} + u_i^{n+1}\left[1 + (\gamma_i + \delta_i - c_i)\Delta\tau\right], \end{aligned}$$

where, as for the explicit case:

$$\gamma_i = \frac{-b_i \Delta\tau}{\Delta x_i} - \frac{a_i \Delta\tau}{\Delta x_{i-1} \Delta x_i / 2}$$
$$\delta_i = \frac{a_i \Delta\tau}{\Delta x_{i-1} \left[\Delta x_i + \Delta x_{i-1}\right] / 2}.$$

The last row of A has $\gamma_N + \delta_N - c_N$ on the main diagonal, and all of these terms are positive since $b_N < 0$ and $-b_N \gg a_N$.[2] The term to the left is $-\gamma_N$ and the second one is $-\delta_N$, so $I + A\Delta t$ is an M-matrix. However, it is no longer tridiagonal.

If desired, tridiagonality can be restored as follows. The last two equations of the linear system (for the general Crank-Nicolson scheme) are:

$$\begin{aligned} -\alpha_{N-1}\theta\Delta\tau u_{N-2}^{n+1} + \left[1 + (\alpha_{N-1} + \beta_{N-1} - c_{N-1})\theta\Delta\tau\right] u_{N-1}^{n+1} - \beta_{N-1}\theta\Delta\tau u_N^{n+1} &= y_{N-1} \\ -\delta_N\theta\Delta\tau u_{N-2}^{n+1} - \gamma_N\theta\Delta\tau u_{N-1}^{n+1} + \left[1 + (\gamma_N + \delta_N - c_N)\theta\Delta\tau\right] u_N^{n+1} &= y_N, \end{aligned}$$

where

$$\begin{aligned} y_{N-1} &= \alpha_{N-1}(1-\theta)\Delta\tau u_{N-2}^n + \left[1 - (\alpha_{N-1} + \beta_{N-1} - c_{N-1})(1-\theta)\Delta\tau\right] u_{N-1}^n \\ &\quad + \beta_{N-1}(1-\theta)\Delta\tau u_N^n \\ y_N &= \delta_N(1-\theta)\Delta\tau u_{N-2}^n + \gamma_N(1-\theta)\Delta\tau u_{N-1}^n + \left[1 - (\gamma_N + \delta_N - c_N)(1-\theta)\Delta\tau\right] u_N^n. \end{aligned}$$

Subtract δ_N/α_{N-1} times the first equation from the second equation to get:

$$\underbrace{\left[-\delta_N\theta\Delta\tau + \frac{\delta_N}{\alpha_{N-1}}\alpha_{N-1}\theta\Delta\tau\right]}_{0} u_{N-2}^{n+1} +$$
$$\left[-\gamma_N\theta\Delta\tau - \frac{\delta_N}{\alpha_{N-1}}\left(1 + (\alpha_{N-1} + \beta_{N-1} - c_{N-1})\theta\Delta\tau\right)\right] u_{N-1}^{n+1} +$$
$$\left[1 + (\gamma_N + \delta_N - c_N)\theta\Delta\tau + \frac{\delta_N}{\alpha_{N-1}}\beta_{N-1}\theta\Delta\tau\right] u_N^{n+1} = y_N - \frac{\delta_N}{\alpha_{N-1}} y_{N-1}.$$

[2] As with the previous case, it is more accurate to note that the terms might not all be positive because γ_i depends on the grid spacing. If necessary, however, the spacing can be changed so that all of the terms are positive.

5.1.11 Stability and Monotonicity Considerations

Much of the discussion above has focussed on coefficients being positive, M-matrices, etc. Here we give an informal discussion as to why this matters. For simplicity, we will ignore boundary conditions here. Recall that an explicit method can be written as

$$u_i^{n+1} = \alpha_i \Delta\tau u_{i-1}^n + [1 - (\alpha_i + \beta_i - c_i)\Delta\tau)]u_i^n + \beta_i \Delta\tau u_{i+1}^n.$$

Assuming that we have used upstream weighting, then $\alpha_i \geq 0$ and $\beta_i \geq 0$. If we make the further assumption that $\Delta\tau$ is small enough that $1 - (\alpha_i + \beta_i - c_i)\Delta\tau \geq 0$, then

$$\left|u_i^{n+1}\right| \leq \alpha_i \Delta\tau \left|u_{i-n}^n\right| + [1 - (\alpha_i + \beta_i - c_i)] \left|u_n^n\right| + \beta_i \Delta\tau \left|u_{i+1}^n\right|.$$

Let $\|u^n\|$ be the maximum value of u_i^n across all nodes i, and substitute in the above to get

$$\left|u_i^{n+1}\right| \leq \|u^n\| (1 + c_i)\Delta\tau \leq \|u^n\|.$$

Since this will hold for any i, we have

$$\left\|u^{n+1}\right\| \leq \|u^n\|.$$

Assuming that the initial condition u^0 and boundary conditions are bounded, this means that our method will be stable, so that small errors will not become very large over time. Note that this requires that $\Delta\tau$ be sufficiently small.

Now consider a fully implicit method

$$u_i^{n+1}[1 + (\alpha_i + \beta_i - c_i)\Delta\tau] = \alpha_i \Delta\tau u_{i-1}^{n+1} + u_i^n + \beta_i \Delta\tau u_{i+1}^n,$$

so

$$\begin{aligned}\left|u_i^{n+1}\right| [1 + (\alpha_i + \beta_i - c_i)] &\leq \alpha_i \Delta\tau \left|u_{i-1}^n\right| + \left|u_i^n\right| + \beta_i \Delta\tau \left|u_{i+1}^n\right| \\ &\leq \|u^n\| + (\alpha_i + \beta_i)\Delta\tau \left\|u^{n+1}\right\|.\end{aligned}$$

Let j be the index for which $\left|u_i^{n+1}\right|$ attains its maximum value. Then

$$\begin{aligned}\left\|u^{n+1}\right\| [1 + (\alpha_j + \beta_j - c_j)\Delta\tau] &\leq \|u^n\| + (\alpha_j + \beta_j)\Delta\tau \left\|u^{n+1}\right\| \\ \Rightarrow \left\|u^{n+1}\right\| (1 - c_j\Delta\tau) &\leq \|u^n\| \\ \Rightarrow \left\|u^{n+1}\right\| &\leq \frac{\|u^n\|}{1 - c_j\Delta\tau} \leq \|u^n\|,\end{aligned}$$

showing that a fully implicit method is unconditionally stable (i.e. there is no time step size restriction). More sophisticated methods can be used to show that Crank-Nicolson schemes are also unconditionally stable.

Moreover, we can write an implicit method in the form

$$u_i^{n+1} = \frac{\alpha_i \Delta\tau u_{i-1}^{n+1} + u_i^n + \beta_i \Delta\tau u_{i+1}^{n+1}}{1 + (\alpha_i + \beta_i - c_i)\Delta\tau}.$$

The coefficients multiplying the terms on the right hand side are all non-negative and add up to a number that is at most one (assuming $r = -c \geq 0$). We will refer to this property as being a "positive coefficient discretization". This is important because it implies that the values of u_i^{n+1} are bounded by the maximum and minimum values of the nodes on either side of i at the new time step and the value of u_i^n. This means that our numerical solution will be free of any spurious oscillations, which is a very desirable feature. In matrix form, the counterpart to a positive coefficient scheme is an M-matrix. It is important to note that an explicit scheme will also be a positive coefficient scheme, provided that the time step stability limitation is satisfied. However, because Crank-Nicolson is an average of explicit and implicit schemes, it is not a positive coefficient discretization unless the time step size is less than double the maximum stable explicit time step size.

In the following, we will ignore explicit methods due to the time step size stability restriction. However, there are advantages and disadvantages to Crank-Nicolson and fully implicit, so we will consider both of them. In particular, Crank-Nicolson methods are unconditionally stable and second order accurate in time, but not positive coefficient, whereas fully implicit schemes are unconditionally stable and positive coefficient, but only first order accurate in time. We want to use Crank-Nicolson to exploit its higher convergence rate, but there are times when this can be costly. The fact that it is not positive coefficient means that spurious oscillations can form in our numerical solution, making it less accurate and slowing down our rate of convergence.

5.2 Numerical Solution in the Presence of Jumps

Remember that as we showed in Chapter 4, the bond pricing equation (4.11) in the presence of jumps has the following form:

$$F_t^T + \frac{1}{2}\sigma^2 F_{rr}^T + (\mu - \lambda_W \sigma)F_r^T - (r + \lambda^*)F^T + \lambda^* \int_0^\infty p\{j(t)\}F^T(j(t)r, t)dj(t) = 0. \quad (5.4)$$

All terms which do not involve the integral will to be handled according to the methods described for the no-jump case in Section 5.1. The methods used here for the jump integral closely follow those developed in [13]. In order to discretize the new integral term in equation 5.4, we will transform it into a correlation integral. This allows us to use computationally efficient Fast Fourier Transform (FFT) techniques. Let

$$I(r) = \int_0^\infty p\{j(t)\}F^T(j(t)r, t)dj(t),$$

or, more compactly,

$$I(r) = \int_0^\infty p\{j\}F^T(jr)dj.$$

Using the change of variables

$$y = \log(jr), \quad j = \frac{\exp(y)}{r}, \quad \text{and } dj = \frac{\exp(y)}{r}dy, \tag{5.5}$$

we obtain:

$$I = \int_{-\infty}^\infty F^T(\exp(y))p\left(\frac{\exp(y)}{r}\right)\frac{\exp(y)}{r}dy. \tag{5.6}$$

Now let $f(u) = p(u)u$. Hence equation (5.6) becomes:

$$I = \int_{-\infty}^\infty F^T(\exp(y))f\left(\frac{\exp(y)}{r}\right)dy. \tag{5.7}$$

Let $x = \log(r)$, $\bar{F}^T(x, \tau) = F^T(\exp(x), \tau) = F^T(r, \tau)$, and $\bar{f}(\log y) = f(y)$. By substituting into equation (5.7), we have:

$$\begin{aligned} I &= \int_{-\infty}^\infty \bar{F}^T(y)\bar{f}(y - \log(r))dy \\ &= \int_{-\infty}^\infty \bar{F}^T(y)\bar{f}(y - x)dy \\ &= \int_{-\infty}^\infty \bar{F}^T(x + y)\bar{f}(y)dy. \end{aligned} \tag{5.8}$$

Note that $\bar{f}(x)$ is the probability density function of a jump of size $x = \log(j)$. As shown in [33], equation 5.8 corresponds to the correlation product $\bar{\otimes}$ of $\bar{F}(y)$ and $\bar{f}(y)$. This can be written as:

$$I = \bar{F}\bar{\otimes}\bar{f} \tag{5.9}$$

Recall that the correlation integral is defined as

$$I(x) = \int_{-\infty}^{\infty} \bar{F}(x+y)\bar{f}(y)dy. \tag{5.10}$$

Now let

$$x_i = x_{\min} + i\Delta x$$
$$\Delta x = \frac{x_{\max} - x_{\min}}{N-1},$$

where $x_{\min} = \log(r_{\min})$ and $x_{\max} = \log(r_{\max})$. $r_{\max}$ is the largest value of r on the grid, while $r_{\min}$ is a point selected near $r = 0$. Let

$$I_i = I(x_{\min} + i\Delta x) \quad i = 0, \ldots, N-1$$
$$\bar{F}_i = \bar{F}(x_{\min} + i\Delta x) \quad i = 0, \ldots, N-1,$$

and

$$\bar{f}_j = \int_{x_j - \Delta x/2}^{x_j + \Delta x/2} \bar{f}(x)dx$$
$$x_j = j\Delta x, \qquad j = -N/2+1, \ldots, N/2. \tag{5.11}$$

The correlation integral is then approximated by:

$$I_i = \sum_{j=-N/2+1}^{j=N/2} \bar{F}_{i+j}\bar{f}_j\Delta y. \tag{5.12}$$

where $\Delta y = \Delta x$. Note that in equation (5.12), there is an approximation error of order $(\Delta y)^2$. Moreover, we have assumed that the number of points N on this log-spaced grid is large enough so that $f[\pm(N/2)\Delta y] \simeq 0$. The next step is to write equation (5.12) in a way that is suitable for evaluation using FFT methods. Begin by breaking it up into two parts:

$$I_i = \sum_{j=-N/2+1}^{-1} \bar{F}_{i+j}\bar{f}_j\Delta y + \sum_{j=0}^{N/2} \bar{F}_{i+j}\bar{f}_j\Delta y. \tag{5.13}$$

Concentrating on the sum over values where $j < 0$, we can write:

$$\sum_{j=-N/2+1}^{-1} \bar{F}_{i+j}\bar{f}_j\Delta y = \sum_{k=N/2+1}^{N-1} \bar{F}_{i+k-N}\bar{f}_{k-N}\Delta y \tag{5.14}$$

by letting $k = j + N$. FFT methods effectively assume that the input data is periodic. We therefore need to consider a periodic extension to $\bar{F}_j$:

$$\bar{F}_{j-N} = \bar{F}_{j+N} = \bar{F}_j. \tag{5.15}$$

We also need to define

$$\bar{f}' = \begin{cases} \bar{f}_i & \text{for } i = 0, \ldots, N/2 \\ \bar{f}_{i-N} & \text{for } i = N/2+1, \ldots N-1 \end{cases}. \tag{5.16}$$

Then we can express equation (5.14) in the form:

$$\sum_{k=N/2+1}^{N-1} \bar{F}_{i+k-N} \bar{f}_{k-N} \Delta y = \sum_{j=N/2+1}^{N-1} \bar{F}_{j+k} \bar{f}'_j \Delta y. \tag{5.17}$$

This means that by combining (5.14) and (5.17) we have:

$$\sum_{j=N/2+1}^{-1} \bar{F}_{i+j} \bar{f}_j \Delta y = \sum_{j=N/2+1}^{N-1} \bar{F}_{i+j} \bar{f}'_j \Delta y, \tag{5.18}$$

and equation (5.12) can be written in the form:

$$I_i = \sum_{j=0}^{N-1} \bar{F}_{i+j} \bar{f}'_j \Delta y, \tag{5.19}$$

which is in the standard form for use in an FFT routine.

Three issues need to be considered at this point, however. First, the input data is not periodic, but we have assumed that it is. This can cause "wrap-around pollution". Following [13], this issue can be addressed by defining three regions for our original grid in r-space. For large values of r, we extend the grid past our original value $r_{\max}$ to some new value r_{extended}. Recall also that we have selected some point $r_{\min}$ as a point near $r = 0$. Values near the extremes of the grid (i.e. above $r_{\max}$ and below $r_{\min}$ will be most affected by wrap-around. In these regions, we will assume that the jump intensity $\lambda^* = 0$, and discard any computed values of the integral. In other words:

- In the region below $r_{\min}$, suppress jumps by setting $\lambda^* = 0$.
- In the region above $r_{\max}$, suppress jumps by setting $\lambda^* = 0$.

- In the region, between $r_{\min}$ and $r_{\max}$, use the FFT approach described here to evaluate the jump integral.

The second issue to be addressed is the fact that we have two different grids: an unevenly-spaced grid for r, and an evenly-spaced grid for $\log(r)$. The r grid is efficient for the parts of equation (5.4) that do not involve jumps, but the FFT requires an evenly spaced (log) grid. Again following [13], we simply use interpolation as required to go back and forth between these two grids.

The third issue is that, as noted in [13], using an implicit method involves a nonlinear iterative scheme. A detailed discussion of the stability and convergence of this type of approach to handling the jump integral term can be found in [13].

Chapter 6

Numerical Tests

This chapter provides a series of tests to evaluate the performance of the numerical algorithms described above in Chapter 5. Our focus is on two simple types of contracts: pure discount (i.e. zero coupon) bonds and European call options on these bonds. From these call options, European put option values can easily be obtained using put-call parity. Moreover, these contracts serve as basic building blocks for a wide variety of interest rate derivative contracts [21]. For example, coupon-paying bonds are just portfolios of pure discount bonds. Interest rate caps and collars are portfolios of European bond options. Interest rate swaps effectively involve the exchange of fixed rate coupon bonds for floating rate bonds. (Floating rate bonds have coupon rates that are reset periodically so that the bonds are worth their par values, so these are easy to value.) Options to enter swaps ("swaptions") are just options to exchange fixed and floating rate bonds.

The tests will proceed as follows. First, we consider three models in the absence of jumps in the context of pure discount bonds. Following this, we will examine the pricing of European call options. The models used are the CIR [11], CKLS [8], and nonlinear drift (QTS) [3] models. The evolution of the short rate under each of these models is shown in Table 6.1, along with the values of the input parameters. In order to provide a fair basis for comparison of the pricing results across the models, the input parameters chosen are approximately equal to those estimated for each of the models in [3], using monthly U.S. Federal Funds interest rate data from 1963 to 1998. The reason for this choice is that this is one of the only studies available which has estimated the parameters of each of these models using the same data set. Note that we will not be using different values of γ, rather we just keep the value of 0.5 in the CIR case and 1.5 in the other two models, as in [3]. Also note that the parameters shown were estimated under the P measure. Recall from Chapter 3 that our pricing PDE includes an adjustment from the P measure to the Q

Model	Short Rate Evolution (P measure)	Parameters
CIR	$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dW(t)$	$\alpha = .014$, $\beta = .20$, $\sigma = .065$, $\gamma = .5$
CKLS	$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dW(t)$	$\alpha = .0085$, $\beta = .10$, $\sigma = .80$, $\gamma = 1.5$
QTS	$dr = (a_{-1}/r + a_0 + a_1 r + a_2 r^2)dt$ $+ \sigma r^{\gamma} dW(t)$	$a_{-1} = .001$, $a_0 = -.035$, $a_1 = .70$, $a_2 = -4.00$, $\sigma = .80$, $\gamma = 1.50$

Table 6.1: Short rate evolution and input parameters for the various models.

measure which involves adding a term involving the market price of risk λ_W. This will end up changing the effective drift of the different models. For simplicity, we will set $\lambda_W = 0$ throughout. For each test, we start with an initial grid for r and a specified number of time steps per year. We then examine the convergence of our calculated values as the initial r-grid is refined by inserting new nodes midway between existing nodes and the number of time steps per year is doubled (i.e. the time step size is divided by two). Having examined bond prices and bond option prices in the no-jump context, we then repeat the numerical tests for bond and bond-option prices in the jump-diffusion context.

6.1 Bond Prices Without Jumps

For the numerical tests presented here, we used the following unevenly-spaced grid with 43 nodes for the short rate r:

0.0, .001, .01, .02, .03, .035, .04, .045,
.05, .052, .054, .056, .058, .06, .062, .064,
.066, .068, .07, .072, .074, .076, .078, .08,
.084, .088, .092, .096, .10, .105, .11, .115,
.12, .13, .14, .16, .18, .20, .24, .28, .35, .50, .75

Note that, for the QTS model, due to the a_{-1}/r term we replaced the first grid point of 0 by 0.0001. We will report calculated bond prices for three different current levels of the short rate ($r = .04$, $r = .07$, and $r = .10$), and three different bond maturities (1 year, 5 years, and 10 years).

We begin by considering the CIR model. In this case, we have an analytic solution available, as indicated in equation (3.18). Table 6.2 reports the analytic values for the cases to be considered here. In other cases, such analytic solutions will not be available,

r	1-year maturity	5-year maturity	10-year maturity
4%	0.958118	0.776373	0.571187
7%	0.932433	0.706853	0.503394
10%	0.907438	0.643557	0.443648

Table 6.2: Analytic bond prices for the CIR model.

	$r = .04$		$r = .07$		$r = .10$	
Maturity	Value	Standard error	Value	Standard error	Value	Standard error
1 year	.958124	.000021	.932443	.000027	.907450	.000031
5 years	.776331	.000155	.706822	.000173	.643538	.000181
10 years	.571276	.000252	.503491	.000256	.443748	.000252

Table 6.3: Monte Carlo simulation of bond prices for the CIR model. The values reported are based on 100,000 simulations with 100 time steps per year.

so as a further check we calculate bond prices using Monte Carlo simulation. We report these values for the CIR model in Table 6.3. These prices are calculated using 100,000 simulation runs with 100 time steps per year. A comparison of Tables 6.2 and 6.3 shows that the Monte Carlo numbers are quite accurate, closely corresponding to the analytic solutions with small standard errors. Based on this, our Monte Carlo results for other cases reported below will use the same number of simulation runs and time steps as here.

Bond values calculated using our PDE algorithm for the fully implicit and Crank-Nicolson cases are reported in Table 6.4. The first two columns of the table report the number of grid nodes and time steps per year. The next three columns contain the calculated bond value, its change (the absolute value of the difference between the value at the current grid refinement and the previous grid refinement), and its ratio (the ratio of successive changes) for the case of $r = 4\%$. Note that the change is calculated using more digits than is reported in the table for the value. The remaining columns of the table repeat this for the cases of $r = 7\%$ and $r = 10\%$. To interpret the ratio numbers, note that with first order (i.e. linear) convergence, we should expect to see a ratio of around 2, since doubling the number of time-steps and grid points reduces our computational error by a factor of two. With second order (i.e. quadratic) convergence, doubling the number of time-steps and grid points will reduce the error by a factor of four, so the ratio should be 4. As can be seen from Table 6.4, the reported values are consistent with the expected first

order convergence for the fully implicit method. We also observe approximate quadratic convergence for the Crank-Nicolson method in all cases. A comparison with the analytic values from Table 6.2 shows that our PDE algorithm results are quite accurate. Figure 6.1 plots the PDE solution surface over the entire ten year horizon and across the entire grid of interest rate values.[1] As to be expected, the solution starts off at a value of 1 for all values of r and declines throughout the entire 10-year period, with the rate of decline being steepest for higher values of r. Note that the imposition of boundary conditions using one-sided derivatives at the extreme high and low values of r has not caused any apparent problems.

We next consider the CKLS model. In this case, a possible concern arises because our one-sided boundary condition when r gets large assumes that the drift term is much larger than the volatility term. Because the drift is linear in r, whereas the volatility involves $r^{3/2}$, the assumption we have made is violated. Figure 6.2 plots the solution surface over the 10-year horizon and shows that the imposition of the one-sided boundary condition in this case has caused problems. For long enough maturities, the bond price increases with the short rate, which does not make sense.

[1]This plot and all subsequent plots are based on the fully implicit approach. This choice was made mainly because fully implicit is less susceptible to computational problems than Crank-Nicolson since it is a positive coefficient scheme (in general) whereas Crank-Nicolson is not. This implies that if we see problems with a fully implicit scheme, they will also be present (and much worse) in a Crank-Nicolson method. However, problems observed with a Crank-Nicolson approach may not show up in a fully implicit method. By plotting the fully implicit solution, we therefore have a good sense of whether or not we have a scheme available which works well across our entire grid.

Nodes	Time steps per year	$r = .04$ Value	Change	Ratio	$r = .07$ Value	Change	Ratio	$r = .10$ Value	Change	Ratio
1-year maturity bond, fully implicit method										
43	50	.957992			.932670			.908069		
85	100	.958053	.00006180		.932552	.00011824		.907755	.00031412	
169	200	.958085	.00003172	1.94	.932492	.00005929	1.99	.907597	.00015862	1.98
337	400	.958101	.00001607	1.97	.932463	.00002967	1.99	.907517	.00007971	1.99
1-year maturity bond, Crank-Nicolson method										
43	50	.958094			.932481			.907562		
85	100	.958112	.00001800		.932445	.00003543		.907469	.00009316	
169	200	.958116	.00000458	3.93	.932436	.00000887	3.99	.907445	.00002343	3.97
337	400	.958118	.00000116	3.96	.932434	.00000222	3.99	.907439	.00000588	3.98
5-year maturity bond, fully implicit method										
43	50	.776640			.707733			.644967		
85	100	.776508	.00013120		.707295	.00043876		.644266	.00070112	
169	200	.776441	.00006723	1.95	.707074	.00022073	1.98	.643912	.00035362	1.98
337	400	.776407	.00003403	1.97	.706963	.00011000	1.99	.643735	.00017746	1.99
5-year maturity bond, Crank-Nicolson method										
43	50	.776376			.706884			.643604		
85	100	.776374	.00000172		.706861	.00002347		.643569	.00003482	
169	200	.776373	.00000038	4.55	.706855	.00000588	3.99	.643560	.00000875	3.97
337	400	.776373	.00000009	4.15	.706853	.00000147	3.99	.643558	.00000219	3.98
10-year maturity bond, fully implicit method										
43	50	.572128			.504555			.444951		
85	100	.571666	.00046135		.503979	.00057620		.444305	.00064607	
169	200	.571429	.00023766	1.94	.503688	.00029156	1.97	.443977	.00032713	1.97
337	400	.571308	.00012064	1.96	.503541	.00014666	1.98	.443813	.00016461	1.98
10-year maturity bond, Crank-Nicolson method										
43	50	.571172			.503401			.443652		
85	100	.571183	.00001060		.503396	.00000535		.443649	.00000307	
169	200	.571186	.00000277	3.82	.503395	.00000132	4.05	.443648	.00000076	4.06
337	400	.571186	.00000069	3.99	.503394	.00000033	3.99	.443648	.00000019	3.98

Table 6.4: CIR bond prices for PDE methods. Parameters are given in Table 6.1.

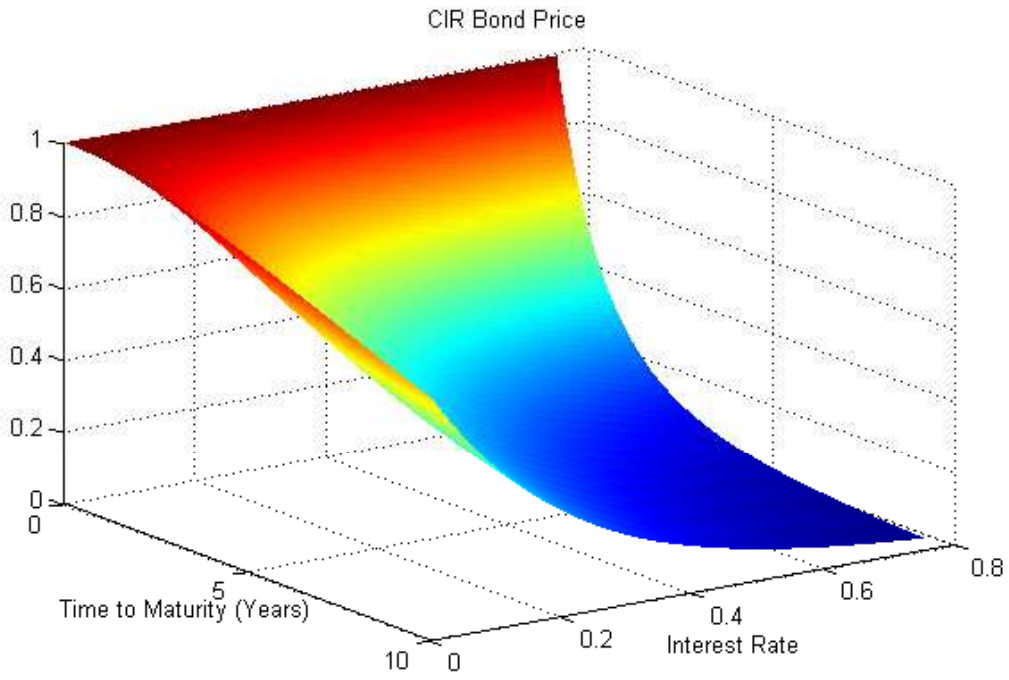

Figure 6.1: Bond price surface for the CIR model. Fully implicit method.

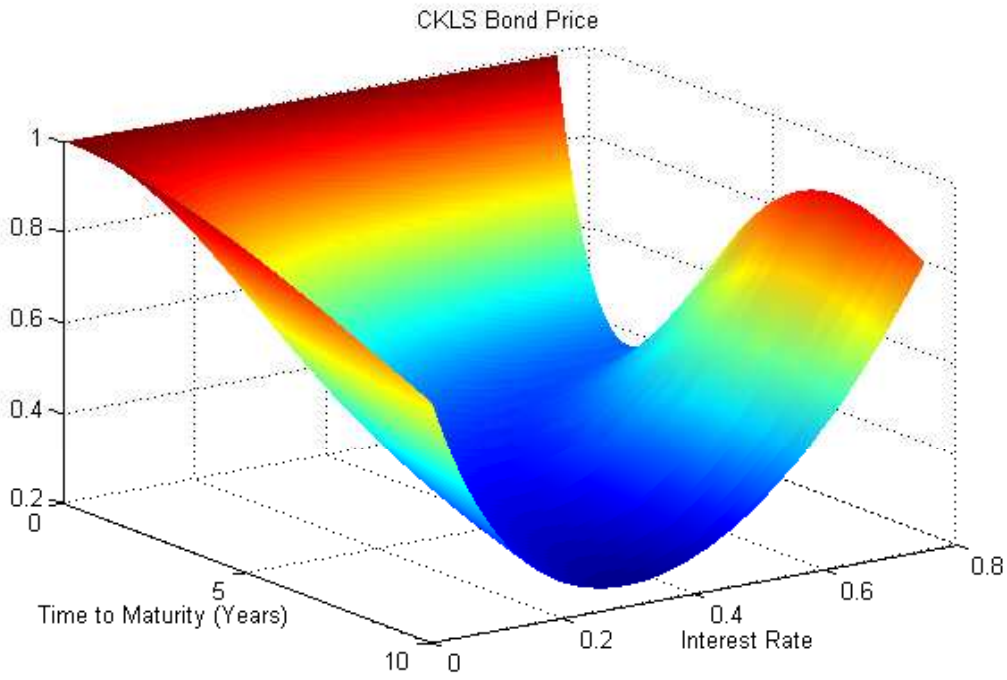

Figure 6.2: Bond price surface for the CKLS model. Fully implicit method.

This indicates that our method of imposing a boundary condition at a high value of r will not work for the CKLS model. One solution would be to extend the grid to much higher values of r and impose an approximate Dirichlet condition that the bond is worth zero at such rates. Alternatively, we can modify the volatility specification of the model so that it does not increase beyond a certain level. For example, we can change the volatility from $\sigma r^{3/2}$ to $\sigma[\min(r, \bar{r})]^{3/2}$ for some relatively high value of $\bar{r}$. Of course, $\bar{r}$ must be low enough that the drift dominates the volatility in order for this to be consistent with our approach using one-sided derivatives. Figure 6.3 plots the solution surface for the choice of $\bar{r} = .15$, and shows that our algorithm with capped volatility appears to perform well. Of course, this will only be an approximation to the actual CKLS model.

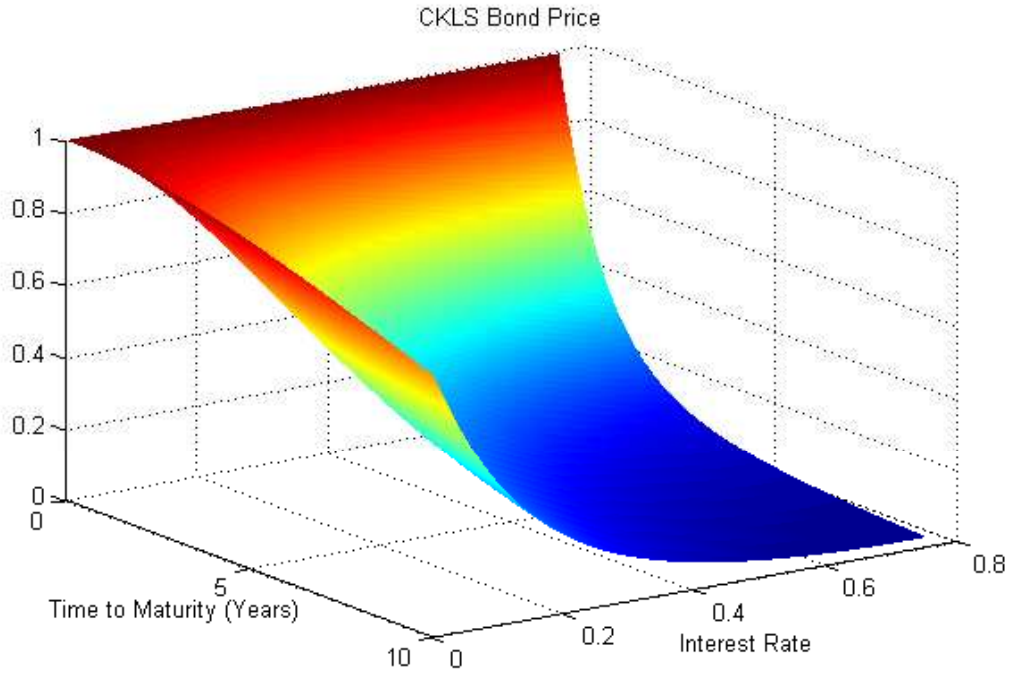

Figure 6.3: Bond price surface for the CKLS model, modified volatility. Fully implicit method.

Nodes	Time steps per year	$r = .04$ Value	Change	Ratio	$r = .07$ Value	Change	Ratio	$r = .10$ Value	Change	Ratio
1-year maturity bond, fully implicit method										
43	50	.958690			.931787			.905691		
85	100	.958706	.00001600		.931755	.00003200		.905600	.00009100	
169	200	.958707	.00000100	16.00	.931751	.00000400	8.00	.905589	.00001100	8.27
337	400	.958708	.00000100	1.00	.931753	.00000200	2.00	.905587	.00000200	5.50
1-year maturity bond, Crank-Nicolson method										
43	50	.958710			.931751			.905586		
85	100	.958708	.00000171		.931751	.00000001		.905587	.00000045	
169	200	.958707	.00000080	2.14	.931751	.00000000	30.13	.905587	.00000012	3.90
337	400	.958707	.00000018	4.46	.931751	.00000000	1.42	.905587	.00000003	3.90
5-year maturity bond, fully implicit method										
43	50	.781190			.696425			.622986		
85	100	.781172	.00001800		.696272	.00015300		.622744	0.00024200	
169	200	.781170	.00000200	9.00	.696254	.00001800	8.50	.622721	0.00002300	10.52
337	400	.781171	.00000100	2.00	.696252	.00000200	9.00	.622719	0.00000200	11.50
5-year maturity bond, Crank-Nicolson method										
43	50	.781303			.696233			.622647		
85	100	.781215	.00008750		.696247	.00001427		.622699	.00005206	
169	200	.781173	.00004202	2.08	.696251	.00000387	3.68	.622715	.00001558	3.34
337	400	.781171	.00000265	15.86	.696252	.00000108	3.58	.622709	.00000557	2.79
10-year maturity bond, fully implicit method										
42	50	.572992			.482203			.411415		
85	100	.572883	.00010900		.481987	.00021600		.411397	.0001800	
169	200	.572871	.00001200	9.08	.481974	.00001300	16.61	.411398	.00000100	18.00
337	400	.572873	.00000200	6.00	.481973	.00000100	13.00	.411393	.00000500	0.20
10-year maturity bond, Crank-Nicolson method										
43	50	.573257			.481892			.411457		
85	100	.573006	.00025098		.481942	.00005010		.411346	.00011064	
169	200	.572873	.00013275	1.89	.481965	.00002268	2.20	.411390	.00004398	2.51
336	400	.572870	.00000266	49.92	.481971	.00000578	3.92	.411391	.00000130	33.85

Table 6.5: CKLS bond prices for PDE methods. Parameters are given in Table 6.1.

Using this modified volatility approach, we report values for the CKLS model using the numerical PDE approach in Table 6.5. Compared with Table 6.4 for the CIR model, we find much more erratic rates of convergence. As we do not have an analytic solution here, we check the results with Monte Carlo simulations. The simulations also used the modified volatility. The Monte Carlo values reported in Table 6.6 are generally in very close agreement with the numerical PDE values, with the exception of the 10-year bond at an interest rate of 10%.

Next we examine the QTS model. Although this has the same volatility as the CKLS model of $\sigma r^{3/2}$, the term in the drift involving r^2 makes the drift sufficiently higher than the volatility (for our parameters) that we do not expect to find the same problems using one-sided derivatives for high values of r as we did in the CKLS case. Table 6.7 reports the results, showing more erratic convergence than seen for the CIR case in Table 6.4, but

	$r = .04$		$r = .07$		$r = .10$	
Maturity	Value	Standard error	Value	Standard error	Value	Standard error
1 year	.958711	.00001127	.931763	.00002465	.905610	.00004047
5 years	.781162	.00010684	.696192	.00018436	.622040	.00025220
10 years	.572894	.00021621	.481171	.00029075	.407574	.00034127

Table 6.6: Monte Carlo simulation of bond prices for the CKLS model. The values reported are based on 100,000 simulations with 100 time steps per year.

better behaviour than observed for the CKLS model in Table 6.5. Monte Carlo results are provided in Table 6.8, and these are very close to the numerical PDE values. Figure 6.4 illustrates the solution surface over the 10-year horizon, and shows no apparent issues related to the use of one-sided derivatives near the boundaries of the grid.

Nodes	Time steps per year	$r = .04$			$r = .07$			$r = .10$		
		Value	Change	Ratio	Value	Change	Ratio	Value	Change	Ratio
1-year maturity bond, fully implicit method										
43	50	.955431			.928529			.903163		
85	100	.955527	.00009518		.928539	.00000914		.903094	.00006892	
169	200	.955533	.00000635	14.99	.928540	.00000112	8.15	.903085	.00000859	8.02
337	400	.955534	.00000059	10.74	.928540	.00000013	8.91	.903084	.00000107	8.05
1-year maturity bond, Crank-Nicolson method										
43	50	.955503			.928541			.903084		
85	100	.955536	.00003234		.928540	.00000056		.903084	.00000014	
169	200	.955534	.00000147	22.03	.928540	.00000012	4.60	.903084	.00000005	2.64
337	400	.955534	.00000039	3.79	.928540	.00000003	4.02	.903084	.00000001	3.79
5-year maturity bond, fully implicit method										
42	50	.727464			.651988			.595897		
85	100	.727643	.00017886		.651803	.00018487		.595714	.00018250	
169	200	.727620	.00002301	7.77	.651783	.00002064	8.95	.595708	.00000598	30.51
336	400	.727615	.00000543	4.23	.651781	.00000201	10.27	.595711	.00000318	1.87
5-year maturity bond, Crank-Nicolson method										
43	50	.727415			.651760			.595565		
85	100	.727637	.00022174		.651775	.00001430		.595673	.00010794	
169	200	.727619	.00001775	12.49	.651779	.00000432	3.30	.595703	.00003034	3.55
337	400	.727615	.00000477	3.72	.651780	.00000111	3.88	.595711	.00000772	3.92
10-year maturity bond, fully implicit method										
43	50	.460206			.398464			.356845		
85	100	.460096	.00010997		.398256	.00020836		.356760	.00008530	
169	200	.460067	.00002849	3.85	.398250	.00000565	36.87	.356787	.00002759	3.09
337	400	.460065	.00000221	12.91	.398254	.00000403	1.40	.356799	.00001233	2.23
10-year maturity bond, Crank-Nicolson method										
43	50	.459830			.398078			.356469		
85	100	.460048	.00021899		.398208	.00012910		.356713	.00024379	
169	200	.460061	.00001272	17.21	.398244	.00003653	3.53	.356781	.00006868	3.54
337	400	.460064	.00000295	4.31	.398253	.00000930	3.92	.356799	.00001746	3.93

Table 6.7: QTS bond prices for PDE methods. Parameters are given in Table 6.1.

Maturity	$r = .04$		$r = .07$		$r = .10$	
	Value	Standard Error	Value	Standard Error	Value	Standard Error
1 year	.955535	.00001176	.928550	.00002555	.903099	.00003910
5 years	.727584	.00012739	.651757	.00017370	.595711	.00019304
10 years	.460184	.00020993	.398388	.00021891	.356958	.00021494

Table 6.8: Monte Carlo simulation of bond prices for the QTS model. The values reported are based on 100,000 simulations with 100 time steps per year.

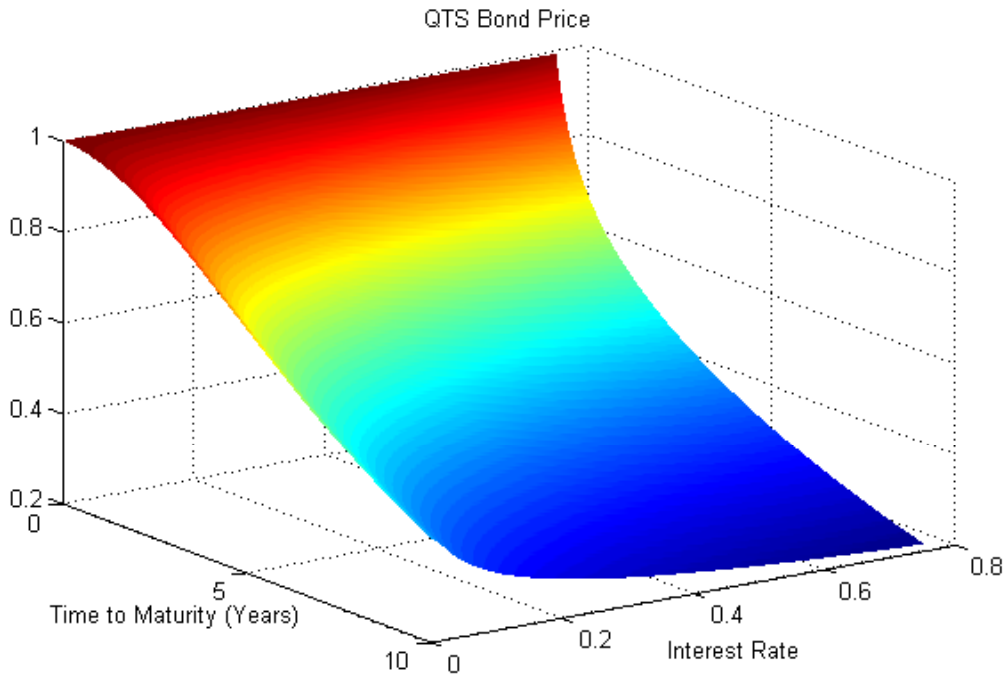

Figure 6.4: Bond price surface for the QTS model. Fully implicit method.

6.2 Bond Option Prices Without Jumps

The next set of numerical tests involve calculating prices of European call options on pure discount bonds. Two different options were used: a 6-month option on a 1-year bond, and a 2-year option on a 10-year bond. The general solution method is as follows. Consider the second case where the maturity of the underlying bond is 10 years. We start off solving (with time running backwards) for the value of a pure discount bond with a payoff of \$1 for all values of r. We let this solution advance for 8 years. This gives us the price, P, of an 8-year bond. We then apply the call option payoff constraint that the contract value is $\max(P - K, 0)$, where K is the strike price, and solve back for an additional two years. Figure 6.5 plots the solution surface obtained for one case (the CIR model) using this procedure. This figure clearly shows the application of the option payoff constraint after the solution has advanced for 8 years. It also shows that the constraint makes the solution value zero for all but relatively low values of r on the grid, suggesting that we could try a grid with fewer nodes than in the cases considered above for the bond price itself. In fact, we found that using the bond price grid from above frequently gave answers that were accurate to 6 digits even on the coarsest grid, making it almost impossible to assess the convergence of the algorithm. Accordingly, we switched to the following initial grid for these tests:

$$
\begin{aligned}
&0,\ .01,\ .02,\ .035,\ .04,\ .05,\ .054,\ .058,\ .062,\\
&.066,\ .07,\ .074,\ .078,\ .084,\ .092,\ .10,\ .11,\ .12,\\
&.14,\ .16,\ .18,\ .20,\ .24,\ .28,\ .35,\ .50,\ .75
\end{aligned}
$$

As for the bond price case, for the QTS model the lowest node of $r = 0$ was replaced by .0001. Finally, the strike price was set at 0.9 in all cases for the option expiring in 6-months, but was set to different values across the models for the 2-year option. The reason for this was that the implied eight-year bond prices at the option expiration were quite different, so setting a fixed strike across all the models lead to some uninteresting cases because the options were so far out-of-the-money as to be virtually worthless.

In the CIR case, there is an analytic solution available for European call options on pure discount bonds (see [11] or [21] for details). Table 6.9 provides these values as a basis for comparison. Note that analytic solutions are not available for either the CKLS or the QTS case.

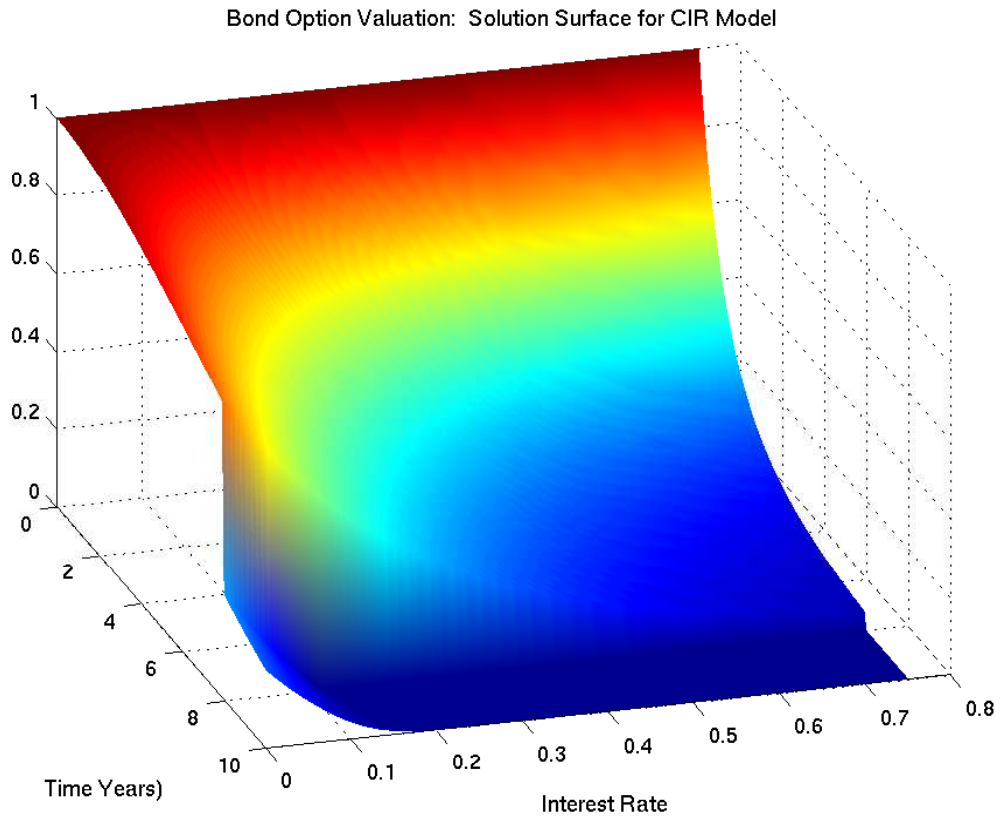

Figure 6.5: Option price surface for the CIR model. Fully implicit method.

r	6-month call option on 2-year bond, $K = .90$	2-year call option on 10-year bond, $K = .52$
4%	0.076576	0.096233
7%	0.063384	0.053236
10%	0.050703	0.023509

Table 6.9: Analytic bond option prices for the CIR model.

Nodes	Time steps per year	$r = .04$ Value	Change	Ratio	$r = .07$ Value	Change	Ratio	$r = .10$ Value	Change	Ratio
6-month option, 1-year bond, $K = 0.90$, fully implicit method										
27	50	.076565			.063409			.050765		
53	100	.076573	.00000810		.063390	.00001920		.050718	.00004670	
105	200	.076575	.00000210	3.85	.063385	.00000480	4.00	.050707	.00001170	3.99
209	400	.076576	.00000050	4.20	.063384	.00000120	4.00	.050704	.00000300	3.90
6-month option, 1-year bond, $K = 0.90$, Crank-Nicolson method										
27	50	.076576			.063384			.050701		
53	100	.076576	.00000010		.063384	.00000010		.050702	.00000090	
105	200	.076576	.00000010	1.00	.063384	.00000004	2.27	.050703	.00000030	3.00
209	400	.076576	.00000010	1.00	.063384	.00000001	4.40	.050703	.00000007	4.28
2-year option, 10-year bond, $K = 0.52$, fully implicit method										
27	50	.096284			.053276			.023307		
53	100	.096255	.00002910		.053208	.00006820		.023406	.00009940	
105	200	.096223	.00003120	0.93	.053225	.00001740	3.91	.023484	.00007760	1.28
209	400	.096231	.00000730	4.27	.053234	.00000870	2.00	.023504	.00001980	3.91
2-year option, 10-year bond, $K = 0.52$, Crank-Nicolson method										
27	50	.096074			.053123			.023247		
53	100	.096202	.00012830		.053188	.00006478		.023398	.00015077	
105	200	.096210	.00000832	15.41	.053222	.00003434	1.88	.023483	.00008496	1.77
209	400	.096227	.00001722	0.48	.053233	.00001079	3.18	.023504	.00002065	4.11

Table 6.10: CIR bond option prices for PDE methods. Parameters are given in Table 6.1.

Tables 6.10, 6.11, and 6.12 provide the results respectively for our numerical PDE approach using the CIR, CKLS (with modified volatility), and QTS models. Overall, two features stand out from these tables. First, the answers are extremely accurate, as the solution values change by very small amounts as the grid is refined. In the CIR case, the 6-month option price under Crank-Nicolson is correct to 6 digits on the coarsest grid. Second, the estimated convergence rates are quite erratic, particularly for Crank-Nicolson. This is most likely due to the application of the option payoff constraint. In general, we would expect that this may lead to uneven convergence in the Crank-Nicolson case because it is not a monotone scheme. Techniques for remedying this in the equity option context are discussed in [27], but, given the accuracy of the calculated values, it does not seem worth investigating the use of such methods here.

As a final point in this option valuation context, it is worth noting that in practice it can be of considerable importance when pricing bond options to ensure that the model prices are consistent with a given initial yield curve, i.e. a given set of pure discount bond prices of different maturities. While a variety of such methods have been discussed in the literature, one of the most generally applicable techniques is due to Dybvig [16]. While this is not the main focus here, for sake of completeness we provide an illustration in Appendix B of how this can be done.

Nodes	Time steps per year	$r = .04$ Value	$r = .04$ Change	$r = .04$ Ratio	$r = .07$ Value	$r = .07$ Change	$r = .07$ Ratio	$r = .10$ Value	$r = .10$ Change	$r = .10$ Ratio
6-month option, 1-year bond, $K = 0.90$, fully implicit method										
27	50	.077006			.062853			.049284		
53	100	.077011	.00000521		.062837	.00001734		.049238	.00004619	
105	200	.077011	.00000040	13.09	.062833	.00000237	7.32	.049231	.00000648	7.12
209	400	.077010	.00000059	.67	.062833	.00000039	6.14	.049230	.00000066	9.80
6-month option, 1-year bond, $K = 0.90$, Crank-Nicolson method										
27	50	.077015			.062834			.049232		
53	100	.077012	.00000261		.062833	.00000037		.049231	.00000070	
105	200	.077011	.00000139	1.87	.062833	.00000025	1.49	.049230	.00000079	0.87
209	400	.077010	.00000071	1.95	.062833	.00000012	2.05	.049230	.00000020	3.97
2-year option, 10-year bond, $K = 0.45$, fully implicit method										
27	50	.157497			.080306			.020160		
53	100	.156777	.00072027		.079981	.00032529		.019613	.00054681	
105	200	.156465	.00031187	2.30	.079890	.00009077	3.58	.019407	.00020588	2.65
209	400	.156311	.00015385	2.02	.079878	.00001211	7.49	.019407	.00010134	2.03
2-year option, 10-year bond, $K = 0.45$, Crank-Nicolson method										
27	50	.157316			.080067			.019909		
53	100	.156754	.00056183		.079951	.00011633		.019581	.00032810	
105	200	.156462	.00029194	1.92	.079886	.00006462	1.80	.019401	.00017940	1.82
209	400	.156311	.00015136	1.92	.079878	.00000884	7.30	.019406	.00010134	1.77

Table 6.11: CKLS bond option prices for PDE methods. Parameters are given in Table 6.1.

Nodes	Time steps per year	$r = .04$ Value	$r = .04$ Change	$r = .04$ Ratio	$r = .07$ Value	$r = .07$ Change	$r = .07$ Ratio	$r = .10$ Value	$r = .10$ Change	$r = .10$ Ratio
6-month option, 1-year bond, $K = 0.90$, fully implicit method										
27	50	.074502			.060420			.047501		
53	100	.074583	.00008088		.060418	.00000148		.047457	.00004391	
105	200	.074592	.00000844	9.58	.060418	.00000037	3.97	.047451	.00000569	7.72
209	400	.074594	.00000205	4.12	.060418	.00000010	3.88	.047451	.00000020	28.42
6-month option, 1-year bond, $K = 0.90$, Crank-Nicolson method										
27	50	.074537			.060420			.047452		
53	100	.074587	.00005071		.060418	.00000147		.047451	.00000144	
105	200	.074592	.00000476	10.64	.060418	.00000039	3.75	.047450	.00000037	3.93
209	400	.074594	.00000159	3.00	.060418	.00000010	3.96	.047450	.00000020	1.82
2-year option, 10-year bond, $K = 0.40$, fully implicit method										
27	50	.099082			.056645			.033368		
53	100	.098563	.00051876		.056295	.00034921		.033196	.00017200	
105	200	.098524	.00003967	13.07	.056255	.00004043	8.63	.033168	.00002790	6.16
209	400	.098523	.00000091	43.77	.056253	.00000210	19.27	.033173	.00010134	0.27
2-year option, 10-year bond, $K = 0.40$, Crank-Nicolson method										
27	50	.098668			.056282			.033124		
53	100	.098511	.00015709		.056250	.00003246		.033164	.00004030	
105	200	.098517	.00000596	26.37	.056249	.00000043	76.31	.033164	.00000040	99.56
209	400	.098522	.00000480	1.24	.056252	.00000293	0.14	.033172	.00010134	.00

Table 6.12: QTS bond option prices for PDE methods. Parameters are given in Table 6.1.

6.3 Bond Prices Under Jump-Diffusion

We now proceed to conduct further numerical tests, but in the context of jump-diffusion. The diffusion parameters remain the same as in Table 6.1 for the CIR, CKLS, and QTS specifications. As for the jump term, we use the same set of parameters across the three specifications. In particular, we assume that jumps follow a lognormal distribution with mean $\mu_J = 0$ and standard deviation $\gamma_J = .05$. The jump intensity λ_J is assumed to be 25. These parameters are roughly equivalent to those estimated in [24] from 3-month U.S. Treasury bill data over the time period from 1965 to 1999. As with the market price of diffusion risk λ_W, we will assume that jump risk premia (which could affect either the distribution of jump sizes or the jump intensity) are zero.

The following grid was used to calculate bond prices in the presence of jumps:

0.0000, 0.0050, 0.0100, 0.0150, 0.0200, 0.0250, 0.0300, 0.0325, 0.0350,
0.0375, 0.0400, 0.0425, 0.0450, 0.0475, 0.0500, 0.0525, 0.0550, 0.0575,
0.0600, 0.0625, 0.0650, 0.0675, 0.0700, 0.0725, 0.0750, 0.0775, 0.0800,
0.0825, 0.0850, 0.0875, 0.0900, 0.0925, 0.0950, 0.0975, 0.1000, 0.1025,
0.1050, 0.1075, 0.1100, 0.1150, 0.1200, 0.1250, 0.1300, 0.1350, 0.1400,
0.1500, 0.1600, 0.1800, 0.2100, 0.2500, 0.3000, 0.3500, 0.4000, 0.5000, 0.7500.

As in the no-jump case, due to the presence of a_{-1}/r in the drift term of the QTS model, we have changed the first grid point to 0.0001 for that model.

Given the lack of analytic solutions to serve as a check on our numerical PIDE approach, bond prices calculated using Monte Carlo simulation are provided for the CIR, CKLS (modified volatility), and QTS specifications under jump-diffusion in Tables 6.13, 6.14, and 6.15 respectively.

Results for the numerical PIDE method are provided in Tables 6.16, 6.17, and 6.18 respectively. Overall, the results are reasonably close to the Monte Carlo estimates, with slightly higher differences for the longer term bonds. The numbers in the ratio column are fairly noisy, but this is in large part due to the high accuracy—in many cases, there is very little change in price as the grid is refined, particularly for the Crank-Nicolson method. Plots of the bond price solution surfaces over the 10 year horizon are given in Figure 6.6 (CIR), Figure 6.7 (CKLS - modified volatility), and Figure 6.8 (QTS) respectively. Again, these figures show no apparent problems resulting from the use of one-sided derivatives near the boundaries of the computational grid.

	$r = .04$		$r = .07$		$r = .10$	
Maturity	Value	Standard error	Value	Standard error	Value	Standard error
1 year	.957551	.00002792	.931520	.00003960	.906205	.00005012
5 years	.767074	.00021552	.695578	.00025873	.631243	.00028909
10 years	.550438	.00035169	.482370	.00037318	.423406	.00038015

Table 6.13: Monte Carlo simulation of bond prices for the CIR model with jumps. The values reported are based on 100,000 simulations with 100 time steps per year.

	$r = .04$		$r = .07$		$r = .10$	
Maturity	Value	Standard error	Value	Standard error	Value	Standard error
1 year	.958146	.00002134	.930816	.00003924	.904362	.00005841
5 years	.771100	.00020140	.683864	.00029425	.609151	.00036189
10 years	.550074	.00037657	.459456	.00044003	.388444	.00047005

Table 6.14: Monte Carlo simulation of bond prices for the CKLS model with jumps. The values reported are based on 100,000 simulations with 100 time steps per year.

	$r = .04$		$r = .07$		$r = .10$	
Maturity	Value	Standard error	Value	Standard error	Value	Standard error
1 year	.954825	.00002192	.927684	.00004011	.902194	.00005520
5 years	.717934	.00020339	.647045	.00024700	.593781	.00025856
10 years	.451173	.00029172	.396345	.00029353	.357719	.00028376

Table 6.15: Monte Carlo simulation of bond prices for the QTS model with jumps. The values reported are based on 100,000 simulations with 100 time steps per year.

Nodes	Time steps per year	$r = .04$ Value	$r = .04$ Change	$r = .04$ Ratio	$r = .07$ Value	$r = .07$ Change	$r = .07$ Ratio	$r = .10$ Value	$r = .10$ Change	$r = .10$ Ratio
1-year maturity bond, fully implicit method										
55	50	.957505			.931549			.906322		
109	100	.957523	.00001773		.931531	.00001769		.906265	.00005689	
217	200	.957532	.00000888	2.00	.931522	.00000884	2.00	.906237	.00002851	2.00
433	400	.957536	.00000444	2.00	.931518	.00000442	2.00	.906222	.00001428	2.00
1-year maturity bond, Crank-Nicolson method										
55	50	.957541			.931514			.906208		
109	100	.957541	.00000005		.931514	.00000001		.906208	.00000022	
217	200	.957541	.00000002	3.18	.931514	.00000000	5.71	.906208	.00000007	3.00
433	400	.957541	.00000000	3.86	.931514	.00000000	5.56	.906208	.00000002	3.64
5-year maturity bond, fully implicit method										
55	50	.766869			.695413			.631158		
109	100	.766846	.00002307		.695305	.00010778		.630938	.00021988	
217	200	.766835	.00001100	2.10	.695260	.00004455	2.42	.630870	.00006778	.24
433	400	.766828	.00000609	1.81	.695238	.00002278	1.90	.630835	.00003533	1.92
5-year maturity bond, Crank-Nicolson method										
55	50	.766813			.695223			.630853		
109	100	.766818	.00000499		.695210	.00001289		.630786	.00006685	
217	200	.766821	.00000296	1.69	.695213	.00000269	4.79	.630794	.00000829	8.06
433	400	.766821	.00000089	3.32	.695214	.00000085	3.17	.630797	.00000271	3.06
10-year maturity bond, fully implicit method										
55	50	.550285			.482067			.423363		
109	100	.550121	.00016421		.481755	.00031191		.422784	.00057908	
217	200	.550076	.00004471	3.67	.481697	.00005758	5.42	.422722	.00006222	9.31
433	200	.550052	.00002410	1.86	.481668	.00002905	1.98	.422692	.00002968	2.10
10-year maturity bond, Crank-Nicolson method										
55	50	.550061			.481804			.423075		
109	100	.550061			.481623	.00018039		.422640	.00043566	
217	200	.550020	.00001113	4.69	.481632	.00000820	22.00	.422650	.00000995	43.80
433	400	.550024	.00000383	2.91	.481635	.00000384	2.14	.422656	.00000640	1.55

Table 6.16: CIR-jump bond prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

Nodes	Time steps per year	$r = .04$ Value	$r = .04$ Change	$r = .04$ Ratio	$r = .07$ Value	$r = .07$ Change	$r = .07$ Ratio	$r = .10$ Value	$r = .10$ Change	$r = .10$ Ratio
1-year maturity bond, fully implicit method										
55	50	.958081			.930783			.904295		
109	100	.958095	.00001427		.930772	.00001083		.904253	.00004262	
217	200	.958102	.00000718	1.99	.930767	.00000539	2.01	.904231	.00002130	2.00
433	400	.958106	.00000360	1.99	.930764	.00000269	2.00	.904221	.00001065	2.00
1-year maturity bond, Crank-Nicolson method										
55	50	.958111			.930762			.904211		
109	100	.958110	.00000087		.930762	.00000031		.904210	.00000043	
217	200	.958110	.00000043	2.04	.930762	.00000015	2.09	.904210	.00000021	2.03
433	400	.958110	.00000021	2.01	.930762	.00000007	2.05	.904210	.00000011	2.00
5-year maturity bond, fully implicit method										
55	50	.769719			.680423			.602553		
109	100	.769640	.00007905		.680290	.00013333		.602275	.00027860	
217	200	.769601	.00003851	2.05	.680225	.00006413	2.08	.602149	.00012574	2.22
433	400	.769582	.00001921	2.00	.680194	.00003172	2.02	.602088	.00006125	2.05
5-year maturity bond, Crank-Nicolson method										
55	50	.769699			.680240			.602221		
109	100	.769630	.00006911		.680199	.00004164		.602109	.00011149	
217	200	.769596	.00003373	2.05	.680180	.00001855	2.25	.602066	.00004290	2.60
433	400	.769580	.00001686	2.00	.680171	.00000896	2.07	.602046	.00001990	2.16
10-year maturity bond, fully implicit method										
55	50	.543090			.446150			.369431		
109	100	.542644	.00044571		.445620	.00053022		.368426	.00100566	
217	200	.542438	.00020576	2.17	.445411	.00020898	2.54	.368087	.00033852	2.97
433	400	.542337	.00010114	2.03	.445313	.00009728	2.15	.367940	.00014780	2.29
10-year maturity bond, Crank-Nicolson method										
55	50	.542877			.445835			.369058		
109	100	.542538	.00033886		.445462	.00037247		.368239	.00081934	
217	200	.542385	.00015256	2.22	.445332	.00013037	2.86	.367994	.00024496	3.34
433	400	.542311	.00007455	2.05	.445274	.00005799	2.25	.367893	.00010096	2.43

Table 6.17: CKLS-jump bond prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

Nodes	Time steps per year	$r = .04$ Value	$r = .04$ Change	$r = .04$ Ratio	$r = .07$ Value	$r = .07$ Change	$r = .07$ Ratio	$r = .10$ Value	$r = .10$ Change	$r = .10$ Ratio
1-year maturity bond, fully implicit method										
55	50	.954728			.927625			.902188		
109	100	.954779	.00005115		.927635	.00001042		.902150	.00003811	
217	200	.954801	.00002189	2.34	.927640	.00000499	2.09	.902130	.00001916	1.99
433	400	.954812	.00001098	1.99	.927643	.00000246	2.03	.902121	.00000966	1.98
1-year maturity bond, Crank-Nicolson method										
55	50	.954814			.927644			.902110		
109	100	.954822	.00000819		.927645	.00000082		.902111	.00000091	
217	200	.954822	.00000051	16.03	.927645	.00000014	5.93	.902111	.00000033	2.73
433	400	.954823	.00000028	1.80	.927645	.00000002	8.21	.902111	.00000010	3.50
5-year maturity bond, fully implicit method										
55	50	.717534			.646812			.593640		
109	100	.717659	.00012510		.646680	.00013245		.593385	.00025497	
217	200	.717656	.00000325	38.52	.646626	.00005425	2.44	.593311	.00007460	3.42
433	400	.717655	.00000054	6.05	.646598	.00002708	2.00	.593276	.00003486	2.14
5-year maturity bond, Crank-Nicolson method										
55	50	.717534			.646812			.593640		
109	100	.717613	.00007941		.646555	.00025767		.593215	.00042557	
217	200	.717633	.00001960	4.05	.646563	.00000835	30.85	.593225	.00001069	39.81
433	400	.717644	.00001088	1.80	.646567	.00000423	1.98	.593233	.00000779	1.37
10-year maturity bond, fully implicit method										
55	50	.450785			.396223			.357823		
109	100	.450698	.00008719		.395900	.00032255		.357391	.00043227	
217	200	.450629	.00006888	1.27	.395815	.00008498	3.80	.357307	.00008447	5.12
433	400	.450599	.00002980	2.31	.395779	.00003639	2.34	.357274	.00003230	2.61
10-year maturity bond, Crank-Nicolson method										
55	50	.450398			.395842			.357456		
109	100	.450504	.00010624		.395709	.00013229		.357207	.00024893	
217	200	.450532	.00002787	3.81	.395720	.00001036	12.77	.357215	.00000752	33.12
433	400	.450551	.00001859	1.50	.395731	.00001129	0.92	.357228	.00001370	0.55

Table 6.18: QTS-jump bond prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

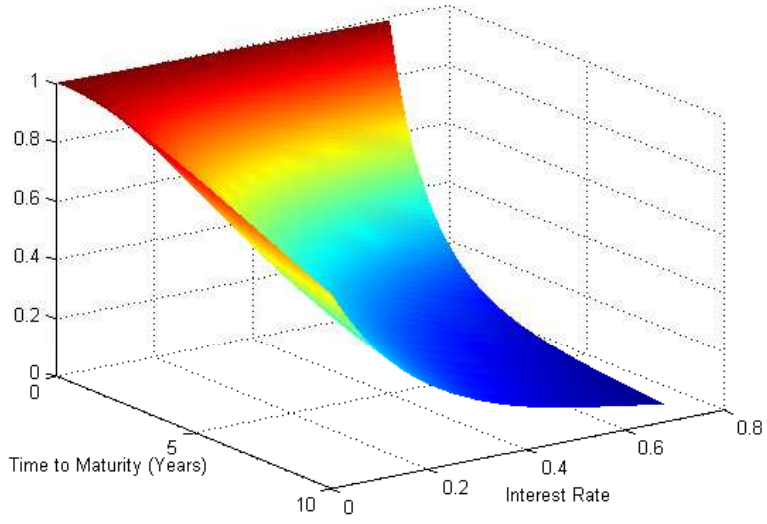

Figure 6.6: Bond price surface for the CIR-jump model. Fully implicit method.

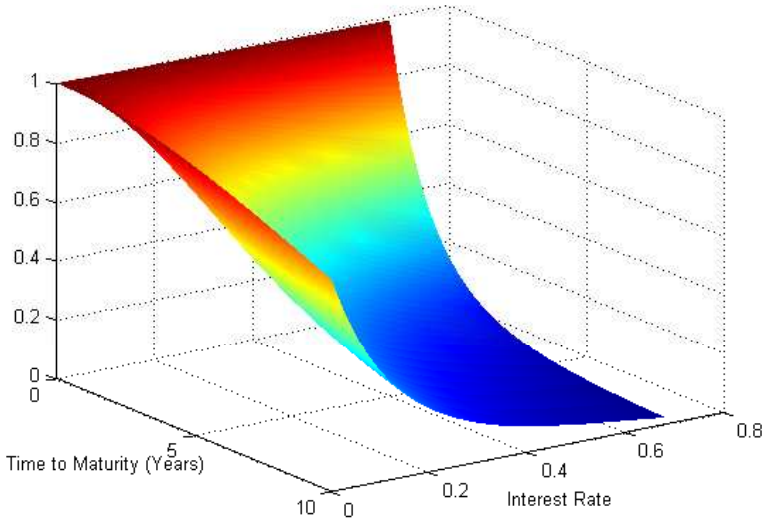

Figure 6.7: Bond price surface for the CKLS-jump model. Fully implicit method.

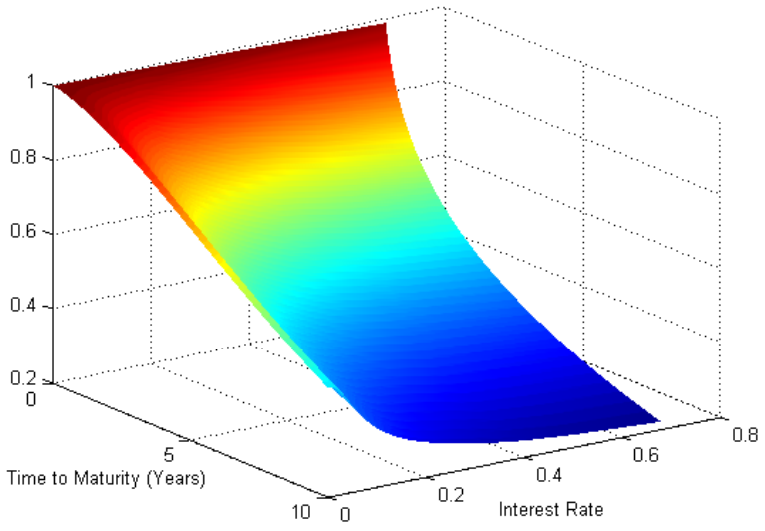

Figure 6.8: Bond price surface for the QTS-jump model. Fully implicit method.

Nodes	Time steps per year	$r = .04$ Value	Change	Ratio	$r = .07$ Value	Change	Ratio	$r = .10$ Value	Change	Ratio
6-month option, 1-year bond, $K = 0.9$, fully implicit method										
55	50	.076119			.062711			.049842		
109	100	.076127	.00000756		.062700	.00001047		.049815	.00002699	
217	200	.076130	.00000378	2.00	.062695	.00000524	2.00	.049800	.00001461	1.85
433	400	.076132	.00000189	2.00	.062692	.00000262	2.00	.049793	.00000744	1.96
6-month option, 1-year bond, $K = 0.9$, Crank-Nicolson method										
55	50	.076134			.062690			.049782		
109	100	.076134	.00000007		.062690	.00000001		.049785	.00000275	
217	200	.076134	.00000002	3.38	.062690	.00000000	8.63	.049785	.00000032	8.59
433	400	.076134	.00000001	3.89	.062690	.00000000	5.85	.049785	.00000004	8.78
2-year option, 10-year bond, $K = 0.40$, fully implicit method										
55	50	.185744			.135952			.096220		
109	100	.185576	.00016819		.135692	.00025987		.095796	.00042385	
217	200	.185527	.00004890	3.44	.135643	.00004909	5.29	.095753	.00004255	9.96
433	400	.185501	.00002617	1.87	.135618	.00002410	2.04	.095738	.00001512	2.81
2-year option, 10-year bond, $K = 0.40$, Crank-Nicolson method										
55	50	.185504			.135725			.096051		
109	100	.185456	.00004778		.135579	.00014613		.095713	.00033813	
217	200	.185467	.00001114	4.29	.135586	.00000711	20.55	.095712	.00002112	30.59
433	400	.185471	.00000386	2.88	.135590	.00000190	3.58	.095712	.00000588	3.79

Table 6.19: CIR-jump bond option prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

6.4 Bond Option Prices Under Jump-Diffusion

We now look at the convergence of our methods for pricing European call options on pure discount bonds under jump-diffusions. As for the no-jump case, we consider a 6-month option on a 1-year bond and a 2-year option on a 10-year bond. The strike price is set at 0.90 for the short term option for all models, but to varying different levels for the long term option across the models. Tables 6.19, 6.20, and 6.21 present the results. They are quite similar to those reported already above in that the PIDE method appears to be very accurate. Solution values on the initial coarse grid generally do not change much as the grid is refined, especially for the short term option. Estimated convergence rates are noisy, especially for Crank-Nicolson. As with the no-jump case, this is likely due to the application of the payoff constraint at the option expiry time.

Nodes	Time steps per year	$r = .04$			$r = .07$			$r = .10$		
		Value	Change	Ratio	Value	Change	Ratio	Value	Change	Ratio
6-month option, 1-year bond, $K = 0.9$, fully implicit method										
55	50	.076543			.062121			.048311		
109	100	.076549	.00000634		.062114	.00000698		.048288	.00002231	
217	200	.076552	.00000320	1.98	.062111	.00000347	2.01	.048277	.00001167	1.91
433	400	.076554	.00000161	1.99	.062109	.00000173	2.00	.048271	.00000589	1.98
6-month option, 1-year bond, $K = 0.9$, Crank-Nicolson method										
55	50	.076557			.062108			.048264		
109	100	.076556	.00000076		.062107	.00000026		.048265	.00000072	
217	200	.076556	.00000038	2.02	.062107	.00000013	2.06	.048265	.00000007	10.33
433	400	.076556	.00000019	2.01	.062107	.00000006	2.03	.048265	.00000009	0.78
2-year option, 10-year bond, $K = 0.35$, fully implicit method										
55	50	.223501			.144290			.087096		
109	100	.223051	.00044995		.143796	.00049410		.086236	.00085993	
217	200	.222843	.00020835	2.16	.143600	.00019562	2.53	.085971	.00026567	3.24
433	400	.222740	.00010252	2.03	.143508	.00009235	2.12	.085855	.00011623	2.29
2-year option, 10-year bond, $K = 0.35$, Crank-Nicolson method										
55	50	.223273			.143995			.086843		
109	100	.222937	.00033571		.143648	.00034645		.086108	.00073505	
217	200	.222786	.00015140	2.22	.143527	.00012153	2.85	.085908	.00020022	3.67
433	400	.222712	.00007405	2.04	.143471	.00005563	2.18	.085823	.00008487	2.36

Table 6.20: CKLS-jump bond option prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

Nodes	Time steps per year	$r = .04$			$r = .07$			$r = .10$		
		Value	Change	Ratio	Value	Change	Ratio	Value	Change	Ratio
6-month option, 1-year bond, $K = 0.9$, fully implicit method										
55	50	.073988			.059747			.046814		
109	100	.074016	.00002788		.059748	.00000137		.046790	.00002403	
217	200	.074027	.00001073	2.60	.059749	.00000044	3.14	.046777	.00001275	1.88
433	400	.074032	.00000540	1.99	.059749	.00000018	2.49	.046770	.00000667	1.91
6-month option, 1-year bond, $K = 0.9$, Crank-Nicolson method										
55	50	.074030			.059748			.046759		
109	100	.074037	.00000726		.059749	.00000080		.046763	.00000333	
217	200	.074037	.00000047	15.31	.059749	.00000013	5.99	.046763	.00000075	4.43
433	400	.074038	.00000027	1.76	.059749	.00000002	7.83	.046763	.00000013	6.00
2-year option, 10-year bond, $K = 0.35$, fully implicit method										
55	50	.135458			.097613			.073870		
109	100	.135337	.00012185		.097298	.00031498		.073476	.00039421	
217	200	.135257	.00008005	1.52	.097215	.00008266	3.81	.073406	.00006963	5.66
433	400	.135221	.00003547	2.26	.097180	.00003459	2.39	.073383	.00002300	3.03
2-year option, 10-year bond, $K = 0.35$, Crank-Nicolson method										
55	50	.135032			.097244			.073575		
109	100	.135123	.00009140		.097113	.00013084		.073328	.00024652	
217	200	.135150	.00002660	3.44	.097123	.00000939	13.94	.073332	.00000399	61.81
433	400	.135168	.00001788	1.49	.097134	.00001152	0.81	.073347	.00001405	0.28

Table 6.21: QTS-jump bond option prices for PIDE methods. Diffusion parameters are given in Table 6.1. Jump parameters: $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$.

Chapter 7

Sensitivity Analysis

Having extensively explored the convergence of our numerical algorithms in Chapter 6, it is worth briefly exploring the implications of the various models for bond and bond option pricing. Other studies in the literature (e.g. [24]) have reported that jumps generally do not have much effect on bond prices, but can significantly impact bond option prices. The economic intuition for the relative lack of impact on bond prices goes back to the fact that we can represent bond prices as expectations (under the pricing measure) of the integral of the sample path of the short rate. Over time, this amounts to a significant averaging effect, by which positive and negative jumps tend to offset each other. This averaging influence matters somewhat less for bond options. Due to the asymmetry of the bond option payoff, jumps near the expiry time of the option can have much stronger effects on bond option values, especially for short term options.

The general idea described above that bond values are significantly dependent on the average value of the short rate means that we should expect the drift parameters to affect bond prices much more. As an illustration, Figure 7.1 plots the estimated yield curves for the CIR, CKLS (with modified volatility), and QTS models in the (no-jump) diffusion context using the parameters from Table 6.1 for initial values of the short rate of 4%, 7%, and 10%. Figure 7.2 repeats this, using the yield curves implied by the corresponding jump models. Generally speaking, a comparison of each panel of Figure 7.1 with the corresponding panel of Figure 7.2 does not show too much difference, except for bonds of relatively long maturity. On the other hand, within each individual panel of either figure the biggest differences tend to be between the QTS model and the other two (except for the case with jumps and an initial interest rate of 10%). Note that the volatility due to either diffusion or jumps and diffusion is exactly the same between the CKLS and QTS models, provided r is below its cap level of 15% for the volatility modification. On the

other hand, the QTS model is fundamentally different from the other two models in that its drift is nonlinear. This leads to some very different behaviour in terms of the implied yield curves, even though the parameters used here for the diffusion model terms were all estimated in [3] from exactly the same data.

The balance of this chapter will be spent investigating the sensitivity of the QTS-jump model to various changes in its input parameters, both in terms of bond prices and bond option prices. All results reported here use the fully implicit method.

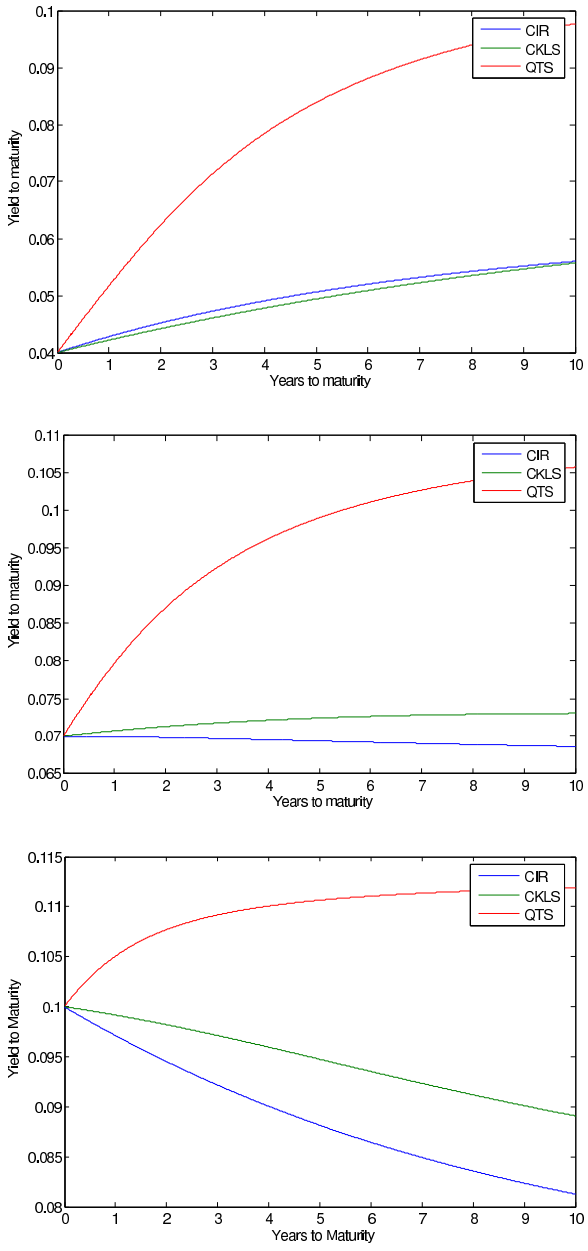

Figure 7.1: Yield curves for the CIR, CKLS, and QTS models. Fully implicit method. Top: initial $r = .04$. Middle: initial $r = .07$. Bottom: initial $r = .10$.

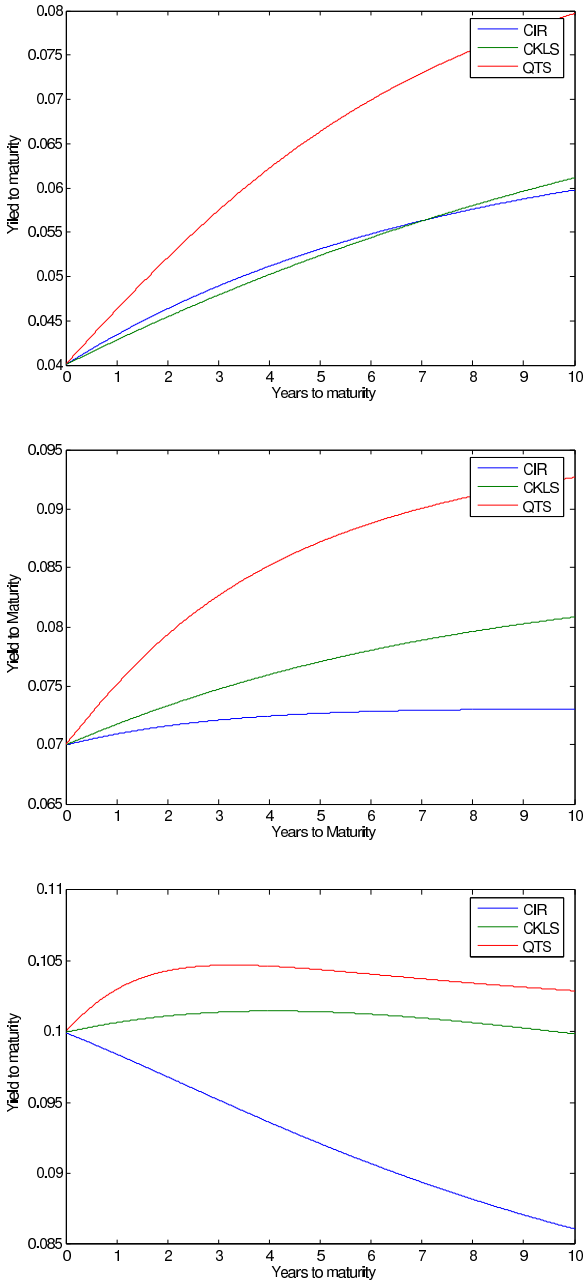

Figure 7.2: Yield curves for the CIR-jump, CKLS-jump, and QTS-jump models. Fully implicit method. Top: initial $r = .04$. Middle: initial $r = .07$. Bottom: initial $r = .10$.

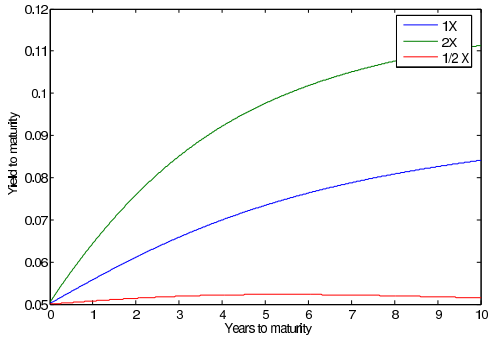

Figure 7.3: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_{-1}.

7.1 Sensitivity of the Yield Curve Under the QTS-Jump Model

As noted above, the yield curves implied by term structure models are quite sensitive to the drift parameters. We now consider how the yield curve varies when we change parameters, using only the QTS-Jump model since it is the most general specification considered in this work.

We begin with the drift parameters. Figure 7.3 plots the yield curve at 5% when we vary a_{-1}, keeping all other parameters at their base levels (given in Table 6.1 for the diffusion parameters, and setting $\lambda_J = 25$, $\mu_J = 0$, and $\gamma_J = .05$ for the jump parameters). In this figure and the following figures, "1X" in the legend means that the parameter under consideration is left at its base value, "2X" means that it is doubled relative to its base value, and "1/2 X" means that the base value is divided by 2. In the case of a_{-1}, increases in this parameter give rise to much higher drift when r is low, pushing bond prices down and the yield curve up significantly. Figure 7.4 deals with a_0, which is the drift tendency independent of the level of r. Note that the base value of $a_0 < 0$, so doubling it makes it more negative, pushing interest rates down. Figure 7.5 plots the results for a_1. Increases in this parameter shift the yield curve substantially higher. The effects of the last drift parameter, a_2 are plotted in Figure 7.6. This parameter drives the interest rate down, especially when r is high. As a result, increases in a_2 lead to higher bond prices and lower yields.

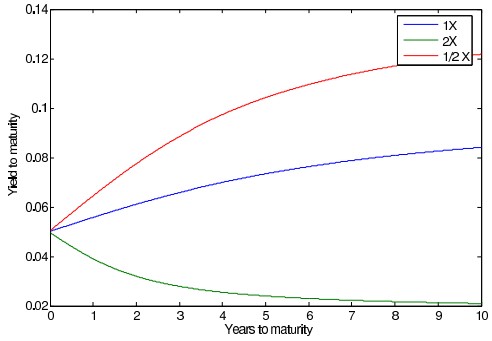

Figure 7.4: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_0.

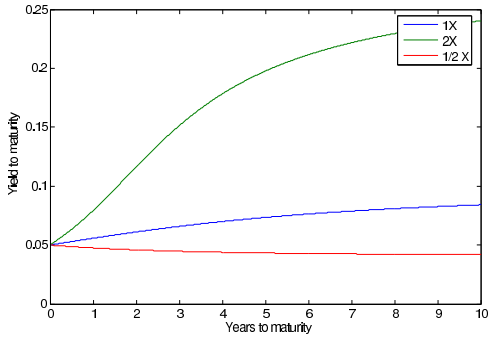

Figure 7.5: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_1.

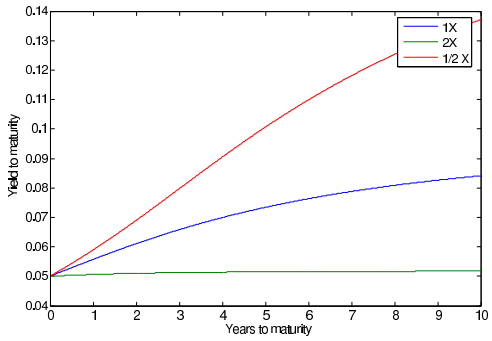

Figure 7.6: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to a_2.

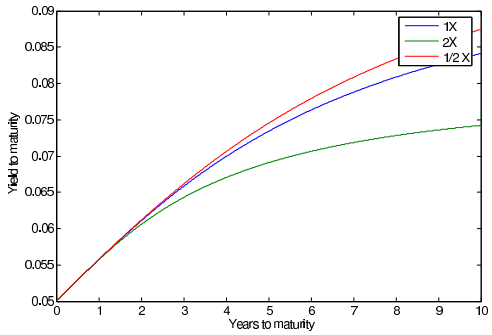

Figure 7.7: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to σ.

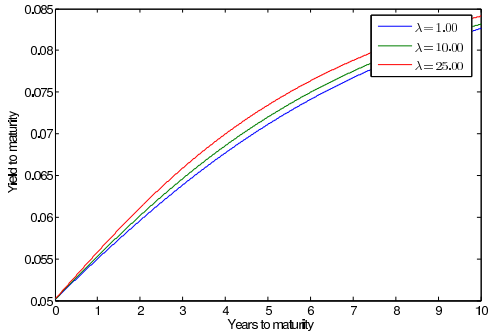

Figure 7.8: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to λ_J.

As shown in Figure 7.7, the volatility parameter σ has very little effect for bonds with maturities shorter than about 2 years, but somewhat more influence for longer term bonds. Since the jump intensity λ_J is already quite high, we look at its effect by changing it to some different values in Figure 7.8 rather than by doubling/halving it. Given the mean jump size of zero, this is effectively another volatility parameter, so it does not have much effect. On the other hand, μ_J does have a significant effect. This is because when it is positive (negative), the interest rate tends to drift up (down), moving the yield curve in the same direction. This is reflected in the plot of Figure 7.9. Conversely, γ_J does not affect the drift significantly, so it does not affect the yield curve much, as shown in Figure 7.10.

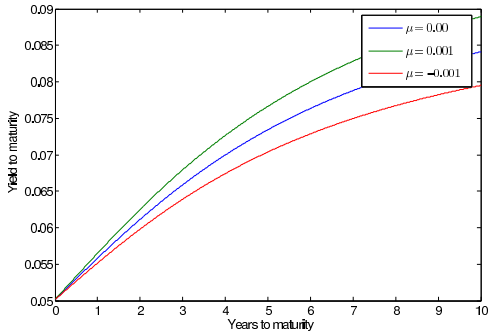

Figure 7.9: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to μ_J.

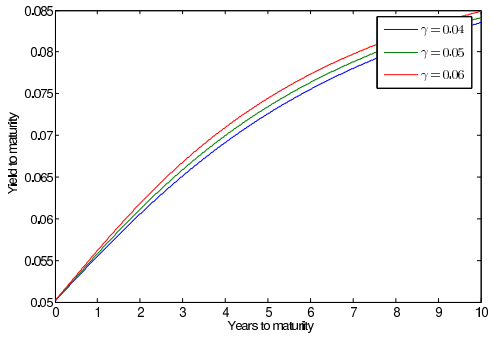

Figure 7.10: Sensitivity of the yield curve under the QTS-jump model at $r = 5\%$ to γ_J.

7.2 Sensitivity of Bond Option Prices Under the QTS-Jump Model

We conclude this chapter by briefly looking at the effects on bond option prices of changing the parameters of the QTS-jump model. We consider a 2-year European call option on a 10-year bond, with a strike price $K = 0.35$. The results are given in Table 7.1. The intuition is straightforward—any change in a parameter that tends to make r lower leads to higher bond prices and a higher option value. This would include increases in the magnitudes of a_{-1}, a_0, and a_2. The drift parameter a_1 has the opposite impact. Increases in the interest rate volatility parameter σ lead to somewhat higher option values. The mean jump size μ_J is, as above, effectively a drift parameter so increases in it reduce option values. The jump intensity λ_J and the jump size standard deviation γ_J do not have much effect, but this is conditional on $\mu_J = 0$. While the effects considered here of the jumps seem to be relatively small compared to those reported in [24], in that study the focus is on options which are

expiring much earlier (a maximum of 3 months). In fact, the effects of jumps on options with a maturity of 3 months are reported to be much smaller than for options expiring in less than 1 month. This is similar to the effects of jumps on equity option prices—one explanation for the "volatility smile" is the strong effect of jumps on short term option prices.

Parameter	Value of Parameter	Price of Bond Option
a_{-1}	0.001	0.121218
	0.002	0.028155
	0.0005	0.283086
a_0	-0.035	0.121218
	-0.070	0.485750
	-0.0175	0.005730
a_1	0.70	0.121218
	1.40	0.000000
	0.35	0.339853
a_2	-4.00	0.121218
	-8.00	0.280134
	-2.00	0.003433
σ	0.80	0.121218
	1.60	0.165312
	0.40	0.106484
λ_J	25.00	0.121218
	10.00	0.124849
	1.00	0.126780
μ_J	0.000	0.121218
	0.001	0.101680
	-0.001	0.140854
γ_J	0.05	0.121218
	0.06	0.118391
	0.04	0.123458

Table 7.1: Sensitivity of bond option prices under the QTS-jump model.

Chapter 8

Conclusions and Future Research

Several empirical studies have shown that models of changes in short-term interest rates should be able to incorporate three features: (i) the volatility of those changes should increase with the level of interest rates; (ii) the drift of the short rate may be stronger when interest rates are either very high or very low; and (iii) changes in interest rates seem to have relatively frequent discontinuous jumps. The main focus of this work has been to develop numerical techniques that can be used to price interest rate derivative contracts given a short rate model that has these features. Extensive numerical tests showed that the proposed methods seem to work reasonably well, being quite accurate on grids with a relatively small number of nodes.

Some possible directions for future research include:

- The existing methods could be refined in several ways. At this point, some aspects (such as grid generation) are *ad hoc*. It would be good to have better methods for placing nodes. Also, there is a tradeoff that should be looked at between time step size and jump intensity. Using implicit type methods (either fully implicit or Crank-Nicolson) involves using a nonlinear iterative scheme. The number of nonlinear iterations increases with the jump intensity for a fixed time step size, so we would want to take bigger steps. However, we also have the constraint that the number of jumps should not be more than one in any given time step. The computational efficiency of our methods could be improved by having a better understanding of the best time step size to use. This is particularly important in the interest rate context because many contracts are long term.

- The methods could be extended in two significant ways. First, many interest rate models involve more than one factor, besides the short rate. Extending our approach

to that type of setting would be challenging. Second, studies such as [24] and [23] have reported that statistical models in which the jump intensity is stochastic (e.g. depending on the level of interest rates) improve the fit to the data. This would be another interesting extension.

APPENDICES

Appendix A

Proof of Risk-Neutral Expectation Formulation for Bond Prices

This appendix provides a proof of the formula given in equation (3.15), showing that bond prices can be represented as expectations under the risk-neutral probability measure Q. The proof is taken from [6].

Let (t, r) be a fixed point of time and space respectively, and let

$$\Psi(s) = \exp\left\{-\int_t^s r(u)du\right\} F^T(s, r(s)).$$

Applying Itô's lemma to $\Psi(s)$ and integrating from t to T, we get:

$$\Psi(T) - \Psi(t) = \int_t^T \left(\frac{\partial \Psi}{\partial s} + (\mu - \sigma\lambda)\frac{\partial \Psi}{\partial r} + \frac{1}{2}\sigma^2 \frac{\partial^2 \Psi}{\partial r^2}\right) ds + \int_t^T \sigma \frac{\partial \Psi}{\partial r} dW(s) \tag{A.1}$$

where

$$\begin{aligned}
\frac{\partial \Psi}{\partial s} &= -r(s)\exp\left\{-\int_t^s r(u)du\right\} F^T(s, r(s)) + F_s^T(s, r(s))\exp\left\{-\int_t^s r(u)du\right\} \\
\frac{\partial \Psi}{\partial r} &= \exp\left\{-\int_t^s r(u)du\right\} F_r^T(s, r(s)) \\
\frac{\partial^2 \Psi}{\partial r^2} &= \exp\left\{-\int_t^s r(u)du\right\} F_{rr}^T(s, r(s)).
\end{aligned}$$

By substituting the latter relations in (A.1) we find:

$$\Psi(T) - \Psi(t) = \int_t^T \exp\left\{-\int_t^s r(u)du\right\}(F_t^T + \{\mu - \lambda\sigma\}F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T - rF^T)ds$$
$$+ \int_t^T \exp\left\{-\int_t^s r(u)du\right\}\left(\sigma F_{rr}^T\right)dW(s) \tag{A.2}$$

The integral on the first line of (A.2) reduces to zero because the integrand includes the term structure equation (3.14). If we take expectations under measure Q, the stochastic integral in the second line of (A.2) beomes zero, leaving:

$$E_t^Q[\Psi(T) - \Psi(t)] = 0 \Rightarrow E_t^Q[\Psi(T)] = E_t^Q[\Psi(t)]. \tag{A.3}$$

According to the definition of $\Psi(t)$:

$$\Psi(t) = \exp\left\{-\int_t^t r(u)du\right\}F^T(t, r(t)) = F^T(t, r(t)). \tag{A.4}$$

Moreover, the boundary condition of the term structure equation $F^T(T, r) = 1$ implies that:

$$\Psi(T) = \exp\left\{-\int_t^T r(u)du\right\}F^T(T, r(T)) = \exp\left\{-\int_t^T r(u)du\right\}. \tag{A.5}$$

Substituting (A.4) and (A.5) in (A.3) gives

$$F^T(t, r(t)) = E_t^Q \exp\left\{-\int_t^T r(u)du\right\},$$

completing the proof.

Appendix B

Pricing Bond Options Consistent with the Initial Yield Curve

This appendix describes how we can price options consistent with a given initial yield curve. The methodology was originally described by Dybvig [16].

Let r_t^a be any interest rate process and let r_t^b be an interest rate process that is a function of time alone, and let the interest rate process $r_t^c \equiv r_t^a + r_t^b$. Define the discount bond prices for interest rate processes a and b by

$$P_{t,s}^a = E_t\left[\exp\left(-\int_{w=t}^{s} r_w^a dw\right)\right]$$
$$P_{t,s}^b = E_t\left[\exp\left(-\int_{w=t}^{s} r_w^b dw\right)\right],$$

where $P_{t,s}$ denotes the price at time t of a bond paying \$1 at time $s > t$. The first useful result is that the discount bond price for process c is just

$$\begin{aligned} P_{t,s}^c &= E_t\left[\exp\left(-\int_{w=t}^{s} r_w^c dw\right)\right] \\ &= E_t\left[\exp\left(-\int_{w=t}^{s} \left(r_w^a + r_w^b\right) dw\right)\right] \\ &= E_t\left[\exp\left(-\int_{w=t}^{s} r_w^a dw\right)\right] \exp\left(-\int_{w=t}^{s} r_w^b dw\right) \\ &= P_{t,s}^a P_{t,s}^b. \end{aligned}$$

The second useful result relates to the pricing of bond options. Consider valuing at time t a European call option with a strike price of K which expires at time T on a discount bond which matures at time s, with $s > T > t$. Suppose we want to know the price of this option under interest rate model c. The payoff at time T is

$$\max\left(P^c_{T,s} - K, 0\right).$$

The value of the option at time t is then

$$\begin{aligned}
&E_t\left[\max\left(P^c_{T,s} - K, 0\right)\exp\left(-\int_{w=t}^{T} r^c_w dw\right)\right] \\
&= E_t\left[\max\left(P^a_{T,s}P^b_{T,s} - K, 0\right)\exp\left(-\int_{w=t}^{T}\left(r^a_w + r^b_w\right)dw\right)\right] \\
&= E_t\left[P^b_{T,s}\max\left(P^a_{T,s} - K/P^b_{T,s}, 0\right)\exp\left(-\int_{w=t}^{T} r^a_w dw\right)\exp\left(-\int_{w=t}^{T} r^b_w dw\right)\right] \\
&= E_t\left[P^b_{T,s}\max\left(P^a_{T,s} - K/P^b_{T,s}, 0\right)\exp\left(-\int_{w=t}^{T} r^a_w dw\right)P^b_{t,T}\right] \\
&= P^b_{t,s}E_t\left[\max\left(P^a_{T,s} - K/P^b_{T,s}, 0\right)\exp\left(-\int_{w=t}^{T} r^a_w dw\right)\right]
\end{aligned}$$

The implication is that the price in model c of a call option with a strike price of K is equal to the price in model a of a call option with a modified strike price of $K/P^b_{T,s}$ multiplied by $P^b_{t,s}$.

These results can be used to ensure that the prices of bond options in a given term structure model a are consistent with a given initial term structure (from model c). Here is an illustration. Consider the CIR model with parameters as in Table 6.1. Suppose that the initial yield curve is flat at 5% for all maturities, and that we want to price a six month European call option on a two year pure discount bond. Given the yield curve, the price of a six month pure discount bond is

$$P^c_{0,0.5} = \$1\exp(-.05 \times 0.5) = .9753,$$

and the price of a two year pure discount bond is

$$P^c_{0,2} = \$1\exp(-.05 \times 2) = .9048.$$

This implies that the forward bond price today is

$$P^c_{0.5,2} = .9048/.9753 = .9277.$$

An "at-the-money" bond option is often regarded as one for which the strike price is equal to the forward bond price, so we will set the strike price $K = .9277$. If we simply calculate the bond option price using the analytic formula for the CIR model, we obtain a value of 0.0038. However, this does not ensure consistency with the observed yield curve. In fact, for the parameter values above, the price of a six month pure discount bond is $P^a_{0,0.5} = 0.9748 \neq P^c_{0,0.5}$. Similarly, the price of a two year pure discount bond is $P^a_{0,2} = 0.8987 \neq P^c_{0,2}$. As a result, the implied forward bond price is 0.9219, not 0.9277.

We can calculate the corresponding prices in model b using the first of Dybvig's results above, i.e. $P^b_{t,s} = P^c_{t,s}/P^a_{t,s}$. We obtain:

	Model c	Model a	Model b
Price of 0.5 year pure discount bond	0.9753	0.9748	1.0005
Price of 2.0 year pure discount bond	0.9048	0.8987	1.0068
Forward price of bond	0.9277	0.9219	1.0064

Now we use the second result to obtain the price of European bond option under model c. Recall that this involves calculating the option price under model a with a modified strike price and then multiplying by $P^b_{0,2}$. The modified strike is $K/P^b_{0.5,2} = 0.9219$.[1] When we calculate the price under model a of an option with a strike price of 0.9219, we obtain a value of 0.006496. We then multiply by $P^b_{0,2} = 1.0068$ to get an option value under model c of 0.006540. This value, which is consistent with today's observed yield curve, is quite different from the value of 0.0038 which was found without imposing this term structure consistency condition. Of course, this procedure is not limited to the CIR model case used for illustration here. We can do this for any term structure model, including jump-diffusion models.

[1]Note that this implies that we are setting the modified strike for model a equal to the forward bond price for model a. This is because the original strike is equal to the observed (model c) forward bond price.

References

[1] C. M. Ahn and H. E. Thompson. Jump-diffusion processes and the term structure of interest rates. *Journal of Finance*, 43:155–174, 1988.

[2] Y. Aït-Sahalia. Nonparametric pricing of interest rate derivative securities. *Econometrica*, 64:527–560, 1996.

[3] Y. Aït-Sahalia. Transition densities for interest rates and other nonlinear diffusions. *Journal of Finance*, 54:1361–1395, 1999.

[4] T. G. Andersen and J. Lund. Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics*, 77:343–377, 1997.

[5] P. Balduzzi, E. J. Elton, and C. Green. Economics news and bond prices: Evidence from U.S. Treasury market. *Journal of Financial and Quantitative Analysis*, 36:523–543, 2001.

[6] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 1998.

[7] D. Brigo and F. Mercurio. *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit*. Springer Finance, 2006.

[8] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders. An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47:1209–1227, 1992.

[9] D. A. Chapman and N. D. Pearson. Is the short rate drift actually nonlinear? *Journal of Finance*, 55:355–388, 2000.

[10] T. G. Conley, L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman. Short-term interest rates as subordinated diffusions. *Review of Financial Studies*, 10:525–577, 1997.

[11] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[12] S. R. Das. The surprise element: Jumps in interest rates. *Journal of Econometrics*, 106:27–65, 2002.

[13] Y. d'Halluin, P. A. Forsyth, and K. R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25:87–112, 2005.

[14] Y. d'Halluin, P. A. Forsyth, K. R. Vetzal, and G. Labahn. A numerical PDE approach for pricing callable bonds. *Applied Mathematical Finance*, 8:49–77, 2001.

[15] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

[16] P. H. Dybvig. Bond and bond option pricing based on the current term structure. In M. A. H. Dempster and S. R. Pliska, editors, *Mathematics of Derivative Securities*, pages 271–293. Cambridge University Press, 1997.

[17] NYSE Euronext. NYSE Euronext and China. Available at `http://www.nyse.com/pdfs/NYSEEuronext_China_factsheet.pdf`, 2008.

[18] Bank for International Settlements. BIS Quarterly Review. Available at `http://www.bis.org`, June 2009.

[19] P. A. Forsyth. An introduction to computational finance without agonizing pain. Unpublished note, available at `http://www.cs.uwaterloo.ca/~paforsyt/agon.pdf`, 2008.

[20] C. Green. Economics news and the impact of trading on bond prices. *The Journal of Finance*, 59:1201–1233, 2004.

[21] J. C. Hull. *Options, Futures, and Other Derivatives*. Pearson, Upper Saddle River, NJ, 7th edition, 2009.

[22] J. James and N. Webber. *Interest Rate Modelling*. Wiley, West Sussex, England, 2000.

[23] G. Jiang and S. Yan. Linear-quadratic term structure models: Toward the understanding of jumps in interest rates. *Journal of Banking and Finance*, 33:473–485, 2009.

[24] M. Johannes. The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance*, 59:227–260, 2004.

[25] F. J. Jones. Yield curve strategies. *Journal of Fixed Income*, 1:43–48, September 1991.

[26] R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1:54–61, June 1991.

[27] D. M. Pooley, K. R. Vetzal, and P. A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40, 2003.

[28] R. Stanton. A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance*, 52:1973–2002, 1997.

[29] H. Takamizawa. Is nonlinear drift implied by the short end of the term structure? *The Review of Financial Studies*, 21:311–346, 2008.

[30] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[31] K. R. Vetzal. Stochastic volatility, movements in short term interest rates, and bond option values. *Journal of Banking and Finance*, 21:169–196, 1997.

[32] K. R. Vetzal. An improved finite difference approach to fitting the initial term structure. *Journal of Fixed Income*, 7:62–81, March 1998.

[33] S.A. Teukolsky W.H. Press, B.P. Flannery and W.T. Vetterling. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, second edition, 1992.