

**ESSAYS ON THE INTEREST RATE MODEL SPECIFICATION,
ESTIMATION, AND VALUATION OF DEFAULTABLE SECURITIES**

by

Xin Li

B.Eng., University of Science and Technology of China, 1995

M.A., University of Guelph, 1997

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

In the Department
of
Economics

© Xin Li 2002

SIMON FRASER UNIVERSITY

November, 2002

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: **Xin Li**
Degree: **PhD (Economics)**
Title of Thesis **Essays on Interest Rate Model Specification,
Estimation, and Valuation of Defaultable Securities**

Examining Committee:

Chair: Gordon Myers

Robert Jones
Senior Supervisor

Ken Kasa
Supervisor

Marie Rekkas
Internal Examiner

Angelo Melino University of Toronto
External Examiner

Date Approved: November 15, 2002

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis *Essays On Interest Rate Model Specification, Estimation,
And Valuation Of Defaultable Securities*

Author:

Xin

Li

August 14, 2003

ABSTRACT

The first chapter compares two different panel-data estimation methods, the Kalman filter and the Chen and Scott methods, on the Cox, Ingersoll, and Ross (CIR, 1985) term structure model through Monte Carlo simulation. Both methods utilize all information available but have different assumptions on the structure of the variance-covariance matrix of the measurement errors, Σ . My main findings are that the Kalman filter method with a diagonal Σ assumption has the best performance with large samples and with both monthly and weekly data. For the small sample case, the Kalman filter with a non-diagonal Σ assumption dominates.

The second chapter presents a re-examination of Chan, Karolyi, Longstaff, and Sanders (CKLS, 1992) based on a panel-data approach. It is assumed that all zero-coupon yields are observed with measurement errors but imposing linear restrictions on the errors. I find that by redefining the regime period, there is strong evidence of a structural break between the 1979-1982 period. Furthermore, I find evidence that interest rate volatility is not as sensitive to level of the interest rates as stated in the CKLS paper. Finally, I find that the Brennan and Schwartz (1980) model is superior to others when the 1979-1982 period is included in the data, whereas the Cox, Ingersoll and Ross (CIR, 1985) model is the best for data excluding the 1979-1982 period. This last finding suggests that the decision to allow or not to allow for a structural break can have a statistically and economically significant impact on the short-rate volatility estimation and model selection.

The third chapter studies the valuation of defaultable, callable bonds and credit default swaps when both interest rates and default intensity are stochastic. The model I adopt in the paper follows the framework of Duffie and Singleton (1999) and I determine

the prices of these two defaultable securities numerically. I allow for non-zero correlation between the market and the credit risk risks and examine the effect of this correlation on valuation and term structures of callable bonds and on default spreads. In addition, for defaultable, callable bonds, I examine the effects of different assumptions regarding recovery rate and the notice period on the valuation of callable bonds.

DEDICATION

To Zhonghua, Emily and My Parents

ACKNOWLEDGEMENTS

I am greatly indebted to my supervisor Professor Robert A. Jones for carefully and diligently going through the various drafts of this thesis, and his infallible economic intuition helped remove errors masked by technique. Thanks to Professor Kenneth Kasa for comments, corrections and support. Thanks to my sister, Professor Kai Li for the comments on earlier drafts. I would also like to thank my husband, Zhonghua Jiang and my parents for their patience and support throughout my period of study at Simon Fraser University. Any errors are the sole responsibility of the author.

TABLE OF CONTENTS

Approval.....	ii
Abstract.....	iii
Dedication	v
Acknowledgements	vi
Table of Contents.....	vii
List of Tables.....	x
List of Figures.....	xi
Chapter One ESTIMATION OF TERM STRUCTURE MODELS WITH LATENT STATE VARIABLES	1
1.1 Introduction.....	1
1.2 Preliminaries.....	4
1.2.1 Absence of arbitrage opportunities and risk-neutral pricing	4
1.2.2 Ito's Lemma	5
1.2.3 The Feynman-Kac formula.....	6
1.3 The one-factor translated CIR model	7
1.3.1 Distribution of the state variable	9
1.3.2 The state-space representation.....	10
1.4 Estimation of the CIR model: overview	11
1.5 Estimation of the CIR model: methods	13
1.5.1 The Chen and Scott method	13
1.5.2 The Kalman filter method.....	14
1.6 The Monte Carlo analysis	16
1.7 Empirical application	23

1.7.1	Data description	23
1.7.2	Distribution of estimated parameters.....	24
1.7.3	Estimation results.....	25
1.8	Conclusion.....	28
Chapter 1 Appendix A.....		30
Chapter 1 Appendix B.....		32
Chapter Two RE-EXAMINATION OF CHAN, KAROLYI, LONGSTAFF, AND SANDERS MODEL: A PANEL-DATA APPROACH.....		34
2.1	Introduction.....	34
2.2	The one-factor CKLS model.....	37
2.3	The estimation method.....	39
2.4	Numerical implementation.....	43
2.4.1	Finite difference methods.....	43
2.4.2	Optimization.....	45
2.5	Simulation analysis.....	46
2.6	Empirical application.....	50
2.6.1	Data description.....	50
2.6.2	Empirical results.....	52
2.6.3	Structural breaks.....	55
2.7	Conclusion.....	59
Chapter 2 Appendix.....		60
Chapter Three VALUING DEFAULTABLE SECURITIES UNDER INTEREST RATE AND DEFAULT RISK CORRELATION.....		67
3.1	Introduction.....	67
3.2	The model structure.....	69

3.3	My model specification.....	73
3.4	Valuation of defaultable, callable bonds.....	76
3.5	Valuation of default swaps.....	84
3.6	Conclusion.....	89
	Chapter 3 Appendix: The Feynman-Kac formula.....	106
	References.....	107

LIST OF TABLES

Table 1. 1 Estimated parameters for the one-factor translated CIR model.....	19
Table 1. 2 Estimation results with different initial values	20
Table 1. 3. Estimated parameters for the one-factor translated CIR model.....	21
Table 1. 4 Estimated parameters for the one-factor translated CIR model.....	22
Table 1. 5. Summary statistics for Treasury yields.....	24
Table 1. 6 The Kalman filter estimates of a one-factor square-root model of Treasury bond yields, January 1985-December 2000.	26
Table 1. 7 Estimated mean, standard deviation of the residuals	28
Table 2. 1 he relationship between alternative one-factor short-term interest rate models and parameter values in equation (2.1)	38
Table 2. 2 Monte Carlo results with different values for γ	48
Table 2. 3 Sensitivity of the estimates with respect to different values of σ_ϵ	49
Table 2. 4 Summary statistics for Treasury yields.....	51
Table 2. 5 Estimates of alternative models for the short-term interest rate ($\lambda = 0$)	53
Table 2. 6 Estimates of alternative models for the short-term interest rate ($\lambda \neq 0$)	54
Table 2. 7 Tests for regime shift from 1979.10 through 1982.09.....	57
Table 2. 8 Estimates of alternative models for the short-term interest rate ($\lambda \neq 0$) excluding 1979.10 –1982.09	58
Table 2. 9 Estimated means, standard errors and correlation matrix of the residuals for the Brennan and Schwartz model	62
Table 2. 10 Estimated means, standard errors and correlation matrix of the residuals for the CIR model	63
Table 3. 1 Summary of model parameters.....	91
Table 3. 2 Spread differences under the RMV and RFV assumptions	91
Table 3. 3 Comparison of total spreads of callable bonds with and without the notice period.....	92

LIST OF FIGURES

Figure 1. 1 The historical U.S. Treasury yields	24
Figure 2. 1 Time series of U.S. interest rates.....	51
Figure 2. 2 Time series of a 3-month T-bill yield.	64
Figure 2. 3 Time series of a 5-year T-bond yield.	64
Figure 2. 4 Time series of a 10-year T-bond yield.....	65
Figure 2. 5 Time series of a 3-month T-bill yield.	65
Figure 2. 6 Time series of a 5-year T-bond yield.	66
Figure 2. 7 Time series of a 10-year T-bond yield.....	66
Figure 3. 1 Total spreads of a 20-year defaultable, callable bond with a 10-year “lock-out” period.....	93
Figure 3. 2 Term structures of par-coupon total spreads of defaultable, callable bonds under RFV.....	94
Figure 3. 3 Term structures of par-coupon total spreads of defaultable, non-callable bonds under RFV.	95
Figure 3. 4 Par-coupon credit spreads of 20-year callable and non-callable bonds.	96
Figure 3. 5 Par-coupon credit spreads of 20-year callable and non-callable bonds.	97
Figure 3. 6 Par-coupon credit spreads of 20-year callable and non-callable bonds.	98
Figure 3. 7 Term structures of par-coupon total spreads of a defaultable, callable bond under RMV and RFV.....	99
Figure 3. 8 Term structures of par-coupon total spreads of defaultable, non-callable bonds under RMV and RFV.	100
Figure 3. 9 Par-coupon total spreads of 20-year defaultable, callable bonds with and without the notice period.	101
Figure 3. 10 Credit derivatives notional volumes.	102
Figure 3. 11 Credit default spreads of 10-year contract under RFV.....	103
Figure 3. 12 Term structures of credit default spreads under RFV.....	104
Figure 3. 13 Term structures of credit default spreads under RFV.....	105

CHAPTER ONE

ESTIMATION OF TERM STRUCTURE MODELS WITH LATENT STATE VARIABLES

In this paper I compare two different panel-data estimation methods, the Kalman filter and the Chen and Scott methods, on the well-known Cox, Ingersoll, and Ross (CIR, 1985) term structure model through Monte Carlo simulation. Both methods utilize all information available but have different assumptions on the structure of the variance-covariance matrix of the measurement errors, Σ . For the Kalman filter method, some authors make the simplifying assumption that Σ is diagonal. To test the eligibility of this simplification, I also perform Monte Carlo simulation on both diagonal and non-diagonal Σ assumptions when employing the Kalman filter method. My main findings are that the Kalman filter method with a diagonal Σ assumption has the best performance with large samples and with both monthly and weekly data. For the small sample case, the Kalman filter with a non-diagonal Σ assumption dominates. In the empirical section of the paper, I apply the Kalman filter method with a diagonal Σ matrix to estimate the parameters of one-factor translated CIR model and find that although my one-factor translated CIR model is not totally satisfactory, it is promising.

1.1 Introduction

In the past three decades, tremendous progress has been made in modeling the term structure of interest rates, which is the key in determining prices of fixed-income derivative securities. The most common approach to model the term structure of interest rates starts with the specification of a time series process, which describes the behavior of the instantaneous interest rate over time. Once a time series process for the spot rate has been specified, one can work out the time series processes for bonds of all maturities, and determine the implied term structure at every time period. Some well-known models using this approach are Merton (1973), Vasicek (1977), Brennan and Schwartz (1980), Cox, Ross and Ingersoll (CIR, 1985), Black, Derman, and Toy (1990),

Hull and White (1990), the translated CIR model in Pearson and Sun (1994), and the empirical models estimated by Chan, Karolyi, Longstaff, and Sanders (CKLS, 1992).

The challenge faced by all practitioners is how to estimate the key parameters of the various models. There are three common approaches adopted in the literature: cross-sectional, time-series, and panel-data. The cross-sectional approach estimates the parameters by using bond yields at a single instant in time. The time-series approach focuses on the dynamic implications of the model and ignores the cross-sectional information, using the observable data as an approximation of the unknown state variable. The panel-data approach, however, takes the dynamic and cross-sectional implications into account simultaneously. There are several advantages of using panel data. First, the panel-data approach fully uses all the information available and is therefore expected to give more accurate estimates of the dynamics of the term structure. Second, the combined use of time series and cross-sectional data allows for the identification of the market price of risk, which is not identified from either the time-series approach or the cross-sectional approach separately. The main objective of this paper is to compare the different estimation methods based on the panel-data approach.

Notice that real data is not always coincident with its theoretical counterpart due to exogenous factors such as rounding of prices, bid-ask spreads, etc. I commonly refer to these deviations as measurement errors. Depending on the structure of Σ , which is a variance-covariance matrix of measurement errors, there are different estimation methods used within the panel-data approach framework. For the case where Σ is assumed to be full rank, the Kalman filter is a well-known method. Geyer and Pichler (1999) employ this method to estimate multifactor CIR models by assuming a diagonal structure of Σ . A second approach is to assume that some of the yields are observed without any measurement errors. In this case, Σ has less than full rank. The state

variables can be uniquely determined and the inversion approach can be used to obtain the joint density functions and therefore the log-likelihood function. Chen and Scott (1993) utilize this method to estimate the parameters for one-, two-, and three-factor CIR models.

Usually there are two main criteria needed to be considered in empirical work, one is accuracy, the other is computation time. Unfortunately, the achievement of one objective is usually at the cost of the other. The more information we use, the more accurate the results we obtain, however, the more computation time we have to spend. The two panel-data approaches mentioned above have some advantages and disadvantages according to these two criteria. The advantages of Chen and Scott (1993) method are that it is relatively easy to implement and can reduce the computation time drastically. However, there is a disadvantage associated with applying their method. Note that some of the yields are assumed to be observed without any error, but the basis of how to choose these yields is unclear. The Kalman filter method, on the other hand, does a better job in capturing the complexity of the real world, but involves heavy computation time. My objective in this paper is to compare these methods through Monte Carlo simulations, and try to provide some evidence when choosing among different methods.

The theoretical framework for the analysis in this paper is the model of CIR (1985), where a general equilibrium model of asset pricing is used to examine the behavior of the term structure and related issues. The CIR model is a single-factor equilibrium model of the term structure that is consistent with no arbitrage opportunities and non-negativity of interest rates. My paper focuses on this one-factor CIR model because of its structural simplicity, although extension to multi-factor models is quite straightforward. Moreover, as documented in Litterman and Scheinkman (1991),

although at least three factors are needed to fully capture the variability of interest rates, almost 90 percent of the variation in U.S. Treasury rates is attributable to the variation in the first factor, which is considered to correspond to the level of interest rates.

The rest of the paper is organized as follows. Section 1.2 presents preliminaries that are needed in the following sections. Section 1.3 reviews the one-factor CIR model for the instantaneous rate. Section 1.4 presents an overview of different estimation methods. The estimation methods are derived in Section 1.5, and Section 1.6 examines the properties of different panel-data approaches with Monte Carlo studies. The empirical results are reported in Section 1.7. Section 1.8 concludes the paper.

1.2 Preliminaries

I assume that the capital markets are perfect and complete and that no arbitrage opportunities are allowed. Trading is continuous in the finite time interval $[0, \tau]$ for a fixed $\tau > 0$. The uncertainty in the economy is characterized by the probability space $(\Omega, \mathfrak{S}, P)$ where Ω is the state space, \mathfrak{S} is the σ -algebra of all subsets of Ω , and P is the probability measure defined on (Ω, \mathfrak{S}) . The information structure is given by an increasing family of sub- σ -algebras $\{\mathfrak{S}_t, t \in [0, \tau]\}$, which is generated by $n \geq 1$ independent Brownian motions $\{Z_1(t), Z_2(t), \dots, Z_n(t), t \in [0, \tau]\}$ initialized at zero. In other words, \mathfrak{S}_t represents the information available at time t , and $\mathfrak{S}_s \subset \mathfrak{S}_t$ for $s \leq t$.

1.2.1 Absence of arbitrage opportunities and risk-neutral pricing

The no-arbitrage condition and risk-neutral valuation are concepts of fundamental importance in the analysis of contingent claims. The no-arbitrage condition states that the cost of a strictly positive contingent claim must be a strictly positive number. This condition implies that a contingent claim whose payoffs can be replicated

by a portfolio of securities must have a price equal to the value of the replicating portfolio. The feasibility of the replication enables application of a powerful methodology--equivalent martingale measure (risk-neutral measure). This method of using the equivalent martingale measure involves calculating the expectation of the discounted payoffs from a security with a particular probability as if all investors were risk neutral.

The existence of an equivalent martingale measure is guaranteed by the no-arbitrage condition. Under the risk-neutral measure, Q ,¹ the current value, V_t , of a security that pays off V_T at time T is the expectation of the discounted future payoff

$$V_t = E_t^Q[\exp(-\int_t^T r(s)ds) \cdot V_T],$$

where $r(s)$ is the spot rate at time $s < T$.

The connection between the risk-neutral measure and the real-world measure is established through the market price of risk,² namely, the required compensation in expected excess return over the risk-free rate for bearing a unit of risk as measured by the volatility of returns.

1.2.2 Ito's Lemma

The main tool used in continuous time finance is Ito's lemma. Suppose X is an n -dimensional Ito process with $dX_i = \mu_i dt + \sum_{j=1}^n \sigma_{ij} dW^j$, let $F(X, t)$ be continuously differentiable in t and twice continuously differentiable real-valued function in X , then

¹ The probability Q has the following properties: 1) discounted asset prices are martingales under Q and 2) Q is equivalent to P in the sense that for any event $A \in \mathcal{F}$, $P(A) = 0$ iff $Q(A) = 0$.

² The exact relationship between those two measures can be found by Girsanov's theorem. (see Duffie, 1992, p237-238)

$$dF(X, t) = \frac{\partial}{\partial t} F(X, t)dt + \sum_{i=1}^n \frac{\partial}{\partial X_i} F(X, t)dX_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial X_i \partial X_j} F(X, t)dX_i dX_j,$$

Note that, formally, I have $dt dt = 0$, $dt dW = 0$, $dW^i dW^j = \begin{cases} dt, i = j \\ 0, i \neq j \end{cases}$.

For one-dimension, Ito's lemma has the simpler form

$$\begin{aligned} dF(X, t) &= \frac{\partial}{\partial t} F(X, t)dt + \frac{\partial}{\partial X} F(X, t)dX + \frac{1}{2} \frac{\partial^2}{\partial X^2} F(X, t)(dX)^2 \\ &= F_t dt + F_x dX + \frac{1}{2} \sigma^2 F_{xx} dt. \end{aligned}$$

1.2.3 The Feynman-Kac formula

Suppose that W^1, W^2, \dots, W^M are M independent standard Brownian motions on a probability space $(\Omega, \mathfrak{F}, P)$. The standard filtration $\{\mathfrak{F}_t, t \geq 0\}$ is generated by these M independent Brownian motions. Let X be an Ito process in \mathfrak{R}^N of the form

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$

where b and σ are valued in \mathfrak{R}^N and $\mathfrak{R}^{N \times M}$, respectively. The corresponding dynamics under the risk-neutral probability measure Q are

$$dX_t = (b(X_t, t) + \sigma(X_t, t)\lambda)dt + \sigma(X_t, t)dW_t^Q,$$

where λ is valued in \mathfrak{R}^M , which represents market price of risk associated with the M different sources of uncertainty.

I next define V by

$$V(X_t, t) = E_t^Q \left[\int_t^T e^{-\phi(s)} u(X_s, s) ds + e^{-\phi(T)} g(X_T, T) \right],$$

where V denotes the value of a contingent claims at time t that has maturity payment $g(X_T, T)$. $u(X_t, t)$ denotes the net cash flow at time t and ϕ_t is defined as $\phi_s = \int \rho(X_u, u) du$, where $\rho(X_t, t)$ is the short-rate process. Then V solves the partial differential equation

$$DV(X_t, t) - \rho(X_t, t)V(X_t, t) + u(X_t, t) = 0,$$

with boundary condition $V(X_T, T) = g(X_T, T)$, and

$$DV(X_t, t) = V_t(X_t, t) + V_x(X_t, t)(b + \sigma\lambda) + \frac{1}{2}tr[\sigma(X_t, t)^T V_{xx}\sigma(X_t, t)].$$

1.3 The one-factor translated CIR model

I follow Pearson and Sun (1994) and Gregory Duffee (1999) by assuming that r_t equals the sum of a constant and one factor, s_t . The one-factor translated CIR model is the following

$$r_t = \alpha + s_t, \quad (1.1)$$

$$ds_t = \kappa(\theta - s_t)dt + \sigma\sqrt{s_t}dw_t^P, \quad (1.2a)$$

where $\kappa > 0$, $\theta > 0$ and $\sigma^2 > 0$ are constants. The parameter θ represents the long-run mean, the parameter κ determines the speed of adjustment toward the long-term mean. dw_t^P is a Wiener process under the P-measure. The parameter α can be viewed as the lower bound of the instantaneous rate. This translation provides a more flexible specification of the process and enables us to test the original CIR model.

Under the equivalent martingale measure, this process can be represented as

$$ds_t = (\kappa(\theta - s_t) - \lambda s_t)dt + \sigma\sqrt{s_t}dw_t^Q, \quad (1.2b)$$

where dw_t^Q is a Wiener process under the Q-measure, λ is the market value of risk, which is associated with the risk in the economy, s_t . The time t price of a pure discount bond that pays off a dollar at time T is given by the expectation, under the equivalent martingale measure

$$\begin{aligned} P(s_t, t, T) &= E^Q[\exp(-\int_t^T (s(\tau) + \alpha)d\tau)] \\ \Rightarrow P(s_t, t, T)e^{\alpha(T-t)} &= E^Q[\exp(-\int_t^T s(\tau)d\tau)] \end{aligned} \quad (1.3)$$

Applying the Feynman-Kac formula, $V(s_t, t, T) = P(s_t, t, T)e^{\alpha(T-t)}$ satisfies the following partial differential equation (PDE)

$$\frac{1}{2}\sigma^2 s V_{ss} + [\kappa(\theta - s) - \lambda s]V_s + V_t = sV, \quad (1.4)$$

with boundary condition: $V(s_t, T, T) = P(s_t, T, T) = 1$.

For my one-factor translated CIR model, I have a closed-form solution for the nominal price of pure discount bond³

$$P(s_t, t, T) = A(t, T)e^{-B(t, T)s_t - (T-t)\alpha}, \quad (1.5)$$

where

$$A(t, T) = \left\{ \frac{2\gamma \exp[(\kappa + \lambda + \gamma)(T-t)/2]}{(\gamma + \kappa + \lambda)(e^{\gamma(T-t)} - 1) + 2\gamma} \right\}^{2\kappa\theta/\sigma^2}. \quad (1.6)$$

³ In this non-Gaussian model, u_t needs not be normally distributed.

$$B(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + \kappa + \lambda)(e^{\gamma(T-t)} - 1) + 2\gamma}, \quad (1.7)$$

$$\gamma = ((\kappa + \lambda)^2 + 2\sigma^2)^{1/2}. \quad (1.8)$$

For the discount bonds, the yield to maturity, $R(s_t, t, T)$, is defined by

$$R(s_t, t, T) = -\frac{\log P(s_t, t, T)}{T - t}. \quad (1.9)$$

Thus, I have

$$R(s_t, t, T) = \frac{s_t B(t, T) - \log A(t, T)}{T - t} + \alpha t, \quad (1.10)$$

where $t = (1, \dots, 1)$.

1.3.1 Distribution of the state variable

It is known that the exact transition density of the state variable for the CIR model is a non-central chi-square, $\chi^2[2cs_t; 2q + 2, 2u]$, with $2q + 2$ degrees of freedom and parameter of non-centrality $2u$ (see CIR, 1985). The probability density of the interest rate at time t , conditional on its value at the time, $t - 1$, is given by

$$f(s_t, t; s_{t-1}, t-1) = ce^{-u-v} \left(\frac{v}{u}\right)^{q/2} I_q(2(uv)^{1/2}), \quad (1.11)$$

where

$$c = \frac{2\kappa}{\sigma^2(1 - e^{-\kappa\Delta t})}, \quad u = cs_{t-1}e^{-\kappa\Delta t}, \quad v = cs_t, \quad q = \frac{2\kappa\theta}{\sigma^2} - 1,$$

and $I_q(\cdot)$ is the modified Bessel function of the first kind with order q .

The expected value and variance of $s(t)$ are the following

$$\hat{s}_{t|t-1} = E(s_t | \mathfrak{S}_{t-1}) = \mu_t = \hat{s}_{t-1} \exp(-\kappa \Delta t) + \theta(1 - \exp(-\kappa \Delta t)), \quad (1.12)$$

$$\begin{aligned} \sum_{t|t-1} &= (E(s_t - \hat{s}_{t|t-1})(s_t - \hat{s}_{t|t-1})' | \mathfrak{S}_{t-1}) = Q_t \\ &= \hat{s}_{t-1} \left(\frac{\sigma^2}{\kappa} \right) (\exp(-\kappa \Delta t) - \exp(-2\kappa \Delta t)) + \theta \left(\frac{\sigma^2}{2\kappa} \right) (1 - \exp(-\kappa \Delta t))^2, \end{aligned} \quad (1.13)$$

where \mathfrak{S}_t represents the information available at time t , Δt is the size of the time interval between t and $t-1$.

1.3.2 The state-space representation

A natural way to estimate the term structure model of interest rates using panel data is the state-space model. In the state-space model, there is a transition equation for the unobservable state variables and a measurement equation for the bond yield to maturities on an arbitrary number of maturities. Under the usual assumption that measurement errors are additive and normally distributed, the measurement equation is given by

$$y_t(s_t) = R_t(s_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma) \quad (1.14)$$

$$= a + bs_t + \varepsilon_t,$$

where $a = -\frac{\log A(t, T)}{T-t} + \alpha t$, $b = \frac{B(t, T)}{T-t}$, Σ is the variance-covariance matrix of ε_t . In

my application, the number of observed bonds and the associated maturities do not change over time. Therefore Σ has a constant dimension. Depending on the structure of Σ , I need different methods for estimating the parameters.

The transition equation is

$$s_t = d + Fs_{t-1} + u_t,^4 \quad E(u_t | \mathcal{S}_{t-1}) = 0, \quad \text{var}(u_t | \mathcal{S}_{t-1}) = Q, \quad (1.15)$$

where $d = \theta(1 - \exp(-\kappa/12))$, $F = \exp(-\kappa/12)$.

1.4 Estimation of the CIR model: overview

In the recent financial literature there are three approaches used for estimating the CIR model: cross-sectional, time-series and panel-data.

In the cross-sectional approach, only information on the yields of bonds of different maturities at a point in time is employed to implement the estimation. The parameters of the CIR model are estimated by fitting the equation (1.14) to minimize the sum of squared errors, ε_t^2 , using yields observed at a specific period of time. By employing this method, one obtains a different set of parameters for each time period. The state variable s_t , treated as an additional unknown parameter, is estimated jointly with the structural parameters. Example of this approach can be found in Brown and Schaefer (1994).

The time-series approach, on the other hand, focuses on the dynamic implications of the model and ignores the cross-sectional information. This approach is based on fitting equation (1.2a) to estimate the parameters, using observable data (e.g., yield of one-month T-bills or money market rates) as an approximation of the unknown state variable. This approach gives one set of parameter estimates. One shortcoming of this approach is that it is impossible to obtain an estimate for the market price of risk, λ ,

⁴ In this non-Gaussian model, u_t needs not be normally distributed.

which is necessary for valuation purposes. Examples of this approach include CKLS (1992) and Longstaff and Schwartz (1993).

The third approach using panel data, takes the dynamic and cross-sectional information into account simultaneously. Therefore, this method in general provides more efficient estimates of the model parameters. Under this approach, it is assumed that the data are observed with some measurement errors. In addition, these errors are assumed to be additive and normally distributed. Depending on the structure of the variance-covariance matrix of the errors, different methods for estimating the parameters are developed.

Chen and Scott (1993) utilize the conditional density of the state variable to estimate the parameters for one-, two-, and three-factor CIR models. They consider measurement errors in the model, but assume that at least one yield is observed without error. Then the change of variable technique is used to obtain the joint density functions and the log-likelihood function for a sample of discount bond yields. Note that they allow the number of bond yields observed to exceed the number of state variables.⁵

The Chen and Scott method mentioned above assumes that the variance-covariance matrix of the errors has less than full rank. An alternative is to assume the variance-covariance matrix of the measurement errors has full rank. In this case, it is convenient to cast the term-structure model in the state space form, and the Kalman filter can be used to estimate the model parameters as well as the state variables.

⁵ Pearson and Sun (1994) use a similar method to estimate a two-factor CIR model. In contrast to Chen and Scott (1993), they assume that two zero-coupon yields are observed without errors and an inversion approach is applied to replace the state variable by the observable yields. Then the maximum likelihood estimation is used. One of the shortcomings of this paper is that it fails to make use of the full cross-sectional information.

1.5 Estimation of the CIR model: methods

1.5.1 The Chen and Scott method

Suppose that I have discretely sampled observations of an N -dimensional yield vector y at time $t = 1, \dots, T$. I assume that the first element of vector y , $y_{1,t}$, is modeled without any error. According to equation (1.14), I can find the exact relationship between $y_{1,t}$ and s_t . Then the change of variable technique can be used to obtain the conditional density of $y_{1,t}$

$$f(y_{1,t} | \mathfrak{S}_{t-1}) = f(s_t | \mathfrak{S}_{t-1}) |J|, \quad (1.16)$$

where $|J^{-1}| = \left| \frac{dy_{1,t}}{ds_t} \right| = |b_1|$ and b_1 is the first element of the vector b .

The rest of the elements in vector y are assumed to be modeled with measurement errors. The modified measurement equation is given by

$$y_{(2),t} = a_{(2)} + b_{(2)}s_t + \varepsilon_{(2),t}, \quad \varepsilon_{(2),t} \sim N(0, \Sigma_{(22)}) \quad (1.17)$$

where $y_{(2),t}$, $a_{(2)}$, $b_{(2)}$ and $\varepsilon_{(2),t}$ denote the vectors with first coordinate eliminated. $\Sigma_{(22)}$ denotes the variance-covariance matrix Σ with the first row and column eliminated. According to equation (1.17), I can verify that $y_{(2),t} | y_{1,t}$ is normally distributed, $y_{(2),t} | y_{1,t} \sim N(a_{(2)} + b_{(2)}s_t, \Sigma_{(22)})$. Then the density of vector y_t can be expressed as

$$\begin{aligned} f(y_t | \mathfrak{S}_{t-1}; \gamma) &= f(y_{(2),t} | y_{1,t}, \mathfrak{S}_{t-1}; \gamma) \times f(y_{1,t} | \mathfrak{S}_{t-1}; \gamma) \\ &= f(y_{(2),t} | s_t, \mathfrak{S}_{t-1}; \gamma) \times f(s_t | \mathfrak{S}_{t-1}; \gamma) \times |J|, \end{aligned} \quad (1.18)$$

where γ is the parameter vector. The likelihood function for the sample of observation on a state variable for $t = 1, \dots, T$ is

$$\log L = \sum_{t=1}^T f(y_{(2),t}; (a_{(2)} + b_{(2)}s_t), \Sigma_{(22)}) + \sum_{t=2}^T \log(f(s_t; \mu_t, Q_t) + (T-1) \log |J|, \quad (1.19)$$

where

$$f(y_{(2),t}; a_{(2)} + b_{(2)}s_t, \Sigma_{(22)}) = (2\pi)^{-(N-1)/2} |\Sigma_{(22)}|^{-1/2} e^{(y_{(2),t} - (a_{(2)} + b_{(2)}s_t))' \Sigma_{(22)}^{-1} (y_{(2),t} - (a_{(2)} + b_{(2)}s_t)) / 2},$$

$$f(s_t; \mu_t, Q_t) = (2\pi Q_t)^{-1/2} e^{(s_t - \mu_t)^2 / 2Q_t},$$

and μ_t and Q_t are defined in equations (1.12) and (1.13), respectively.

Note that the exact distribution of the state variable, s_t , is non-central chi-square. However, under certain regularity conditions,⁶ a non-central chi-square distribution can be approximated by a normal distribution.

1.5.2 The Kalman filter method

In this subsection I present a panel-data estimator of the CIR model which is based on Geyer and Pichler (1998).

Since my term structure model is non-Gaussian model, the linear Kalman filter is no longer optimal and I do not obtain the exact likelihood function (see Lund; 1997).

⁶ If the size of the time interval Δt is small enough, the distribution of non-central chi-square variable is approximately normally distributed. In this paper, I use both monthly and weekly data, $\Delta t = (1/12)$ and $\Delta t = (1/52)$.

Estimation proceeds using quasi-maximum-likelihood (QML) estimation where the exact transition density is replaced by a normal density.

The Kalman filter recursion is a set of equations that allows a projection and conditional variance-covariance matrix to be updated once a new observation becomes available. It consists of a sequence of prediction and correction steps. First, the prediction step is given by

$$\hat{s}_{t|t-1} = \mu_t = d + F\hat{s}_{t-1}, \quad (1.20)$$

$$\Sigma_{t|t-1} = Q_t = \hat{s}_{t-1} \left(\frac{\sigma^2}{\kappa} \right) \left(\exp\left(-\frac{\kappa}{12}\right) - \exp\left(-\frac{2\kappa}{12}\right) \right) + \theta \left(\frac{\sigma^2}{2\kappa} \right) \left(1 - \exp\left(-\frac{\kappa}{12}\right) \right)^2, \quad (1.21)$$

$$\hat{y}_{t|t-1} = a + b\hat{s}_{t|t-1}, \quad (1.22)$$

$$E_{t|t-1} = \text{cov}(y_t | \mathfrak{S}_{t-1}) = bQ_t b' + \Sigma. \quad (1.23)$$

Therefore, the quasi-likelihood of y_t can be expressed as

$$y_t | y_{t-1} \sim N(\hat{y}_{t|t-1}, E_{t|t-1}). \quad (1.24)$$

Second, in the correction step, the additional information provided by \mathfrak{S}_t is used to obtain a more precise estimate of s_t ⁷.

$$\hat{s}_t = E(s_t | \mathfrak{S}_t) = \hat{s}_{t|t-1} + P_t v_t, \quad (1.25)$$

⁷ See Harvey (1989, ch.3) for details of the derivations.

where $P_t = Q_t b'(bQ_t b' + \Sigma)^{-1}$ and $v_t = y_t - \hat{y}_{t|t-1}$,

with the MSE matrix

$$\Sigma_t = Q_t - P_t(bQ_t b' + \Sigma)P_t'. \quad (1.26)$$

The Kalman filter provides all the necessary information to calculate the quasi-log-likelihood function

$$\log L = -\frac{1}{2} \log 2\pi(T-1)T - \frac{1}{2} \sum_{t=1}^T \log |E_{t|t-1}| - \frac{1}{2} \sum_{t=1}^T (v_t' E_{t|t-1}^{-1} v_t). \quad (1.27)$$

1.6 The Monte Carlo analysis

To evaluate the estimation methods described in Section 1.5 and to determine the small sample properties of the parameter estimates, hypothetical data sets are created by Monte Carlo simulation. The simulated N -dimensional yield vectors y_t for $t = 1, \dots, T$ are obtained using a three-step procedure.

First, I simulate the observations s_t for $t = 1, \dots, T$. Notice that the state variable s_t used here follows a mean-reverting, square-root process (CIR, 1985). It is well known that the conditional distribution for s_t is the non-central chi-square, $\chi^2[2cs_t; 2q + 2, 2u]$, with $2q + 2$ degrees of freedom and parameter of non-centrality $2u$. The definitions of c , q , and u can be found in Section 1.3.1.

In this paper, the exact conditional distribution is used to obtain the simulated values. Based on the discussion in Duan and Simonato (1999), the simulated value of the process at time t , is obtained by the following steps:

1. Simulate the degrees of freedom of the central chi-square using

$$df = 2q + 2 + 2j,$$

where j is a Poisson random variable with mean $u = cs_{t-\Delta t} e^{-k\Delta t}$. The time interval $\Delta t = 1/12$ or $\Delta t = 1/52$ for weekly data and $s_0 = \theta$.

2. Let g denote the random variable drawn from the central chi-square with df degrees of freedom. To obtain this random variable, I draw df random variables,

$$X_i, i = 1, \dots, df, \text{ from the standard normal distribution, then } g = \sum_{i=1}^{df} X_i^2 \sim \chi^2(df).$$

3. Compute $s_t = \frac{g}{2c}$.

Second, the theoretical yields, $R(s_t, t, T)$, are calculated on the basis of the simulated path of s and the selected maturities for the zero-coupon bonds. Assuming that there are N different maturities. The formula has been described in equation (1.10), Section 1.3.

Third, I add the noise, ε_t to $R(s_t, t, T)$, and obtaining the observed yields, $y(s_t, t, T)$, for $t = 1, \dots, T$.

$$y(s_t, t, T) = R(s_t, t, T) + \varepsilon_t.$$

Notice that ε_t follows a normal distribution with mean zero and variance-covariance matrix Σ , which has dimension $N \times N$. To generate the error terms, I need to find G^* such that $\Sigma = G^* G^{*T}$ by using the Cholesky decomposition, then draw M errors e from the standard normal distribution, and ε_t is obtained as $\varepsilon_t = G^* e$.

Notice that Geyer and Pichler (1998) assume that all yields are observed with some measurement errors, which are both serially and cross-sectionally uncorrelated, i.e., the variance-covariance matrix of error terms, Σ , is diagonal. It is difficult to believe that the true Σ matrix is diagonal because of the complexity in the real world, so this assumption is just a simplification. In addition, Lund (1997) points out that a diagonal error variance-covariance matrix is not robust under linear transformations of the data. One of the purposes of this paper is to examine the pros and cons associated with the diagonal Σ through Monte Carlo simulation, that is, I want to find out how bad the estimates could be by making such an assumption or how good a simplification it is if the estimates are acceptable.

In this section, I conduct a total of three Monte Carlo simulations that differ according to the assumptions made about Σ by the researchers. In all of the cases, the true Σ matrix is a full-rank, non-diagonal matrix.

1. The Chen and Scott method: the Σ matrix is not a full rank, at least one yield is observed without error.

2. The Kalman filter method: i). the Σ matrix is a full-rank, non-diagonal matrix.

ii). the Σ matrix is a full-rank, diagonal matrix.

In theory the true Σ matrix should be symmetric, positive definite. However, the estimation procedure does not guarantee this condition, which causes a serious problem --- log of a non-positive number, for example. To deal with this situation, I estimate the G matrix instead of the Σ matrix. The advantage of this method is that there is no restriction

on the elements in G , while Σ has to be positive definite. The true Σ matrix and the relationship between these two matrices can be found in the Appendix A.

**Table 1. 1 Estimated parameters for the one-factor translated CIR model
(large sample, monthly data)**

	κ	θ	σ	λ	α	s.e. of Est. State Variable
True values	0.30000	0.10000	0.03340	-0.01500	0.00000	
KF(general)	0.29280 (0.06173)	0.09856 (0.00491)	0.03989 (0.00451)	-0.01398 (0.01172)	0.00200 (0.00484)	0.005621
KF(diagonal)	0.32260 (0.05899)	0.09779 (0.00714)	0.04394 (0.00473)	-0.01598 (0.01458)	0.00239 (0.00656)	0.007330
C & S	1.07600 (1.35400)	0.12340 (0.04596)	0.18670 (0.05806)	-0.01157 (0.22210)	-0.02659 (0.03315)	0.044795

The true variance-covariance matrix of the measurement errors Σ is assumed to be a full-rank, non-diagonal matrix. The first row in each category is the estimates of the parameters. The numbers in parentheses are the standard errors of the estimated average values. The last column reports the standard error of the estimated state variable. The results are based on 500 simulated samples each with 200 observations of yields of 4 zero-coupon bonds of different maturities. The time interval between two observations is one month. 4 different maturities are used: 6-month, 12-month, 1-, and 2-year.

The Monte Carlo results are shown in Table 1.1 and 1.3 where the average estimates of 500 simulated samples are reported as the panel-data estimates. The benchmark parameter values used for the mean-reverting case are $\kappa = 0.3$, $\theta = 0.1$, $\sigma = 0.0334$, $\lambda = -0.015$, $\alpha = 0.0$. In all the simulations I assume using monthly observations, $\Delta = 1/12$.

To evaluate a large sample performance, the Monte Carlo results are reported in Table 1.1. As can be seen, the standard errors of these estimates are lower as I employ the Kalman filter method and the estimates are unbiased compared to the Chen and

Scott method. This implies that in general, the Kalman filter method gives us better estimates than the Chen and Scott method. Since all of the methods estimate the state variable as well, I also report the standard errors of the estimated state variables. The result seems consistent with the conclusion regarding to the performance of the kalman filter method I get so far. Since more information is employed, as expected, the Kalman filter method with a general Σ matrix is superior to all other methods, both in estimating the model parameters and in estimating the state variable. However, this advantage is considerably weakened once I take into account the computation burden. Although I only use four different maturities in the simulation, the Kalman filter method with a diagonal Σ matrix decreases computation time drastically, and the estimates from this method do not differ much from those obtained from the Kalman filter method with a general Σ matrix. Moreover, since the starting values of the parameters are arbitrarily chosen, in order to investigate the robustness of the Kalman filter method with a diagonal Σ matrix with respect to the initial values, I start the estimation procedure several times using different initial values of parameters. The results are reported in Table 1.2. I find that the estimates of the CIR parameters converge reliably to the same values. Thus, I conclude that the diagonal assumption of the variance-covariance matrix for the measurement error is a reliable and good simplification and can be used as a substitute of the general variance-covariance matrix.

Table 1. 2 Estimation results with different initial values

	κ	θ	σ	λ	α
True values	0.30000	0.10000	0.03340	-0.01500	0.00000
Estimation 1	0.32260	0.09779	0.04394	-0.01598	0.00239
Estimation 2	0.31680	0.09897	0.04349	-0.01439	0.00258
Estimation 3	0.32390	0.09623	0.04430	-0.01930	0.00224

Diagonal-Kalman filter method is used in each case. The same realization of the simulated yield curve time series is used in all three cases. They differ only in the starting vector of parameter values used.

Table 1.3 examines the small sample properties of the parameter estimates; 50 observations are used this time instead of previous 200, keeping everything else being the same as before. Aside from the understandably larger standard errors as the number of observations shrinks, the most notable feature is that the Kalman filter method both with the general Σ matrix and with the diagonal Σ matrix are quite robust with regarding to the sample size, and performance of the Kalman filter with the general Σ matrix is still better than that of the Kalman filter with the diagonal Σ matrix. Since computation burden is not a problem in this case, the results suggest that the Kalman filter method with the general Σ matrix is the best candidate to deal with small sample issues.

**Table 1. 3. Estimated parameters for the one-factor translated CIR model
(small sample, monthly data)**

	κ	θ	σ	λ	α	s.e. of Est. State Variable
True values	0.30000	0.10000	0.03340	-0.01500	0.00000	
KF(general)	0.28400 (0.18910)	0.09930 (0.00599)	0.03745 (0.00967)	-0.01097 (0.01708)	0.00135 (0.00706)	0.007206
KF(diagonal)	0.30360 (0.19980)	0.09825 (0.00853)	0.04246 (0.01198)	-0.01257 (0.02160)	0.00216 (0.009630)	0.009885
C&S	2.00700 (8.52900)	0.13030 (0.14040)	0.37060 (0.74290)	0.28700 (0.57430)	0.01199 (0.14360)	0.147380

The true variance-covariance matrix of the measurement errors Σ is assumed to be a full rank, non-diagonal matrix. The first row in each category is the estimates of the parameters. The numbers in parentheses are the standard errors of the estimated average values. The last column reports the standard error of the estimated state variable. The results are based on 500 simulated samples each with 50 observations of yields of 4 zero-coupon bonds of different maturities. The time interval between two observations is one month. 4 different maturities are used: 6-month, 12-month, 1-, and 2-year.

To ensure the validity of these methods at the frequencies commonly used in practice, I also compare these methods using weekly data. Since for weekly data,

sample sizes are usually large, I only consider large sample performance here. The results are displayed in Table 1.4.

One result emerging from Table 1.4 is that even both methods have volatile estimates of σ ; overall the Kalman filter method with the diagonal Σ matrix provides better estimates than the Kalman filter method with the general Σ matrix. Another finding is that using weekly data noticeably improves the Chen and Scott method.

**Table 1. 4 Estimated parameters for the one-factor translated CIR model
(large sample, weekly data)**

	κ	θ	σ	λ	α	s.e. of Est. State Variable
True values	0.50000	0.10000	0.02340	-0.01500	0.00000	
KF(general)	0.42180 (0.18150)	0.09924 (0.00309)	0.04061 (0.00856)	-0.01304 (0.01279)	0.00100 (0.00385)	0.00430
KF(diagonal)	0.49050 (0.16410)	0.09944 (0.00298)	0.04616 (0.00845)	-0.01388 (0.01209)	0.00088 (0.00369)	0.00417
C&S	0.43410 (3.87500)	0.13430 (0.83730)	0.19210 (0.26800)	-0.01399 (0.10240)	-0.29150 (0.09867)	0.57620

The true variance-covariance matrix of the measurement errors is assumed to be full rank, non-diagonal matrix. The first row in each category is the estimates of the parameters. The numbers in parentheses are the standard errors of the estimated average values. The last column reports the standard error of the estimated state variable. The results are based on 200 simulated samples each with 400 observations of yields of 4 zero-coupon bonds of different maturities. The time interval between two observations is one week. 4 different maturities are used: 6-month, 12-month, 1-, and 2-year.

In summary, I conclude that the Kalman filter method with the diagonal Σ matrix is a good candidate to estimate the parameters in large samples, both with monthly and weekly data, and the Chen and Scott method is a good alternative when weekly data is used. For small samples, the Kalman filter method with the general Σ matrix is the best

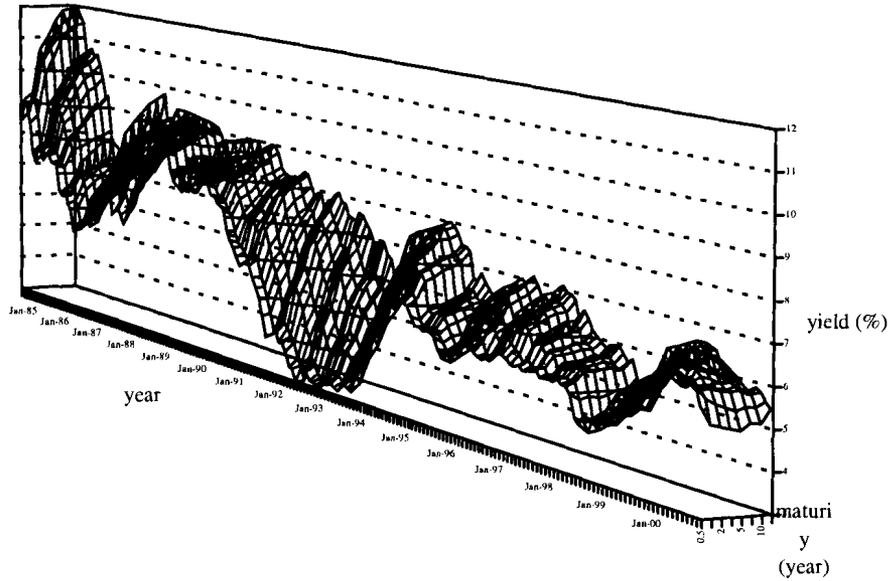
choice. Thus, in the empirical section of the paper, where monthly data is used, the Kalman filter method with the diagonal Σ matrix is employed.

1.7 Empirical application

1.7.1 Data description

The data set consists of 192 monthly observations of the yield curve for the US government Treasury bills, notes, and bonds from January 1985 to December 2000, obtained from the FRED database. The original FRED data are daily quoted. For my purpose, I pick each month-end observation as an approximation of the monthly data. The observations are not actual price quotes, but rather are estimates of the coupon that would be required for bonds of various maturities to trade at par. Eight different maturities are considered, they are 6-month and 1-year simple interest rates, and 2, 3, 5, 7, 10 and 30-year semi-annual compounded par-coupon yields. To get a feel for the data, I provide in Figure 1.1 the surfaces of the Treasury yield curves for the entire sample period, and in Table 1.5 the summary statistics for the yields. Since the CIR model corresponds to continuous compounded zero-coupon yields rather than simple interest rates or semi-annual compounded par-coupon yields, some modifications are necessary. For the six-month and one-year simple interest rates, only simple transformation is needed. For the rest of the data, first, I use a linear interpolation to get par-coupon yields every six months, then apply the recursion method to find corresponding zero-coupon yields. Details are contained in the Appendix B.

Figure 1. 1 The historical U.S. Treasury yields



This figure presents the historical movement of yields to maturity of eight Treasury bills, notes, and bonds. The sample contains monthly observations in period from January 1985 to December 2000. Eight different maturities are considered, they are 6 month, 1-, 2-, 3-, 5-, 7-, 10-, and 30-year bonds.

Table 1. 5. Summary statistics for Treasury yields

Maturity	0.5	1	2	3	5	7	10	30
Mean	5.79	5.90	6.51	6.70	6.96	7.18	7.27	7.51
Std	1.48	1.44	1.51	1.49	1.46	1.46	1.47	1.39
Maximum	9.32	9.28	10.66	11.03	11.55	11.87	11.91	11.90
Minimum	2.90	3.01	3.80	4.20	4.23	4.38	4.44	4.98
Skewness	0.16	0.16	0.34	0.45	0.58	0.65	0.60	0.59

This table presents the sample means, standard deviations (Std), maximum, minimum and skewness of the yields to maturity of eight Treasury bills, notes, and bonds. The sample contains 192 monthly observations in the period between January 1985 and December 2000.

1.7.2 Distribution of estimated parameters

Under certain regularity conditions, the maximum likelihood estimates of the parameter vector γ are asymptotically distributed around the true γ as follows:

$$\hat{\gamma} \xrightarrow{a} N(\gamma, (I(\gamma))^{-1}), \quad (1.28)$$

where I denotes the information matrix

$$I = -E\left(\frac{\partial^2 \ln L}{\partial \gamma \partial \gamma'}\right). \quad (1.29)$$

L is the likelihood function evaluated at the true γ . However, in practice, the second derivatives can be quite complicated to derive and the expected value of the second derivatives of the log likelihood is unknown. To get around these problems, I adopt a method of Berndt, Hall, Hall and Hausmann (the BHHH estimator) as given in Greene (1997, p.139, eq.4-52). The BHHH estimator for $I(\gamma)$ is:

$$\left[\sum_{t=1}^T \left(\frac{\partial \ln L_t}{\partial \gamma} \right) \left(\frac{\partial \ln L_t}{\partial \gamma} \right)' \right]_{\gamma=\hat{\gamma}}, \quad (1.30)$$

where L_t denotes the probability density of the one-period observation y_t . Thus, for each observation, I determine numerically the partial derivatives of the log-likelihood with respect to the thirteen parameters $\gamma = (\kappa, \theta, \sigma, \lambda, \alpha, g_1, \dots, g_8)'$ evaluated at the maximum likelihood estimate $\hat{\gamma}$, and then accumulate the outer product of that vector with itself. Standard errors for the parameters are the square roots of the corresponding diagonal elements of the inverse of this matrix. This estimator is extremely convenient, in most cases, because it does not require any computations beyond those required to solve the likelihood equation. In addition, it is always positive definite.

1.7.3 Estimation results

Estimation results of the spot rate process are displayed in Table 1.6. The estimated standard errors for the QML estimates are obtained as described in the previous subsection.

Table 1. 6 The Kalman filter estimates of a one-factor square-root model of Treasury bond yields, January 1985-December 2000.

κ	θ	σ	λ	α
0.04935	0.03146	0.09195	-0.10290	0.03537
(0.00351)	(0.00109)	(0.00378)	(0.00489)	(0.00121)

The instantaneous interest rate is

$$r_t = \alpha + s_t,$$

where $ds_t = \kappa(\theta - s_t)dt + \sigma\sqrt{s_t}dw_t^P,$ (true measure)

$$ds_t = (\kappa(\theta - s_t) - \lambda s_t)dt + \sigma\sqrt{s_t}dw_t^Q. \quad (\text{martingale measure})$$

Month-end yields of the 6-month and 12-month Treasury bills and 2-, 3-, 5-, 10-, and 30-year coupon bonds are observed with normally distributed measurement errors independent across time and instruments. The numbers in parentheses are asymptotic standard errors of the estimated parameters.

I obtain highly significant parameter estimates (at the 5% significance level). The significant mean reversion parameter, $\kappa = 0.04935$, implies mean reversion in the underlying instantaneous rate. The estimate of 0.04935 implies a mean half life of 14 years,⁸ which implies a very slow mean reversion for interest rates, comparing to 4 years in Geyer and Pichler (1999). Similar to Geyer and Pichler (1999), the risk premium parameter λ has the expected negative sign. The mean reversion parameter under the risk-neutral measure $\kappa+\lambda$ is negative as well, which implies that the risk premium for holding long-term bonds is positive. The lower bound of the nominal interest rate, α , is 0.03537, which seems reasonable given that the estimated long-run mean of instantaneous rate is $\alpha+\theta = 0.06683$. Notice that the estimate of α is statistically significant different from zero, which suggests that my data is in favor of the one-factor

⁸ The mean half life is the expected time for the process to return halfway to its long-run average mean, θ . $e^{-\kappa t} = 0.5 \Rightarrow t = -\ln(0.5) / \kappa$. See Chen and Scott (1993).

translated CIR model instead of the original CIR model; moreover, comparing to -1 of the estimate of α in Duffee (1999), my result rules out the possibility of negative instantaneous spot rate.

Notice that the state variable is also estimated. In what follows, I ignore this estimation error and stay the properties of the filtered s_t . To learn more about this CIR model and the estimated time series, I perform unit root tests on the estimated state variable. Dickey and Fuller (1981) have developed several tests for unit roots in time series. Specifically, I use the augmented Dicky-Fuller (ADF) unit root test with four lags and assume that there is no time trend but a constant in interest rates. The ADF test statistic is -2.1809 , the critical value is -2.8770 at the 5% significance level. Therefore, the null hypothesis of a unit root cannot be rejected by the data. The existence of unit roots in the state variable series implies no mean reversion, my estimate of κ , however, implies a weak, but significant mean reversion. The reason for this discrepancy in results is that the estimation procedures use both time-series and cross-sectional information, whereas the ADF test deals with only time-series, so the test of unit root has a low power in this case.

The CIR model implies no autocorrelation between the changes of the state variable over sufficiently short intervals, since the Weiner process has independent increments. To examine the validity of my CIR model, I also test whether this is the case. The estimation equation is

$$\Delta s_t = \alpha + \beta \Delta s_{t-1} + \varepsilon_t. \quad (1.31)$$

The ordinary least squares estimation method is used, and result shows that the estimated autocorrelation coefficient, $\hat{\beta}$, is 0.1749, the associated t-value is 2.4657, the

critical values are 1.960 and 2.576 at the 5% and 1% significance levels, respectively. Thus the null hypothesis of no autocorrelation can be rejected at the 5% level but cannot be rejected at the 1% level. Even with rejection at 5%, quantitatively, 0.1749 implies quite weak autocorrelation. This result implies that my one-factor translated CIR model is not totally satisfactory, but it is promising.

Table 1. 7 Estimated mean, standard deviation of the residuals

	0.5	1	2	3	5	7	10	30
Mean	-0.0049	-0.0046	-0.0008	-0.0003	-0.0002	-0.0000	-0.0017	-0.0000
Est. Std	0.0085	0.0075	0.0060	0.0049	0.0041	0.0044	0.0051	0.0079

The element of the estimated variance-covariance matrix of the residuals, $\hat{\sigma}_{ij}$, $i = 1, \dots, N$, and $j = 1, \dots, N$ can be calculated as

$$\hat{\sigma}_{ij} = \frac{\sum_{t=1}^T (e_{it} - \bar{e}_i)(e_{jt} - \bar{e}_j)}{T},$$

where N is the number of maturities, T is the number of observations, e_{it} is the residual at time t with maturity i , which equals $(Y_{it} - \hat{R}_{it})$, and \bar{e}_i is the mean of the residual with maturity i . The estimated standard deviation of the residuals is just the square root of the diagonal elements.

Given the parameter estimates, I construct the residuals of the model, defined as the difference between the observed yields and the predicted yields and therefore equal to the prediction errors generated by the Kalman filter. It appears that my model provides quite a nice fit to the observed bond yields: 38.9% of the residuals are below 25bp, 65.0% are below 50bp, and 88.1% are below 100bp.

1.8 Conclusion

In this paper, I compare two different estimation methods for the one-factor translated CIR model. The methods utilize both the cross-sectional and time-series information but have different assumptions on the structure of the variance-covariance

matrix of the measurement errors. The Monte Carlo study shows that based on the criteria of accuracy, robustness, and computation time, the Kalman filter method with the diagonal Σ matrix has the best performance in large samples, both with monthly and weekly data. The Chen and Scott method is a good alternative when dealing with weekly data. For the small sample case, since computation burden is not an issue, the Kalman filter method with a general Σ matrix dominates. In the empirical section, I apply the Kalman filter method with the diagonal Σ matrix to estimate the parameters of the one-factor translated CIR model and find that the model gives a nice fit to the observed bond yields. Given the model's simplicity, this result implies that although my one-factor translated CIR model is not totally satisfactory, it is promising.

Chapter 1 Appendix A

The true Σ matrix, which is used to generate the observed yields, is described as follows:

$$\Sigma = \begin{bmatrix} 0.0000988408 & & & \\ 0.0000066431 & 0.0000251146 & & \\ -0.0000048936 & -0.000002105 & 0.0000998776 & \\ 0.0000065611 & 0.0000008547 & -0.0000034215 & 0.0000040135 \end{bmatrix}.$$

Since the Σ matrix is symmetric and positive definite (p.d.), there exists a unique representation of the form

$$\Sigma = ADA',$$

where A is a lower triangular matrix with 1s along the principal diagonal,

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix},$$

and D is a diagonal matrix

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix},$$

where $d_i > 0$ for all i .

In my case, I parameterize variance-covariance matrix Σ as follows

$$G = \begin{bmatrix} g_{11} & 0 & \cdots & 0 \\ g_{21} & g_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix},$$

where $g_{ii} = \ln d_i$ and $g_{ij} = a_{ij}$ for $i > j$. Notice that the range of the elements of G is unbounded.

To create Σ matrix from G , I make the following transformation:

$$(1). g_{ii}^* = (e^{g_{ii}})^{1/2},$$

$$(2). g_{ij}^* = g_{ij} \cdot g_{jj}^*, j < i,$$

$$(3). g_{ij}^* = 0, j > i,$$

$$(4). \Sigma = G^* G^{*T}.$$

The elements of G can take any values, but Σ is assured to be symmetric and positive definite and G is unique.

Chapter 1 Appendix B

This appendix explains how to process the data so that continuous compounded zero-coupon yields can be constructed from simple interest rates and semi-annual compounded par-coupon yields.

1). Transformation from simple interest rates to continuous compounded zero-coupon yields:

Suppose that r is the simple annual interest rate, T is the time to maturity of the corresponding bond, y is the resulting continuous compounded zero-coupon yield. Then the relationship between r and y can be expressed as:

$$e^{-yT} = \frac{1}{1 + r * T}$$
$$\Rightarrow y = \frac{1}{T} \ln(1 + r * T) \quad (1.32)$$

2). Transformation from semi-annual compounded par-coupon yields to continuous compounded zero-coupon yields:

In order to obtain continuous compounded zero-coupon yields, first, I need to find semi-annual compounded par-coupon yields every six months starting with six-month yield and ending at thirty-year yield; this can be done by applying linear interpolation method. Second, the recursive method is used to find continuous compounded zero-coupon yields every six months. Third, I pick the yields with the eight maturities that I have mentioned in the text. The recursive method is as follows.

Suppose that y_t stands for the continuous compounded zero-coupon yields with maturity t . c_t is the annual par coupon payment given maturity t . Then for the t -maturity bond, the relation between y_t and c_t is:

$$\begin{aligned}
100 &= \frac{c_t}{2} * e^{-y_{0.5} * 0.5} + \frac{c_t}{2} * e^{-y_1 * 1} + \dots + \frac{c_t}{2} * e^{-y_{t-0.5} * (t-0.5)} + (100 + \frac{c_t}{2}) * e^{-y_t * t} \\
\Rightarrow e^{-y_t * t} &= \frac{100 - (\frac{c_t}{2}) * (e^{-y_{0.5} * 0.5} + e^{-y_1 * 1} + \dots + e^{-y_{t-0.5} * (t-0.5)})}{100 + \frac{c_t}{2}} \\
\Rightarrow y_t &= -\frac{1}{t} \ln\left(\frac{100 - (\frac{c_t}{2}) * (e^{-y_{0.5} * 0.5} + e^{-y_1 * 1} + \dots + e^{-y_{t-0.5} * (t-0.5)})}{100 + \frac{c_t}{2}}\right) \tag{1.33}
\end{aligned}$$

Notice that the values of $y_{0.5}$ and y_1 can be obtained from equation (1.32), then by applying equation (1.33), I can directly calculate all other continuous compounded zero-coupon yields as needed.

CHAPTER TWO

RE-EXAMINATION OF CHAN, KAROLYI, LONGSTAFF, AND SANDERS MODEL: A PANEL-DATA APPROACH

In this paper I present a re-examination of Chan, Karolyi, Longstaff, and Sanders (CKLS, 1992) based on a panel-data approach. It is assumed that all zero-coupon yields are observed with measurement errors. By imposing linear restrictions on the errors, the underlying state variables are uniquely identified. Hence, the likelihood function is directly available without the need for a filtering algorithm. I find that by redefining the regime period, there is strong evidence of a structural break during the Federal Reserve Experiment period between the 1979-1982 period. Furthermore, I find evidence that interest rate volatility is not as sensitive to level of the interest rates as stated in the CKLS paper. Finally, I find that the Brennan and Schwartz (1980) model is superior to others when the 1979-1982 period is included in the data, whereas the Cox, Ingersoll and Ross (CIR, 1985) model is the best for data excluding the 1979-1982 period. This last finding suggests that the decision to allow or not to allow for a structural break can have a statistically and economically significant impact on the short-rate volatility estimation and model selection.

2.1 Introduction

The short-term interest rate is important in many financial economics models, such as models of the term structure of interest rates, bond pricing models, and derivatives security pricing models. It also plays an important role in the development of tools for effective risk management and in many empirical studies analyzing term premiums and yield curves.

As a first step in modeling short-term interest rates, one-factor models of the term structure of interest rates are widely discussed by both academic researchers and practitioners. Some of the examples are Merton (1973), Vasicek (1977), Cox, Ingersoll, and Ross (1980), (1985), Dothan (1978), Brennan and Schwartz (1980), and Cox (1975). However, until relatively recently, these models have not been formally compared. Chan, Karolyi, Longstaff, and Sanders (1992) derive a model, which nests

these well-known interest rate models. They apply the generalized method of moments (GMM) to estimate the parameters and compare these models to explain the U.S. 1-month Treasury bill yields. They conclude that models, which allow the conditional volatility of interest rate changes to be highly dependent on the level of the interest rate, capture the dynamic behavior of short-term interest rates more successfully. Furthermore, because between October 1979 and September 1982, the U.S. Federal Reserve conducted an experiment in targeting monetary aggregates rather than targeting interest rate levels, CKLS also perform a structural break test over the 1964-1989 sample period, and conclude that there is no evidence of a structural regime shift after October 1979.

Bliss and Smith (1998) also use GMM to reexamine the CKLS model by redefining the possible regime shift period. They find that there is strong evidence of a structural break over the 1979-1982 period. Furthermore, they find evidence that, contrary to CKLS's claim, a moderate elastic interest rate process can capture the dependence of volatility on the level of interest rates, while highly elastic models cannot. In particular, this study finds support for the square-root CIR model.

As far as the estimation methodology is concerned, GMM has been the main tool for most of the empirical studies. Interestingly, Nowman (1997) applies the Gaussian estimation techniques developed by Bergstrom (1983, 1985, 1986, 1990) for continuous time stochastic differential equations on both British and U.S. data, and he finds that the volatility of the short-term interest rate is not highly sensitive to the level of interest rates in the United Kingdom, whereas it is in the United States. Episcopos (1999) also uses the Gaussian methodology to examine the stochastic behavior of the 1-month interbank rate in ten countries, and finds that the constant elasticity variance (CEV) model

outperforms other competing models, and the estimate of elasticity of interest rate volatility parameter is also much lower than the CKLS study suggests.

This paper contributes to the literature by re-examining the CKLS model using a panel-data approach, which takes into account the dynamics of the interest rates and the shape of the yield curve simultaneously. The objective of this paper is to examine the performance of alternative models using different datasets. Papers using the panel-data approach are Chen and Scott (1993), Duffie and Singleton (1997), and Geyer and Pichler (1999). All of these authors work within the exponential-affine framework, in which a closed-form solution for the term structure can be derived. My panel-data approach, based on Jones and Wang (1996), however, can be used to examine a wide variety of term-structure models, including non-linear interest rate models. I assume that all zero-coupon yields are observed with measurement errors. However, I impose some linear restrictions on these errors. That is, the linear combinations of the measurement errors are zero at each point in time. By imposing these restrictions, the state variables can be uniquely identified, and the exact likelihood function can be derived.

The paper's main conclusion is that, depending on whether one includes the 1979—1982 data, the Brennan and Schwartz and CIR models are superior to other competing models, respectively. The second result is that there is strong evidence of a regime shift during the Federal Reserve Experiment period. The third one is that the volatility in interest rates is not as sensitive to the level of interest rates as stated in the CKLS paper.

The rest of this paper is organized as follows. Section 2.2 describes the short-term interest rate models examined in the paper. Section 2.3 derives the estimation methodology. Section 2.4 describes the numerical implementation. Section 2.5 examines the properties of the panel-data approach with a Monte Carlo study. Section

2.6 presents the empirical results by first re-estimating the CKLS model and all alternative restricted models, then by testing for the structural break by using different definitions of the regime shift period. Section 2.7 summarizes the paper.

2.2 The one-factor CKLS model

CKLS present the following general model for short-term interest rates

$$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P, \quad (2.1)$$

where dw_t^P is a standard Brownian motion, and α , β , σ and γ are unknown parameters. I consider a continuous trading economy with a trading interval $[0, \tau]$. The uncertainty in the economy is represented by a filtered probability space $(\Omega, P, \mathcal{S}, (\mathcal{S}_t)_{0 \leq t \leq \tau})$.⁹ In this model, r_t moves towards the unconditional mean $-\alpha/\beta$, $-\beta$ measures the speed of the reversion, and γ determines the sensitivity of the variance to the level of r_t . The specification of equation (2.1) allows for a possible nonlinear diffusion term. The most noteworthy feature of this model is that many well-known interest rate models can be derived from the above model by imposing restrictions on the values of α , β , σ and γ . Table 2.1 summarizes the specifications and the corresponding parameter restrictions.

The Merton (1973) model is simply a Brownian motion with drift. Model 2 is an Ornstein-Uhlenbeck process used by Vasicek (1977) in deriving a non-arbitrage model of discount bond prices. The model of Cox, Ingersoll, and Ross (1985) is frequently referred to as the square-root process. It has been used extensively in developing valuation models for interest-rate-sensitive contingent claims. Model 4 is used by Dothan

⁹ The detail description of the filtered probability space can be found in my first chapter, Section 1.2.

(1978) in valuing discounted bonds. The Geometric Brownian Motion (GBM) is used by Black and Scholes (1973) to derive the price of options. The model of Brennan and Schwartz (1980) is used to derive a numerical model for convertible bond prices. Finally, model 7 is the constant elasticity of variance (CEV) model, which is introduced by Cox (1975).

Table 2. 1 he relationship between alternative one-factor short-term interest rate models and parameter values in equation (2.1)

Model	Specification	Restrictions			
		α	β	σ	γ
Merton (1973)	$dr_t = \alpha dt + \sigma dw_t^P$		0		0
Vasicek (1977)	$dr_t = (\alpha + \beta r_t)dt + \sigma dw_t^P$				0
CIR (1985)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^{1/2} dw_t^P$				1/2
Dothan (1978)	$dr_t = \sigma r_t dw_t^P$	0	0		1
GBM	$dr_t = \beta r_t dt + \sigma r_t dw_t^P$	0			1
Brennan-Schwartz (1980)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t dw_t^P$				1
CEV (1975)	$dr_t = \beta r_t dt + \sigma r_t^\gamma dw_t^P$	0			
CKLS (1992)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$				

Typically the continuous-time model, equation (2.1), is discretized as follows

$$r_{t+1} - r_t = (\alpha + \beta r_t)\Delta + \eta_{t+1}, \quad (2.2)$$

$$E_t(\eta_{t+1}) = 0, \quad E_t(\eta_{t+1}^2) = \sigma^2 r_t^{2\gamma} \Delta. \quad (2.3)$$

Although the model in equation (2.2) neglects errors introduced as a result of time aggregation, Nowman (1997) finds out that the bias resulting from using this discrete approximation is very small. Therefore, in my Monte Carlo study, I use the above mechanism in equations (2.2) and (2.3) to generate r_t with different maturities and I use weekly observations, with $\Delta = 1/52$.

Methods to estimate this model are divided into two groups. The first is a time-series approach and the second is a panel-data approach. The time-series approach is based on equations (2.2) and (2.3), using a yield with short time to maturity as a proxy for r_t . Depending on the assumptions made on the error term, η , the parameters of the model are then estimated using either maximum likelihood methods, for example, Nowman (1997), or the general method of moments technique, for example, CKLS (1992), and Bliss and Smith (1998). The panel-data approach takes the cross-sectional aspect into account as well and additional information from the shape of the yield curve can be gained. Moreover, it enables identification of the parameter λ , the market price of risk. Specifically, the drift of the CKLS interest rate model under the Q-measure is given by $\mu(r(t)) = \alpha + \beta r(t) + \lambda \sigma r(t)^\gamma$. Some examples include Chen and Scott (1993), Pearson and Sun (1994), Duan and Simonato (1999), and Geyer and Pichler (1999).

In this paper, I focus on the panel-data method to estimate the CKLS model. I follow the idea of Jones and Wang (1996), where the measurement errors on the yields satisfy certain conditions so that the maximum likelihood method can be employed. Furthermore, since my model does not have closed-form solutions, numerical methods are used to calculate the zero-coupon yields. In the next sections, I describe the estimation method and how to implement the numerical method.

2.3 The estimation method

In this subsection I present a panel-data estimator, which is based on Jones and Wang's (1996) approach.

Suppose that I have discretely observed yields $Y_{i,t}$ at time $t = 1, \dots, T$ with maturities τ_i , $i = 1, \dots, N$. Given an interest rate model, a corresponding theoretical

value of the yield, $R_i(r_t) = R(r_t, \tau_i; \Phi)$, can be found by arbitrage principles. However, the observed yields are contaminated by measurement errors, which I assume to be additive. Hence, I have the following relationship

$$Y_t = R(r_t) + \varepsilon_t, \quad (2.4)$$

$$\varepsilon_t \sim N(0, \Sigma),$$

where Σ is a variance-covariance matrix of error term, which has an $N \times N$ dimension. To identify the state variable r_t at each point in time, Jones and Wang (1996) assume that linear combinations of the measurement errors are zero. That is,

$$\sum_{i=1}^N c_{1,i} \varepsilon_{i,t} = 0, \quad t = 1, \dots, T. \quad (2.5)$$

Here, the $c_{1,i}$'s are the constraint coefficients that provide the linkage between Y_t and r_t . Then the combination of equation (2.4) and (2.5) gives us a way to obtain the estimates of r_t ,

$$\sum_{i=1}^N c_{1,i} Y_{i,t} = \sum_{i=1}^N c_{1,i} R_i(\hat{r}_t), \quad t = 1, \dots, T. \quad (2.6)$$

Equation (2.4) is then modified as

$$Y_{(2),t} = R_{(2)}(r_t) + \varepsilon_{(2),t}, \quad (2.7)$$

$$\varepsilon_{(2),t} \sim N(0, \Sigma_{(22)}),$$

where $Y_{(2),t}$, $R_{(2)}(r_t)$ and $\varepsilon_{(2),t}$ denote the vectors with first coordinate eliminated, $\Sigma_{(22)}$ denotes the variance-covariance matrix Σ with the first row and column eliminated.

To simplify my expression later, I put equation (2.6) and (2.7) into matrix form, that is,

$$CY_t = CR(r_t) + C\varepsilon_t \Rightarrow \tilde{Y}_t = \tilde{R}(\hat{r}_t) + \tilde{\varepsilon}_t$$

$$\Rightarrow \begin{pmatrix} \tilde{Y}_{1,t} \\ Y_{(2),t} \end{pmatrix} = \begin{pmatrix} \tilde{R}_1(\hat{r}_t) \\ R_{(2)}(r_t) \end{pmatrix} + \begin{pmatrix} \tilde{\varepsilon}_{1,t} \\ \varepsilon_{(2),t} \end{pmatrix}, \quad (2.8)$$

where $\tilde{Y}_{1,t} = \sum_{i=1}^N c_{1,i} Y_{i,t}$, $\tilde{R}_1(\hat{r}_t) = \sum_{i=1}^N c_{1,i} R_i(\hat{r}_t)$, $\tilde{\varepsilon}_{1,t} = \sum_{i=1}^N c_{1,i} \varepsilon_{i,t} = 0$ and the matrix C is the linear transformation of the measurement errors

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,N} \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix}_{N \times N}. \quad (2.9)$$

Notice that the linear transformation matrix C must have full rank, and it allows for a wide variety of noise structures. For example, consider the constraint coefficients: $c_{1,1} = 1$, all other $c_{1,i}$'s are zero. This is the specification used by Chen and Scott (1993). An alternative choice could be: $c_{1,i} = 1$ for $i = 1, \dots, N$. This specification requires that the residuals for different maturities sum to zero at each point of time.

In the following, I can obtain the likelihood function given that the state variable has been specified. The density function of Y_t is given by

$$f(Y_t; \Phi) = |\det(C)| f(\tilde{Y}_t | \mathcal{S}_{t-1}; \Phi)$$

$$= |\det(C)| \times f(Y_{(2),t} | \tilde{Y}_{1,t}, \mathcal{S}_{t-1}; \Phi) \times f(\tilde{Y}_{1,t} | \mathcal{S}_{t-1}; \Phi)$$

$$= |\det(C)| \times f(Y_{(2),t} | \tilde{Y}_{1,t}, \mathfrak{S}_{t-1}; \Phi) \times f(\hat{r}_t | \mathfrak{S}_{t-1}; \Phi) \times |J_t|, \quad (2.10)$$

where Φ denotes the vector of parameters to be estimated. The first equality follows from the transformation of Y_t to $\tilde{Y}_t = CY_t$. The second equality is implied by the definition of conditional density. The third equality is the transformation from \hat{r}_t to $\tilde{Y}_{1,t}$, which explains the Jacobian term. That is

$$|J_t^{-1}| = \left| \frac{d\tilde{Y}_{1,t}}{dr_t} \right| = (c_{1,1}, c_{1,2}, \dots, c_{1,N}) \frac{dY_t}{dr_t} \Big|_{r_t = \hat{r}_t}. \quad (2.11)$$

The density function of \hat{r}_t conditional on \mathfrak{S}_{t-1} is approximated by a normal density, and the rest of yields are also normally distributed. The likelihood function for the sample of observations on the state variable at $t = 1, \dots, T$ is thus

$$L = \prod_{t=1}^T |\det(C)| f(Y_{(2),t}; R_{(2)}(\hat{r}_t), \Sigma_{(22)}) \times \prod_{t=2}^T f(\hat{r}_t; \hat{\mu}_{t|t-1}, \hat{Q}_{t|t-1}) \times |J_t|, \quad (2.12)$$

where

$$f(Y_{(2),t}; R_{(2)}(\hat{r}_t), \Sigma_{(22)}) = (2\pi)^{-(N-1)/2} |\Sigma_{(22)}|^{-1/2} e^{(Y_{(2),t} - R_{(2)}(\hat{r}_t))' \Sigma_{(22)}^{-1} (Y_{(2),t} - R_{(2)}(\hat{r}_t)) / 2},$$

$$f(\hat{r}_t; \hat{\mu}_{t|t-1}, \hat{Q}_{t|t-1}) = (2\pi \hat{Q}_{t|t-1})^{-1/2} e^{(\hat{r}_t - \hat{\mu}_{t|t-1})^2 / 2 \hat{Q}_{t|t-1}},$$

and the corresponding log-likelihood function is

$$\log L(\Phi) = \sum_{t=1}^T \log(f(Y_{(2),t}; R_{(2)}(\hat{r}_t), \Sigma_{(22)})) + \sum_{t=2}^T \log(f(\hat{r}_t; \hat{\mu}_{t|t-1}, \hat{Q}_{t|t-1})) + \sum_{t=2}^T \log |J_t|, \quad (2.13)$$

where $\hat{\mu}_{t|t-1} = \hat{r}_{t-1} + (\alpha + \beta \hat{r}_{t-1})\Delta$, $\hat{\sigma}_{t|t-1} = \sigma^2 \hat{r}_{t-1}^{2\gamma} \Delta$, $\Delta = t - (t-1)$ and $|\det(C)| = 1$.¹⁰

2.4 Numerical implementation

In this paper I use Crank-Nicholson method to calculate zero-coupon bond prices since I do not have a closed-form solution for them with a non-affine interest rate model. In this section, I briefly describe how the Crank-Nicholson method works and how the log-likelihood function is calculated numerically.

2.4.1 Finite difference methods

Three different varieties of finite difference methods have been employed in solving partial differential equations (PDEs) numerically. They include the explicit finite difference method, the fully implicit finite difference method, and the Crank-Nicholson method. All of these three methods work backwards from maturity T to current time. In this paper, I use the Crank-Nicholson method to calculate the zero-coupon bond price. This method is similar to implementing the implicit finite difference method, but it has faster convergence than either the explicit or implicit method.

To illustrate the Crank-Nicholson method, I consider how it is used to value a zero-coupon bond. Suppose that the spot rate process follows

$$dr = \mu(r,t)dt + \sigma(r,t)dw_t^P, \quad (2.14)$$

where dw_t^P is a standard Brownian motion. The functions $\mu(r,t), \sigma^2(r,t)$ are the instantaneous drift and variance, respectively, of the process $r(t)$.

¹⁰ The only condition on the transformation matrix is that it must be non-singular, and choices on $c_{1,i}, i = 1, \dots, N$ are arbitrary. For simplification, I assume $c_{1,1} = 1$ without loss of generality.

The PDE¹¹ that the bond price must satisfy is

$$\frac{\partial P(r,t,T)}{\partial t} + (\mu(r,t) + \lambda\sigma(r,t))\frac{\partial P(r,t,T)}{\partial r} + \frac{1}{2}\sigma^2(r,t)\frac{\partial^2 P(r,t,T)}{\partial r^2} - rP(r,t,T) = 0 \quad (2.15)$$

with boundary condition $P(r,T,T) = 1$,

where $P(r,t,T)$ is the price of a discount bond at time t with a maturity T , λ is the market price of risk.

Numerically solving equation (2.15) involves the following steps. First, I impose a 2-dimensional grid in spot rate-time (r, t) space. Let $i = 0, 1, 2, \dots, N$ index the state variable (the spot rate) axis, and $n = 0, 1, 2, \dots, T$ index the time axis, where N and T are the numbers of intervals respectively in each axis. I assume the distance between adjacent points on the state variable axis is h , while that on the time axis is k . The price of a zero-coupon bond at any grid point (i, n) is written as $P_{i,n}$. Second, I approximate the partial derivatives of P at each gridpoint. The approximations to the partial derivatives are given by

$$P = \frac{P_{i,n} + P_{i,n+1}}{2},$$

$$\frac{\partial P}{\partial t} = \frac{P_{i,n+1} - P_{i,n}}{k},$$

$$\frac{\partial P}{\partial r} = \frac{P_{i+1,n} - P_{i-1,n} + P_{i+1,n+1} - P_{i-1,n+1}}{4h},$$

¹¹ The derivation of the PDE for a zero-coupon bond can be found in Vasicek (1977).

$$\frac{\partial^2 P}{\partial r^2} = \frac{P_{i+1,n} - 2P_{i,n} + P_{i-1,n} + P_{i+1,n+1} - 2P_{i,n+1} + P_{i-1,n+1}}{2h^2}.$$

Third, I substitute the above expressions into the equation (2.15), and collect all the terms involving the unknown $P_{\cdot,n+1}$ on the left side and known $P_{\cdot,n}$'s on the right side. Note that the $P_{\cdot,n+1}$'s cannot individually be written as simple linear combinations of the $P_{\cdot,n}$'s, but are simultaneously determined as the solution to this system of linear equations.¹² Fourth, I then move backwards and recursively solve for P for the entire grid. Once I have these P values taking on a grid of r , and since P is assumed to be smooth almost everywhere, I can interpolate within this grid to get values for arbitrary (r, t) .

2.4.2 Optimization

In this subsection, I briefly describe how to numerically calculate the log-likelihood function. The procedure involves the following steps. First, based on the discussion above, the Crank-Nicholson method can be used to find the zero-coupon bond prices at each gridpoint of r , given a certain parameter values Φ . Then the zero-coupon yields are calculated at these discrete points of r , but I require the full function of $R(r)$. Therefore, I use a cubic interpolation for R . Second, the Newton-Raphson method is employed to solve equation (2.6) for \hat{r}_t . The Jacobian term $|J|$ can then be easily calculated on the basis of the first derivative of the cubic interpolation for each maturity evaluated at \hat{r}_t . Third, notice that I need to estimate Σ in addition to Φ , but with Σ , there

¹² Substitution of the difference expressions into the equation (15) only gives us a linear equation for each interior point in the grid. That gives $N - 1$ equations at each time step, which is not sufficient to determine the $N + 1$ unknowns. The missing two equations must be provided by boundary conditions applied at each time step.

are $N(N-1)/2$ additional parameters to be estimated, this is far too many to estimate simultaneously. To simplify my estimation procedure, given the parameter values Φ , I estimate the $\Sigma_{(22)}$ by the following expression

$$\hat{\Sigma}_{(22)}(\Phi) = \frac{1}{T} \sum_{t=1}^T (Y_{(2),t} - R_{(2)}(\hat{r}_t))(Y_{(2),t} - R_{(2)}(\hat{r}_t))'.$$

Fourth, I build the log-likelihood function in equation (2.13) and iterate Φ until $-\log L$ converges to a minimum.¹³

2.5 Simulation analysis

Before I get to the empirical application, to validate the estimation procedure, check for errors and explore the characteristics of the Jones and Wang method, I apply the estimation procedure to simulated histories of interest rates. The simulated N -dimensional yield vectors Y_t for $t = 1, \dots, T$ are obtained in a three-step procedure.

First, I generate time series for the state variable, r_t , for $t = 1, \dots, T$, based on equation (2.2) and (2.3), which takes the form

$$r_{t+\Delta} = r_t + (\alpha + \beta r_t)\Delta + \sigma r_t^\gamma \Delta W_t,$$

where the time interval $\Delta = 1/52$ and $\Delta W_t \sim iidN(0, \Delta)$. For my Monte Carlo studies $T = 500$.

¹³ The multi-dimensional minimization algorithm I use in this paper is called Powell's method (see, for example, Numerical Recipes for details of this algorithm)

Second, the theoretical yields, $R(r_t)$, are computed on the basis of the simulated path of r and the selected maturities for the zero-coupon bonds. Since there is no close-form solution for the CKLS model, the numerical method has to be used to find the zero-coupon yields. This is done by combining the Crank-Nicholson method and the cubic interpolation between grid points.

Third, the yield curve noise, ε_t , is simulated. This noise is added to the theoretical yields to give simulated yield curve observations, Y_t , for $t = 1, \dots, T$. Notice that my estimation procedure assumes that the state variable is completely determined from the yield curve and model parameters. The linkage between the state variable and the yield curve is established through the following linear constraint

$$\sum_{i=1}^N c_{1,i} \varepsilon_{i,t} = 0,$$

and my simulation of ε_t must satisfy this constraint. Throughout the Monte Carlo studies in this paper, I assume that $c_{1,i} = 1$, $i = 1, \dots, N$ and all the errors of the observed yields behave identically. This means that $\text{var}(\varepsilon_{t,i}) = \sigma_\varepsilon^2$, and $\text{cov}(\varepsilon_{t,i}, \varepsilon_{t,j}) = -\sigma_\varepsilon^2 / (N - 1)$ for $i \neq j$.

The Monte Carlo results are shown in Table 2.2 where the average estimates of 500 simulated samples are reported as the panel-data estimator. In all the simulations I assume weekly observation, $\Delta=1/52$. To check for robustness of the Jones and Wang method, I pick four different gamma values, that is, $\gamma = 0, 0.5, 1, 1.3$. The average estimates based on a time-series approach are also included to compare it with the panel-data approach.

Table 2. 2 Monte Carlo results with different values for γ

	α	β	σ	γ
Model 1				
True	0.016	-0.2	0.02	0
Time-series	0.01717 (0.01228)	-0.26919 (0.21566)	0.02019 (0.00246)	-0.01770 (0.10461)
Panel-data	0.01531 (0.00501)	-0.18350 (0.02294)	0.01964 (0.00015)	0.00008 (0.00126)
Model 2				
True	0.016	-0.2	0.07	0.5
Time-series	0.02036 (0.01218)	-0.29979 (0.26062)	0.07483 (0.00545)	0.49208 (0.03141)
Panel-data	0.01611 (0.00282)	-0.20522 (0.05382)	0.06971 (0.00049)	0.50015 (0.00126)
Model 3				
True	0.016	-0.2	0.25	1
Time-series	0.02270 (0.01449)	-0.30497 (0.28744)	0.27679 (0.04027)	0.97653 (0.07499)
Panel-data	0.01410 (0.00166)	-0.18453 (0.03303)	0.24255 (0.00562)	1.00037 (0.00140)
Model 4				
True	0.016	-0.2	0.5	1.3
Time-series	0.02349 (0.01584)	-0.29740 (0.28694)	0.57241 (0.11588)	1.26100 (0.11603)
Panel-data	0.01480 (0.00131)	-0.19395 (0.01262)	0.48823 (0.00937)	1.30038 (0.00155)

The general model to be estimated is given by: $dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$. These four models are distinguished by different values for γ . In each category, I report results from both the time-series approach and the panel-data approach. The numbers in parentheses are the standard errors of the estimated average values. The results are based on 500 simulated samples each with 500 observations of zero-coupon yields with $\lambda=0$. The measurement errors on all the points are set to 10 basis points. The time interval between two observations is one week. Nine different maturities are used: 1-month, 3-month, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year. 1-month simulated data is used to approximate the spot rate in the time-series approach.

The results of Table 2.2 indicate that it is necessary to use the panel-data approach to obtain reliable estimates of the parameters, especially the estimate of the drift. The average estimates of the drift based on the time-series approach are much less precise compared to the estimates obtained from the panel-data method. On the other hand, the estimates of the volatility parameters are less affected by the estimation method. Table 2.2 also shows that the Jones and Wang method is considerably robust with respect to different interest rate models. To find out how the size of the

measurement error, σ_ε , affects the estimates, I also report the estimates with different sizes of σ_ε in Table 2.3. It shows that they are unaffected by the size of σ_ε . Thus, I conclude that the estimates of the parameters are improved by combining time-series and cross-sectional information.

Table 2. 3 Sensitivity of the estimates with respect to different values of σ_ε

	α	β	σ	γ
Model 1				
True	0.01600	-0.20000	0.02000	0.00000
1bp	0.01538 (0.00499)	-0.18337 (0.02298)	0.01963 (0.00014)	0.00008 (0.00122)
10bps	0.01531 (0.00501)	-0.18350 (0.02294)	0.01964 (0.00015)	0.00008 (0.00126)
20bps	0.01526 (0.00511)	-0.18355 (0.02295)	0.01964 (0.00015)	0.00008 (0.00131)
Model 2				
True	0.01600	-0.20000	0.07000	0.50000
1bp	0.01610 (0.00284)	-0.20512 (0.05384)	0.06970 (0.00048)	0.50016 (0.00122)
10bps	0.01611 (0.00282)	-0.20522 (0.05382)	0.06971 (0.00049)	0.50015 (0.00126)
20bps	0.01612 (0.00285)	-0.20517 (0.05347)	0.06971 (0.00049)	0.50014 (0.00129)
Model 3				
True	0.01600	-0.20000	0.25000	1.00000
1bp	0.01511 (0.00168)	-0.19069 (0.02753)	0.24622 (0.00533)	1.00022 (0.00133)
10bps	0.01510 (0.00166)	-0.18453 (0.03303)	0.24255 (0.00562)	1.00037 (0.00140)
20bps	0.01513 (0.00170)	-0.19030 (0.02910)	0.24630 (0.00561)	1.00021 (0.00145)
Model 4				
True	0.01600	-0.20000	0.50000	1.30000
1bp	0.01556 (0.00134)	-0.19489 (0.01011)	0.49610 (0.00843)	1.30016 (0.00136)
10bps	0.01480 (0.00131)	-0.19395 (0.01262)	0.48823 (0.00937)	1.30038 (0.00155)
20bps	0.01557 (0.00132)	-0.19492 (0.01096)	0.49613 (0.00885)	1.30019 (0.00148)

The unrestricted model to be estimated is given by: $dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$. These four models are distinguished by different values for γ . In each category, I report the true values of the parameters and the average values of the panel-data estimates with different sizes of

measurement error on the yields, σ_ε , where $\text{var}(\varepsilon_{t,i}) = \sigma_\varepsilon^2$, and $\text{cov}(\varepsilon_{t,i}, \varepsilon_{t,j}) = -\sigma_\varepsilon^2 / (N - 1)$ for $i \neq j$. The numbers in parentheses are the standard errors of the estimated average values. The results are based on 500 simulated samples each with 500 observations of zero-coupon yields with $\lambda=0$. The time interval between two observations is one week. Nine different maturities are used: 1-month, 3-month, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year.

2.6 Empirical application

2.6.1 Data description

The data set consists of 1246 weekly observations of the yield curve for the US government Treasury Bills and Bonds from February 1977 to December 2000, obtained from the FRED database. The original FRED data are quoted daily. For my purpose, I pick each Wednesday quote as an approximation of the weekly data. If the Wednesday data are not available, I use the Thursday data instead, and so forth. The observations are not actual price quotes, but rather are estimates of the coupon that would be required for bonds of various maturities to trade at par. Eight different maturities are considered, they are 3-month, 6-month and 1-year simple interest rates, and 2-, 3-, 5-, 7- and 10-year semi-annual compounded par-coupon yields. For my estimation purpose, I transfer these semi-annual compounded par-coupon yields into continuously compounded zero-coupon yields.¹⁴ To get a feel for the data, I provide in Figure 2.1 the movements of the Treasury yield curves for the entire sample period, and in Table 2.4 the summary statistics for the yields.

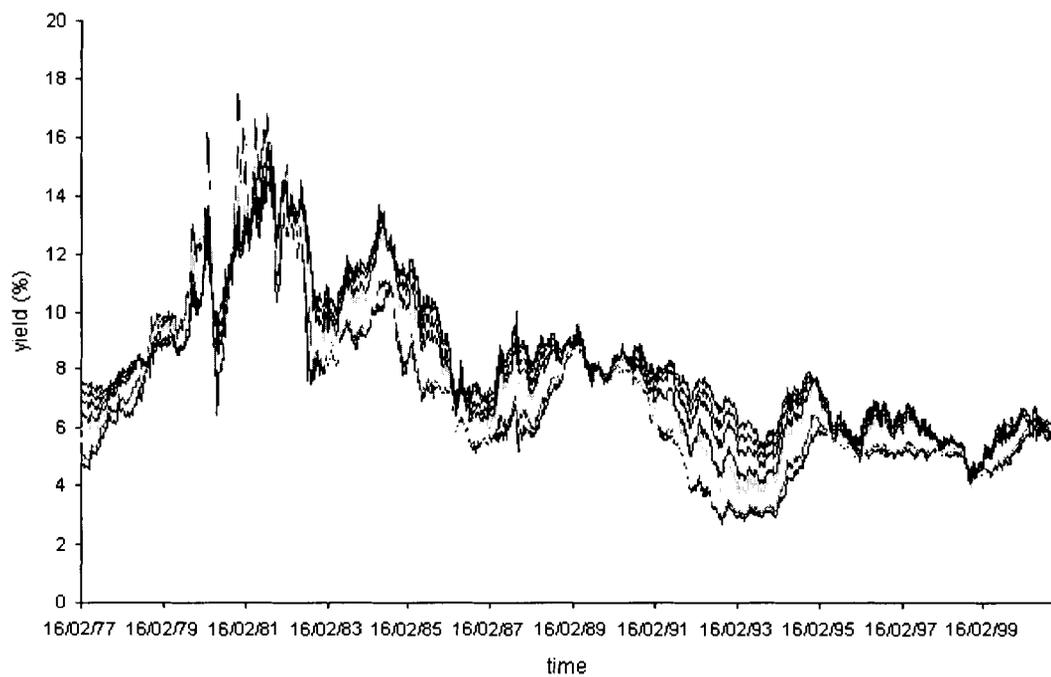
¹⁴ The details of transformation can be found in my first chapter, Appendix A.

Table 2. 4 Summary statistics for Treasury yields

Maturity	0.25	0.5	1	2	3	5	7	10
Mean	0.071	0.073	0.076	0.079	0.082	0.085	0.089	0.092
Std	0.029	0.029	0.031	0.027	0.026	0.025	0.024	0.023
Maximum	0.176	0.168	0.182	0.162	0.158	0.155	0.152	0.149
Minimum	0.027	0.029	0.030	0.037	0.040	0.041	0.043	0.043
Skewness	1.206	1.106	1.196	0.892	0.890	0.862	0.808	0.721
Kurtosis	1.323	0.933	1.209	0.223	0.159	-0.005	-0.160	-0.292
$\rho_1(r_t)$	0.994	0.995	0.996	0.996	0.996	0.997	0.997	0.997
$\rho_1(r_t - r_{t-1})$	0.087	0.101	0.138	0.115	0.096	0.082	0.065	0.057

This table presents the sample means, standard deviations (Std), maximum, minimum, skewness, kurtosis, the first autocorrelation of the yields, and the first autocorrelation of the first difference of the yields. The sample contains 1246 weekly observations in the period from February 1977 to December 2000.

Figure 2. 1 Time series of U.S. interest rates.



This graph shows the movement of U.S. interest rates with 3m-, 6m-, 12m-, 2-, 3-, 5-, 7-, and 10-year maturities in the period from February 16, 1977 to December 28, 2000.

2.6.2 Empirical results

In this section, I present my empirical results. I begin by estimating the unrestricted and the seven restricted interest rate processes with the panel-data approach. Then I compare my results with CKLS (1992) and Bliss and Smith (1998). My analysis consists of two components: first, I assume that the agents are risk neutral, which implies that the market price of risk, λ , is zero; second, I assume that the agents are risk averse with a non-zero λ . Table 2.5 presents the estimates of these eight models with $\lambda = 0$.

The results presented in Table 2.5 lead to the following observations. First, I find that the parameter γ , which measures the degree of volatility dependence on the level of interest rates, is statistically significant in the unrestricted model. Moreover, the estimate of γ in this model is 0.5041, which is much less than unity, comparing to 1.4999 of the CKLS result. This discrepancy could be due to the different data sets or the different estimation methods used, or both. Second, CKLS claim that there is only weak evidence of mean reversion in the short-term rate, as the parameter estimate of β is insignificant in the unrestricted model. My results in Table 2.5 confirm their findings. Third, my likelihood ratio test shows that two models, CIR and Brennan-Schwartz models are not rejected at the 5% significance level in their unrestricted forms.

There are some additional observations worth noting about the restricted models. Whenever γ is constrained, the t-statistics of σ are generally higher compared with those of α and β . Interpreted loosely, this result shows that the volatility coefficient is the second most important parameter after γ .

Table 2. 5 Estimates of alternative models for the short-term interest rate ($\lambda = 0$)

Parameters				LLF	χ^2 -Stat.	d.f
α	β	σ	γ			
CKLS						
0.00206 (0.00224)	-0.00328 (0.01726)	0.05383 (0.00592)	0.50410 (0.05302)	61351.81		
Merton						
0.00250 (0.00717)	0.0	0.01832 (0.00035)	0.0	61102.66	498.30	2
Vasicek						
0.00378 (0.00302)	-0.01894 (0.02422)	0.01786 (0.00063)	0.0	61210.71	282.20	1
CIR						
0.00241 (0.00310)	-0.00804 (0.03049)	0.05390 (0.00233)	0.5	61350.23	3.16	1
Dothan						
0.0	0.0	0.01273 (0.00007)	0.0	60603.42	1496.78	3
GBM						
0.0	0.00000 (0.00000)	0.13620 (0.00120)	1.0	61066.50	570.63	2
BS						
0.00222 (0.00074)	-0.01404 (0.00645)	0.15310 (0.00186)	1.0	61350.61	2.40	1
CEV						
0.0	0.00000 (0.00010)	0.13890 (0.01030)	1.02200 (0.03318)	61066.69	570.24	1

The unrestricted model to be estimated is given by: $dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$. The alternative models are distinguished by the restrictions on the parameter values. The estimation results are based on weekly 3-, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year yields. The numbers in parentheses are standard errors of the estimated parameters.¹⁵ The identification of r is based on $c_1 = c_2 = \dots = c_8 = 1$. The risk premium, λ , is set to zero.

¹⁵ The estimated standard errors of the estimated parameters are calculated based on the method of Berndt, Hall, Hall and Hausmann (the BHHH estimator) as given in Greene (1997, P.139, eq. 4-52). Details can be found in my first chapter, Section 1.7.2.

Table 2. 6 Estimates of alternative models for the short-term interest rate ($\lambda \neq 0$)

Parameters					LLF	χ^2 -Stat.	d.f
α	β	σ	γ	λ			
CKLS							
0.0013 (0.0015)	-0.0185 (0.0551)	0.2625 (0.0299)	1.1860 (0.0536)	0.1470 (1.3210)	61472.86		
Merton							
0.0001 (0.0065)	0.0	0.0182 (0.0003)	0.0	0.1282 (0.3954)	61102.78	740.16	2
Vasicek							
0.0017 (0.0068)	-0.0193 (0.0357)	0.0183 (0.0011)	0.0	0.1179 (0.5512)	61212.16	521.40	1
CIR							
0.0012 (0.0050)	-0.0128 (0.0633)	0.0554 (0.0046)	0.5	0.1086 (0.6793)	61359.12	227.48	1
Dothan							
0.0	0.0	0.0183 (0.0002)	0.0	0.1366 (8.0360)	61102.70	740.32	3
GBM							
0.0	0.0000 (0.0243)	0.1842 (0.0084)	1.0	0.1346 (2.5040)	61455.8	34.12	2
BS							
0.0008 (0.0017)	-0.0100 (0.0628)	0.1760 (0.0107)	1.0	0.1309 (2.7050)	61471.24	1.24	1
CEV							
0.0	-0.0401 (0.0609)	0.1990 (0.0224)	1.0430 (0.0501)	0.3559 (0.8680)	61456.11	33.50	1

The unrestricted model to be estimated is given by: $dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$. Under the Q -measure, the drift changes to $\mu^Q(r(t)) = \alpha + \beta r(t) + \lambda \sigma r(t)^\gamma$.¹⁶ The alternative models are distinguished by the restrictions on the parameter values. The estimation results are based on weekly 3-, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year yields. The numbers in parentheses are

¹⁶ Note that from the CKLS model, the drift under the Q -measure for the CIR model is $(\alpha + \beta r(t) + \lambda \sigma r(t)^{1/2})$, whereas the drift derived from the CIR paper is $(\alpha + \beta r(t) + \lambda^* r(t))$. In this paper, I actually estimate λ for the CIR model.

standard errors of the estimated parameters. The identification of r is based on $c_1 = c_2 = \dots = c_8 = 1$. The risk premium, λ , is assumed to be nontrivial.

One of the advantages of the panel-data approach is that it allows us to estimate the risk premium as well, which is necessary for valuation purposes. Table 2.6 presents the results with the non-zero risk premium, λ . It shows that we also get a significant estimate of γ but with a much higher value, $\gamma = 1.1860$, compared to 0.5041 in the previous case without λ . However, this value is still smaller than 1.5, which is obtained by CKLS in their 1992 paper. The likelihood ratio test shows that all the restricted models except the Brennan and Schwartz model are rejected at the 5% significance level. The discussion of the goodness of fit of this model can be found in Appendix. Recall that the Brennan and Schwartz model also cannot be rejected in the case of $\lambda = 0$. All these results provide some supporting evidence of this model at this point. Turning to the unrestricted model in the two cases with and without λ , I find that the null hypothesis of $\lambda=0$ is strongly rejected by applying the likelihood ratio test. This result leads us to use the unrestricted model with a non-zero λ to carry out the regime shift test.

2.6.3 Structural breaks

Many empirical studies of the term structure have concluded that the shift in Federal Reserve monetary policy in October 1979 resulted in a structural break in the interest rate process. CKLS (1992) test for the structural break by adding a dummy variable in their model. They assume that the change of monetary policy has permanent effect on the interest rate process so that they allow for the dummy variable equaling unity for all observations following October 1979 and zero otherwise. They conclude that there is no evidence of a structural break after October 1979. Bliss and Smith (1998),

however, by redefining the regime period, lead to a reverse conclusion. They assume that the structural shift period is temporary and coincides with the Federal Reserve Experiment of October 1979 through September 1982. They find that there is strong evidence of a structural break.

My paper follows Bliss and Smith's definition of the regime period, but using the panel-data approach to re-examine the CKLS model. The test criterion used is the likelihood ratio test. The restricted model is the one with no structural break. Then I use the whole dataset to implement the estimation and get $L(\hat{\Phi}_r)$, where $L(\cdot)$ denotes the log-likelihood function evaluated at the maximum, $\hat{\Phi}_r$, the restricted parameter estimates. The unrestricted model allows different parameter estimates for different regimes, that is, it has two components. One uses dataset between 1979-1982, the other uses dataset from 1977 to 1979 and from 1982 to 2000. Then I get two log-likelihood functions: $L(\hat{\Phi}_{u,79-82})$ and $L(\hat{\Phi}_{u,77-79,82-00})$. The likelihood ratio test statistics are constructed as follows:

$$-2(L(\hat{\Phi}_r) - (L(\hat{\Phi}_{u,79-82}) + L(\hat{\Phi}_{u,77-79,82-00}))) \sim \chi^2(f),$$

where f is degree of freedom, which equals the number of restrictions. The results are presented in Table 2.7.

Table 2.7 presents summaries of the temporary regime shift tests by using the CKLS model. The hypothesis that there was a regime shift, is supported by this data. I also note that the data, which includes the 1979-1982 period, support a relatively strong relationship between the volatility and the level of interest rates, $\gamma = 1.1860$, whereas if this period is excluded, a much weaker relationship is suggested with $\gamma = 0.4696$.

Table 2. 7 Tests for regime shift from 1979.10 through 1982.09

Parameters					LLF	χ^2 -Stat.	d.f.
α	β	σ	γ	λ			
CKLS Unrestricted (1977.2 – 1979.9 and 1982.10 – 2000. 12)							
0.0013 (0.0026)	-0.0080 (0.0334)	0.0359 (0.0062)	0.4696 (0.0694)	0.0521 (4.1950)	55581.31		
CKLS Unrestricted (1979.10 – 1982.9)							
0.0530 (0.0693)	-0.4457 (0.4508)	0.2921 (0.1713)	0.7721 (0.3946)	0.1153 (2.6780)	7100.084		
CKLS Restricted (1977.2 – 2000.12)							
0.0013 (0.0015)	-0.0185 (0.0551)	0.2625 (0.0299)	1.1860 (0.0536)	0.1470 (1.3210)	61472.86	1208.534	5

The estimation results are based on weekly 3-, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year yields. The numbers in parentheses are standard errors of the estimated parameters. The identification of r is based on $c_1 = c_2 = \dots = c_8 = 1$. The risk premium, λ , is assumed to be nontrivial ($\lambda \neq 0$). The unrestricted model has two components: one using data from 1977.02 through 1979.09 and from 1982.10 to 2000.12, the other using data from 1979.10 to 1982.09. The restricted model uses the whole dataset, that is, from 1977.02 to 2000.12.

The results of Table 2.7 suggest that model specification is quite sensitive to the data being used. Given the evidence from the CKLS model that there was a structural shift from October 1979 through September 1982, I next examine whether any of the restricted variants of the CKLS model are able to fit the data with the 1979-1982 period removed. The purpose of doing this is that if the Federal Reserve Experiment of October 1979 through September 1982 is treated as an anomalous event and unlikely to be repeated, I want to know what could be the best model to fit the data by removing this period. The results are shown in Table 2.8. One model, the CIR model, is not rejected at the 5% significance level. This result is consistent with that obtained by Bliss and Smith (1998). Their study also finds support for the square-root CIR process. The discussion of the goodness of fit of this CIR model can be found in Appendix.

**Table 2. 8 Estimates of alternative models for the short-term interest rate ($\lambda \neq 0$)
excluding 1979.10 –1982.09**

Parameters					LLF	χ^2 -Stat.	d.f
α	β	σ	γ	λ			
CKLS							
0.0013 (0.0026)	-0.0080 (0.0334)	0.0359 (0.0062)	0.4696 (0.0694)	0.0521 (4.1950)	55581.31		
Merton							
0.0003 (0.0049)	0.0	0.0108 (0.0002)	0.0	0.1060 (20.000)	55540.34	81.94	2
Vasicek							
0.0018 (0.0033)	-0.0091 (0.0345)	0.0111 (0.0006)	0.0	0.0274 (0.4025)	55552.33	57.96	1
CIR							
0.0012 (0.0027)	-0.0108 (0.0326)	0.0392 (0.0017)	0.5	0.0801 (4.5580)	55580.48	1.68	1
Dothan							
0.0	0.0	0.0108 (0.0001)	0.0	0.1301 (22.360)	55540.33	81.96	3
GBM							
0.0	0.0000 (0.0003)	0.1454 (0.0046)	1.0	0.1278 (3.6010)	55570.74	21.14	2
BS							
0.0008 (0.0014)	-0.0078 (0.0560)	0.1415 (0.0062)	1.0	0.1032 (3.8150)	55577.87	6.88	1
CEV							
0.0	0.0000 (0.0000)	0.0318 (0.0054)	0.4200 (0.0707)	0.1280 (6.1900)	55572.13	18.36	1

The unrestricted model to be estimated is given by: $dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dw_t^P$. Under the Q-measure, the drift changes to $\mu^Q(r(t)) = \alpha + \beta r(t) + \lambda \sigma r(t)^\gamma$. The alternative models are distinguished by the restrictions on the parameter values. The estimation results are based on weekly 3-, 6-month, 1-, 2-, 3-, 5-, 7-, and 10-year yields. The numbers in parentheses are standard errors of the estimated parameters. The identification of r is based on $c_1 = c_2 = \dots = c_8 = 1$. The risk premium, λ , is assumed to be nontrivial ($\lambda \neq 0$).

2.7 Conclusion

In this paper I have presented a panel-data approach to obtain the maximum likelihood estimates of the dynamics of interest rates. The estimation method does not depend on an analytical expression for the yields of zero-coupons. Thus, I am able to analyze a wide variety of term structure models. The Monte Carlo study demonstrates the efficiency of the panel-data approach compared to the traditional time-series approach.

I then use this panel-data approach to reexamine Chan, Karolyi, Longstaff, and Sanders (1992) and have three main findings. First, by defining the structural shift as coincident with the policy shift, I find that there is strong evidence of a structure break. Second, I find evidence that a relatively moderate- γ interest rate process can capture the dependence of volatility on the level of interest rates. That is, when the 1979-1982 data are included, γ is estimated to be 1.1860; when this period is excluded, the estimate of γ is 0.4696, compared to 1.5 in CKLS. Third, the Brennan and Schwartz (1980) model is superior in terms of data fit when the 1979-1982 period data is included, whereas the CIR (1985) model is the best without the 1979-1982 data. These last two findings suggest that the decision to allow or not to allow for a structural break can have a statistically and economically significant impact on the short-rate volatility estimation and model selection.

Chapter 2 Appendix

From Sections 2.6.2 and 2.6.3, I find that the Brennan and Schwartz (1980) model is superior in terms of data fit when the 1979-1982 period data is included, whereas the CIR (1985) model is the best without the 1979-1982 data. To examine further the quality of fit, given the parameter estimates, I construct residuals of the two models and report the estimated means, estimated standard errors and estimated correlation matrices of the residuals for both models.

The element of the estimated variance-covariance matrix of the residuals, $\hat{\sigma}_{ij}$, $i = 1, \dots, N$, and $j = 1, \dots, N$ can be calculated as

$$\hat{\sigma}_{ij} = \frac{\sum_{t=1}^T (e_{it} - \bar{e}_i)(e_{jt} - \bar{e}_j)}{T},$$

where N is the number of maturities, T is the number of observations, e_{it} is the residual at time t with maturity i , which equals $(Y_{it} - \hat{R}_{it})$, and \bar{e}_i is the mean of the residual with maturity i .

The estimated correlation matrix is then given by

$$\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i \hat{\sigma}_j},$$

where $\hat{\sigma}_i, \hat{\sigma}_j$ are square roots of the diagonal entries in the estimated variance-covariance matrix of the residual with maturities i and j , respectively.

Tables 2.9 and 2.10 show properties of the predicted errors from the Brennan and Schwartz (1980) model and the CIR (1985) model, respectively.

Furthermore, I also plot actual yields and predicted yields for 3-month T-bill, 5-year, and 10-year T-bonds from the Brennan and Schwartz model and the CIR model, respectively. These are shown in from Figure 2.2 to Figure 2.7. It can be seen from these graphs that the one-factor Brennan and Schwartz and CIR models do reasonably good job of fitting the short rates and medium rates in sample, even though these two models tend to overestimate the short rates and underestimate the medium rates.

Table 2. 9 Estimated means, standard errors and correlation matrix of the residuals for the Brennan and Schwartz model

Maturity <i>i</i>	Mean \bar{e}_i	Est. Std $\hat{\sigma}_i$	Correlation matrix ($\hat{\rho}_{ij}$)											
			3m	6m	1yr	2yrs	3yrs	5yrs	7yrs	10yrs				
3m	-0.0050	0.0078	1.0000											
6m	-0.0033	0.0064	0.9547	1.0000										
1yr	-0.0013	0.0075	0.7859	0.8833	1.0000									
2yrs	0.0007	0.0019	-0.1434	0.0286	0.0100	1.0000								
3yrs	0.0012	0.0027	-0.8577	-0.8154	-0.7812	0.4894	1.0000							
5yrs	0.0021	0.0049	-0.9439	-0.9848	-0.9226	-0.0087	0.8332	1.0000						
7yrs	0.0030	0.0065	-0.9190	-0.9797	-0.9339	-0.1052	0.7693	0.9835	1.0000					
10yrs	0.0027	0.0074	-0.8962	-0.9658	-0.8947	-0.2198	0.6854	0.9559	0.9812	1.0000				

\bar{e}_i is the mean of the residual with maturity *i*, $\hat{\sigma}_i$ represents the square roots of the diagonal entry in the estimated variance-covariance matrix of the residuals, and $\hat{\rho}_{ij}$ is the correlation between residuals. The residuals are the result of the difference between Y_t and \hat{R}_t , where Y_t is the actual yield and \hat{R}_t is the predicted yield estimated by using the Jones and Wang method.

Table 2. 10 Estimated means, standard errors and correlation matrix of the residuals for the CIR model

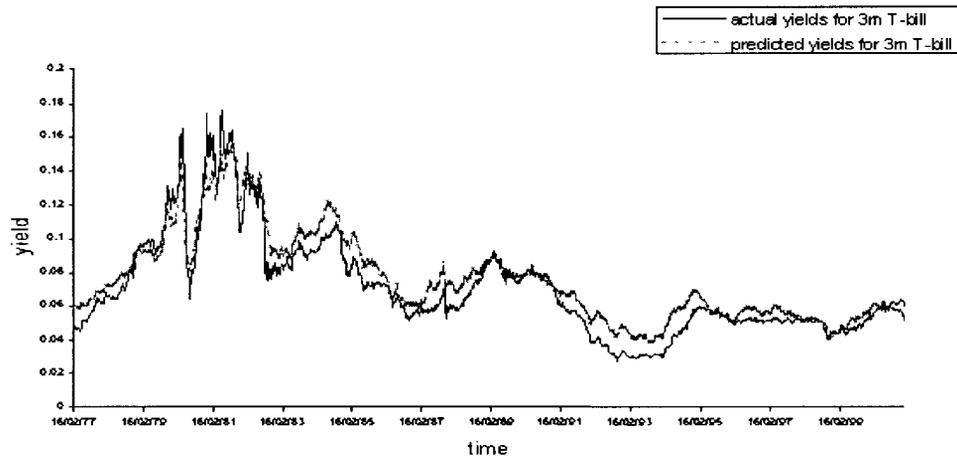
Maturity i	Mean \bar{e}_i	Est. Std $\hat{\sigma}_i$	Correlation matrix ($\hat{\rho}_{ij}$)												
			3m	6m	1yr	2yrs	3yrs	5yrs	7yrs	10yrs					
3m	-0.0066	0.0060	1.0000												
6m	-0.0048	0.0052	0.9583	1.0000											
1yr	-0.0033	0.0053	0.7517	0.8722	1.0000										
2yrs	0.0006	0.0017	0.1590	0.3175	0.3619	1.0000									
3yrs	0.0017	0.0018	-0.7656	-0.7163	-0.6577	0.3632	1.0000								
5yrs	0.0033	0.0039	-0.9370	-0.9827	-0.9144	-0.3240	0.7279	1.0000							
7yrs	0.0046	0.0051	-0.9103	-0.9768	-0.9232	-0.4384	0.6324	0.9776	1.0000						
10yrs	0.0045	0.0063	-0.8910	-0.9615	-0.8800	-0.5361	0.5341	0.9493	0.9804	1.000					

\bar{e}_i is the mean of the residual with maturity i , $\hat{\sigma}_i$ represents the square roots of the diagonal entry in the estimated variance-covariance

matrix of the residuals, and $\hat{\rho}_{ij}$ is the correlation between residuals. The residuals are the result of the difference between Y_i and \hat{R}_i , where Y_i

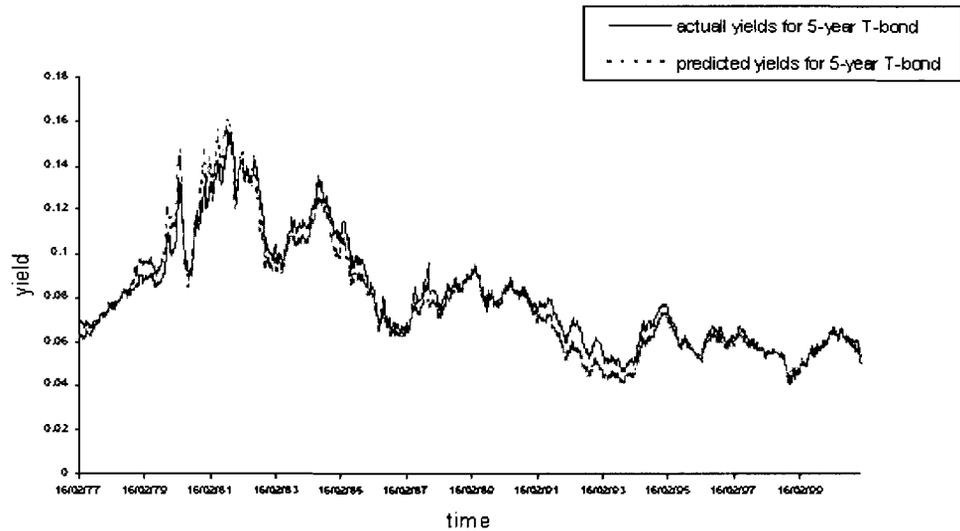
is the actual yield and \hat{R}_i is the predicted yield estimated by using the Jones and Wang method.

Figure 2. 2 Time series of a 3-month T-bill yield.



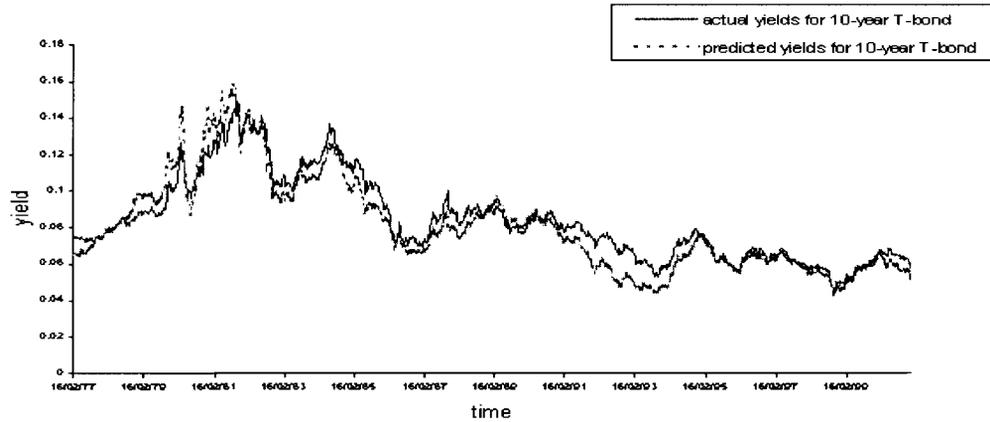
This Figure shows actual yields for a 3-month T-bill and predicted yields for the same T-bill from the one-factor Brennan and Schwartz model. Weekly data is used from 1977.02 to 2000.12.

Figure 2. 3 Time series of a 5-year T-bond yield.



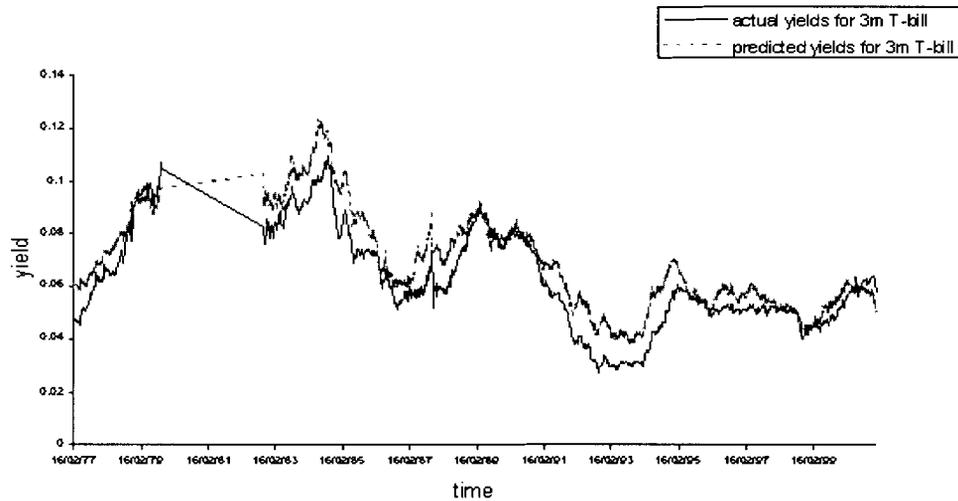
This Figure shows actual yields for a 5-year T-bond and predicted yields for the same T-bond from the one-factor Brennan and Schwartz model. Weekly data is used from 1977.02 to 2000.12.

Figure 2. 4 Time series of a 10-year T-bond yield.



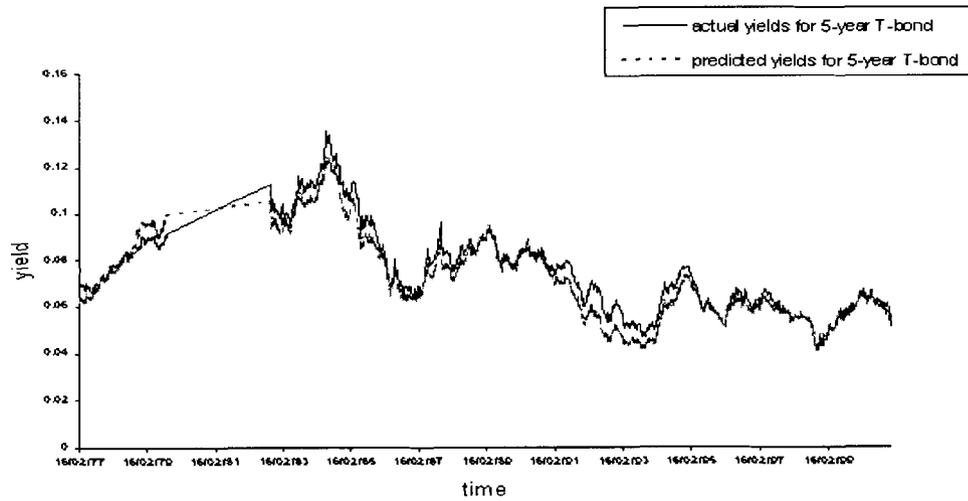
This Figure shows actual yields for a 10-year T-bond and predicted yields for the same T-bond from the one-factor Brennan and Schwartz model. Weekly data is used from 1977.02 to 2000.12.

Figure 2. 5 Time series of a 3-month T-bill yield.



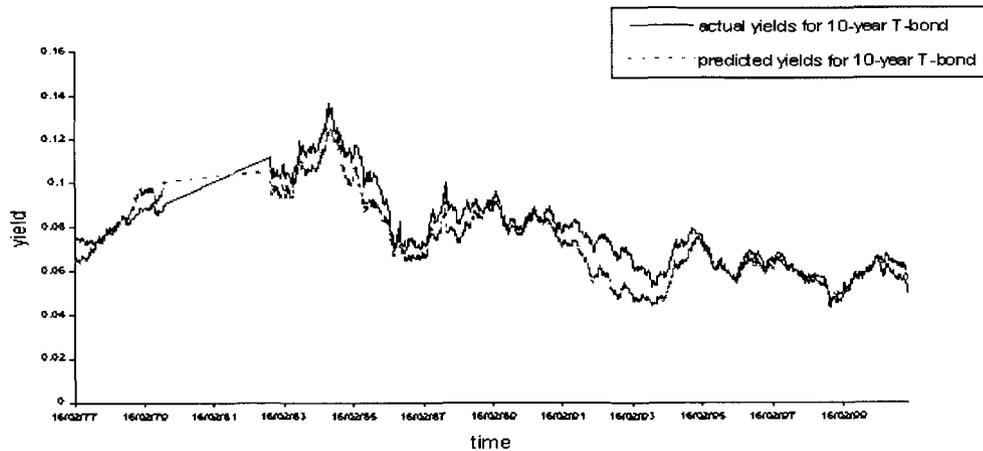
This Figure shows actual yields for a 3-month T-bill and predicted yields for the same T-bill from the one-factor CIR model. Weekly data is used from 1977.02 to 2000.12, excluding data from 1979.10 to 1982.09.

Figure 2. 6 Time series of a 5-year T-bond yield.



This Figure shows actual yields for a 5-year T-bond and predicted yields for the same T-bond from the one-factor CIR model. Weekly data is used from 1977.02 to 2000.12, excluding data from 1979.10 to 1982.09.

Figure 2. 7 Time series of a 10-year T-bond yield.



This Figure shows actual yields for a 10-year T-bond and predicted yields for the same T-bond from the one-factor CIR model. Weekly data is used from 1977.02 to 2000.12, excluding data from 1979.10 to 1982.09.

CHAPTER THREE

VALUING DEFAULTABLE SECURITIES UNDER INTEREST RATE AND DEFAULT RISK CORRELATION

This paper studies the valuation of defaultable, callable bonds and credit default swaps when both interest rates and default intensity are stochastic. The model I adopt in the paper follows the framework of Duffie and Singleton (1999) and I determine the prices of these two defaultable securities numerically. Most work in the literature assumes zero correlation between the market and the credit risk. In my paper I allow for non-zero correlation between the two risks and examine the effect of this correlation on valuation and term structures of callable bonds and on default spreads. In addition, for defaultable, callable bonds, I examine the effects of different assumptions regarding recovery rate and the notice period on the valuation of callable bonds.

3.1 Introduction

There are two basic approaches to model the dynamic behavior of default risk. One approach -- so called structural models -- pioneered by Black and Scholes (1973) and Merton (1974) and extended by Black and Cox (1976), Shimko, Tejima, and Deventer (1993), Kim, Ramaswamy, and Sundaresan (1993), Longstaff and Schwartz (1995), Briys and De Varenne (1997), and others, explicitly models the evolution of firm value which is assumed to be observed by investors. Default is triggered when the value of the firm's assets falls below or hits a pre-specified boundary. Although these models have proven very useful in examining credit risk, this class of models has been criticized for several reasons. First, since in most models firm value is described as a continuous diffusion process, default time is predictable. Therefore, credit spreads tend to be zero for the short-term debt of a solvent firm. This feature obviously contradicts empirical observations where credit spreads for short-term maturities of highly-rated firms remain strictly positive. Second, the issuer's assets and liabilities are typically not traded in

financial markets. So that their value is not directly observable and estimation of the firm value becomes problematic.

The second approach – also called the reduced-form approach -- adopted by Jarrow and Turnbull (1995), Duffie and Singleton (1997), Jarrow, Lando, and Turnbull (1997), Duffie (1998), Duffie (1999), Duffie and Singleton (1999), Madan and Unal (2000), and others, differs from the structural approach in the sense that the relation between default and firm value is not considered in a structural or explicit way. Instead, the default time is directly modelled as an unpredictable Poisson event. In particular, whenever the Poisson event occurs, the firm experiences a sudden loss in market value which could precipitate bankruptcy. The reduced-form approach is viewed as a viable alternative to the structural approach for several reasons. First, it is more tractable and easy to implement than the latter. Specifically, given an arbitrage-free setting and certain assumptions, defaultable bonds can be priced in the same way as default-free bonds. Second, the reduced-form approach does not require specifying the structure of the firm's liabilities.

One of the major criticisms of the reduced-form approach is that since the default time is modelled as exogenous unpredictable Poisson event, this modelling approach has difficulty in answering questions such as: what cause firms to default? Duffie and Lando (2001) show that under the assumption of imperfect accounting information, which is more realistic than the assumption of perfect information, the structural and the reduced-form models show amazing similarity. Specifically, one may formally view the structural model with imperfect information as equivalent to the reduced-form model. This view helps remove the major objection to reduced-form models in the literature.

This paper follows the approach proposed by Duffie and Singleton (1999). Such a model, being similar to those commonly used to price default-free bonds and

derivatives, allows for easy evaluation of defaultable claims. My objective in this paper is to apply the Duffie and Singleton model to price defaultable, callable bonds and credit default swaps. In addition, I explore the effect of market and credit risk correlation on the valuation of these two claims.

The rest of the paper is organized as follows. Section 3.2 describes the model structure. Section 3.3 specifies the model to be applied in the paper. Section 3.4 presents the valuation results of defaultable, callable bonds. Section 3.5 presents the pricing results credit default swaps. Section 3.6 summarizes the paper.

3.2 The model structure

The reduced-form credit risk model in Duffie and Singleton (1999) is the basis for valuing defaultable securities and derivatives in my paper. Trading can take place any time during the interval $[0, \bar{T}]$. Traded are default-free bonds and defaultable bonds of all maturities. Markets are assumed to be complete and frictionless, with no arbitrage opportunities. Under the assumptions of no arbitrage and complete markets, there exists a unique equivalent martingale measure \mathbf{Q} under which the market value of each security is the expectation of the discounted present value of its cash flows, using the compounded default-free short rate for discounting. For example, the value of a zero-coupon default-free bond, issued at date t and maturing at date T , with promised payoff of 1 at maturity is

$$\delta(t, T) = E_t^{\mathbf{Q}} \left[e^{-\int_t^T r_u du} \right], \quad (3.1)$$

where $E_t^{\mathbf{Q}}$ denotes risk-neutral expectation conditional on information known at date t . On the other hand, if the issuer defaults prior to the maturity date T , then both the magnitude and timing of the payoff to investors may be uncertain. Let τ denote the first

time that this firm defaults, and let $1_{\{\tau > t\}}$ be the indicator of the event that $\tau > t$, which takes the value of 1 if the issuer has not defaulted prior to time t , and 0 otherwise. Then the value of this risky zero-coupon bond, with the promised payoff of 1 at maturity is

$$V_0(t, T) = E_t^Q \left[e^{-\int_t^T r_u du} 1_{\{\tau > T\}} + e^{-\int_t^T r_u du} W 1_{\{\tau \leq T\}} \right], \quad (3.2)$$

where W is the value of recovery. The value of this risky debt is composed of two parts. The first part is the present value of the promised payment if default does not occur. The second part is the present value of the promised payment in default.

Depending on how the default time τ ¹⁷ is modelled, and how the recovery amount is specified, equation (3.2) could lead to different valuation formulas. This paper follows the model in Duffie (1998) and Duffie and Singleton (1999). First, I adopt the reduced-form approach, in which the timing of default is modelled as an unpredictable Poisson process with stochastic intensity, $h(t)$, and $h(t)\Delta$ approximates the probability of default over the next time period of length Δ , given that the firm has not defaulted yet at time t . Then the conditional probability at time s , given all available information at that time, of survival to time t , is given by

$$p(t | s) = E_s^Q [1_{\{\tau > t\}}] = E_s^Q \left[e^{-\int_s^t h_u du} \right]. \quad (3.3)$$

¹⁷ How to define the time of default τ depends on which model is being used. The time of default τ in structural-form models is the first hitting time of a diffusion process at a fixed barrier. The time of default τ in reduced-form models is the time of the first jump of a Poisson process.

Second, given that the commonly used reduced-form models differ mainly in their treatment of the recovery, I will focus on two assumptions on recovery: recovery of face value (RFV) and recovery of market value (RMV).¹⁸

Recovery of face value (RFV)

The RFV assumption, which is studied in Duffie (1998), assumes that the recovery amount is a fraction of the face value of the claim. The value of recovery is expressed as

$$W_t = (1 - L_t) \cdot 1, \quad (3.4)$$

where L_t is the fractional loss in face value at time t , and the value of risky zero-coupon bond is given by¹⁹

$$\begin{aligned} V_0^{RFV}(t, T) &= E_t^Q \left[e^{-\int_t^T \phi_u du} + \int_t^T 1 \cdot (1 - L_u) e^{-\int_t^u \phi_s ds} h_u du \right] \\ &= E_t^Q \left[e^{-\Phi(T)} + \int_t^T (1 - L_u) e^{-\Phi(u)} h_u du \right], \end{aligned} \quad (3.5)$$

where $\phi_t = r_t + h_t$ and $\Phi(\tau) = \int \phi_u du$. Furthermore, by adding deterministic and continuous coupon rate $c(t)$ in my model, the value of this risky bond is

$$V^{RFV}(t, T) = V_0^{RFV}(t, T) + E_t^Q \left[\int_t^T c(u) e^{-\int_t^u \phi_s ds} du \right]$$

¹⁸ In addition to RFV and RMV, there is a third specification in modeling the recovery rate: recovery of treasury (RT). This approach assumes that, when default occurs, the debtholders recover a fraction of the value of an otherwise equivalent, default-free bond. See Jarrow and Turnbull (1995) for an example.

¹⁹ Please refer to Duffie and Singleton (1999) for details.

$$\begin{aligned}
&= E_t^Q \{ e^{-\Phi(T)} + \int^T (1-L_u) e^{-\Phi(u)} h_u du + \int^T c(u) e^{-\Phi(u)} du \} \\
&= E_t^Q \{ \int^T e^{-\Phi(u)} (c(u) + (1-L_u) h_u) du + e^{-\Phi(T)} \}. \tag{3.6}
\end{aligned}$$

Recovery of market value (RMV)

The RMV assumption assumes that the recovery amount is a fraction of the market value of the same bond right before the default. Thus the value of recovery, W , is given by

$$W_t = (1 - L_t) v(t_-, T), \tag{3.7}$$

where L_t denotes the expected fractional loss in market value at time t , t_- represents an instant before default. Combining equation (3.2), (3.3) and (3.7) and under certain technical conditions, I obtain²⁰

$$V_0^{RMV}(t, T) = E_t^Q [\exp(-\int^T R_s ds)] = E_t^Q [e^{-\Psi(T)}], \tag{3.8}$$

where $R_t = r_t + h_t L_t$ and $\Psi(\tau) = \int^T R_u du$. Equation (3.8) implies that a defaultable zero-coupon bond may be priced as if it were default-free by replacing the usual short-term interest rate process r with the default-adjusted short-rate process R . Discounting at this default-adjusted short-rate process R therefore accounts for the probability of default, timing of default, and the effect of loss on default. In the case of deterministic and continuous coupon rate $c(t)$, the value of defaultable debt is given as

²⁰ For derivation details, please see Duffie and Singleton (1999).

$$\begin{aligned}
V^{RMV}(t, T) &= V_0^{RMV}(t, T) + \int_t^T c(u) V_0^{RMV}(t, u) du \\
&= E_t^Q \left\{ e^{-\Psi(T)} + \int_t^T e^{-\Psi(u)} c(u) du \right\}. \tag{3.9}
\end{aligned}$$

One may wonder about the implications of choosing one recovery assumption over the other. The assumption of RMV is easier to implement, because prices of defaultable bonds can be computed in a RMV model by the same formulas used for default-free bonds, using the default-adjusted short-rate process R instead of the usual short-term interest rate process r . If, however, one assumes liquidation at default and that absolute priority applies, then the assumption of RFV is more realistic, as it assumes the same recovery rate for bonds of equal seniority by the same issuer. Duffie and Singleton (1999) and Skinner and Diaz (2000) point out that for the estimation of par bond spreads, or for the estimation of risk-neutral default intensities from par spreads, these two assumptions make little difference. In this paper, I examine whether these two assumptions make any difference when applied to the valuation of callable bonds.

3.3 My model specification

In this section, I describe a specific model which is used to value defaultable, callable bonds and credit default swaps in my paper later. Notice that this model allows for both default risk and interest rate risk and the correlation between these two risks is not necessarily zero.

Assumption 1: The interest rate r_t follows a continuous, adapted Cox-Ingersoll-Ross (CIR, 1985) process

$$dr_t = \kappa_r (\theta_r - r_t) dt + \sigma_r \sqrt{r_t} dZ_{1t}, \tag{3.10}$$

where Z_{1t} is a one-dimensional Brownian motion under the P-measure, $\kappa_r, \theta_r > 0$, and $2\kappa_r\theta_r \geq \sigma_r^2$. For pricing purpose I assume a constant market price of risk λ , which transforms equation (3.10) by the means of the Girsanov theorem into

$$dr_t = [\kappa_r(\theta_r - r_t) - \lambda_r r_t]dt + \sigma_r \sqrt{r_t} dZ_{1t}^Q, \quad (3.11)$$

where Z_{1t}^Q is a one-dimensional Brownian motion under the Q-measure.

Assumption 2: The intensity h_t evolves according to the equation

$$dh_t = \kappa_h(\theta_h - h_t)dt + \sigma_h \sqrt{h_t} dZ_{2t}, \quad (3.12)$$

where $\kappa_h, \theta_h > 0$, and $2\kappa_h\theta_h \geq \sigma_h^2$. The instantaneous correlation between dZ_{1t} and dZ_{2t} is ρdt . Under the Q-measure, I assume that the process can be written as

$$dh_t = [\kappa_h(\theta_h - h_t) - \lambda_h h_t]dt + \sigma_h \sqrt{h_t} dZ_{2t}^Q, \quad (3.13)$$

where Z_{2t}^Q is a one-dimensional Brownian motion under the Q-measure.

Assumption 3: The loss rate L_t is assumed to be a constant, $L_t = L$.

Note that Assumption 3 is not necessary. The loss rate can be deterministic or even random.²¹ One of my objectives is to examine how different assumptions on the recovery affect the valuation of defaultable contingent claims. By making such a simplified assumption, I can concentrate on this issue.

²¹ For example, see Duffie and Singleton (1999) and Skinner and Diaz (2000).

Assumption 4: I assume that coupon is paid continuously, and the annualized coupon payment is c .

The reason for us to assume continuous coupon payment instead of semi-annual coupon payment is that it is relatively easy to solve partial differential equations (PDE) using numerical methods.

In this paper, I rely on numerical methods to value several contingent claims. Specifically, alternating direction implicit (ADI) method is used to deal with this two-factor model. Although the CIR-type process used in this paper belongs to the “affine” family,²² in which a closed-form solution can then be derived, most interest rate and default risk models are not of the “affine” type, therefore an analytical solution is unavailable. Moreover, numerical methods allow us to consider a wide range of models used in the literature. To be able to apply numerical methods, I need to find the PDEs under different recovery assumptions. This can be done by applying the Feynman-Kac representation (see the Appendix for details).

Recovery of face value (RFV)

According to equation (3.6) and the Feynman-Kac formula, $V^{RFV}(t, T, c, L)$ is the unique solution to the following PDE

$$\begin{aligned} & \frac{1}{2} \sigma_r^2 r V_{rr} + \frac{1}{2} \sigma_h^2 h V_{hh} + [\kappa_r (\theta_r - r) - \lambda_r r] V_r + [\kappa_h (\theta_h - h) - \lambda_h h] V_h \\ & + \rho \sigma_r \sigma_h \sqrt{r} \sqrt{h} V_{hr} - (r + h)V + c + (1 - L)h + V_t = 0 \end{aligned} \quad (3.14)$$

²² The affine term structural model is a class of models in which the yields to maturity are affine (constant-plus-linear) functions in some state variable vector X_t . Examples of the affine processes in the term-structure literature are the Gaussian (Vasicek, 1977) and square-root diffusion models (Cox, Ingersoll, and Ross, 1985). Detail discussions of affine models can be found in Dai and Singleton (2000), and Duan and Simonato (1999).

with the boundary condition

$$V(T, T, c, L) = 1. \quad (3.15)$$

Recovery of market value (RMV)

According to equation (3.9) and the Feynman-Kac formula, $V^{RMV}(t, T, c, L)$ is the unique solution to the following PDE

$$\begin{aligned} & \frac{1}{2} \sigma_r^2 r V_{rr} + \frac{1}{2} \sigma_h^2 h V_{hh} + [\kappa_r (\theta_r - r) - \lambda_r r] V_r + [\kappa_h (\theta_h - h) - \lambda_h h] V_h \\ & + \rho \sigma_r \sigma_h \sqrt{r} \sqrt{h} V_{hr} - (r + hL)V + c + V_t = 0 \end{aligned} \quad (3.16)$$

with the boundary condition

$$V(T, T, c, L) = 1.$$

3.4 Valuation of defaultable, callable bonds

The majority of corporate bonds are callable. The call option gives the issuer the right to call the bond at a fixed call price any time before the bond maturity, after an initial “lock-out” period.²³ The issuer may call back the bond under the following two situations: when the interest rate falls or when the credit quality of the issuer improves. In either case, the issuer calls the bond and replaces it with lower-cost debt. According to Buttler and Waldvogel (1996), there are three types of callable bonds:

European callable bonds: the issuer has the right to call the bond at only one date (typically the last coupon date before maturity).

²³ The “lock-out” period is defined as the length of time from issuance until the first possible call date. The range of the “lock-out” period is from as short as a month to more than ten years.

American callable bonds: the issuer may call the bond at any time after an initial “lock-out” period.

Semi-American (Bermudan) callable bonds: the issuer has the right to call the bond at one of a set of pre-specified dates (usually coinciding with coupon dates) after a “lock-out” period.

In this section, I apply my pricing model to examine the valuation of defaultable callable bonds when both interest rates and intensity are stochastic. I assume that the issuer follows rule for calling bond so as to minimize the market value of that bond. This rule implies that the issuer will exercise the option to call in the bond at time τ if and only if its market price, if not called, is higher than the strike price on the call. I thus have another boundary condition

$$V(\tau, T, c, L) = \min(V(\tau, T, c, L), K_\tau), \quad (3.17)$$

where $V(\tau, T, c, L)$ is the bond price at time τ , assuming that the bond has not defaulted by τ , and K_τ is the call price at time τ .

Then the valuation equations (3.14) and (3.16), subject to equations (3.15) and (3.17), can be solved numerically by the ADI method, respectively.

Several papers in the literature have looked at this issue. However, because of the complexity of the default and call options, much of the existing work has treated interest rates as constant.²⁴ From the perspective of my paper, the two most relevant papers are Acharya and Carpenter (2000) and D’halluin, Forsyth, Vetzal, and Labahn

²⁴ Examples include Merton (1974), Brennan and Schwartz (1977), and Black and Cox (1976), etc.

(2001). Acharya and Carpenter consider both interest rates and default risk, but they employ the structural approach to model default risk. In addition, they do not take into account the notice period. In practice, however, most callable bonds require that the issuer provide an advance notice of a decision to exercise the embedded call. D'halluin et al. (2001) use a numerical PDE approach to price callable bonds with the notice period. However, they focus on only default-free contracts.

The valuation model applied in this paper incorporates all the following features: interest rate risk, default risk, correlation between these two risks, call provisions, optimal call policies, and the notice period. To clarify the interaction of the call and default options, I also look at several simpler counterparts to defaultable, callable bonds: a defaultable, non-callable bond, and a default-free non-callable bond with the same face value and maturity, and a default-free callable bond with the same face value, maturity and call scheme.

My benchmark is an American defaultable, callable bond without the notice period. I assume that the coupon is continuously paid and the RFV assumption is used. Table 1 gives a summary of the parameter values and call scheme. The parameter values for the interest rate process are $\kappa_r = 0.55$, $\theta_r = 0.035$, $\sigma_r = 0.38$, $\lambda_r = -0.41$, which are taken directly from D'halluin et al (2001). The parameter values for the intensity process are $\kappa_h = 0.22$, $\theta_h = 0.0058$, $\sigma_h = 0.073$, and $\lambda_h = -0.25$. These values are chosen based on the work of Duffee (1999). In his paper, Duffee estimates these parameter values based on the credit ratings of the bonds' issuers. A total of 161 firms are included and 5 different credit ratings are used to implement the estimation. For my purpose, I need only one set of parameter values that are applicable to all credit ratings, so I compute weighted average values of the parameters across different credit ratings. The recovery rate used throughout this paper is 0.4, which is suggested in Altman and

Kishore (1996). The base case correlation between interest rates and intensity, ρ , is zero. The defaultable, callable bond I consider here is a 20-year bond with a 10-year “lock-out” period.

Sensitivity analysis to ρ

The effect of stochastic interest rates and stochastic intensity on total spreads²⁵ depends on the correlation between these two stochastic variables, ρ . I consider the cases when $\rho = -0.8, 0$, and 0.8 . Figure 3.1 illustrates how total spreads of a 20-year defaultable, callable bond with a 10-year “lock-out” period change when the correlation ρ takes different values, given initial interest rate r_0 . Two types of callable bonds are considered, one with low credit quality, the other with high credit quality, based on initial default intensity h_0 . Here is what I find from Figure 3.1:

- For a callable bond with given maturity, as the initial interest rate r_0 goes up, the total spread goes up. Moreover, the total spread goes up as ρ decreases. This result is consistent with Acharya and Carpenter (2000).²⁶

The first part of the results can be explained intuitively. As the initial interest rate r_0 goes up, because of the nature of mean-reversion of the interest rate process, the probability of decrease in interest rates in the future is high, then the probability of calling the bond is high, so investors ask for higher spread to compensate for it.

²⁵ The total spread in our paper is defined as the difference between the par-coupon yield of defaultable, callable bond and the par-coupon yield of default-free non-callable bond with the same maturities. This definition is used to distinguish traditional definition of par-coupon credit spread of non-callable bonds.

²⁶ Acharya and Carpenter (2000) claim that spreads move in the same direction as the correlation between interest rates and firm value. Since the higher the firm value, the less likely the firm defaults, it is equivalent to say that spreads widen as the correlation between interest rates and intensity decreases.

- The effect of the correlation measure on spreads is greater on low quality bonds than on high quality bonds.

In Figure 3.2, I describe how term structures of par-coupon total spreads of defaultable, callable bonds move as ρ changes. I find:

- For a short-term bond, the total spread increases as ρ goes up; for a long-term bond, however, the total spread increases when ρ decreases. Furthermore, the effect of ρ on the total spreads is much more stronger for low quality bonds than for high quality bonds. For a 10-year bond, the effect of changing ρ from -0.8 to 0.8 is approximately 67bps for low quality bonds, whereas the difference is only around 16bps for high quality bonds.

For comparison purposes, I also calculate spreads for defaultable, non-callable bonds with different ρ . The results are shown in Figure 3.3. I find:

- The spreads of defaultable, non-callable bonds are also quite sensitive to the value of ρ , and the effect of ρ on spreads is stronger as the bond's quality is lower. All these results are quite similar to those of defaultable, callable bonds. The effect of ρ on spreads of non-callable bonds are consistent; that is, the total spread always increases as ρ goes down, and this result is independent of the maturity.

Call premiums and credit spreads

People in risk management may be interested in answering the following questions:

1. Should investors require the same credit spreads for both callable and non-callable bonds?

2. Should investors require the same call premiums for both defaultable and default-free bonds?

Interestingly, the answers to these two questions are basically the same. There are two ways to calculate total spreads: one way is to calculate the sum of credit spreads of callable bonds and call premiums of default-free bonds; the other way is to calculate the sum of credit spreads of non-callable bonds and call premiums of defaultable bonds. Then different credit spreads of callable and non-callable bonds imply different call premiums of defaultable and default-free bonds. Figure 3.4 shows that under assumption of zero correlation, the par-coupon credit spreads for callable and non-callable bonds are not the same for a 20-year bond, and are not even close at high interest rates. For example, when the interest rate is 15% per year, the par-coupon credit spreads of 20-year callable bonds are 77.59 basis points for high quality bonds and 472.30 basis points for low quality bonds; for non-callable bonds the spreads are 109.99 basis points for high quality and 516.18 basis points for low quality bond. Figures 3.5 and 3.6 show the cases with positive and negative ρ 's, respectively. My results show that investors indeed require different spreads for callable and non-callable bonds and different call premiums for defaultable and default-free bonds.

Comparison of total spreads under the RFV and RMV assumptions

Duffie and Singleton (1999) show that for non-callable bonds, the term structures of par-coupon yield spreads for RMV and RFV are rather similar even when the recovery rates in both cases are the same. To find out whether this result applies to defaultable, callable bonds as well, I calculate the term structures of par-coupon total spreads of defaultable, callable bonds under these two assumptions.

- The assumption of RMV gives a larger spread. The higher the initial value of intensity, h_0 , the larger gap of the spread between these two assumptions. This result is consistent with the patterns that Duffie and Singleton (1999) find empirically for non-callable bonds.
- For defaultable, callable bonds, the term structures of par-coupon total spreads under RMV and under RFV are not quite the same, especially for longer-term, low quality bonds. For example, under the assumption of zero correlation, consider a 10-year defaultable, callable bond with a long-run mean equals 58 basis points, the total spread is 509.16 basis points under RMV, and 485.57 basis points under RFV, the difference is 23.59 basis points, which is quite significant. Figure 3.7 shows the results.

For comparison purposes, I also calculate the term structure of par-coupon yield spreads for defaultable, non-callable bonds under both the RMV and RFV assumptions by using exactly the same parameter values.

- For defaultable, non-callable bonds, the two recovery assumptions give us rather similar term structures of par-coupon yield spreads for short-term and medium-term bonds. This result is quite consistent with Duffie and Singleton (1999) even though when I use different models and parameter values. For example, the gap for a 10-year defaultable, non-callable bond under these two different assumptions is only 15.37 basis points. For long-term bonds, however, the gap is quite noticeable. It is shown in Figure 3.8. Table 3.2 gives the comparison.

The effect of the notice period on valuation of callable bonds

In practice, most callable bonds require that the issuer provide an advance notice of its intention to exercise the embedded call. As noted by Bliss and Ronn (1998), the

standard description of the optimal call policy for the issuer is no longer correct when the advance notice must be given. To find out the effect of notice periods on the valuation of callable bonds, in this section, I consider the valuation of a 20-year American callable bond with a 3-month notice period, everything else being the same as in the case without the notice period. My calling rule follows Jordan and Jorgensen (1996) where a bond is called when the price of the callable bond exceeds the price of a 3-month Treasury bill on the notification date. More specifically, since my call price is not a single value but a scheme, the face values of these 3-month Treasury bills change as the call price changes. I calculate the total spreads of the callable bond with the notice period and compare the results to the bond without-notice-period. Table 3.3 shows the results of the comparison. Several conclusions can be drawn from Table 3.3:

- As the initial interest rate r_0 goes up, the total spread required by investors also goes up regardless of the initial quality of the bond.
- For any given r_0 , the total spread of callable bond without the notice period is higher than that of callable bond with the notice period.

The above result is quite intuitive as that the notice period is valuable to investors, and they would like to pay more for this option. The total spread for the bond with the notice period thus is lower than the one without the notice period.

- Although the total spreads are different by adding the notice period, the change is not significant. For example, consider a 20-year defaultable, callable bond with a long-run mean equals 58 basis points and $\rho = 0$, the total spread of this bond with a 3-month notice period is 226.98 basis points when $r_0 = 15\%$ and $h_0 = 50$ basis points; the total spread of the same bond without the notice period is 231.89 basis points. This can be seen in Figure 3.9.

3.5 Valuation of default swaps

Credit derivatives are contracts that transfer an asset's risk and return from one counterpart to another without transferring ownership of the underlying asset. There are four major types of credit derivatives: default swaps, total return swaps, credit spread put options, and credit linked notes.²⁷ The global market for credit derivatives is still quite small compared with other derivatives markets. It represents only 1% of the global derivatives market, but it is growing rapidly. Figure 3.10 shows the increased trading volume of credit derivatives. The exponential growth as seen in Figure 3.10 has generated significant interests in the fair valuation of credit derivatives in both the academic and practitioner communities.

Among these four types of credit derivatives, the default swap is the most common credit derivative. According to the British Bankers Association (BBA) 2001 survey, 40% of the market notional amounts outstanding come from credit default swaps. A default swap is a contract that provides insurance against the risk of a default by a particular company. The protection seller pays the protection buyer a given contingent amount if there is a credit event, such as default.²⁸ In return, the protection buyer makes periodic premium payments to the protection seller until the time of the credit event, or the maturity date of the credit swap, whichever is first. In the case of a credit event the default swap can be settled by either physical delivery or in cash. In a physically settled swap, the protection seller receives the underlying and pays the face

²⁷ A good introduction on credit derivatives can be found in Bomfim (2001).

²⁸ Other commonly used credit events that are defined by the International Swaps and Derivatives Association (ISDA) are: failure to pay, bankruptcy, cross-default, restructuring, cross-acceleration, repudiation, merger, regulatory suspension, and downgrading.

value to the protection buyer. In a cash settled swap, the protection seller pays the difference between par and the recovery value of the underlying.

Because of the popularity of default swaps, there are quite many papers working on the valuation of swaps. Most of them assume that the credit and the market risk are statistically independent. Examples are given in Bomfim (2001), Dwlianedis and Lagnado (2000), and Hull and White (2000). Based on this simplified assumption, these papers develop models for pricing default swaps. One noticeable exception is Jarrow and Yildirim (2002), who relax this independence assumption and still provide a simple analytic formula for the valuation of default swaps. Specifically, to obtain their simple while realistic empirical formulation of the model, Jarrow and Yildirim assume that the economy is Markovian in a single state variable – the spot interest rate, and that the intensity is a linear function of the spot rate, which is used to incorporate the correlation between the credit and the market risks in their model. An analytic expression is then derived in the context of a reduced-form credit risk model.

My approach, in some ways, follows Jarrow and Yildirim's (2002) work. I also use the reduced-form credit risk model and incorporate the correlation between intensity and spot rate in my model. However, one major difference of my work from theirs is that, I relax the assumption of a linear relationship between the credit and the market risks; instead, I model stochastic processes for intensity and spot rates separately. The correlation is specified in the random part. The numerical procedure is then used in my paper to value default swaps.

The pricing of default swaps is fundamentally linked to three factors: (1) the credit risk of the reference entity, (2) the expected recovery rate associated with the reference entity, (3) the credit risk of the protection seller. The first and third factors highlight the two types of risk faced by the protection buyer: issuer default risk, and counterparty

default risk. A fourth factor may also affect the pricing of default swaps: the default correlation between the reference entity and the protection seller. However, such a joint event is typically much less likely than a default by the reference entity or the protection seller alone, so I do not consider the joint default in this paper. In addition, I also assume that there is no counterparty default risk, so I do not need to consider the pricing impact of default by the protection seller.

There are two pricing issues associated with a default swap:

1. At the beginning of the contract, the standard default swap involves no exchange of cash flows, and therefore has zero market value. One must determine the annuity premium for which the market value of the default swap is indeed zero. This premium is sometimes called the default swap spread.
2. After origination, given the annuity premium, one must determine the current market value of the default swap, which is generally non-zero.

In this section, I focus on the valuation of the default swap spread.

According to the discussion above, a default swap typically has a zero market value when it is set up, and thus pricing such a contract is equivalent to finding the value of the default swap spread, s , that makes the expected present value of payments made by the buyer have the same value as the expected present value of payments received by the buyer in the case of default.

Assumption 5: I assume that the default swap spread, s , is paid continuously.

This assumption simplifies my procedure in the sense that I do not need calculate the accrued portion of payment in the case when the credit event happens between the payment dates.

Under the RFV assumption of recovery, the expected present value of payments made by the buyer is given by

$$\begin{aligned} E_t^Q \left[\int_0^{\tau} s e^{-\int_t^u \phi_s ds} du \right] \\ = E_t^Q \left[\int_0^{\tau} s e^{-\Phi(u)} du \right], \end{aligned} \quad (3.18)$$

where $\phi_t = r_t + h_t$, and $\Phi(\tau) = \int_0^{\tau} \phi_u du$.

The expected present value of payments received by the buyer in case of default is

$$\begin{aligned} E_t^Q \left[e^{-\int_t^{\tau} r_u du} \cdot L \cdot 1_{\{\tau \leq T\}} \right] \\ = E_t^Q \left[\int_0^{\tau} L \cdot e^{-\Phi(u)} \cdot h_u du \right]. \end{aligned} \quad (3.19)$$

The credit default spread, s , is the value that makes expressions in equations (3.18) and (3.19) equal

$$s = L \cdot \frac{E_t^Q \left[\int_0^{\tau} e^{-\Phi(u)} h_u du \right]}{E_t^Q \left[\int_0^{\tau} e^{-\Phi(u)} du \right]}. \quad (3.20)$$

In this subsection, I implement numerical methods to estimate the credit default spread, s , assuming both the spot rate and the intensity processes are stochastic. My benchmark is a 10-year credit default swap contract with zero correlation ($\rho = 0$) between interest rates and intensity. To evaluate the effect of the correlation on the valuation of the credit default spread, I also consider two other cases: $\rho > 0$ and $\rho < 0$. In addition, I explore the term structure of the credit default spread and trace the effect of

changing ρ on the shape of the term structure of the credit default spread. The values of model parameters are summarized in Table 3.1. Through examining Figure 3.11, I have the following findings:

- The choice of the initial interest rate r_0 has no significant effect on the valuation of the credit default spread given the initial intensity h_0 , and this result is not sensitive to the choice of ρ ; especially for the high quality reference entities. For example, consider a case where ρ is zero and $h_0 = 50$ basis points, the credit default spread is 68.36 basis points when $r_0 = 0.0\%$, and 65.84 basis points when $r_0 = 15.0\%$.
- As ρ goes up, the credit default spread goes down.

Figures 3.12 and 3.13 show the credit default spread term structure for different initial credit qualities. I find:

- For short-term swaps, the change of ρ has little effect on the valuation of the credit default spread, regardless of the credit quality. For example, for a 2-year swaps with high quality reference entities, the credit default spread is 39.38 basis points when $\rho = -0.8$, 38.45 basis points when $\rho = 0.0$, and 37.28 basis points when $\rho = 0.8$. For a 2-year swaps with low quality reference entities, the credit default spread is 380.15 basis points when $\rho = -0.8$, 376.84 basis points when $\rho = 0.0$, and 373.33 basis points when $\rho = 0.8$. The differences are less than 1 basis point for high quality reference entities and less than 4 basis points for low quality reference entities.
- For longer-term swaps, the values of credit default spreads differ significantly as ρ changes. For example, consider a 10-year swaps with low quality reference

entities, the credit default spread is 449.41 basis points when $\rho = -0.8$, 412.12 basis points when $\rho = 0.0$, and 376.49 basis points when $\rho = 0.8$. The differences are more than 30 basis points, which are quite noticeable.

- The percentage changes of credit default spreads are not symmetric for a high quality reference entity as ρ changes from a negative value to zero and from zero to a positive value, especially for long-term swaps. For example, the percentage change of the credit default spread for a 10-year swaps with high quality reference entities is 18.7% when ρ changes from a negative value to zero; and the percentage change increases to 31.1% as ρ changes from zero to a positive number.

3.6 Conclusion

This paper studies the valuation of defaultable, callable bonds and credit default swaps when both interest rates and default intensity are stochastic, and the correlation between these two variables is nontrivial. I determine the prices of these two contingent claims and examine the effect of different correlation on valuation.

For the defaultable, callable bond, my main findings are as follows. First, the total spread has negative relationship with the correlation, ρ , and the magnitude of this correlation effect changes with the initial credit quality of the bond. Second, the influence of changing ρ on the term structures of par-coupon total spreads of defaultable, callable bond is quite different from its defaultable, non-callable counterpart. Third, investors require different spreads for callable and non-callable bonds and different call premiums for defaultable and default-free bonds. Fourth, different assumptions on the recovery make noticeable differences on the total spreads of medium-term and long-term

defaultable, callable bonds. Fifth, adding the notice period makes no significant change of the total spread.

For the credit default swaps, I find that first, the choice of interest rate, r_0 , has no significant effect on the valuation of the credit default spread. Second, as the correlation ρ goes up, the credit default spread goes down. Third, the effect of changing ρ on the absolute changes of credit default spreads is stronger for low quality reference entities. Fourth, the change of ρ has little effect on the valuation of the credit default swaps for short-term swaps, for longer-term swaps, the values of credit default spreads diverge drastically as ρ changes.

Table 3. 1 Summary of model parameters

			Interest rate	Intensity	Year from T	Call price
Maturity T	20 years	κ	0.55	0.22	1-5	\$100
Coupon C		θ	0.035	0.0058	6	\$100.5
Principal	\$100	σ	0.38	0.073	7	\$101
Loss rate	0.6	λ	-0.41	-0.25	8	\$101.5
					9	\$102
					10	\$102.5

This table gives my choice of parameter values used in the paper. Input data used for the models considered. The interest rate process parameter values are from D'halluin et al. (2001), and the intensity process parameter values are chosen based on Duffee (1999). Note that the years for the call prices run backwards in time, so for example the bond is callable at a price of \$100.5 six years before maturity. Coupon payments are on an annual basis.

Table 3. 2 Spread differences under the RMV and RFV assumptions

Maturity	Type of bonds	H_mean = 58bps, $\rho = 0$	
		$h_0 = 50\text{bps}$	$h_0 = 600\text{bps}$
5 year	Callable	1.52	11.76
	Non-callable	0.61	5.63
10year	Callable	3.53	23.59
	Non-callable	2.19	15.73
20year	Callable	9.96	43.09
	Non-callable	7.82	33.81

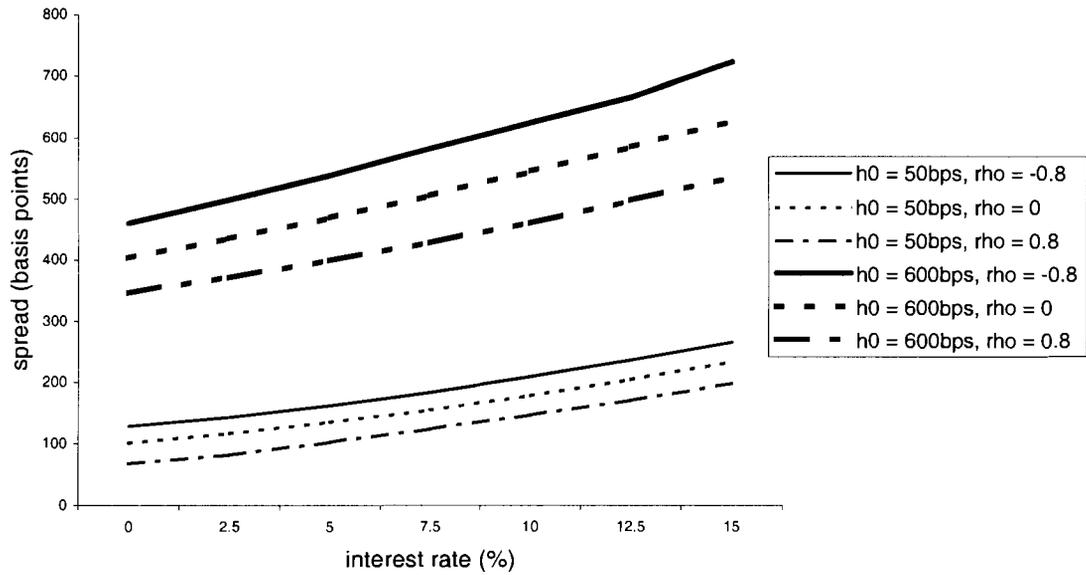
The spread difference is given in basis points. I consider 5-, 10-, and 20-year callable and non-callable bonds with different recovery assumptions. Two types of callable bonds are considered here: one with low quality ($h_0=600\text{bps}$), the other with high quality ($h_0=50\text{bps}$). The parameter values are given in Table 3.1.

Table 3. 3 Comparison of total spreads of callable bonds with and without the notice period

H_mean = 58bps, $\rho = 0$				
	With the notice period		Without the notice period	
r_0 (%)	$h_0 = 50\text{bps}$	$h_0 = 600\text{bps}$	$h_0 = 50\text{bps}$	$h_0 = 600\text{bps}$
0	98.35	402.20	100.92	404.40
2.5	111.92	433.27	116.48	435.75
5.0	130.53	466.80	135.31	469.59
7.5	151.21	502.61	156.34	505.85
10.0	173.92	540.88	179.19	544.36
12.5	199.14	581.31	204.15	584.70
15.0	226.98	623.36	231.89	626.60

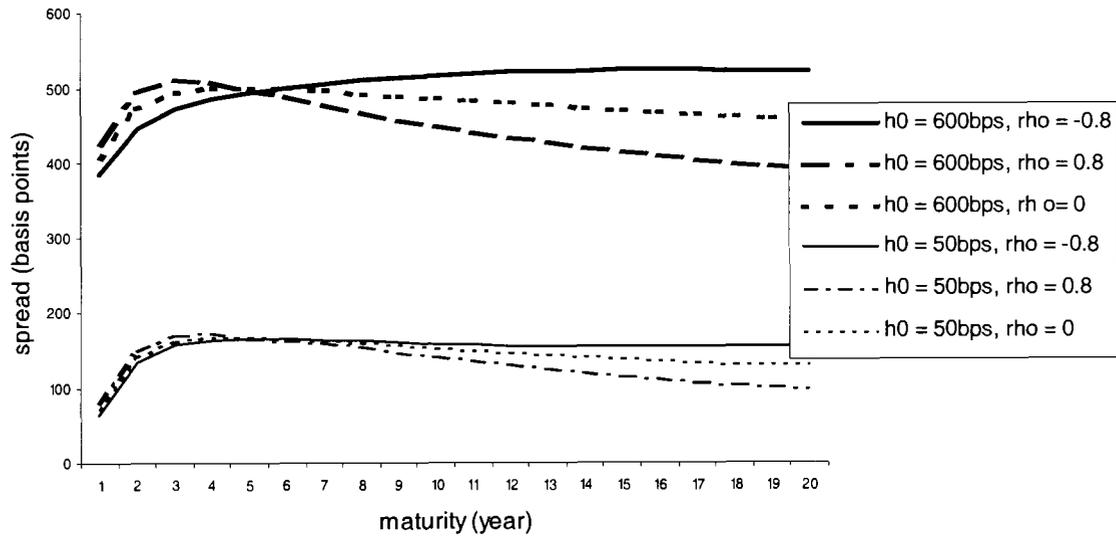
The total spreads are given in basis points. I calculate the total spreads of a 20-year American callable bond with 10-year lockout period with and without the notice period. The notice period is assumed to be 3 months. Two types of callable bonds are considered here: one with low credit quality ($h_0=600\text{bps}$), the other with high credit quality ($h_0=50\text{bps}$). The correlation, ρ , is assumed to be zero. The parameter values are given in Table 3.1.

Figure 3. 1 Total spreads of a 20-year defaultable, callable bond with a 10-year “lock-out” period.



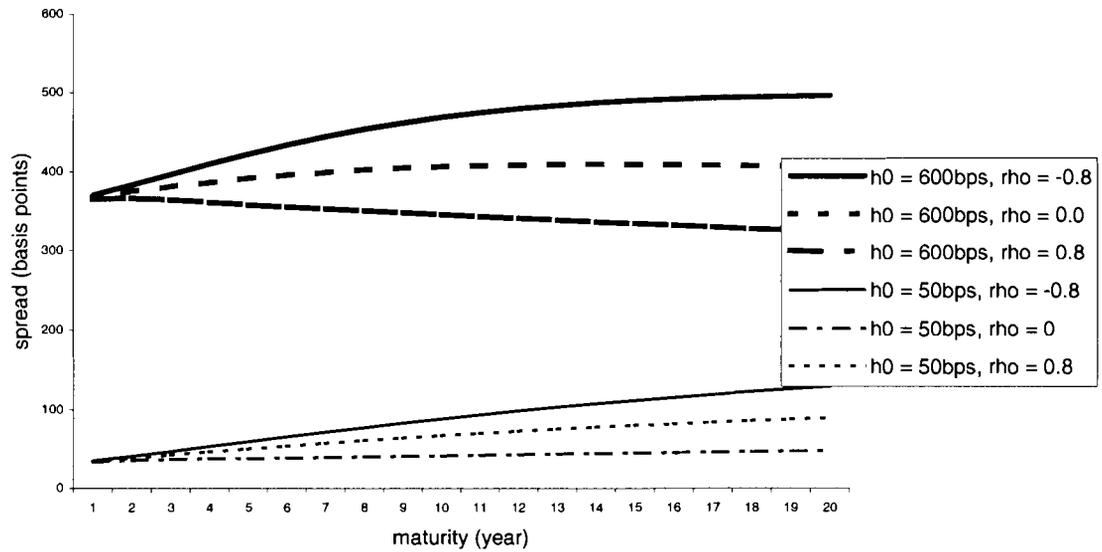
The recovery rate is 0.4 and the call scheme is shown in Table 3.1. Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). Three different values of the correlation variable, ρ , are considered: -0.8, 0, and 0.8.

Figure 3. 2 Term structures of par-coupon total spreads of defaultable, callable bonds under RFV.



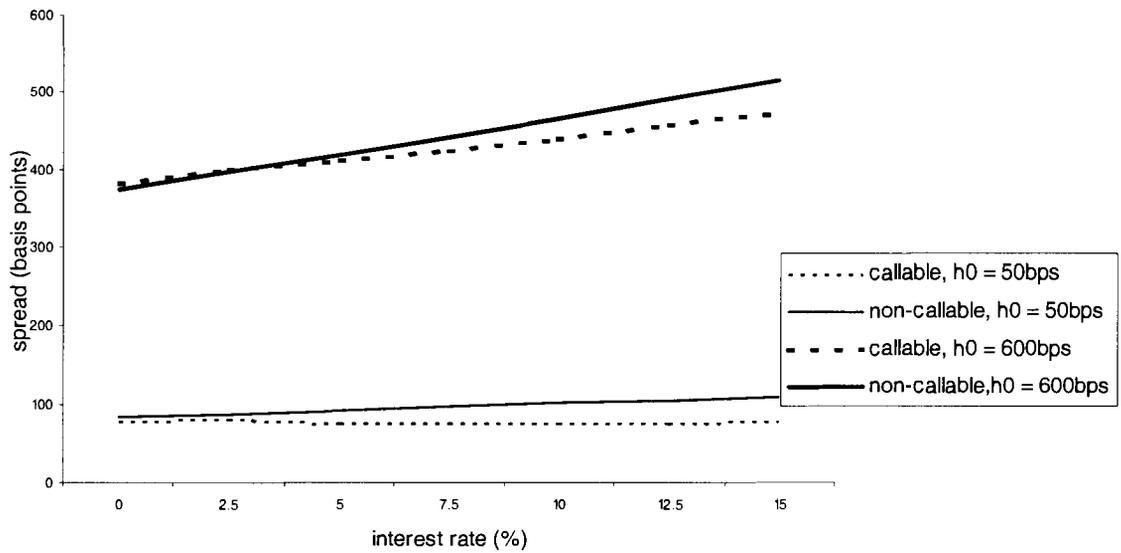
Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). Three different values of the correlation variable, ρ , are used: -0.8, 0, and 0.8. The values of model parameters are given in Table 3.1. Note that to simplify my procedure, I use a single call price, \$101.00, instead of a call scheme, and I assume that the first call dates of all maturities are one half of their maturities, respectively. The initial interest rate, r_0 , is chosen to be equal to 3.75%.

Figure 3. 3 Term structures of par-coupon total spreads of defaultable, non-callable bonds under RFV.



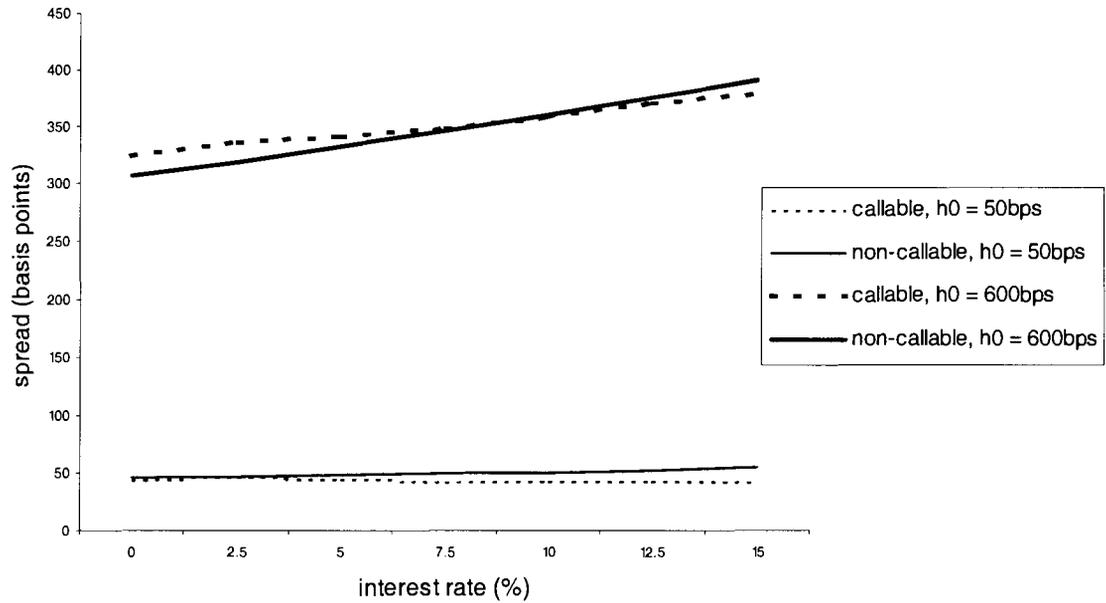
Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). Three different values of the correlation variable, ρ , are used: -0.8, 0, and 0.8. The values of model parameters are contained in Table 3.1.

Figure 3. 4 Par-coupon credit spreads of 20-year callable and non-callable bonds.



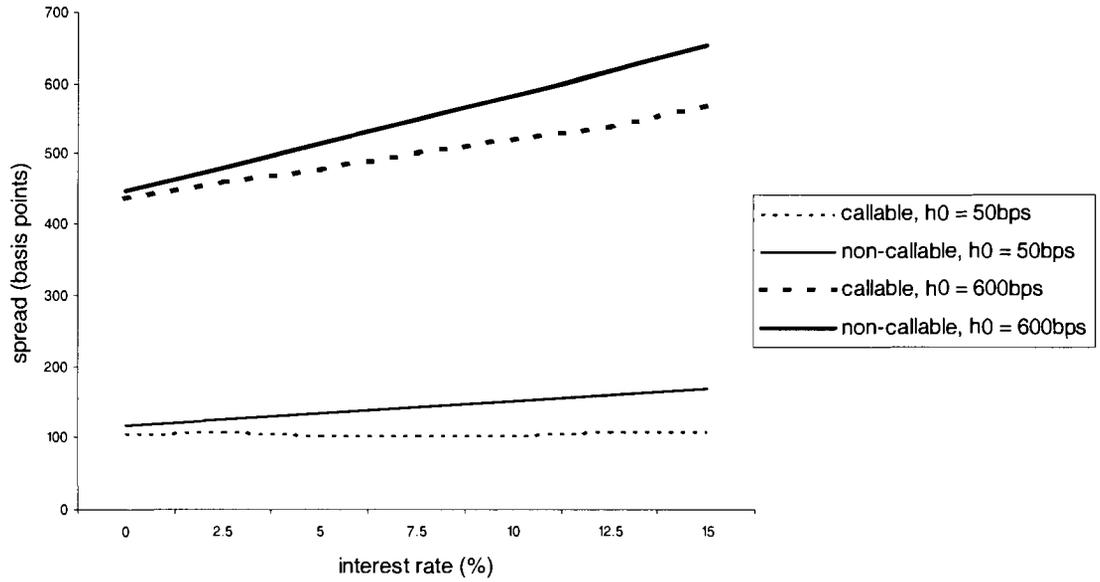
These two bonds have the same face value, maturity. For the callable bond, the “lock-out” period is 10 years and the call scheme is shown in Table 3.1. The recovery rate is 0.4 in both cases. Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). The correlation parameter, ρ , is assumed to be zero.

Figure 3. 5 Par-coupon credit spreads of 20-year callable and non-callable bonds.



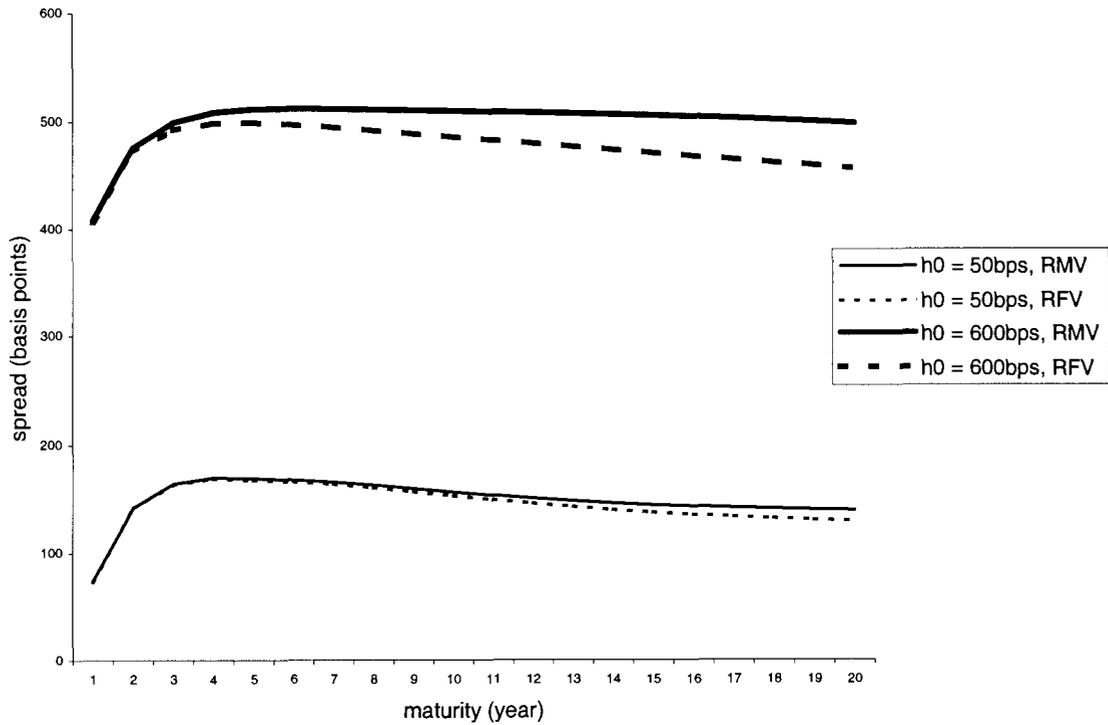
These two bonds have the same face value, maturity and issuer. For the callable bond, the “lock-out” period is 10 years and the call scheme is shown in Table 3.1. The recovery rate is 0.4 in both cases. Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). The correlation parameter, ρ , is assumed to be 0.8.

Figure 3. 6 Par-coupon credit spreads of 20-year callable and non-callable bonds.



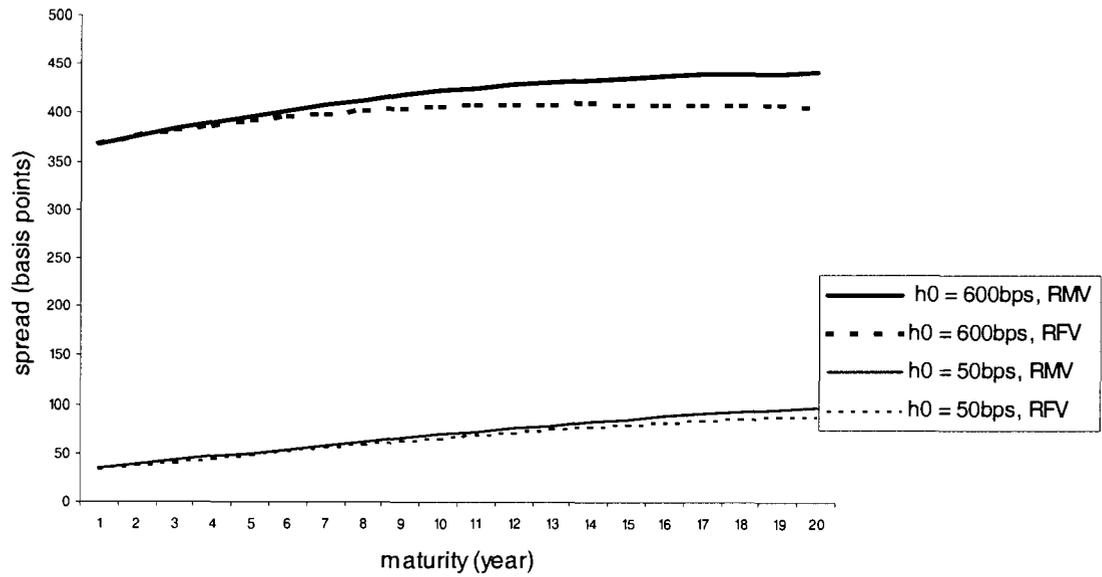
These two bonds have the same face value, maturity and issuer. For the callable bond, the “lock-out” period is 10 years and the call scheme is shown in Table 3.1. The recovery rate is 0.4 in both cases. Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). The correlation parameter, ρ , is assumed to be -0.8.

Figure 3. 7 Term structures of par-coupon total spreads of a defaultable, callable bond under RMV and RFV.



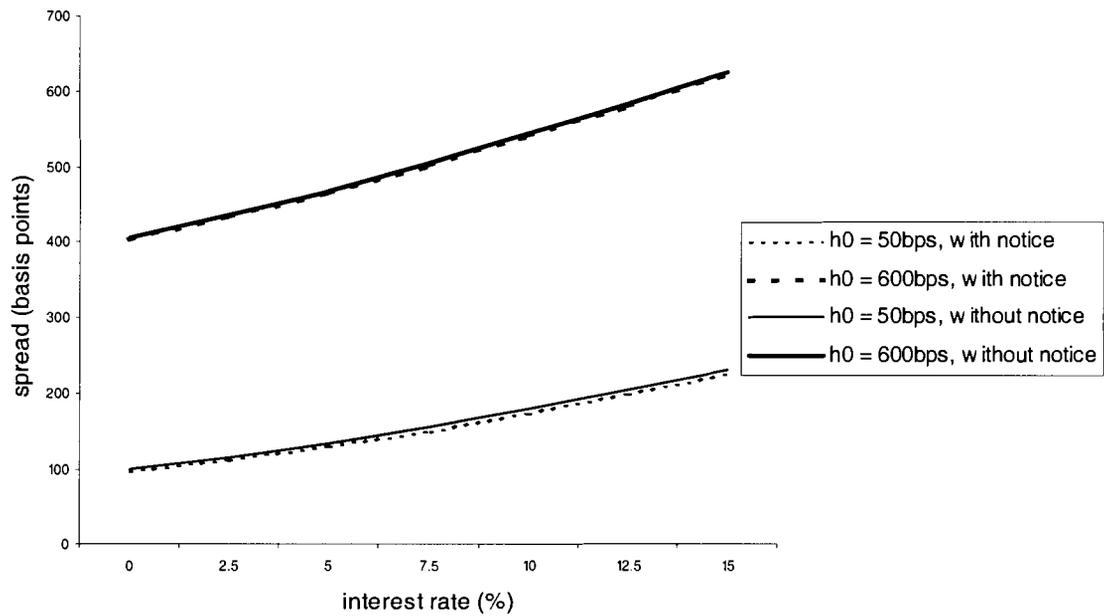
Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). The correlation parameter, ρ , is assumed to be zero. The values of model parameters are given in Table 3.1. Note that to simplify my procedure, I use a single call price, \$101.00, instead of a call scheme, and I assume that the first call dates of all maturities are one half of their respective maturities. The initial interest rate, r_0 , is chosen to be equal to 3.75%. The recovery rate is 0.4 in both cases.

Figure 3. 8 Term structures of par-coupon total spreads of defaultable, non-callable bonds under RMV and RFV.



Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). The correlation parameter, ρ , is assumed to be zero. The values of model parameters are given in Table 3.1. The recovery rate is 0.4 in both cases.

Figure 3. 9 Par-coupon total spreads of 20-year defaultable, callable bonds with and without the notice period.



The notice period is 3 months. The recovery rate is 0.4 in both cases. Two types of callable bonds are considered here: one with low credit quality ($h_0 = 600$ bps), the other with high credit quality ($h_0 = 50$ bps). The correlation parameter, ρ , is assumed to be zero.

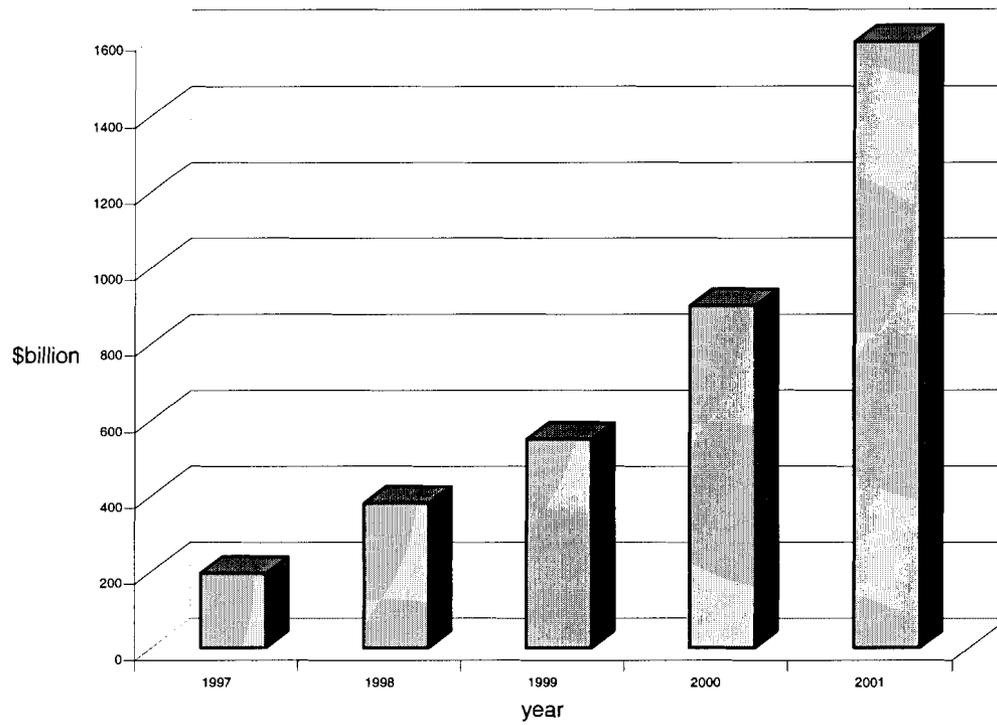
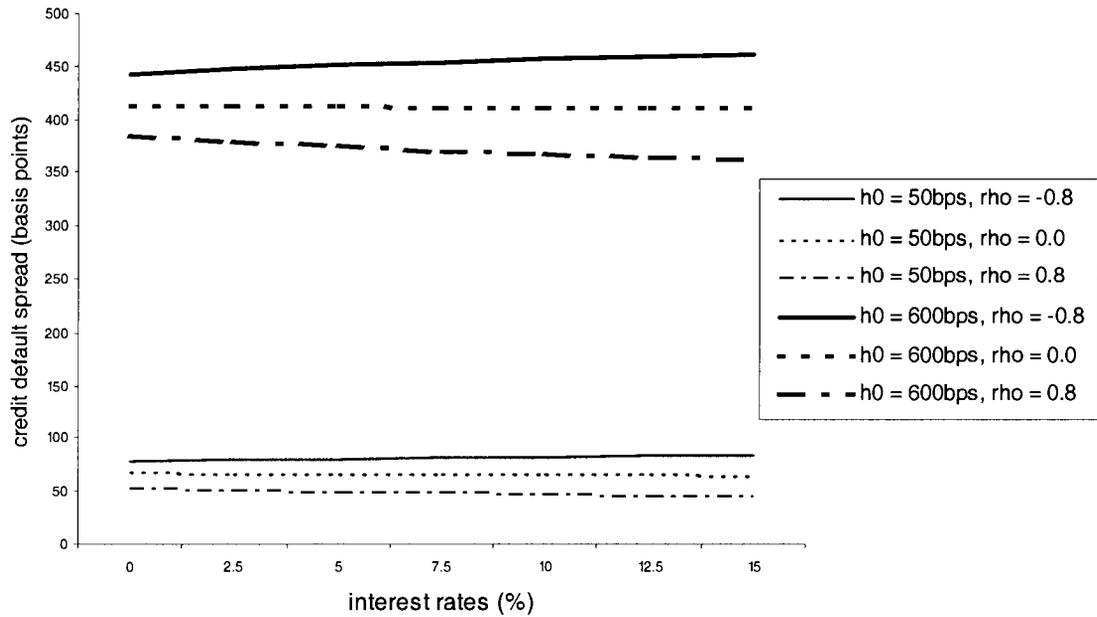


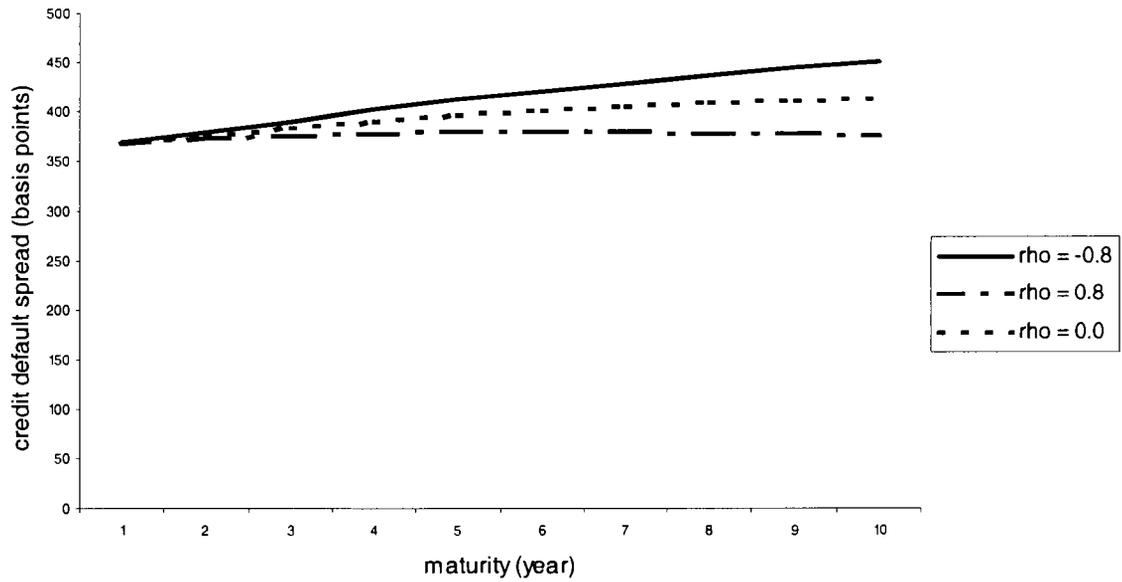
Figure 3. 10 Credit derivatives notional volumes.

Figure 3. 11 Credit default spreads of 10-year contract under RFV.



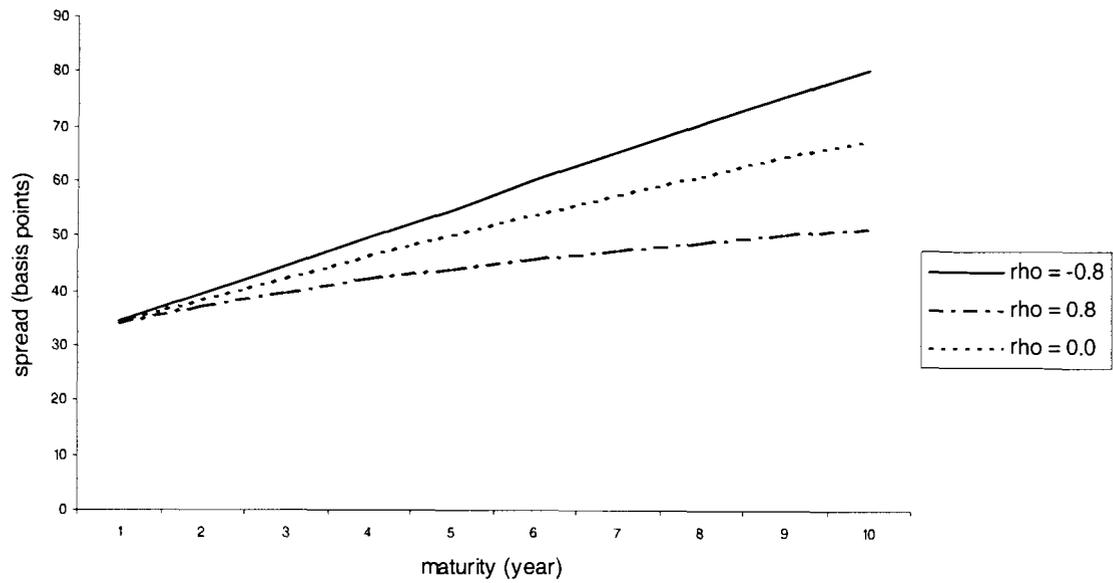
The recovery rate is 0.4. Two types of reference entities are considered here: one with low credit quality ($h_0 = 600\text{bps}$), the other with high credit quality ($h_0 = 50\text{bps}$). Three different values of the correlation variable, ρ , are considered: -0.8, 0, and 0.8.

Figure 3. 12 Term structures of credit default spreads under RFV.



The values of model parameters are given in Table 3.1. The recovery rate is 0.4. I consider only low credit quality reference entities ($h_0 = 600\text{bps}$). Three different values of the correlation variable, ρ , are considered: -0.8, 0, and 0.8.

Figure 3. 13 Term structures of credit default spreads under RFV.



The values of model parameters are given in Table 3.1. The recovery rate is 0.4. I consider only high credit quality entities ($h_0 = 50\text{bps}$). Three different values of the correlation variable, ρ , are considered: -0.8, 0, and 0.8.

Chapter 3 Appendix: The Feynman-Kac formula

Let X_t be the following diffusion process

$$dX_s^{t,x} = x + \int_s^t b(X_u^{t,x}, u) du + \int_s^t \sigma(X_u^{t,x}, u) dW_u.$$

Proposition: Feynman-Kac formula if b and σ satisfy the Lipschitz condition, if the real-valued functions f , g and ρ satisfy the Lipschitz condition on $\mathfrak{R}^K \times [0, T]$ for a scalar $T > 0$, and if the functions $b, \sigma, f, g, \rho, b_x, \sigma_x, u_x, f_x, \rho_x, b_{xx}, \sigma_{xx}, u_{xx}, f_{xx}$ and ρ_{xx} are continuous and satisfy the growth conditions, then the (twice continuously-differentiable) function $V: \mathfrak{R}^K \times [0, T] \rightarrow \mathfrak{R}$ defined by

$$V(x, t) = E\left[\int_t^T e^{-\phi(u)} f(X_u^{t,x}, u) du + e^{-\phi(T)} g(X_T^{t,x}, T)\right],$$

where ϕ is the discount factor and

$$\phi(s) = \int_t^s \rho(X_u^{t,x}, u) du$$

is the unique solution to the following partial differential equation

$$DV(x, t) - \rho(x, t)V(x, t) + f(x, t) = 0 \in \mathfrak{R}^K \times [0, T],$$

with limit condition

$$V(x, T) = g(x, T), x \in \mathfrak{R}^K,$$

where

$$DV(x, t) = V_t(x, t) + V_x(x, t)b(x, t) + \frac{1}{2}tr[\sigma^T(x, t)V_{xx}(x, t)\sigma(x, t)].$$

REFERENCES

- Acharya, V.V. and J. Carpenter, 2000, Corporate Bonds: Valuation, Hedging, and Optimal Call and Default Policies, *Working paper*, New York University.
- Altman, E.I. and V.M. Kishore, 1996, Almost Everything You Wanted to Know About Recoveries on Defaulted Bonds, *Financial Analysts Journal*, 52, 57-65.
- Ammann, M., 1999, *Pricing Derivative Credit Risk*, Springer.
- Barone, E., Giovanni Barone-Adesi and A. Castagna, 1998, Pricing Bonds and Bond Options with Default Risk, *European Financial Management*, Vol. 4, No. 2, 231-282.
- Bergstrom, A.R., 1983, Gaussian Estimation of Structural Parameters in Higher-Order Continuous Time Dynamic Models, *Econometrica*, 51, 117-152.
- Bergstrom, A.R., 1985, The Estimation of Parameters in Nonstationary Higher-Order Continuous Time Dynamic Models, *Econometric Theory*, 1, 369-385.
- Bergstrom, A.R., 1986, The Estimation of Open Higher-Order Continuous Time Dynamic Models with Mixed Stock and Flow Data, *Econometric Theory*, 2, 350-373.
- Bergstrom, A.R., 1990, *Continuous Time Econometric Modelling*, Oxford University Press, Oxford.
- Bhar, R., C. Chiarella, and W.J. Runggaldier, 2001, Estimation in Models of the Instantaneous Short Term Interest Rate by Use of a Dynamic Bayesian Algorithm, *Working Paper*, University of New South Wales.
- Black, F., E. Derman, and W. Toy, 1990, A One-Factor Model of Interest Rates and its Application to Treasury Bond Options, *Financial Analysts Journal*, 46, 33-39.
- Black, F. and J.C. Cox, 1976, Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, *Journal of Finance*, 31, 351-367.
- Black, F. and M.Schole, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637-654.
- Bliss, R.R., and D.C. Smith, 1998, The Elasticity of Interest Rate Volatility: Chan, Karolyi, Longstaff, and Sanders Revised, *Journal of Risk*, vol. 1, no. 1 (Fall):21-46.
- Bomfim, A.N., 2000, Understanding Credit Derivatives and their Potential to Synthesize Riskless Assets, *Federal Reserve Board*.
- Brennan, M.J., and E.S. Schwartz, 1980, Analyzing Convertible Bonds, *Journal of Financial and Quantitative Analysis*, 15, 907-929.

- Briys, E. and F. de Varenne, 1997, Valuing Risky Fixed Rate Debt: An Extension, *Journal of Finance and Quantitative Analysis*, 32, 239-248.
- Brown, R.H. and P.H. Dybvig, 1986, The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance*, 41, 183-193.
- Brown, R.H. and S.M. Schaefer, 1994, The Term Structure Estimation Using the Cox, Ingersoll, and Ross Model, *Journal of Financial Economics*, 35, 3-42.
- Buttler, H.-J. and J. Waldvogel, 1996, Pricing Callable Bonds by Means of Green's Function, *Mathematical Finance*, 6, 53-88.
- Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, 47, 1209-1227.
- Chapman, D.A. and N.D. Pearson, 2001, Recent Advances in Estimating Models of Term Structure, *Financial Analysts Journal*, Vol. 57, No. 4, 77-95.
- Chen, R.R. and B.J. Soprano, 1999, The Valuation of Default-Triggered Credit Derivatives, *Working paper*, Rutgers Business School.
- Chen, R.R. and L. Scott, 1993, Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income*, December, 14-31.
- Cox, J.C., 1975, Notes on Option Pricing I: Constant Elasticity of Variance Diffusions, *Working paper*, Stanford University.
- Cox, J.C., J. Ingersoll, and S. Ross, 1980, An Analysis of Variable Rate Loan Contract, *Journal of Finance*, 35, 389-403.
- , 1985, A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385-407.
- Delianedis, G. and R. Lagnado, 2000, Recovery Assumptions in the Valuation of Credit Derivatives, *Journal of Fixed Income*, March, 20-30.
- D'halluin, Y., P.A. Forsyth, K.R. Vetzal, and G. Labahn, 2001, A Numerical PDE Approach for Pricing Callable Bonds, *Applied Mathematical Finance*, 8, 49-77.
- Dothan, U.L., On the Term Structure of Interest Rates, *Journal of Financial Economics*, 6, 59-69.
- Duan, J.C., and J.G. Simonato, 1999, Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter, *Review of Quantitative Finance and Accounting*, 13(2), 111-135.
- Duffie, D., 1992, *Dynamic Asset Pricing Theory*, Princeton University Press.

- Duffie, D., 1998, Defaultable Term Structure Models with Fractional Recovery of Par, *Working paper*, Graduate School of Business, Stanford University.
- Duffie, D. and D. Lando, 2001, Term Structures of Credit Spreads with Incomplete Accounting Information, *Econometrica*, 69, 633-664.
- Duffie, D. and K.J. Singleton, 1999, Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies*, 12, 687-720.
- Duffie, D. and K.J. Singleton, 2000, *Credit Risk for Financial Institutions: Management and Pricing*.
- Duffee, G., 1999, Estimating the Price of Default Risk, *Review of Financial Studies*, 12, 197-226.
- Dullmann, K., and M. Windfuhr, 2000, Credit Spreads Between German and Italian Sovereign Bonds: Do One-Factor Affine Models Work?, *Canadian Journal of Administrative Sciences*, 17(2), June, 166-181.
- Episcopos, A., 2000, Further Evidence on Alternative Continuous Time Models of the Short-Term Interest Rate, *Journal of International Financial Markets*, 10(2), 199-212.
- Geyer, Alois L.J. and S. Pichler, 1999, A State-Space Approach to Estimate and Test Multifactor Cox-Ingersoll-Ross Models of the Term Structure, *Journal of Financial Research*, Vol. XXII, No.1, 107-130.
- Gray, S., and S. Treepongkaruna, 2001, On the Robustness of Short-term Interest Rate Models, *Working paper*, University of Queensland.
- Greene, William H., 1997, *Econometric Analysis*, Third Edition, Prentice Hall.
- Hamilton, J.D., 1994, *Time Series Analysis*, Princeton University Press.
- Harrison, J. M., and D.M. Kreps, 1979, Martingale and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory*, 20, 381-408.
- Harvey, A.C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, UK.
- Honore, P., 1998, Panel-Data Estimation of Non-Linear Term-Structure Models, *Ph.D Dissertation*, The Aarhus School of Business.
- Hull, J. and A. White, 1990, Pricing Interest Rate Derivative Securities, *Review of Financial Studies*, 3, 573-592.
- , 2000, Valuing Credit Default Swaps I: No Counterparty Default Risk, *Journal of Derivatives*, 8, 29-40.

Jarrow, R.A., D. Lando and S. Turnbull, 1997, A Markov Model for the Term Structure of Credit Spreads, *Review of Financial Studies*, 10, 481-523.

Jarrow, R.A. and S. Turnbull, 1995, Pricing Options on Financial Securities Subject to Default Risk, *Journal of Finance*, 50, 53-86.

Jarrow, R.A. and S. Turnbull, 2000, The Intersection of Market and Credit Risk, *Journal of Banking and Finance*, 24, 271-299.

Jarrow, R.A. and Y. Yildirim, 2002, Valuing Default Swaps Under Market and Credit Risk Correlation, *Journal of Fixed Income*, March, 7-19.

Jones, R.A., 1999, Estimating Correlated Diffusions, *memo*, Simon Fraser University.

Jones, Robert, A., and G. Wang, 1996, Estimation of Interest Rate Model Parameters, *memo*, Simon Fraser University.

Jordan, B.D., and R. Jorgensen, 1996, Measuring the Benefit of a Bond Refunding, *Working paper*, University of Missouri-Columbia.

J.P. Morgan, 1999, *The J.P. Morgan Guide to Credit Derivatives*, Risk Publications.

Kiff, J. and R. Morrow, 2000, Credit Derivatives, *Bank of Canada Review*, Autumn, 3-11.

Kim, I. J., K. Ramaswamy and S. Sundaresan, 1993, Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model, *Financial Management*, Autumn, 117-131.

Langstaff, F. A. and E. S. Schwartz, 1995, A Simple Approach to Valuing Risky Fixed and Floating Rate Debt, *Journal of Finance*, 50, 789-819.

Litterman, R., and J. Scheinkman, 1991, Common Factors Affecting Bond Returns, *Journal of Fixed Income*, December, 54-61.

Lund, J., 1997, Econometric Analysis of Continuous-time Arbitrage-free Models of the Term Structure of Interest Rates, *Ph.D Dissertation*, The Aarhus School of Business.

Manus, Des Mc, and D. Watt, 1999, Estimating One-factor Models of Short-Term Interest Rate, *Working Paper*, Bank of Canada.

Merton, R.C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica*, 41, 867-887.

-----, 1974, On the Pricing of Corporate Debt: the Risk Structure of Interest Rates, *Journal of Finance*, 29, 449-470.

Munnik, Jeroen F.J.de and P.C. Schotman, 1994, Cross-Sectional Versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market, *Journal of Banking and Finance*, 18, 997-1025.

Musiela, M., and M. Patkowski, 1997, *Martingale Methods in Financial Modelling*, New York: Springer.

Nandi, S., 1998, Valuation Models for Default-Risky Securities: An Overview, *Federal Reserve Bank of Atlanta*, Fourth Quarter.

Nowman, K., 1997, Gaussian Estimation of Single-factor Continuous Time Models of the Term Structure of Interest Rate, *Journal of Finance*, 52, 1695-1703.

Pearson, N.D., and T.S. Sun, 1994, Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model, *Journal of Finance*, 49, 1279-1304.

Phillips, Peter C.B., and Jun Yu, 2000, Gaussian Estimation of Continuous Time Models of the Short Term Interest Rate, *Working Paper*, Yale University.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, 1992, *Numerical Recipes in Fortran, The Art of Scientific Computing*, Second Edition, Cambridge University Press.

Shimko, D., N.Teijima, and D.van Deventer, 1993, The Pricing of Risky Debt When Interest Rates are Stochastic, *Journal of Fixed Income*, 3, 58-65.

Vasicek, O., An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177-188.

Yan, H., 2001, Dynamic Models of the Term Structure, *Financial Analysts Journal*, Vol, 57, No. 4, 60-76.

Zhou, C. S., 1997, A Jump-Diffusion Approach to Modeling Credit Risk and Valuing Defaulting Securities, *Working paper*, Federal Reserve Board.