Essays on the Term Structure of Interest Rates

Magnus Hyll

AKADEMISK AVHANDLING

Som för avläggande av filosofie
doktorsexamen vid Handelshögskolan i
Stockholm
framläggs för offentlig granskning
fredagen den 2 februari 2001, kl. 13.15
i sal 550 Handelshögskolan,
Sveavägen 65

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.
EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization
The research activities are organized in nineteen Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Center for People and Organization (PMO)	Acting Prof Jan Löwstedt
Center for Innovation and Operations Management; (T)	Prof Christer Karlsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Lennart Sjöberg
MARKETING	
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Center for Consumer Marketing; (CCM)	Associate Prof Magnus Söderlund
Marketing, Distribution and Industrial Dynamics; (D)	Prof Lars-Gunnar Mattsson
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMICS STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt,

Adress
EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Essays on the Term Structure of Interest Rates

Essays on the Term Structure of Interest Rates

Magnus Hyll

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics 2000

ISBN 91-7258-548-X

Keywords: Affine term structure, counting process, factor model, forward rate model, inversion of the yield curve, jump-diffusion process, quasi arbitrage-free term structure, short-rate model, short-rate realization, term structure of interest rates

Distributed by:
EFI, Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden

Printed by: Elanders Gotab, Stockholm 2000.

Contents

[1] Co-authored with Tomas Björk.

List of Figures

Acknowledgments

One day at the end of April 1993, under a sunny sky on a yacht somewhere in the Greek archipelago, I was asked by two good friends (at that time undergraduate students in finance) to help them construct something they called "a black box." By trading without having to take any risks whatsoever, this box was going to make us, or more precisely, them, a fortune.

There was only one problem. They didn't know enough mathematics to build a black box on their own. So they set about trying to persuade me to join the team. I was about a year short of an M.Sc. in engineering physics, so it didn't take me very long to figure out the color and the shape of the black box. But I didn't have a clue about the rest. And, as it turned out, neither did they. Yet their offer was very generous: I would receive 5% of the total earnings, leaving them to share 95%. My survival instinct got the better of me, so I declined. Going sailing with these guys was risky enough, and the little black box soon fell into oblivion.

However, I soon found a number of more important reasons to embark on the study of economics, and set sail a year later towards a Ph.D. in finance. The conditions, as expected, have been challenging, but with the harbor now in sight, I would like to thank all of those who helped me along the way.

I am very much indebted to my thesis advisor, Tomas Björk. His vast and expert knowledge of interest rate theory, invaluable comments and ability to give precise and constructive criticism on key details have helped me write this thesis. The second essay is a joint effort.

This thesis was written while I was a visiting Ph.D. student at the Universitat Pompeu Fabra in Barcelona. I am grateful to Xavier Freixas at the Departament d'Economia i Empresa for giving me that opportunity. I am also very grateful for the generosity shown by Arturo Kohatsu and Manuel Moreno, whose comments and helpful suggestions have improved several parts of this thesis. ¡Muchas gracias!

I would like to thank Hans-Peter Bermin, whom I met at the UPF. We have had many fruitful discussions about the contents of this thesis, but more importantly, HP has been truly supportive and has become a very good friend. He also read a complete draft of the final manuscript, and I therefore hold him at least partly responsible for the errors he didn't point out to me.

Over the last few years, my family and closest friends have given me enormous amounts of support, motivation and consultation (mostly for free). Without them, writing this thesis would have been impossible. I am deeply indebted to you, and dedicate this thesis to you all.

Eva Eklund, Anki Helmer, Kerstin Lindskog and Marita Rosing at the Stockholm School of Economics have generously helped me with all kinds of administrative matters. Thank you!

This thesis contained the word "rate" 452 times, and for a while there was actually a "short rat" in there somewhere too. But then Maria De Liseo proof-read the manuscript and eradicated this pest and all other bugs. The remaining mitsakes are my own. Thank you Maria, I appreciate all your comments and help!

Many thanks to Patrik Säfvenblad, the LaTeXphantom, who deserves credit for his assistance with these matters.

I would also like to express my sincere gratitude to Bankforskningsinstitutet for their financial support.

Finally, writing a thesis can be challenging and requires a number of personal characteristics. In my case, creativity and determination—not to be confused with the perhaps less glamorous factors of luck and stubbornness—have been vital. The inspiration to see this project through largely derived from the blue skies and the Mediterranean coast that Barcelona shares so generously with its inhabitants; from the sublime control and elegance of Rivaldo on the field; from the so masterly death defying and daring José Tomás; from the moving sound and beat of Ketama and Jarabe de Palo, and from the pleasure that comes with a "cortado" at Café Zurich, where the strangest things can happen...

Stockholm, October 2000
Magnus Hyll

Introduction and Summary

In most basic applications in financial economics, one common interest rate is often used to discount all future cash flows. This implies that the yield curve (the relation between the yields implied by default-free discount bonds for different maturities) is completely flat, a fact that most people would not accept. The yield curve, or the term structure of interest rates, can take on a variety of shapes that differ over time and from country to country.

This thesis examines the term structure of interest rates. In certain applications, for instance, when determining the price of a contingent claim whose future payoff depends on the performance of a bond, a model of the term structure is needed. This immediately raises a number of questions:

- What drives the movements of the yield curve, i.e., which variables are significant when designing a model for the term structure?
- How can explanatory factors be included and made use of in a model?
- Are the properties of the model consistent with stylized facts?
- What, then, can we actually use the model for?

Several key issues with regard to modeling the term structure of interest rates are dealt with in this thesis.

The first essay in this thesis, "Affine Term Structures and Short-Rate Realizations of Forward Rate Models Driven by Jump-Diffusion Processes," examines the link between the forward rate models and short-rate models in a setting where the short rate is the only explanatory factor, and where the noise in the model is driven by Wiener processes as well as counting processes. It examines the problem of determining when such a given forward rate model can be realized by a short-rate model, i.e., when the short rate induced by the forward rates is a Markov process for all choices of the initial forward rate curve.

To lay the ground for the main result, the related problem of determining when the short rate gives rise to an affine term structure is studied. Necessary and sufficient conditions for the existence of an affine term structure are given. It is then shown that a forward rate model has a short-rate realization if and only if it is generated by an affine short-rate model, i.e., a short-rate model giving rise to an affine term structure.

The second essay, "On the Inversion of the Yield Curve," co-authored with Tomas Björk, considers a general benchmark short-rate factor model of the term structure of interest rates. It is shown that the benchmark model can be extended so that the implied theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve. It is also shown that the fitted theoretical term structure can be obtained in terms of discount bond prices in a special case of the benchmark model. Further, prices in the extended model of simple contingent claims (not path dependent), such as European options, can be obtained by considering the same contract in a special case of the benchmark model.

The third and fourth essays, "An Efficient Series Expansion Approach to a Two-Factor Model of the Term Structure of Interest Rate," and "An Efficient Series Expansion Approach to The Balduzzi, Das, Foresi and Sundaram Model of the Term Structure of Interest Rates," are, as their titles indicate, similar and based on the same idea. Both papers present a specific model, and a series expansion is used to solve for discount bond prices.

The third essay presents a two-factor model where both factors follow CIR-type diffusion processes. In a comparison of this model and a corresponding Gaussian model, no substantial differences are found regarding the flexibility and shapes of the yield curves and forward rate curves they generate. The fourth essay revisits the model by BDFS, and apart from giving an explicit solution to discount bond prices, it offers an extended model that enables the implied theoretical term structure to be fitted exactly to an arbitrary initially observed yield curve.

The fifth essay, "Quasi Arbitrage-Free Discount Bond Prices in the Cox, Ingersoll and Ross Model," presents an example showing that caution is needed when solving partial differential equations that naturally appear when studying the term structure of interest rates. Discount bond prices can be obtained either from a direct computation of an expectation, or as the solution to the so-called term structure equation. It is shown that when the short rate follows the dynamics proposed by Cox, Ingersoll and Ross (CIR), there exists a solution to the term structure equation that is different from the classic CIR solution. This new and different solution fails to meet one of the standard regularity conditions, but only at one particular point. However, under additional conditions, this solution can be interpreted as a term structure, which is referred to as a "quasi arbitrage-free term structure."

1

Affine Term Structures and Short-Rate Realizations of Forward Rate Models Driven by Jump-Diffusion Processes

ABSTRACT

This essay considers forward rates driven by a multidimensional Wiener process and a multidimensional counting process where the volatilities of the forward rates depend on time, time to maturity and the short rate, and where the intensities of the counting processes depend on time and the short rate. It examines the problem of determining when such a given forward rate model can be realized by a short-rate model, i.e., when the short rate induced by the forward rates is a Markov process for all choices of the initial forward rate curve. To begin with, it studies the related problem of determining when the short rate, driven by a multidimensional Wiener process and a multidimensional counting process, gives rise to an affine term structure. Necessary and sufficient conditions for the existence of an affine term structure are given in terms of the drift and volatilities of the short rate, and the intensities of the counting processes driving the short rate. It is then shown that a forward rate model has a short-rate realization if and only if it is realized by an affine short-rate model, i.e., a short-rate model giving rise to an affine term structure.

1 Introduction

There are essentially two main approaches to modeling the term structure of interest rates. The first approach is to choose a number of explanatory stochastic variables, or factors, that are assumed to explain the behavior of the term structure. The first realistic model of this type is the model by Vasicek [28], who used the short rate as the only factor. Since then, several authors have contributed to what is now a large collection of single- and multi-factor models. The models by Dothan [13], Cox *et al.* [11], Longstaff [24], Black *et al.* [6], Black and Karasinski [7], Fong and Vasicek [16], Longstaff and Schwartz [25], Balduzzi *et al.* [2] and Chen [9] are all of the same framework in that they model their different factors as Markov processes (or systems), which allows the derivation of a partial differential equation (sometimes referred to as the term structure equation) from which the term structure can be obtained.

The other approach is to follow the idea suggested by Heath, Jarrow and Morton (HJM) [18], which requires the specification of the initial forward rate curve and the volatilities of the forward rates. In other words, the whole term structure is modeled. By construction, these models will then generate a term structure that exactly fits the initially observed yield curve. Examples of these models include Ho and Lee [19], Hull and White (Extended Vasicek) [20], Hull and White [21] and Hyll [22]. However, the short rate induced by the forward rates is typically non-Markov. In fact, only in a few special cases of volatility structures of the forward rates will the short rate follow a Markov process.[1]

Traditionally, most models of the term structure of both of the approaches described above have been of diffusion-type, although it is well known that simple diffusion models of the term structure, such as Gaussian models, do not capture all the large variations in interest rates, and therefore models including jump processes may do a better job. A number of events may serve as sources of jumps in prices in bond markets. First, intervention by central banks is to be expected since target rates are used as monetary policy. Second, supply or demand shocks may cause prices to jump, as may unanticipated economic news announcements. From an empirical point of view, as an example of a jump-diffusion process, a Poisson-Gaussian process is more flexible and therefore more likely to generate the kind of skewness and kurtosis effects consistent with market data.[2] Examples of models of the term structure with jumps include the model by Shirakawa [27] who studies term structures driven by a finite number of counting processes, and Babbs and Webber [1] who study a model in which the short rate is driven by a finite number of counting processes where the intensities of the short rate are driven by the short rate itself. Das [12] studies an extension of the Vasicek model that allows for jumps, but the arrival of the jumps and the jump size are independent of the level of the short rate. Instead, the jump

[1] For forward rates driven by a single Wiener process where the volatility may depend on a short rate, these are the models by Ho-Lee [19] and the extensions by Hull and White [20] of the Vasicek [28] and CIR [11] models.

[2] See e.g. Das [12].

is distributed with sign based on a Bernoulli distribution, and size based on an exponential distribution. With this specification, Das solves analytically for discount bond prices.

From a theoretical point of view, it is of interest to study exactly how the two different approaches to modeling the term structure are linked in the presence of a multidimensional Wiener process and a multidimensional counting process. It is of interest to determine when the short rate induced by the forward rates is a Markov process for all initial forward rate curves, i.e., when a given HJM-type forward rate model can be realized by a Markovian short-rate model, and, on the other hand, when a given Markovian short-rate model has a natural forward rate formulation (see below). Clearly, it is also interesting to understand how a short rate-realization is related to the term structure.

This essay considers forward rates driven by a multidimensional Wiener process and a multidimensional counting process where the volatilities of the forward rates depend on time, time to maturity and the short rate, and where the intensities of the counting processes depend on time and the short rate. The main objective is to determine under what conditions on the volatility structure of the forward rates there exists a short-rate realization, i.e., when the short rate induced by the forward rates is a Markov process for all choices of initial forward rate curve. However, the essay first studies the problem of determining when a short-rate model gives rise to an affine term structure. Necessary and sufficient conditions for the existence of an affine term structure are given in terms of the drift and volatilities of the short rate, and of the intensities driving the counting processes, and it is shown that the term structure is affine if and only if the short rate is affine. With this result at hand, the essay's central question can be explored. Necessary and sufficient conditions for the existence of a short-rate realization of a given forward rate model are given in terms of the volatilities of the forward rates and of the intensities driving the forward rates. It is shown that a given forward rate model has a short-rate realization if and only if the forward rate model is affine.

It is a classic problem to determine when a given forward rate model of pure diffusion-type can be realized by a short-rate model. One of the first papers is by Carverhill [8], who considers deterministic volatility structure of forward rates driven by a multidimensional Wiener process. Jeffrey [23] considers a forward rate model with a single driving Wiener process where the volatility may depend on the short rate. Under certain conditions, a forward rate model can be realized by a finite dimensional (Markovian) system. In the case of deterministic volatility structures, this problem is solved by Björk and Gombani [3], whereas the general nonlinear case is studied in detail by Björk and Svensson [5]. Forward rate models with volatility dependent on benchmark forward rates are studied by Chiarella and Kwon [10]. Ritchken and Sankarasubramanian [26] define the integrated variance of the short rate as a state variable, and show that a wide class of volatility structures of the forward rates can be reduced to a two-dimensional state variable Markov process. To sum up, virtually all previous studies assume that the forward rates are driven by one or several Wiener processes.

Examples of studies of the short rate and the existence of affine term structures when jumps are involved include Björk *et al.* [4], who provide an extensive theoretical treatment when interest rates are driven by a general marked point process as well as a Wiener process, and give sufficient conditions for the existence of an affine term structure. A recent independent paper by Filipovic (2000) [15] offers a complete characterization of affine term structure models based on a Markov short rate process, including jumps.

The closest, previous study to this is that by Jeffrey [23], and a similar methodology is used in this essay. In a setting where the forward rates are driven by a single Wiener process, it is shown that a forward rate model can be realized by a Markovian short-rate model for all choices of initial forward rate curve if and only if the volatility of the forward rates is of a certain multiplicative form. This, in turn, implies that a forward rate model has a short-rate realization if and only if it is affine.

The rest of this essay is organized as follows: Section 2 presents the basic assumptions about the forward rates and the short rate, and section 3 derives the natural consistency relations that these objects must satisfy. Section 4 examines affine term structures, and gives necessary and sufficient conditions for the existence of an affine term structure. Short-rate realizations of forward rate models are then studied in section 5, which gives necessary and sufficient conditions for the existence of a short-rate realization of a given forward rate model. Section 6 summarizes and concludes.

2 Preliminaries

We denote by $P(t,T)$ the price at time $t \geq 0$ of a discount bond maturing at $T \geq t$, and define, as usual, the forward rates $f(t,T)$ and the short rate r_t by the relations $f(t,T) = -\frac{\partial}{\partial T} \log P(t,T)$ and $r_t = \lim_{T \to t} f(t,T)$. We assume that the bond market is frictionless and free of arbitrage in the sense that there exists a martingale measure Q such that the process $\frac{P(t,T)}{B(t)}$ is a martingale under Q, where $B(t)$ is the money market account defined as $B(t) = \exp\left(\int_0^t r_u du\right)$.

Throughout the essay, the following is assumed:

Assumption 1 *It is assumed that the forward rates under Q satisfy the SDE*

$$df(t,T) = \alpha(t,T)dt + \sum_{i=1}^{m} \sigma_i(t,T)dW_t^i + \sum_{j=1}^{n} \eta_j(t,T)dN_t^j, \tag{1}$$

where $W_t^i, i = 1,2,...,m$ are independent standard Wiener processes and $N_t^j, i = 1,2,...,n$ are counting processes with intensity under Q given by the deterministic and continuous functions $\lambda_j(t,r) > 0$, and it is assumed that $\sigma_i\ (t,T)$ and $\eta_j(t,T)$ are given deterministic functions differentiable with respect to T, and that $\eta_j(t,T)$ is continuous.

In the presence of the counting processes, we recall that the Heath-Jarrow-Morton drift condition is given by

$$\alpha(t,T)=\sum_{i=1}^{m}\sigma_i(t,T)\int_t^T\sigma_i(t,u)du-\sum_{j=1}^{n}\lambda_j(t,r)\eta_j(t,T)\exp\left(-\int_t^T\eta_j(t,u)du\right), \quad (2)$$

with the obvious property[3]

$$\alpha(t,t)=-\sum_{j=1}^{n}\lambda_j(t,r)\eta_j(t,t). \quad (3)$$

In what follows, the volatilities of the forward rates are allowed to depend on the short rate, although for notational ease this r-dependence is sometimes omitted. Furthermore, the following assumption is made about the dynamics of the short rate:

Assumption 2 *It is assumed that the short rate induced from (1) satisfies the SDE*

$$dr_t=a(t,r_t)dt+\sum_{i=1}^{m}b_i(t,r_t)dW_t^i+\sum_{j=1}^{n}c_j(t,r_t)dN_t^j, \quad (4)$$

where $a(t,r), b_i(t,r)$ and $c_j(t,r)$ are given deterministic functions, and that $c_j(t,r)$ is continuous.

For technical reasons (see below), the following assumption is made:

Assumption 3 *It is assumed that $c_1(t,r)\neq c_2(t,r)...\neq c_n(t,r)\neq 0$ for all t and r.*

Remark 1 *Below, the expression "volatilities of the forward rates (short rate)" refers to the functions $\sigma_i(t,T)$ and $\eta_j(t,T)$ ($b_i(t,r)$ and $c_j(t,r)$). The functions $\sigma_i(t,T)$ ($b_i(t,r)$) are referred to as the Wiener volatilities of the forward rates (short rate). Furthermore, the expression "volatility structure of forward rates" refers to the functions $\sigma_i(t,T)$ and $\eta_j(t,T)$.*

Thus far, we have specified a forward rate model and its corresponding short-rate realization. The assumption of the existence of the short-rate realization allows us to work within the PDE framework, and it is assumed that the price at time t of a discount bond is given by

$$P(t,T)=\exp\left(-H(t,T,r_t)\right), \quad (5)$$

where $H(t,T,r)$ is assumed to be sufficiently differentiable in all variables and to satisfy the terminal condition $H(T,T,r)=0$ for all T and r. For notational ease the variable r in H will

[3] See e.g. Björk *et al.* [4].

often be omitted and H will simply be written as $H(t,T)$. Only if the third argument of H is different from r will all three arguments of H be explicitly written out. Moreover, the term structure is said to be affine if $H(t,T,r)$ is of the form[4]

$$H(t,T,r) = -A(t,T) + B(t,T)r.$$

It is assumed that $A(t,T)$ is differentiable with respect to t and T, that $B(t,T)$ is differentiable with respect to t and that $B(t,T)$ is differentiable infinitely many times with respect to T. The following lemma gives some simple relations, which will be used later in the essay.

Lemma 1 *Suppose that discount bond prices are given by (5). Then,*

$$\begin{aligned}
H_T(t,T,r) &= f(t,T,r) && (6)\\
H_T(t,t,r) &= r && (7)\\
H_{rT}(t,t,r) &= 1 && (8)\\
H_{rrT}(t,t,r) &= 0 && (9)\\
H_{tT}(t,t,r) + H_{TT}(t,t,r) &= 0 && (10)\\
H_r(t,t,r) &= 0 && (11)\\
H_t(t,t,r) + H_T(t,t,r) &= 0. && (12)
\end{aligned}$$

Proof. Relations (6) and (7) follow directly from the definition of the forward rates. By differentiating (7) with respect to r we obtain (8) and (9). Relation (10) is obtained by differentiating (7) with respect to t. Finally, relations (11) and (12) are obtained by differentiating the terminal condition $H(t,t,r) = 0$ with respect to r and t respectively. ■

3 Relations

This section considers the formal relations that must hold between the forward rates, the short rate and discount bond prices (the function $H(t,T,r)$). These objects cannot be specified independently, and it is natural to start to determine these internal relations so that the assumptions above are consistent.

[4] See Duffie and Kan [14].

Proposition 1 *Suppose that the dynamics of the forward rates and the short rate are given by (1) and (4) respectively, and that discount bond prices are given by (5). Then,*

$$\alpha(t,T) = H_{tT}(t,T) + a(t,r)H_{rT}(t,T) + \frac{1}{2}H_{rrT}(t,T)\sum_{i=1}^{m} b_i^2(t,r), \tag{13}$$

$$\sigma_i(t,T,r) = H_{rT}(t,T)b_i(t,r), \qquad i = 1,2,...,m \tag{14}$$

and

$$\eta_j(t,T,r) = H_T(t,T,r + c_j(t,r)) - H_T(t,T,r), \qquad j = 1,2,...,n. \tag{15}$$

Proof. From (6), we have that $H_T(t,T,r) = f(t,T)$. Using Itô's formula on the function $H_T(t,T,r_t)$, we have

$$\begin{aligned} dH_T &= (H_{tT} + a(t,r_t)H_{rT} + \frac{1}{2}H_{rrT}\sum_{i=1}^{m} b_i^2(t,r_t))dt + H_{rT}\sum_{i=1}^{m} b_i(t,r_t)dW_t^i \\ &\quad + \sum_{j=1}^{n} \left(H_T(t,T,r_{t-} + c_j(t,r_{t-})) - H_T(t,T,r_{t-})\right) dN_t^j. \end{aligned} \tag{16}$$

Comparing (1) and (16) gives the desired result. ■

The following useful corollaries are direct consequences of Proposition 1.

Corollary 1 *Under the same assumptions as in Proposition 1, we have*

$$b_i(t,r) = \sigma_i(t,t,r) \tag{17}$$

and

$$c_j(t,r) = \eta_j(t,t,r). \tag{18}$$

Proof. Take $T = t$ in (14) and (15) and use (7) and (8). ■

Remark 2 *We thus note that from (14) and (17) we have the somewhat surprising result that*

$$\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)} = \frac{\sigma_j(t,T,r)}{\sigma_j(t,t,r)}$$

for all $i,j = 1,2,...,m$, i.e., for all volatilities $\sigma_i(t,T,r)$, the quotient $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is independent of i.

Corollary 2 *Under the same assumptions as in Proposition 1, we have*

$$\int_t^T \sigma_i(t,u)du = b_i(t,r)H_r(t,T) \tag{19}$$

and

$$\int_t^T \eta_j(t,u)du = H(t,T,r+c_j(t,r)) - H(t,T,r).$$

Proof. Integrating (14) and (15) with respect to T, and using (11) we have

$$\begin{aligned}\int_t^T \sigma_i(t,u)du &= b_i(t,r)\int_t^T H_{rT}(t,u)du \\ &= b_i(t,r)H_r(t,T)\end{aligned}$$

and

$$\begin{aligned}\int_t^T \eta_j(t,u)du &= \int_t^T H_T(t,u,r+c_j(t,r))du - \int_t^T H_T(t,u,r)du \\ &= H(t,T,r+c_j(t,r)) - H(t,T,r). \blacksquare\end{aligned}$$

Corollary 3 *Under the same assumptions as in Proposition 1, the HJM drift condition can be written as*

$$\alpha(t,T) = H_r(t,T)H_{rT}(t,T)\sum_{i=1}^m b_i^2(t,r) - \sum_{j=1}^n \lambda_j(t,r)\eta_j(t,T)\exp\left(-\int_t^T \eta_j(t,u)du\right).$$

Proof. Using (14) and (19), we have that

$$\sum_{i=1}^m \sigma_i(t,T)\int_t^T \sigma_i(t,u)du = H_r(t,T)H_{rT}(t,T)\sum_{i=1}^m b_i^2(t,r).$$

Using (2) now gives the desired result. ■

Corollary 4 *Under the same assumptions as in Proposition 1, we have*

$$a(t,r) + H_{tT}(t,t) = \alpha(t,t). \tag{20}$$

Proof. Take $T = t$ in (13) and use (9). ■

Remark 3 *More explicitly, Corollary 4 thus states that*

$$H_{TT}(t,t,r) = a(t,r) + \sum_{j=1}^{n} \lambda_j(t,r)c_j(t,r),$$

where (3),(10) and (18) have been used. Now $H_{TT}(t,t,r)$ is the initial slope of the forward rate curve, and a heuristic interpretation of $\lambda_j(t,r)$ is that it represents the expected number of jumps per unit of time at t. Thus $\lambda_j(t,r)c_j(t,r)$ might be interpreted as being the expected jump (per unit of time) at t, and hence Corollary 4 states that the initial slope of the forward rate curve at time t is equal to the drift of the short rate plus the total expected jump at time t.

Other interesting relations may be obtained by turning to the PDE framework. Let $F(t,r)$ denote the price at time t of a (simple) contract with maturity T. Using standard financial theory, $F(t,r)$ solves the PDE:

$$\begin{aligned} F_t(t,r) + \mathcal{A}F(t,r) + \sum_{j=1}^{n} \lambda_j(t,r)\left(F(t,r+c_j(t,r)) - F(t,r)\right) &= rF(t,r) \qquad (21)\\ F(T,r) &= \Phi(r), \qquad (22) \end{aligned}$$

where $\Phi(r)$ specifies the payoff to be received at time T, and the differential operator $\mathcal{A}$ is defined by

$$\mathcal{A}F(t,r) = a(t,r)F_r(t,r) + \frac{1}{2}F_{rr}(t,r)\sum_{i=1}^{m} b_i^2(t,r).$$

By using (5) the following PDE is obtained for $H(t,T)$:

$$\begin{aligned} H_t(t,T) + A^*H(t,T) + \sum_{j=1}^{n} \lambda_j(t,r)\left(1 - e^{-(H(t,T,r+c_j(t,r))-H(t,T))}\right) + r &= 0 \qquad (23)\\ H(T,T) &= 0, \qquad (24) \end{aligned}$$

where the differential operator A^* is defined by

$$A^*H(t,T) = a(t,r)H_r(t,T) - \frac{1}{2}\left(H_r^2(t,T) - H_{rr}(t,T)\right)\sum_{i=1}^{m} b_i^2(t,r). \qquad (25)$$

4 Affine Short-Rate Models

This section gives necessary and sufficient conditions, in terms of the drift and volatilities of the short rate, and the intensities of the counting processes, for the existence of an affine term structure.

Theorem 1 *Suppose that the dynamics of the short rate are given by (4), and that discount bond prices are given by (5). Then, if*

$$\begin{aligned} a(t,r) &= \alpha^{(0)}(t) + \alpha^{(1)}(t)r && (26)\\ \sum_{i=1}^{m} b_i^2(t,r) &= \beta^{(0)}(t) + \beta^{(1)}(t)r && (27)\\ c_j(t,r) &= \gamma_j(t) && (28)\\ \lambda_j(t,r) &= \delta_j^{(0)}(t) + \delta_j^{(1)}(t)r, && (29) \end{aligned}$$

the term structure is affine and discount bond prices are given by $P(t,T) = \exp\left(-H(t,T,r_t)\right)$ *with*

$$H(t,T,r) = -A(t,T) + B(t,T)r, \tag{30}$$

where $A(t,T)$ *and* $B(t,T)$ *satisfy the system*

$$\begin{aligned} B_t(t,T) + \alpha^{(1)}(t)B(t,T) - \frac{1}{2}\beta^{(1)}(t)B^2(t,T) + \sum_{i=1}^{n}\delta_i^{(1)}(t)\left(1 - e^{-\gamma_i(t)B(t,T)}\right) &= -1 && (31)\\ B(T,T) &= 0, && (32) \end{aligned}$$

and

$$\begin{aligned} A_t(t,T) - \alpha^{(0)}(t)B(t,T) + \frac{1}{2}\beta^{(0)}(t)B^2(t,T) - \sum_{i=1}^{n}\delta_i^{(0)}(t)\left(1 - e^{-\gamma_i(t)B(t,T)}\right) &= 0 && (33)\\ A(T,T) &= 0. && (34) \end{aligned}$$

Conversely, suppose that discount bond prices are of the form (30), $B(t,T)$ *is differentiable with respect to* T *infinitely many times,* $B_T(t,T) \neq 0$*, and that* $c_1(t,r) \neq c_2(t,r)... \neq c_n(t,r) \neq 0$ *for all* t *and* r*. Then* $a(t,r)$, $\sum_{i=1}^{m}\left(b_i(t,r)\right)^2$, $c_j(t,r)$ *and* $\lambda_j(t,r)$ *must be of the form (26)–(29).*

Proof. Suppose first that (26)–(29) holds. Then it is easy to see that the system (23)–(24) has a solution of the form (30) where $A(t,T)$ and $B(t,T)$ satisfy the system (31)–(34). To prove the converse, suppose that discount bond prices are given by (30). The PDE (23) then becomes

$$\begin{aligned} & a(t,r)B(t,T) - \frac{1}{2}B^2(t,T)\sum_{i=1}^{m} b_i^2(t,r) - \sum_{j=1}^{n} \lambda_j(t,r)e^{-c_j(t,r)B(t,T)} \\ = \quad & -H_t(t,T) - r - \sum_{j=1}^{n} \lambda_j(t,r). \end{aligned} \tag{35}$$

Differentiating (35) with respect to T and dividing by $B_T(t,T) \neq 0$, we get

$$a(t,r) - B(t,T)\sum_{i=1}^{m} b_i^2(t,r) + \sum_{j=1}^{n} \lambda_j(t,r)c_j(t,r)e^{-c_j(t,r)B(t,T)} = \Gamma_1(t,T,r), \tag{36}$$

where we have defined $\Gamma_1(t,T,r) = -\frac{H_{tT}(t,T)}{B_T(t,T)}$. To simplify the notation, we let $\Gamma_k(t,T,r), k = 1,2,...$ be a sequence of affine functions of the form $\Gamma_k(t,T,r) = C_k(t,T) + D_k(t,T)r$. The exact expression of $C_k(t,T)$ and $D_k(t,T)$ is not important in what follows; what matters is the affine structure of $\Gamma_k(t,T,r)$. Differentiating once more with respect to T and dividing by $B_T(t,T)$, we obtain

$$\sum_{i=1}^{m} b_i^2(t,r) + \sum_{j=1}^{n} \lambda_j(t,r)c_j^2(t,r)e^{-c_j(t,r)B(t,T)} = \Gamma_2(t,T,r). \tag{37}$$

Again, differentiating with respect to T and dividing by $B_T(t,T)$ gives

$$\sum_{j=1}^{n} \lambda_j(t,r)c_j^3(t,r)e^{-c_j(t,r)B(t,T)} = \Gamma_3(t,T,r).$$

Repeating this procedure k times, we have

$$\sum_{j=1}^{n} \lambda_j(t,r)c_j^k(t,r)e^{-c_j(t,r)B(t,T)} = \Gamma_k(t,T,r), \tag{38}$$

and taking $T = t$ in (38), we obtain

$$\sum_{j=1}^{n} \lambda_j(t,r)c_j^k(t,r) = \Gamma_k(t,t,r), \qquad k = 3,4,.... \tag{39}$$

For n consecutive values of k, i.e., $k, k+1, ..., k+n-1$, we write (39) as a system:

$$C(c_1, ..., c_n; k)\lambda(t,r) = \begin{bmatrix} \Gamma_k(t,t,r) \\ \Gamma_{k+1}(t,t,r) \\ \vdots \\ \Gamma_{k+n-1}(t,t,r) \end{bmatrix}, \tag{40}$$

where

$$C(c_1, ..., c_n; k) = \begin{bmatrix} c_1^k(t,r) & c_2^k(t,r) & \cdots & c_n^k(t,r) \\ c_1^{k+1}(t,r) & c_2^{k+1}(t,r) & \cdots & c_n^{k+1}(t,r) \\ \vdots & \vdots & \ddots & \vdots \\ c_1^{k+n-1}(t,r) & c_2^{k+n-1}(t,r) & \cdots & c_n^{k+n-1}(t,r) \end{bmatrix}$$

and $\lambda(t,r) = \begin{bmatrix} \lambda_1(t,r) & \lambda_2(t,r) & \cdots & \lambda_n(t,r) \end{bmatrix}'$. Computing the determinant of C, we have that

$$\det C = \det C^t = (c_1 \cdot c_2 \cdot ... \cdot c_n)^k \det \bar{C},$$

where $\bar{C}$ is a Vandermonde matrix, i.e.,

$$\bar{C}(c_1, ..., c_n) = \begin{bmatrix} 1 & c_1(t,r) & \cdots & c_1^{n-1}(t,r) \\ 1 & c_2(t,r) & \cdots & c_2^{n-1}(t,r) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & c_n(t,r) & \cdots & c_n^{n-1}(t,r) \end{bmatrix}.$$

It is well known that the determinant of a Vandermonde matrix can be calculated as

$$\det \bar{C}(c_1, ..., c_n) = \Pi_{1 \le i < j \le n}(c_i(t,r) - c_j(t,r)),$$

and hence the determinant of C is

$$\det C = (c_1 \cdot c_2 \cdot ... \cdot c_n)^k \, \psi(c_1, ..., c_n),$$

where[5]

$$\psi(c_1, ..., c_n) = \Pi_{1 \le i < j \le n}(c_i(t,r) - c_j(t,r)).$$

[5] See e.g. Hammond and Sydsaeter [17] for a definition of a Vandermonde matrix.

Thus, given the above assumptions, the determinant of C is nonzero for all t and r. Furthermore, we define the matrix $C_j(t, r, c_1, ...c_{j-1}, c_{j+1}, ..., c_n, ; k)$ as the matrix C with the j:th column substituted by the vector $\begin{bmatrix} \Gamma_k(t,t,r) & \Gamma_{k+1}(t,t,r) & \cdots & \Gamma_{k+n-1}(t,t,r) \end{bmatrix}'$. Expanding the determinant of C_j, we have

$$\det C_j = \sum_{i=0}^{n-1} (-1)^{i+j+1} \Gamma_{k+i}(t,t,r) \det C_{(i+1)j},$$

where $C_{(i+1)j}$ arises from deleting row $i+1$ and column j from C. Again, by the structure of $C_{(i+1)j}$, it is easily seen that the determinant of C_j is of the form

$$\det C_j = (c_1 \cdot ...c_{j-1} \cdot c_{j+1} ... \cdot c_n)^k \sum_{i=0}^{n-1} \psi_{ij}(c_1, ..., c_n) \Gamma_{k+i}(t,t,r),$$

where $\psi_{ij}(c_1, ..., c_n)$ is some function independent of k. Now, since C is nonsingular, we can solve for arbitrary λ_j from (40), and thus obtain

$$\begin{aligned} \lambda_j(t,r) &= \frac{\det C_j}{\det C} \\ &= \frac{(c_1 \cdot ...c_{j-1} \cdot c_{j+1} ... \cdot c_n)^k \sum_{i=0}^{n-1} \psi_{ij}(c_1, ..., c_n) \Gamma_{k+i}(t,r)}{(c_1 \cdot c_2 \cdot ... \cdot c_n)^k \psi(c_1, ..., c_n)} \\ &= \frac{\sum_{i=0}^{n-1} \psi_{ij}(c_1, ..., c_n) \Gamma_{k+i}(t,r)}{c_j^k \psi(c_1, ..., c_n)}, \end{aligned} \tag{41}$$

or

$$c_j^k(t,r)\psi(c_1, ..., c_n)\lambda_j(t,r) = \sum_{i=0}^{n-1} \psi_{ij}(c_1, ..., c_n) \Gamma_{k+i}(t,t,r). \tag{42}$$

Now let $k_l, l = 1, 2, ...N$ be distinct positive integers greater than 2, and let $\mu_l(t), l = 1, 2, ...N$ be functions not all identically zero. Take $k = k_l$ in (42), multiply by $\mu_l(t)$ and sum over all l and change the order of summation to obtain

$$\sum_{l=1}^{N} \mu_l(t) c_j^{k_l}(t,r) \psi(c_1, ..., c_n) \lambda_j(t,r) = \sum_{i=0}^{n-1} \psi_{ij}(c_1, ..., c_n) \sum_{l=1}^{N} \mu_l(t) \Gamma_{k_l+i}(t,t,r).$$

Our objective is now to show that it is possible to choose N and the functions $\mu_l(t), l = 1, 2, ...N$ so that

$$\sum_{l=1}^{N} \mu_l(t)\Gamma_{k_l+i}(t,t,r) = 0, \qquad i = 0, 1, ..., n-1.$$

Since $\Gamma_{k_l+i}(t,t,r) = a_{k_l+i}(t)r + b_{k_l+i}(t)$ we have

$$\begin{aligned}\sum_{l=1}^{N} \mu_l(t)\Gamma_{k_l+i}(t,t,r) &= \sum_{l=1}^{N} \mu_l(t)\left(a_{k_l+i}(t)r + b_{k_l+i}(t)\right) \\ &= r\sum_{l=1}^{N} \mu_l(t)a_{k_l+i}(t) + \sum_{l=1}^{N} \mu_l(t)b_{k_l+i}(t). \end{aligned} \tag{43}$$

Now the right hand side of (43) is identically zero for all values of r if for each i

$$\begin{aligned}\sum_{l=1}^{N} \mu_l(t)a_{k_l+i}(t) &= 0 \\ \sum_{l=1}^{N} \mu_l(t)b_{k_l+i}(t) &= 0,\end{aligned}$$

giving $2n$ equations. In matrix form we have

$$\begin{bmatrix} a_{k_1}(t) & a_{k_2}(t) & \cdots & a_{k_N}(t) \\ b_{k_1}(t) & b_{k_2}(t) & \cdots & b_{k_N}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{k_1+n-1}(t) & a_{k_2+n-1}(t) & \cdots & a_{k_N+n-1}(t) \\ b_{k_1+n-1}(t) & b_{k_2+n-1}(t) & \cdots & b_{k_N+n-1}(t) \end{bmatrix} \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \vdots \\ \mu_N(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{44}$$

and the dimension of the matrix in (44) is thus $2n \times N$. We then know from linear algebra that for fixed t and $N > 2n$ there exists a nontrivial (nonzero) solution to the system (44). Consequently, for any fixed $N > 2n$ there exist $\mu_l(t), l = 1, 2, ...N$ not all identically zero, such that

$$\sum_{l=1}^{N} \mu_l(t)c_j^{k_l}(t,r)\psi(c_1, ..., c_n)\lambda_j(t,r) = 0.$$

Since $\lambda_j(t,r) > 0$ and $\psi(c_1,...,c_n) \neq 0$ we have in fact

$$\sum_{l=1}^{N} \mu_l(t) c_j^{k_l}(t,r) = 0. \tag{45}$$

Differentiating (45) with respect to r gives

$$\frac{\partial c_j(t,r)}{\partial r} \sum_{l=1}^{N} \mu_l(t) k_l c_j^{k_l-1}(t,r) = 0. \tag{46}$$

Assuming that $\frac{\partial c_j(t,r)}{\partial r} \neq 0$ we can divide (46) by $\frac{\partial c_j(t,r)}{\partial r}$, and then differentiate once more with respect to r. Repeating the procedure of dividing by $\frac{\partial c_j(t,r)}{\partial r}$ and then differentiating with respect to r sufficiently many times, we arrive at the contradiction that $\frac{\partial c_j(t,r)}{\partial r} = 0$, which then shows that $\frac{\partial c_j(t,r)}{\partial r} = 0$, or $c_j(t,r) = \gamma_j(t)$ for some time dependent function $\gamma_j(t)$. Since c_j is independent of r and Γ_k is affine in r, it then follows from (41) that $\lambda_j(t,r)$ is of the form (29). We can now work our way backwards, and from (37) we see that $\sum_{i=1}^{m} b_i^2(t,r)$ is of the form (27). The proof is complete by noting that from (36) it follows that $a(t,r)$ is of the form (26). ∎

Theorem 1 is thus an extension of a previously known result by Duffie and Kan [14]. This essay considers a short rate driven by a multidimensional Wiener process and a multidimensional counting process, and Theorem 1 shows that we have the previously known conditions on the drift and Wiener volatilities of the short rate for the existence of an affine terms structure. In the presence of a multidimensional counting process, we have now also determined the additional conditions on the jumps and the intensities of the counting processes that are necessary and sufficient for the existence of an affine term structure, namely that the jumps are independent of the short rate, and that the intensities are affine in the short rate.

5 Short-Rate Realizations

This section begins by presenting some definitions.

Definition 1 *We say that a given forward rate model has a* ***short-rate realization*** *if the short rate induced by the forward rates is a Markov process for all choices of the initial forward rate curve.*

Suppose now that we are given a general Markovian short-rate process, and carry out the following hypothetical program:

1. solve the term structure equation to obtain discount bond prices
2. derive volatilities of the forward rates from the discount bond prices

3. initiate a new forward rate model with the volatilities as derived from 2 (above)
4. derive the short rate from the new forward rate model in 3 (above) and an arbitrary initial forward rate curve.

The question is now: What do we obtain by having carried out this program? It is not obvious that we should obtain the Markovian short-rate model that we started with. While this is perhaps what one would expect, the result, as will be shown below, is somewhat surprising. In fact, only in a few special cases do we obtain our original Markovian short-rate process. We therefore make the following definition:

Definition 2 *We say that a given Markovian short-rate model has a* ***natural forward rate formulation*** *if the short rate derived in* 4 *(above) is a Markov process.*

Further, the following definition is made:

Definition 3 *A short-rate model with dynamics of the form (4), where* $a(t,r), b_i(t,r), c_i(t,r)$ *and* $\lambda_i(t,r)$ *are given by (26)–(29), is called an* ***affine short-rate model****, and a forward rate model that has an affine short-rate realization is called an* ***affine forward rate model.***

To lay the ground for the main result in this section, it is convenient to start with an important partial result. Given that a short-rate formulation exists, the following proposition characterizes the affine term structure in that the term structure is affine if and only if the Wiener volatilities of the forward rates are of a certain multiplicative form. However, the proposition does not explicitly determine which volatility structure of the forward rates will give rise to an affine term structure.

Proposition 2 *Suppose that the dynamics of the forward rates and the short rate are given by (1) and (4) respectively, and that discount bond prices are given by (5). Then, the term structure is affine if and only if* $\sigma_i(t,T,r)$ *is of the form*

$$\sigma_i(t,T,r) = \varphi(t,T)\sigma_i(t,t,r), \qquad i = 1,2,...,m, \tag{47}$$

where $\varphi(t,T)$ *is independent of* r*. Furthermore, if the term structure is affine, we also have*

$$\eta_j(t,T,r) = \varphi(t,T)c_j(t,r), \qquad j = 1,2,...,n.$$

Proof. Suppose first that (47) holds. Then, from (14) and (17), we have

$$\begin{aligned} H_{rT}(t,T,r) &= \frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)} \\ &= \varphi(t,T). \end{aligned} \tag{48}$$

Integrating (48) with respect to r gives

$$H_T(t,T,r) = \varphi(t,T)r + \theta(t,T), \tag{49}$$

where $\theta(t,T)$ is an arbitrary function independent of r. Integrating (49) with respect to T and using $H(t,t,r) = 0$ for all t shows that the term structure is affine. To prove the converse, suppose that the term structure is affine. Then $H_{rT}(t,T,r)$ is a time dependent function, say $H_{rT}(t,T,r) = \varphi(t,T)$. Then, again from (14) and (17), it follows that $\sigma_i(t,T,r)$ is of the form (47). Furthermore, if $\sigma_i(t,T,r)$ is of the form (47) (i.e., the term structure is affine), then it follows from (15) and (49) that $\eta_j(t,T,r) = \varphi(t,T)c_j(t,r)$, for $j = 1,2,...,n$. ■

The following theorem is an extension to the jump-diffusion case of a previous result by Jeffrey [23], and determines which forward rate models can be realized by a short-rate model.

Theorem 2 *Suppose that the dynamics of the forward rates are given by (1). Then, the forward rate model has a short-rate realization of the form (4) if and only if it is affine.*

Proof. The proof is slightly different but essentially a copy of the proof in Jeffrey [23]. From (14), we have that

$$H_T(t,T,r) = H_T(t,T,0) + \int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy. \tag{50}$$

Furthermore, differentiating (50) with respect to t, we have that

$$H_{tT}(t,T,r) = H_{tT}(t,T,0) + \frac{\partial}{\partial t}\int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy, \tag{51}$$

and in particular

$$H_{tT}(t,t,r) = H_{tT}(t,t,0) + \frac{\partial}{\partial t}\int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy\bigg|_{T=t}. \tag{52}$$

Using (20) and (52), the drift of the short rate becomes

$$a(t,r) = \alpha(t,t) - H_{tT}(t,t,0) - \frac{\partial}{\partial t}\int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy\bigg|_{T=t}. \tag{53}$$

From (13), we have that

$$H_{tT}(t,T) + a(t,r)\frac{\sigma_i(t,T)}{\sigma_i(t,t)} + \frac{1}{2}\left\{\frac{\partial}{\partial r}\frac{\sigma_i(t,T)}{\sigma_i(t,t)}\right\}\sum_{i=1}^m \sigma_i^2(t,t) = \alpha(t,T), \tag{54}$$

and integrating (54) with respect to t and using $H_T(t,T) = f(t,T)$ and $H_T(T,T) = r$ gives

$$\begin{aligned} f(t,T) &= \int_t^T a(u,r)\frac{\sigma_i(u,T)}{\sigma_i(u,u)}du + \frac{1}{2}\int_t^T \left\{\frac{\partial}{\partial r}\frac{\sigma_i(u,T)}{\sigma_i(u,u)}\right\}\sum_{i=1}^m \sigma_i^2(u,u)du \\ &\quad - \int_t^T \alpha(u,T)du + r, \end{aligned} \tag{55}$$

where we recall that

$$\alpha(t,T) = \sum_{i=1}^m \sigma_i(t,T)\int_t^T \sigma_i(t,u)du - \sum_{j=1}^n \lambda_j(t,r)\eta_j(t,T)\exp\left(-\int_t^T \eta_j(t,u)du\right).$$

Taking $t = 0$ in (55), we get

$$\begin{aligned} f(0,T) &= \int_0^T a(u,r)\frac{\sigma_i(u,T)}{\sigma_i(u,u)}du + \frac{1}{2}\int_0^T \left\{\frac{\partial}{\partial r}\frac{\sigma_i(u,T)}{\sigma_i(u,u)}\right\}\sum_{i=1}^m \sigma_i^2(u,u)du \\ &\quad - \int_0^T \alpha(u,T)du + r, \end{aligned} \tag{56}$$

and hence the drift of the short rate $a(t,r)$ can be determined in terms of the volatility structure of the forward rates, the intensities of the counting processes, and the initial forward rate curve. To be able to fit any initial forward rate curve, from (53) we see that it is necessary that $H_{tT}(t,t,0)$ can be chosen arbitrarily. From (51), (53) and (54) it follows that

$$\begin{aligned} & H_{tT}(t,T,0) + \frac{\partial}{\partial t}\int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy + \frac{1}{2}\left\{\frac{\partial}{\partial r}\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}\right\}\sum_{i=1}^m \sigma_i^2(t,t,r) \\ =\; & \alpha(t,T) + \left(H_{tT}(t,t,0) + \frac{\partial}{\partial t}\int_0^r \frac{\sigma_i(t,T,y)}{\sigma_i(t,t,y)}dy\bigg|_{T=t} - \alpha(t,t)\right)\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}. \end{aligned} \tag{57}$$

Now let $H_{tT}(t,t,0)$ be an arbitrary function such that (57) holds, and suppose that $H^*_{tT}(t,t,0)$ is one particular function such that (57) holds. Then,

$$H_{tT}(t,T,0) - H^*_{tT}(t,T,0) = (H_{tT}(t,t,0) - H^*_{tT}(t,t,0))\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}. \tag{58}$$

Differentiating (58) with respect to r, we find that

$$0 = (H_{tT}(t,t,0) - H^*_{tT}(t,t,0))\frac{\partial}{\partial r}\left\{\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}\right\}. \tag{59}$$

If $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is dependent on r, then from (59) $H^*_{tT}(t,t,0) = H_{tT}(t,t,0)$, and from (53) and (56) it follows that $f^*(0,T) = f(0,T)$, i.e., the theoretical term structure can only fit *one* initial forward rate curve. From (59), it also follows that $H^*_{tT}(t,t,0)$ can be chosen freely if and only if $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is independent of r, and a necessary and sufficient condition for the initial forward rate curve to be arbitrary is thus that $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is independent of r. Hence, a short-rate realization can exist if and only if $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is independent of r. Now, if $\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)}$ is independent of r, $\sigma_i(t,T,r)$ is of the form $\sigma_i(t,T,r) = \varphi(t,T)\sigma_i(t,t,r)$, and from Proposition 1 this is a necessary and sufficient condition for the term structure to be affine. Technicalities aside, from Theorem 1 the term structure is affine if and only if the short rate is affine, and thus, by definition, the forward rate model is affine. This completes the proof. ∎

Thus, a given forward rate model has a short-rate realization if and only if it is affine. Also, we have seen that a short-rate realization exists if and only if $H_{tT}(t,t,0)$ can be chosen arbitrarily. This means that $\alpha^{(0)}(t)$ is given by the solution to a certain integral equation, given in the following corollary.

Corollary 5 *Suppose that the dynamics of the forward rates are given by (1), and that the short-rate realization is given by (4) with $a(t,r), b_i(t,r), c_i(t,r)$ and $\lambda_i(t,r)$ as in (26)-(29). Then $\alpha^{(0)}(t)$ is given as the solution to the following integral equation:*

$$\begin{aligned}
&\int_0^t \alpha^{(0)}(s)B_T(s,t)ds + \int_0^t \sum_{j=1}^n \delta_j^{(0)}(s)\gamma_j(s)B_T(s,t)e^{-\gamma_j(s)B(s,t)}ds \\
= &\int_0^t \beta^{(0)}(s)B(s,t)B_T(s,t)ds + f^*(0,t) - B_T(0,t)r_0,
\end{aligned} \tag{60}$$

where $f^*(0,t)$ denotes the initially observed forward rate curve.

Proof. We let $P^*(0,T), T \geq 0$, denote the initially observed yield curve, and we note that it is equivalent to fit either initial discount bond prices or the initial forward rate curve. Since we have an affine term structure, we have a perfect initial fit if

$$P^*(0,T) = \exp\left(A(0,T) - B(0,T)r_0\right). \tag{61}$$

Taking the logarithm of (61), differentiating with respect to T, and using the definition of the initial forward rates we obtain

$$f^*(0,T) + A_T(0,T) - B_T(0,T)r_0 = 0. \tag{62}$$

We know from above that to be able to fit any initial yield curve, it is necessary that $H_{tT}(t,t,0)$, and thus $A_{tT}(t,t)$, can be chosen arbitrarily. As a consequence, by differentiating (33) with respect to T and evaluating at (t,t), it follows that it is necessary that $\alpha^{(0)}(t)$ can be chosen arbitrarily. Integrating (33) and then differentiating with respect to T, and using (62) then gives (60). ■

We have thus determined the class of forward rate models that have a short-rate realization for the case when the volatilities of the forward rates depend on time, time to maturity and the short rate, and where the intensities of the counting processes depend on time and the short rate. Returning to the hypothetical program in this section, we now know that when executing the last step (step 4), we will obtain a Markov process for the short rate for all choices of an initial yield curve if and only if the volatility structure derived from step 2 is of the form $\sigma_i(t,T,r) = \varphi(t,T)\sigma_i(t,t,r)$. Then, from (14) and (17) it follows that such a volatility structure necessarily was derived from an affine term structure, and Theorem 1 shows that the short-rate model we began with in the first place necessarily was affine. Thus, only the affine short-rate models have a natural forward rate formulation. In other words, a short-rate model has a natural forward rate formulation if and only if the corresponding forward rate model has a short-rate realization.

6 Summary and Conclusions

This essay studies forward rate models driven by a multidimensional Wiener process and a multidimensional counting process, and it is assumed that the volatilities of the forward rates depend on time, time to maturity and the short rate, and that the intensities of the counting processes depend on time and the short rate.

First, the essay set about to determine when a short-rate model gives rise to an affine term structure. Necessary and sufficient conditions for the existence of an affine term structure were given in terms of the drift and volatilities of the short rate, and of the intensities driving the counting processes. The term structure was shown to be affine if and only if the short rate is affine. Next, the existence of a short-rate realization of a given forward rate model was investigated. Necessary and sufficient conditions for the existence of a short-rate realization were given in terms of the volatilities of the forward rates. A given forward rate model was shown to have a short-rate realization if and only if the forward rate model is affine.

In summary, the forward rate models that have a short-rate realization were determined. Perhaps somewhat surprisingly, these forward rate models were all revealed to be affine.

References

[1] Babbs, S. and N. Webber (1993): "A Theory of the Term Structure with an Official Short Rate," Working paper, University of Warwick.

[2] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram (1996): "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income,* December, 43–53.

[3] Björk, T., and A. Gombani (1999): "Minimal Realizations of Forward Rates," *Finance and Stochastics* **3**(4), 413–432.

[4] Björk, T., Y. Kabanov, and W. Runggaldier (1997): "Bond Market Structure in the Presence of Marked Point Processes," *Mathematical Finance 7(2),* 211–239.

[5] Björk, T. and L. Svensson (1999): "On the Existence of Finite Dimensional Realizations for Nonlinear Forward Rate Models," Working paper, Stockholm School of Economics.

[6] Black, F., E. Derman, and W. Toy (1990): "A One-Factor Model of Interest Rates and its Applications to Treasury Bond Options," *Financial Analysts Journal,* January-February, 33–39.

[7] Black, F. and P. Karasinski (1991): "Bond and Option Pricing when Short Rates are Lognormal," *Financial Analysts Journal,* July-August, 52–59.

[8] Carverhill, A. (1994): "When Is the Short Rate Markovian?," *Mathematical Finance* **4**(4), 305–312.

[9] Chen, L. (1996): "Interest Rate Dynamics, Derivatives Pricing, and Risk Management," Springer Verlag.

[10] Chiarella, C. and K. Kwon (1998): "Forward Rate Dependent Markovian Transformations of the Heath-Jarrow-Morton Term Structure Model," Working paper, School of Finance and Economics, University of Technology, Sydney.

[11] Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985b): "A Theory of the Term Structure of Interest Rates," *Econometrica* **53**(2), 385–407.

[12] Das, S.R. (1996): "Poisson-Gaussian Processes and the Bond Markets," Working paper, Harvard Business School.

[13] Dothan, M. (1978): "On the Term Structure of Interest Rates," *Journal of Financial Economics* **6**, 59–69.

[14] Duffie, D. and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance* **6**(4), 379–406.

[15] Filipovic, D. (2000): "A General Characterization of Affine Term Structure Models," Working paper, ETH, Zurich.

[16] Fong, H.G. and O.A. Vasicek (1991): "Interest Rate Volatility as a Stochastic Factor," Unpublished manuscript, Gifford Fong Associates.

[17] Hammond, P. and K. Sydsaeter (1995): "Mathematics For Economic Analysis," Prentice Hall.

[18] Heath, D., R. Jarrow, and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica* **60**(1), 77–105.

[19] Ho, T.S. and S.B. Lee (1986): "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* **41**(5), 1011–1029.

[20] Hull, J. and A. White (1990): "Pricing Interest Rate Derivative Securities," *The Review of Financial Studies* **3**(4), 573–592.

[21] Hull, J. and A. White (1994b): "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models," *The Journal of Derivatives,* Winter, 37–48.

[22] Hyll, M. (2000d): "An Efficient Series Expansion Approach to the Balduzzi, Das, Foresi and Sundaram Model of the Term Structure of Interest Rates," Working paper, Stockholm School of Economics.

[23] Jeffrey, A. (1995): "Single Factor Heath-Jarrow-Morton Term Structure Models Based on Markov Spot Interest Rate Dynamics," *Journal of Financial and Quantitative Analysis* **30**(4), 619–642.

[24] Longstaff, F.A. (1989): "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics* **23**, 195–224.

[25] Longstaff, F.A. and E.S. Schwartz (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance* **47**(4), 1259–1282.

[26] Ritchken, P. and L. Sankarasubramanian (1995): "Volatility Structures of Forward Rates and the Dynamics of the Term Structure," *Mathematical Finance* **5**(1), 55–72.

[27] Shirakawa, H. (1991): "Interest Rate Option Pricing with Poisson-Gaussian Forward Rate Curve Processes," *Mathematical Finance* **1**(4), 77–94.

[28] Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* **5**, 177–188.

2

On the Inversion of the Yield Curve[1]

ABSTRACT

In this essay we consider a rather general benchmark short-rate factor model of the term structure of interest rates, and show that the model can be extended so that the implied theoretical term structure exactly fits the initially observed yield curve. The model is general in the sense that many existing models can be seen as special cases. We also show that the fitted theoretical term structure can be obtained in terms of discount bond prices in a special case of the benchmark model. Furthermore, we show that prices in the extended model of simple contingent claims (not path dependent), such as European options, can be obtained by considering the same contingent claim in a special case of the benchmark model.

1 Introduction

Due to the wider use of interest rate derivatives over the last 25 years or so, there has been a flow of models for valuing these instruments correctly. The early classic models, such as the single-factor models by Vasicek [12] and Cox, Ingersoll and Ross [3], have the advantage of pricing many interest rate derivatives analytically. Two and three-factor models followed. Examples include the two-factor models by Longstaff and Schwartz [11] and Fong and Vasicek [5], and the three-factor model by Balduzzi *et al.* (BDFS) [1]. Models with an increased number of factors can generate more flexible and realistic term structure shapes, and therefore probably better price derivative instruments based on the yield curve. However, these models all share the drawback of not being able to fit the initially observed yield curve exactly.

Ho and Lee [7] pioneered a new approach and proposed a model where the initial yield curve is an input to the model and the term structure generated by the model then automatically fits the initial yield curve. Soon after Hull and White followed with their Extended Vasicek model [8], and a new kind of model, a forward rate model, was then designed by Heath, Jarrow and Morton [6]. Since then, Hull and White [9] have presented a two-factor model that fits the initial yield curve exactly, and Hyll [10] has extended the BDFS three-factor model [1] for the same purpose.[2]

In this essay we consider a general benchmark short-rate factor model of the term structure of interest rates, and propose a method for extending the model so that the theoretical term structure is consistent with the initially observed yield curve, i.e., so that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve. The benchmark model is rather general, and we stress that we do not confine ourselves to the popular class of models giving rise to an affine term structure.[3]

We also study the pricing of a general contingent claim (i.e., not path dependent) in the extended model, and show that the price of the contingent claim in the extended model can be expressed in terms of the price of the same contingent claim in a special case of the benchmark model.

Very little seems to be written on the inversion of the yield curve in general. Apart from the models above mentioned, which are all separate cases of models fitting the initial yield curve, Brigo and Mercurio [2] present a method to extend time-homogenous short-rate models to models that fit an arbitrary initial yield curve exactly. Their method involves a deterministic shift of the short rate, and preserves the analytical tractability of the original model. Our method of inverting the yield curve is the classic one in that we extend the benchmark model by introducing a (particular) time dependent function in the drift of the short rate à la Hull

[2] Hyll [10] shows that discount bond prices in BDFS model [1] can be obtained using a series expansion, and also extends the BDFS model so that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve.

[3] See Duffie and Kan [4].

and White to obtain a perfect initial fit. Like the method of Brigo and Mercurio, our method leads to a deterministic shift of the short rate, but we also cover multi-factor models.

The rest of this essay is organized as follows: In section 2 we state the benchmark model of the term structure and show that it can be fitted exactly to an arbitrary initially observed yield curve. In section 3 we obtain a general pricing formula for a simple contingent claim, and show that the price in the extended model can be determined in terms of the price of the same contingent claim in a special case of the benchmark model. Several fully worked examples are then included in section 4. We show in section 5 that our inversion technique is also suited for different but similar settings, and consider a special case when the short rate is the sum of several stochastic factors. A fully worked example is also given. Section 6 summarizes and concludes.

2 Fitting the Initial Yield Curve Exactly

We take as our starting point a general benchmark short-rate factor model $\mathcal{M}$ of the term structure of interest rates, where the short rate r and the additional n-dimensional factor vector Y under a martingale measure Q satisfy the stochastic differential equations

$$\begin{aligned} dr_s &= (b - ar_s - c(s, Y_s))\, ds + \gamma(s, Y_s) dW_s, \qquad s > t \\ r_t &= r \end{aligned} \tag{1}$$

and

$$\begin{aligned} dY_s &= \mu(s, Y_s) ds + \sigma(s, Y_s) dW_s, \qquad s > t \\ Y_t &= y. \end{aligned} \tag{2}$$

In the dynamics of the short rate in (1), a and b are constants, $c : \mathcal{R}^{n+1} \to \mathcal{R}$ is a deterministic function, $\gamma : \mathcal{R}^{n+1} \to \mathcal{R}^{n+1}$ is a deterministic function and W is a standard Wiener process of dimension $n + 1$. Furthermore, in (2), $\mu : \mathcal{R}^{n+1} \to \mathcal{R}^n$ and $\sigma : \mathcal{R}^{n+1} \to \mathcal{R}^n \times \mathcal{R}^{n+1}$ are deterministic functions. Of particular interest is a special case of $\mathcal{M}$ when $b = 0$. This special model is denoted by $\mathcal{M}^0$, and the corresponding short rate is denoted by r^0 and satisfies the stochastic differential equation

$$\begin{aligned} dr_s^0 &= -\left(ar_s^0 + c(s, Y_s)\right) ds + \gamma(s, Y_s) dW_s, \qquad s > t \\ r_t^0 &= r, \end{aligned} \tag{3}$$

and Y is given by (2). In the model $\mathcal{M}^0$, we denote by $P^0(t,T)$ the price at time $t \geq 0$ of a discount bond with unit payment at time $T \geq t$, i.e., $P^0(t,T) = F^0(t,T,r_t,Y_t)$, where

$$F^0(t,T,r,y) = E^Q_{t,r,y}\left[\exp\left(-\int_t^T r_s^0 ds\right)\right],$$

and where the short hand notation used is $E^Q_{t,r,y}[\cdot] = E^Q[\cdot \mid r_t = r, Y_t = y]$. As usual, we define the instantaneous forward rates by $f^0(t,T) = -\frac{\partial}{\partial T}\log P^0(t,T)$ for $0 \leq t \leq T$.

We now propose a new model, denoted by $\mathcal{M}^\varphi$, which is an extension of $\mathcal{M}$ in that it allows for a time dependent function in the drift of the short rate. More precisely, we assume that the short rate r^φ in the extended model $\mathcal{M}^\varphi$ under Q satisfies the stochastic differential equation

$$\begin{aligned} dr_s^\varphi &= (\varphi(s) - ar_s^\varphi - c(s,Y_s))\,ds + \gamma(s,Y_s)dW_s, \qquad s > t \\ r_t^\varphi &= r, \end{aligned} \tag{4}$$

and that Y is given by (2). We denote by $P^\varphi(t,T)$ the price at time $t \geq 0$ of a discount bond with unit payment at $T \geq t$ in the model $\mathcal{M}^\varphi$, i.e., $P^\varphi(t,T) = F^\varphi(t,T,r_t^\varphi,Y_t)$, where

$$F^\varphi(t,T,r,Y) = E^Q_{t,r,y}\left[\exp\left(-\int_t^T r_s^\varphi ds\right)\right],$$

and consequently, we denote the instantaneous forward rates by $f^\varphi(t,T) = -\frac{\partial}{\partial T}\log P^\varphi(t,T)$ for $0 \leq t \leq T$. We denote the initially observed yield curve by $P^*(0,t)$, $t \geq 0$, and $f^*(0,t) = -\frac{\partial}{\partial t}\log P^*(0,t)$ is the corresponding observed forward rate curve. We assume that $f^*(0,t)$ is differentiable with respect to t so that $f_T^*(0,t) := \frac{\partial}{\partial t}f^*(0,t)$ exists for all $t \geq 0$. Given a fixed r_0 and y_0, our objective is now to choose φ in (4) such that the model $\mathcal{M}^\varphi$ initially (at $t = 0$) can be fitted to an arbitrary observed yield curve exactly. We then have the following proposition:

Proposition 1 *The theoretical term structure implied by the model $\mathcal{M}^\varphi$ can be fitted exactly to an arbitrary initially observed yield curve. A perfect initial fit is obtained by choosing*

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right).$$

Discount bond prices are then given by

$$P^\varphi(t,T) = \frac{P^*(0,T)}{P^*(0,t)}\frac{P^0(0,t)}{P^0(0,T)}P^0(t,T)\exp\left(\left(f^*(0,t) - f^0(0,t)\right)B(t,T)\right), \tag{5}$$

where

$$B(t,T) = \frac{1 - \exp\left(-a(T-t)\right)}{a}.$$

Proof. A key observation is that from the dynamics of r^0 and r^φ is follows that

$$r_s^0 = e^{-a(s-t)}r - \int_t^s e^{-a(s-u)}c(u,Y_u)du + \int_t^s e^{-a(s-u)}\gamma(u,Y_u)dW_u$$

and

$$\begin{aligned} r_s^\varphi &= e^{-a(s-t)}r + \int_t^s e^{-a(s-u)}\varphi(u)du \\ &\quad - \int_t^s e^{-a(s-u)}c(u,Y_u)du + \int_t^s e^{-a(s-u)}\gamma(u,Y_u)dW_u, \end{aligned}$$

and hence

$$r_s^\varphi = r_s^0 + \chi(t,s), \tag{6}$$

where

$$\chi(t,s) = \int_t^s e^{-a(s-u)}\varphi(u)du. \tag{7}$$

As a consequence, discount bond prices $P^\varphi(t,T)$ in the model $\mathcal{M}^\varphi$ can be determined in terms of discount bond prices $P^0(t,T)$ from $\mathcal{M}^0$. We have

$$\begin{aligned} P^\varphi(t,T,r,y) &= E^Q_{t,r,y}\left[\exp\left(-\int_t^T r_s^\varphi ds\right)\right] \\ &= E^Q_{t,r,y}\left[\exp\left(-\int_t^T \left(r_s^0 + \chi(t,s)\right)ds\right)\right] \\ &= \exp\left(-\int_t^T \chi(t,s)ds\right)E^Q_{t,r,y}\left[\exp\left(-\int_t^T r_s^0 ds\right)\right] \\ &= \exp\left(-\int_t^T \chi(t,s)ds\right)P^0(t,T,r,y), \end{aligned} \tag{8}$$

and thus

$$P^\varphi(0,T,r_0,y_0) = \exp\left(-\int_0^T \chi(0,s)ds\right)P^0(0,T,r_0,y_0). \tag{9}$$

Now take the logarithm of (9), differentiate with respect to T and use the definitions of the forward rates. Then

$$f^{\varphi}(0,T) = \chi(0,T) + f^0(0,T).$$

We now want to choose φ such that $f^{\varphi}(0,T) = f^*(0,T)$ for all $T \geq 0$; we then have

$$f^*(0,T) - f^0(0,T) = \int_0^T e^{-a(T-u)}\varphi(u)du, \tag{10}$$

where we have used (7). Differentiating (10) with respect to T, we have

$$f_T^*(0,T) - f_T^0(0,T) = \varphi(T) - a\int_0^T e^{-a(T-u)}(\varphi(u) - b)du,$$

and using (10) we get

$$\varphi(T) = f_T^*(0,T) - f_T^0(0,T) + a\left(f^*(0,T) - f^0(0,T)\right). \tag{11}$$

Then, using(7) and (11), we obtain

$$\begin{aligned} -\int_t^T \chi(t,s)ds &= -\int_t^T\int_t^s e^{-a(s-u)}\varphi(u)duds \\ &= \log\frac{P^*(0,T)P^0(0,t)}{P^*(0,t)P^0(0,T)} + \left(f^*(0,t) - f^0(0,t)\right)B(t,T). \end{aligned} \tag{12}$$

Equations (8) and (12) now prove (5). ■

3 A General Pricing Formula

In this section, we consider a contingent claim Z with maturity T and payoff function $\Phi(r_T, Y_T)$. The contract is simple in the sense that it only depends on the terminal values of the factors, i.e., r_T and Y_T, and not on the paths of the factors. We denote the price of Z at time $t, 0 \leq t \leq T$, in the model $\mathcal{M}^{\varphi}$ by $F^{\varphi}(t,r,y)$, and the corresponding price in $\mathcal{M}^0$ is denoted by $F^0(t,r,y)$. We then have the following proposition:

Proposition 2 *Let Z be a contingent claim with payoff function $\Phi(r_T, Y_T)$. The price of Z in the model $\mathcal{M}^{\varphi}$ for an arbitrary function $\varphi(t)$ is given by*

$$F^{\varphi}(t,r,y) = e^{\int_t^T \psi(s,T))ds}F^0(t, r + \psi(t,T), y),$$

where

$$\begin{aligned}\psi(t,T) &= \int_t^T \frac{\partial B(u,t)}{\partial t}\varphi(u)du \\ \int_t^T \psi(s,T)ds &= -\int_t^T B(u,t)\varphi(u)du \\ B(t,T) &= \frac{1-\exp\left(-a(T-t)\right)}{a}.\end{aligned}$$

Specifically, if the function $\varphi(t)$ in the model $\mathcal{M}^\varphi$ is given by Proposition 1 in the previous section, we have

$$F^\varphi(t,r,y) = \frac{P^*(0,T)}{P^*(0,t)}\frac{P^0(0,t)}{P^0(0,T)}F^0(t,r+\psi(t,T),y)e^{B(T,t)\left(f^0(0,T)-f^*(0,T)\right)}, \tag{13}$$

where

$$\psi(t,T) = f^0(0,t) - f^*(0,t) - e^{a(T-t)}\left(f^0(0,T) - f^*(0,T)\right).$$

Proof. For a fixed choice of t, r, y and φ, define the processes r_s^φ and Y_s by

$$\begin{aligned}dr_s^\varphi &= \left(\varphi(s) - ar_s^\varphi - b(s,Y_s)\right)ds + \gamma(s,Y_s)dW_s, \qquad s > t \\ r_t^\varphi &= r,\end{aligned}$$

and

$$\begin{aligned}dY_s &= \mu(s,Y_s)ds + \sigma(s,Y_s)dW_s, \qquad s > t \\ Y_t &= y.\end{aligned}$$

Now define the process r_s^0 by

$$\begin{aligned}dr_s^0 &= -\left(ar_s^0 + b(s,Y_s)\right)ds + \gamma(s,Y_s)dW_s, \qquad s > t \\ r_t^0 &= r + \int_t^T e^{-a(t-u)}\varphi(u)du.\end{aligned}$$

With these definitions of r_s^φ and r_s^0, calculations similar to those in the proof of Proposition 1 give that $r_T^\varphi = r_T^0$ and that

$$r_s^\varphi = r_s^0 - \psi(s,T),$$

where

$$\psi(s,T) = \int_s^T e^{-a(s-u)}\varphi(u)du. \tag{14}$$

We then have

$$\begin{aligned}
F^\varphi(t,r,y) &= E_t^Q\left[e^{-\int_t^T r_s^\varphi ds}\Phi(r_T^\varphi, y_T) \mid r_t^\varphi = r, y_t = y\right] \\
&= E^Q\left[e^{-\int_t^T (r_s^0 - \psi(s,T))ds}\Phi(r_T^\varphi, y_T) \mid r_t^\varphi = r, y_t = y\right] \\
&= e^{\int_t^T \psi(s,T))ds} E^Q\left[e^{-\int_t^T r_s^0 ds}\Phi(r_T^0, y_T) \mid r_t^0 = r + \psi(t,T), y_t = y\right] \\
&= e^{\int_t^T \psi(s,T))ds} F^0(t, r + \psi(t,T), y).
\end{aligned}$$

With $\psi(s,T)$ as in (14), we have that

$$\begin{aligned}
\int_t^T \psi(s,T)ds &= \int_t^T \int_s^T e^{-a(s-u)}\varphi(u)duds \\
&= -\int_t^T B(u,t)\varphi(u)du,
\end{aligned} \tag{15}$$

and with $\varphi(t)$ given by Proposition 1, we obtain

$$\int_t^T \psi(s,T)ds = \log\frac{P^*(0,T)P^0(0,t)}{P^*(0,t)P^0(0,T)} + B(T,t)\left(f^0(0,T) - f^*(0,T)\right).$$

Differentiating (15) with respect to t, we also have

$$\psi(t,T) = \int_t^T \frac{\partial B(u,t)}{\partial t}\varphi(u)du,$$

which with $\varphi(t)$ given by Proposition 1 gives

$$\psi(t,T) = f^0(0,t) - f^*(0,t) - e^{a(T-t)}\left(f^0(0,T) - f^*(0,T)\right). \blacksquare$$

Remark 1 *We note that if the function $\varphi(t)$ in the model $\mathcal{M}^\varphi$ is given by Proposition 1, and the payoff function is $\Phi(r_T, Y_T) = 1$, i.e., we are considering a discount bond, then the general pricing formula in (13) reduces to (5) in Proposition 1.*

4 Examples

We saw above that a model $\mathcal{M}^{\varphi}$ for a particular choice of $\varphi(t)$ can be fitted exactly to an arbitrary initially observed yield curve. If $\mathcal{M}^{\varphi}$ gives rise to an affine term structure, then performing the inversion of the yield curve (i.e., determining $\varphi(t)$ and calculating the new discount bond prices) is often straightforward, but calculations are often lengthy and the final result somewhat messy. In these circumstances, Proposition 1 proves useful and drastically simplifies the inversion procedure.

In this section we present some examples that illustrate the usefulness of Proposition 1. The first example below is the well-known Hull and White extension [8] of the Vasicek model. The second example is the Hull and White two-factor model [9], where we shall see the importance of the use of Proposition 1. The third example is an extension by Hyll [10] of the Balduzzi *et al.* three-factor model [1].

4.1 The Hull and White Extension of the Vasicek Model

To illustrate the usefulness of Proposition 1, we let the benchmark model $\mathcal{M}$ be the Vasicek model, i.e., we let

$$\begin{aligned} dr_s &= (b - ar_s)\, ds + \sigma dW_s, \qquad s > t \\ r_t &= r, \end{aligned}$$

and hence $\mathcal{M}^0$ is given by

$$\begin{aligned} dr_s^0 &= -ar_s^0 ds + \sigma dW_s, \qquad s > t \\ r_t &= r. \end{aligned}$$

Discount bond prices and forward rates in $\mathcal{M}^0$ are given by

$$P^0(t,T,r) = \exp\left(A(t,T) - B(t,T)r\right) \tag{16}$$

$$f^0(t,T,r) = -\frac{1}{2}\sigma^2 B^2(t,T) + e^{-a(T-t)}r, \tag{17}$$

where

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

Now let $\mathcal{M}^{\varphi}$ be the Hull and White extension of the Vasicek model, i.e.,

$$\begin{aligned} dr_s^{\varphi} &= (\varphi(s) - ar_s^{\varphi})\,dt + \sigma dW_s, \qquad s > t \\ r_s^{\varphi} &= r. \end{aligned}$$

Proposition 1 then gives that the term structure in the Hull and White extension of the Vasicek model fits the initial yield curve exactly if

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right),$$

and that discount bond prices are given by

$$P^{\varphi}(t,T) = \frac{P^*(0,T)}{P^*(0,t)}\frac{P^0(0,t)}{P^0(0,T)}P^0(t,T)\exp\left(\left(f^*(0,t) - f^0(0,t)\right)B(t,T)\right),$$

where P^0 and f^0 are defined in (16)–(17). After some simplifications, we arrive at the well-known expression

$$P^{\varphi}(t,T) = \frac{P^*(0,T)}{P^*(0,t)}e^{-\frac{1}{4a^3}\sigma^2\left(e^{2at}-1\right)\left(e^{-aT}-e^{-at}\right)^2 - B(t,T)(r-f^*(0,t))}.$$

4.2 The Hull and White Two-Factor Model

Now consider the following Gaussian benchmark model $\mathcal{M}$:

$$\begin{aligned} dr_s &= (b + u_s - ar_s)\,ds + \sigma dW_s, \qquad s > t \\ r_t &= r, \end{aligned}$$

where u is a stochastic disturbance term satisfying

$$\begin{aligned} du_s &= -bu_s ds + \gamma dZ_s, \qquad s > t \\ u_t &= u. \end{aligned} \tag{18}$$

The Wiener processes are assumed to satisfy $dW_t dZ_t = \rho dt$. The corresponding model $\mathcal{M}^0$ is given by

$$\begin{aligned} dr_s^0 &= \left(u_s - ar_s^0\right)ds + \sigma dW_s, \qquad s > t \\ r_t^0 &= r, \end{aligned}$$

and u is given by (18). Discount bond prices and forward rates in the model $\mathcal{M}^0$ are given by

$$
\begin{aligned}
P^0(t,T,r,u) &= \exp\left(A(t,T) - B(t,T)r - C(t,T)u\right) && (19)\\
f^0(t,T,r,u) &= A_t(t,T) - B_t(t,T)r - C_t(t,T)u, && (20)
\end{aligned}
$$

where

$$
\begin{aligned}
A(t,T) &= -\frac{1}{2a^2}\left(\sigma^2 + \frac{\gamma^2}{b^2} + \frac{2\rho\gamma\sigma}{b}\right)\left(B(t,T) - (T-t)\right) - \frac{1}{a+b}\left(\frac{\gamma^2}{2b^2} + \frac{\rho\gamma\sigma}{b}\right)C(t,T)\\
&\quad -\frac{\gamma^2}{4b}C^2(t,T) - \frac{1}{4a}\left(\sigma^2 + \frac{\gamma^2}{b}\frac{1}{(a+b)} + \frac{2\rho\gamma\sigma}{a+b}\right)B^2(t,T)\\
&\quad -\frac{1}{a+b}\left(\frac{\gamma^2}{2b} + \rho\gamma\sigma\right)B(t,T)C(t,T)\\
B(t,T) &= \frac{1 - e^{-a(T-t)}}{a}\\
C(t,T) &= \frac{a\left(1 - e^{-b(T-t)}\right) - b\left(1 - e^{-a(T-t)}\right)}{ab(a-b)}.
\end{aligned}
$$

The Hull and White two-factor model may now be seen as an extension of the model $\mathcal{M}$, and hence we let the short rate satisfy the stochastic differential equation

$$
\begin{aligned}
dr_s^\varphi &= \left(\varphi(s) + u_s - ar_s^\varphi\right)ds + \sigma dW_s, \qquad s > t\\
r_t^\varphi &= r,
\end{aligned}
$$

and u is given by (18). From Proposition 1, we have that the term structure in the Hull and White two-factor model fits the initial yield curve exactly if

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right),$$

and discount bond prices are then given by

$$P^\varphi(t,T) = \frac{P^*(0,T)}{P^*(0,t)}\frac{P^0(0,t)}{P^0(0,T)}P^0(t,T)\exp\left(\left(f^*(0,t) - f^0(0,t)\right)B(t,T)\right), \qquad (21)$$

where P^0 and f^0 are given in (19)–(20). Compared to the original expression of discount bond prices (see [9], pp. 44–45), the expression in (21) is much more manageable and easier to implement. It is also more informative, and the ease with which the inversion procedure is carried out in this example shows the strength and usefulness of Proposition 1.

4.3 The Extension of the Balduzzi et al. Three-Factor Model

As a last example, we let the benchmark model $\mathcal{M}$ be the model by Balduzzi *et al.* (BDFS) [1], i.e., for $s > t$ we have

$$\begin{aligned} dr_s &= (\theta_s - \kappa r_s - \lambda V_s)\, ds + \sqrt{V_s} dW_s^1 \\ d\theta_s &= (\beta - \alpha\theta_s)\, ds + \gamma dW_s^2 \\ dV_s &= (b - aV_s)\, ds + \sigma\sqrt{V_s} dW_s^3, \end{aligned}$$

where $r_t = r, \theta_t = \theta$ and $V_t = V$. Furthermore, $\kappa, \lambda, \alpha, \gamma, a, b$ and σ are constants, and the standard Wiener processes W^1, W^2 and W^3 satisfy $dW^1 dW^2 = dW^2 dW^3 = 0$ and $dW^1 dW^3 = \rho dt$. Now the extension $\mathcal{M}^\varphi$ by Hyll [10] reads ($s > t$)

$$\begin{aligned} dr_s^\varphi &= (\varphi(s) + \theta_s - \kappa r_s^\varphi - \lambda V_s)\, ds + \sqrt{V_s} dW_s^1 \\ d\theta_s &= -\alpha\theta_s ds + \gamma dW_s^2 \\ dV_s &= (b - aV_s)\, ds + \sigma\sqrt{V_s} dW_s^3, \end{aligned}$$

where $r_t^\varphi = r, \theta_t = \theta$ and $V_t = V$, and we define $\mathcal{M}^0$ by

$$\begin{aligned} dr_s^0 &= \left(\theta_s - \kappa r_s^0 - \lambda V_s\right) ds + \sqrt{V_s} dW_s^1, \qquad s > t \\ r_t^0 &= r, \end{aligned}$$

and θ and V are as in $\mathcal{M}^\varphi$. Discount bond prices and forward rates in the model $\mathcal{M}^0$ are given by

$$\begin{aligned} P^0(t, T, r, \theta, V) &= \exp\left(A(t,T) - B(t,T)r - C(t,T)\theta - D(t,T)V\right) && (22) \\ f^0(t, T, r, \theta, V) &= A_t(t,T) - B_t(t,T)r - C_t(t,T)\theta - D_t(t,T)V, && (23) \end{aligned}$$

where

$$\begin{aligned} B(t,T) &= \frac{1 - \exp\left(-\kappa(T-t)\right)}{\kappa} \\ C(t,T) &= \frac{\alpha(1 - \exp\left(-\kappa(T-t)\right)) - \kappa(1 - \exp\left(-\alpha(T-t)\right))}{\alpha\kappa\left(\alpha - \kappa\right)}. \end{aligned}$$

The complete expressions for $A(t,T)$ and $D(t,T)$ are given in terms of a series expansion in Hyll [10]. Proposition 1 now gives that the term structure in the extended BDFS model fits the initial yield curve exactly if

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right),$$

and discount bond prices are given by

$$P^{\varphi}(t,T) = \frac{P^*(0,T)}{P^*(0,t)} \frac{P^0(0,t)}{P^0(0,T)} P^0(t,T) \exp\left(\left(f^*(0,t) - f^0(0,t)\right) B(t,T)\right),$$

where P^0 and f^0 are given in (22)–(23).

5 A Different Setting

The above procedure of inverting the yield curve can also be applied to different but similar settings. Indeed, loosely speaking, all that is required is that the drift of one of the factors, say X, is affine in X, and that the volatilities of X do not depend on X. Furthermore, we require that the dynamics of Y do not depended on X.

One frequent setting is to model the short rate as a sum of stochastic variables. Suppose, therefore, that in the benchmark model $\mathcal{M}$ the short rate is given by $r = X + \sum_{i=1}^{n} Y^i$, where X and Y under a martingale measure Q satisfy the stochastic differential equations

$$\begin{aligned} dX_s &= (b - aX_s - c(s,Y_s))\, ds + \gamma(s,Y_s) dW_s, \qquad s > t \\ X_t &= x, \end{aligned}$$

and

$$\begin{aligned} dY_s &= \mu(s,Y_s)ds + \sigma(s,Y_s)dW_s, \qquad s > t \\ Y_t &= y. \end{aligned} \tag{24}$$

The model $\mathcal{M}^0$ is obtained be letting $b = 0$ in $\mathcal{M}$, i.e., $r^0 = X^0 + \sum_{i=1}^{n} Y^i$, where X satisfies the stochastic differential equation

$$\begin{aligned} dX_s^0 &= -\left(aX_s^0 + c(s,Y_s)\right) ds + \gamma(s,Y_s) dW_s, \qquad s > t \\ X_t^0 &= x, \end{aligned}$$

and Y is given by (24). The extension $\mathcal{M}^{\varphi}$ is given by $r^{\varphi} = X^{\varphi} + \sum_{i=1}^{n} Y^i$, where X^{φ} satisfies the stochastic differential equation

$$\begin{aligned} dX_s^{\varphi} &= (\varphi(s) - aX_s^{\varphi} - c(s,Y_s))\, ds + \gamma(s,Y_s) dW_s, \qquad s > t \\ X_t^{\varphi} &= x, \end{aligned}$$

and Y is given by (24). Now let $P^{\varphi}(t,T), f^{\varphi}(t,T), P^0(t,T)$ and $f^0(t,T)$ denote discount bond prices and forward rates in $\mathcal{M}^{\varphi}$ and $\mathcal{M}^0$ respectively, and let $P^*(0,t)$ and $f^*(0,t)$ be the initially observed yield curve and forward rate curve. We then have the following proposition:

Proposition 3 *The theoretical term structure implied by the model $\mathcal{M}^{\varphi}$ can be fitted exactly to an arbitrary initially observed yield curve. A perfect initial fit is obtained by choosing*

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right).$$

Discount bond prices are then given by

$$P^{\varphi}(t,T) = \frac{P^*(0,T)}{P^*(0,t)} \frac{P^0(0,t)}{P^0(0,T)} P^0(t,T) \exp\left(\left(f^*(0,t) - f^0(0,t)\right) B(t,T)\right),$$

where

$$B(t,T) = \frac{1 - \exp\left(-a(T-t)\right)}{a}.$$

Proof. Exactly the same as for Proposition 1. ■

As an example, we consider the model $\mathcal{M}$ where the short rate is the sum of two stochastic variables, i.e., we let $r = X + Y$, where X and Y under a martingale measure Q satisfy the stochastic differential equations

$$\begin{aligned} dX_s &= \left(b - aX_s^{\varphi} - cY_s\right) ds + \sigma dW_s^1, \qquad s > t \\ X_t &= x \end{aligned}$$

and

$$\begin{aligned} dY_s &= \left(\beta - \alpha Y_s\right) ds + \gamma dW_s^2, \qquad s > t \\ Y_t &= y, \end{aligned} \tag{25}$$

where $a, b, c, \sigma, \alpha, \beta$ and γ are constants and the standard Wiener processes W^1 and W^2 satisfy $dW^1 dW^2 = \rho dt$. The extension $\mathcal{M}^{\varphi}$ is $r^{\varphi} = X^{\varphi} + Y$, where X^{φ} satisfies

$$\begin{aligned} dX_s^{\varphi} &= \left(\varphi(s) - aX_s^{\varphi} - bY_s\right) ds + \sigma dW_s^1, \qquad s > t \\ X_t^{\varphi} &= x, \end{aligned}$$

and Y is given by (25). Furthermore, the model $\mathcal{M}^0$ is obtained by setting $b=0$ in $\mathcal{M}$, i.e., $r^0 = X^0 + Y$, where X^0 satisfies

$$\begin{aligned} dX_s^0 &= -\left(aX_s^0 + bY_s\right) ds + \sigma dW_s^1, \qquad s > t \\ X_t^0 &= x, \end{aligned}$$

and Y is given by (25). Discount bond prices and forward rates in $\mathcal{M}^0$ are given by

$$\begin{aligned} P^0(t,T,x,y) &= \exp(A(t,T) - B(t,T)x - C(t,T)y) &(26)\\ f^0(t,T,x,y) &= \beta C(t,T) - \frac{1}{2}\sigma^2 B^2(t,T) - \frac{1}{2}\gamma^2 C^2(t,T) - \gamma\sigma\rho B(t,T)C(t,T) \\ &\quad + e^{-a(T-t)}x + (1 - \alpha C(t,T) - bB(t,T))\,y, &(27) \end{aligned}$$

where

$$\begin{aligned} A(t,T) &= \left(\frac{\sigma^2}{2a^2} + \frac{\gamma^2(a-b)^2}{2\alpha^2 a^2} + \frac{\gamma\sigma\rho(a-b)}{\alpha a^2} - \frac{\beta(a-b)}{\alpha a}\right)(T-t) \\ &\quad + \left(\frac{\gamma^2 b\left((a-b)(a+\alpha) + \alpha a\right)}{2a\alpha^2 a(a+\alpha)} + \frac{\gamma\sigma\rho\left(b(a+\alpha) - a\alpha\right)}{a^2\alpha(a+\alpha)} - \frac{b\beta}{a\alpha} - \frac{\sigma^2}{2a^2}\right) B(t,T) \\ &\quad + \left(\frac{\gamma^2\left(b - a - \alpha\right)}{2\alpha^2(a+\alpha)} + \frac{\beta}{\alpha} - \frac{\gamma\sigma\rho}{\alpha(a+\alpha)}\right) C(t,T) \\ &\quad + \left(\frac{b\gamma\sigma\rho}{2a(a+\alpha)} - \frac{\sigma^2}{4a} - \frac{\gamma^2 b^2}{4a\alpha(a+\alpha)}\right) B^2(t,T) \\ &\quad + \frac{\gamma}{a+\alpha}\left(\frac{\gamma b}{2\alpha} - \sigma\rho\right) B(t,T)C(t,T) - \frac{\gamma^2}{4\alpha} C^2(t,T) \\ B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \\ C(t,T) &= \frac{b}{a-\alpha} B(t,T) + \left(1 - \frac{b}{a-\alpha}\right) \frac{1 - e^{-\alpha(T-t)}}{\alpha}. \end{aligned}$$

Proposition 3 then gives that discount bond prices in the model $\mathcal{M}^\varphi$ fit the initial yield curve if

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + a\left(f^*(0,t) - f^0(0,t)\right),$$

and that discount bond prices are given by

$$P^\varphi(t,T) = \frac{P^*(0,T)}{P^*(0,t)} \frac{P^0(0,t)}{P^0(0,T)} P^0(t,T) \exp\left(\left(f^*(0,t) - f^0(0,t)\right) B(t,T)\right),$$

where P^0 and f^0 are given in (26)–(27).

Thus we see that the inversion procedure can be used in several types of settings, and it is straightforward to obtain a general pricing formula similar to that of Proposition 2 in section 3.

6 Summary and Conclusions

In this essay we identify a class of short-rate factor models of the term structure of interest rates and show that the theoretical term structure generated by these models can be fitted exactly to an arbitrary initially observed yield curve. We have also seen that the fitted term structure can be obtained in terms of discount bond prices in a special case of the benchmark model, which greatly simplifies the inversion procedure. Furthermore, prices in the extended model of simple contingent claims (i.e., not path dependent), such as European options, can be obtained by considering the same contingent claim in a special case of the benchmark model.

Although the class of invertible short-rate factor models of the term structure considered here is rather large, several models have not been it included, among them the classic model by Cox, Ingersoll and Ross [3]. With the inversion technique used in this essay, i.e., by introducing a time dependent function in the drift of the short rate, analytical expressions remain to be found for the fitted term structure in these models.

References

[1] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram (1996): "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income,* December, 43–53.

[2] Brigo, D. and F. Mercurio (2000): "A Deterministic-Shift Extension of Analytically-Tractable and Time-Homogenous Short-Rate Models," Forthcoming in Finance and Stochastics, Banca IMI, Milano.

[3] Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985b): "A Theory of the Term Structure of Interest Rates," *Econometrica* **53**(2), 385–407.

[4] Duffie, D. and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance* **6**(4), 379–406.

[5] Fong, H.G. and O.A. Vasicek (1991): "Interest Rate Volatility as a Stochastic Factor," Unpublished manuscript, Gifford Fong Associates.

[6] Heath, D., R. Jarrow, and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica* **60**(1), 77–105.

[7] Ho, T.S. and S.B. Lee (1986): "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* **41**(5), 1011–1029.

[8] Hull, J. and A. White (1990): "Pricing Interest Rate Derivative Securities," *The Review of Financial Studies* **3**(4), 573–592.

[9] Hull, J. and A. White (1994): "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models," *The Journal of Derivatives,* Winter, 37–48.

[10] Hyll, M. (2000d): "An Efficient Series Expansion Approach to the Balduzzi, Das, Foresi and Sundaram Model of the Term Structure of Interest Rates," Working paper, Stockholm School of Economics.

[11] Longstaff, F.A. and E.S. Schwartz (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance* **47**(4), 1259–1282.

[12] Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* **5**, 177–188.

3

An Efficient Series Expansion Approach to a Two-Factor Model of the Term Structure of Interest Rates

ABSTRACT
This essay considers a two-factor model of the term structure of interest rates where the short rate and the (scaled) mean reversion level of the short rate are the factors. The short rate follows a square root process, where the stochastic (scaled) mean reversion level of the short rate also follows a square root process, and hence the model can be seen as an extension of the Cox, Ingersoll and Ross model. Using a method of series expansion, a solution to discount bond prices is provided along with several examples of yield curves and forward rate curves. The model is also compared with a corresponding model in a Gaussian framework, and no substantial differences are found between the two models regarding the flexibility and shapes of the yield curves and forward rate curves they generate.

1 Introduction

A typical model in continuous time of the term structure of interest rates consists of a number of explanatory factors describing the evolution of the yield curve and a system of stochastic differential equations satisfied by these factors under a martingale measure Q. The early models by Vasicek [14] and Cox, Ingersoll and Ross [6] identify the short rate as the only explanatory factor, and each model specifies a distinct stochastic differential equation satisfied by the short rate. For obvious reasons, these models are referred to as "pure short-rate models". Prices of discount bonds, and hence yields of different maturities, can then be obtained by solving a related partial differential equation, sometimes referred to as the term structure equation. The pure short-rate models seem to explain the movements of the short end of the yield curve rather well, but the rather simple shapes of the generated term structures have, over the years, led various authors to suggest several multi-factor models as possibly better alternatives. Models identifying the reversion level and volatility of the short rate as the second factor are the most common two-factor models. These models are often natural extensions of already existing one-factor models, and these extensions are easily motivated from an economic point of view. Other types of two-factor models are those where the long rate and the spread between the long rate and the short rate are the explanatory factors.

The ease with which one solves the term structure equation in the multi-factor case depends on the choice of explanatory factors and the nature of the system of stochastic differential equations satisfied by these factors. A closed form solution is of course desirable, and therefore one is usually confined to what is known as affine term structure models. A term structure is called affine if the solution to the term structure equation is (exponential-) affine in the factors. Duffie and Kan [7] showed that the term structure is in fact affine if and only if the stochastic differential equations (SDEs) governing the factors satisfy certain conditions. Unfortunately, even if the term structure is affine, it might not be attainable in closed form. Thus, in principle, when proposing a multi-factor model of the term structure of interest rates, the factors and the system of SDEs satisfied by the factors need to be specified not only so that the term structure is affine, but also in such a way that the solution to the term structure equation can be obtained in closed form. As a consequence of this tractability perspective on term structure models, authors seem to favor models in which at least some factors satisfy (a system of) Ornstein-Uhlenbeck processes, i.e., SDEs with Gaussian (normally distributed) noise. Numerical solutions, discrete tree models, Monte-Carlo simulations or approximation techniques are required if a closed form solution to the term structure equation is unattainable. The computational difficulties arising in an affine term structure model consist in determining the solution to a system of ordinary differential equations (ODEs), where time (or time to maturity, depending on the notation chosen) is the variable. Loosely speaking, depending on how we specify the SDEs for the factors, the system of ODEs to be solved may be linear in the simplest cases, but non-linear in more severe cases. There are no standard methods available for solving non-linear ODE systems exactly, and therefore some kind of numerical method is required.

This essay presents a model of the term structure of interest rates where, although the term structure is affine, a closed form solution to the term structure equation is not available. Instead, the technique that is used is the one first employed by Selby and Strickland (SS) [13]. It involves expressing a solution to an ordinary differential equation in the form of a series expansion; with this technique SS greatly simplified the implementation of the Fong and Vasicek model (FV) [8]. Hyll [11] uses the same technique to obtain a solution to discount bond prices in the Balduzzi *et al.* model [1].

In the two-factor model presented below, the short rate and the scaled mean reversion level of the short rate are chosen as the explanatory factors. For the sake of simplicity, the scaled mean reversion level will henceforth be referred to as the mean reversion level. Target rates are used as a monetary policy tool by most central banks, and it is therefore appealing from an economic point of view with a stochastic mean reversion level of the short rate. There is nothing new in using the short rate and its mean reversion level as explanatory factors. This has been done before, but within a Gaussian framework.[1] In what follows, the short rate is assumed to follow a CIR process, or a square root process, where the reversion level of the short rate is stochastic and satisfies a square root process. Thus, negative values for both the factors are precluded, which adds credibility to the model. The cost is the computational difficulties arising when solving the system of ODEs needed for discount bond prices. However, these difficulties are overcome by using series expansions.

The suggested model can be seen as an extension of the Cox, Ingersoll and Ross model [6], but also as a special case of the more general model designed by Chen [4], which in its full version does not fall into the usual category of affine term structure models. Prices of discount bonds in the Chen model are expressed in a rather complicated manner. Even in the special case when the Chen model reduces to the model presented in this essay, expressions for discount bonds are not particularly user-friendly. The alternative solution offered in this essay is an improvement when it comes to implementing the model.

The rest of the essay is organized as follows: Section 2 presents the two-factor model, and the term structure is derived and solved for by using a series expansion in section 3. Examples of yield curves and forward rate curves are given in sections 4 and 5, and in section 6 the flexibility of the yield and forward rate curves are compared to those generated by a corresponding Gaussian model. Section 7 summarizes and concludes.

[1] See e.g. Balduzzi *et al.* [2].

2 Model

It is assumd that the short rate r_t and the mean reversion level θ_t of the short rate under a martingale measure Q satisfy the following stochastic differential equations:

$$dr_t = (\theta_t - \kappa r_t)dt + \sigma\sqrt{r_t}dW_t^1 \tag{1}$$
$$d\theta_t = (\beta - \alpha\theta_t)dt + \eta\sqrt{\theta_t}dW_t^2, \tag{2}$$

where $\kappa, \sigma, \alpha, \beta$ and η are positive constants and W_t^1 and W_t^2 are uncorrelated standard Wiener processes. In (1)–(2), both factors follow what is usually referred to as square root processes, with θ_t being the mean reversion level of the short rate. The above model can therefore be thought of as an extension of the Cox, Ingersoll and Ross model. Note that the independence of the Wiener processes is desirable from a tractability point of view if an affine term structure is to be expected, since otherwise, the necessary and sufficient conditions for the existence of an affine term structure would be violated.[2] Although the Wiener processes are independent, the processes r_t and θ_t are still correlated due to their coupled dynamics. From an economic point of view, it would naturally be desirable if the noises driving r_t and θ_t where correlated, but then we would most certainly have to rely on numerical methods to solve the corresponding term structure equation.

3 The Term Structure and Properties of Discount Bond Prices

This section considers the term structure of interest rates for the model in (1)–(2), and we denote by $P(t,T)$ the price at time $t \geq 0$ of a default-free discount bond with unit payment at $T \geq t$. Using standard financial theory, we have that discount bond prices are given by $P(t,T) = F(t, r_t, \theta_t; T)$, where the pricing function $F(t, r, \theta; T)$ satisfies the so-called term structure equation

$$F_t + (\theta - \kappa r)F_r + \frac{1}{2}\sigma^2 r F_{rr} + (\beta - \alpha\theta)F_\theta + \frac{1}{2}\eta^2\theta F_{\theta\theta} = rF \tag{3}$$
$$F(t, r, \theta; T) = 1. \tag{4}$$

In an affine term structure model, the pricing function is given by

$$F(t, r, \theta; T) = \exp\left(A(t,T) - B(t,T)r - C(t,T)\theta\right), \tag{5}$$

[2] See Duffie and Kan [7] for an introduction to affine term structure models.

where $A(t,T), B(t,T)$ and $C(t,T)$ are time dependent functions satisfying the terminal conditions

$$A(T,T) = B(T,T) = C(T,T) = 0. \tag{6}$$

Substituting (5) into (3) and rearranging, we obtain

$$A_t - \beta C + \left(-B_t + \kappa B + \frac{1}{2}\sigma^2 B^2 - 1\right) r + \left(-C_t - B + \alpha C + \frac{1}{2}\eta^2 C^2\right)\theta = 0. \tag{7}$$

Since (7) must hold for all values of r and θ, we have in fact a system of ordinary differential equations:

$$\begin{aligned} A_t(t,T) - \beta C(t,T) &= 0 & (8)\\ B_t(t,T) - \kappa B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T) + 1 &= 0 & (9)\\ C_t(t,T) - \alpha C(t,T) - \frac{1}{2}\eta^2 C^2(t,T) + B(t,T) &= 0, & (10) \end{aligned}$$

with terminal conditions as in (6). We then have the following proposition:

Proposition 1 *Suppose that the short rate r_t and the mean reversion level θ_t of the short rate under a martingale measure Q satisfy (1) and (2). Then discount bond prices are given by*

$$P(t,T) = \exp\left(A(t,T) - B(t,T)r_t - C(t,T)\theta_t\right), \tag{11}$$

where

$$\begin{aligned} A(t,T) &= \frac{2\beta}{\eta^2}\log\frac{e^{\varepsilon\gamma(T-t)}Q(1)}{Q(e^{-\gamma(T-t)})} & (12)\\ B(t,T) &= 2\frac{e^{\gamma(T-t)} - 1}{(\gamma+\kappa)\left(e^{\gamma(T-t)} - 1\right) + 2\gamma} & (13)\\ C(t,T) &= -\frac{2\gamma}{\eta^2}\left(\varepsilon + \frac{e^{-\gamma(T-t)}Q'(e^{-\gamma(T-t)})}{Q(e^{-\gamma(T-t)})}\right), & (14) \end{aligned}$$

with

$$\begin{aligned} \gamma &= \sqrt{\kappa^2 + 2\sigma^2}\\ \varepsilon &= \frac{\alpha}{2\gamma} - \frac{1}{2\gamma}\sqrt{\alpha^2 + \frac{4\eta^2}{\gamma+\kappa}}. \end{aligned}$$

The function $Q(x)$ is defined by

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2),$$

where

$$\frac{K_2}{K_1} = -\frac{Q'(1, q_1) + \varepsilon Q(1, q_1)}{Q'(1, q_2) + \varepsilon Q(1, q_2)}$$

and

$$U(x, q_i) = \sum_{k=0}^{\infty} c_k^{(i)} x^{k+q_i},$$

with $q_1 = 0$ and $q_2 = 1 - \tilde{\alpha}_0$. The coefficients $c_k^{(i)}$ ($i = 1, 2$ and $k = 1, 2, ...$) of $U(x, q_i)$ are given by

$$c_k^{(i)} = -\frac{\sum_{j=0}^{k-1} \tilde{\beta}_{k-j} c_j^{(i)}}{p(q_i + k)}, \tag{15}$$

where

$$\begin{aligned}
c_0^{(1)} &= 1 \\
c_0^{(2)} &= 1 \\
\tilde{\beta}_k &= -\frac{\eta^2}{\gamma \sigma^2}(-\delta)^k, k = 1, 2, ... \\
\tilde{\alpha}_0 &= 1 + 2\varepsilon - \frac{\alpha}{\gamma} \\
\delta &= \frac{\gamma - \kappa}{\gamma + \kappa} \\
p(q) &= q\,(q - 1 + \tilde{\alpha}_0)\,.
\end{aligned}$$

Proof. See Appendix A and B. ∎

From (13) we see that $B(t, T)$ is a strictly positive function for $0 \le t < T$, and differentiating $B(t, T)$ with respect to T, we find that $B_T(t, T)$ is positive for $0 \le t \le T$. Thus $B(t, T)$ is a strictly increasing function, and hence from (11) it follows that discount bond prices are decreasing and convex in r. To see that bond prices are also decreasing and convex in θ, we rewrite (10) as

$$C_T(t, T) + \alpha C(t, T) + \frac{1}{2}\eta^2 C^2(t, T) = B(t, T), \tag{16}$$

where $C_t(t,T) = -C_T(t,T)$ has been used since $C(t,T)$ is a time homogenous function. Remembering the boundary conditions for $B(t,T)$ and $C(t,T)$, it follows that $C_T(t,t) = 0$. Differentiating (16) with respect to T, we obtain

$$C_{TT}(t,T) + \alpha C_T(t,T) + \eta^2 C(t,T)C_T(t,T) = B_T(t,T), \tag{17}$$

and hence it follows that $C_{TT}(t,t) = B_T(t,t) = 1$, where (9) has been used and that $B_t(t,T) = -B_T(t,T)$ since $B(t,T)$ is a time homogenous function. Thus $C(t,t) = C_T(t,t) = 0$ and $C_{TT}(t,t) = 1$. Now since $C_{TT}(t,t) = 1, C_T(t,T)$ must initially be increasing, and therefore positive, and hence $C(t,T)$ is initially increasing and positive. To see that $C(t,T)$ is non-decreasing for all T, assume there exists $T^* > t$ such that $C(t,T)$ is increasing for $t \le T < T^*$, and that $C_T(t,T^*) = 0$. It then follows from (17) that $C_{TT}(t,T^*) = B_T(t,T^*) > 0$, and hence $C(t,T)$ increases again for $T > T^*$. This argument shows that $C(t,T) \ge 0$ and $C_T(t,T) \ge 0$ for $0 \le t \le T$.

Differentiating (11) with respect to T, and using (8) and the fact that $A(t,T)$ is a time homogenous function gives

$$P_T(t,T) = -\left(\beta C(t,T) + B_T(t,T)r + C_T(t,T)\theta\right) P(t,T),$$

and since $\beta > 0$, it follows that bond prices are decreasing in T.

Since bond prices are decreasing in both r and θ, $P(t,T)$ has the desirable property that it tends to zero as $r \to \infty$ or $\theta \to \infty$ for $0 \le t < T$. Furthermore, it can be shown that $P(t,T)$ tends to zero as T tends to infinity.

4 Term Structure Shapes

The yields are defined as $Y(t,T) = -\frac{\log P(t,T)}{T-t}$, and the yields in the above model thus take the form

$$Y(t,T) = -\frac{A(t,T) - B(t,T)r - C(t,T)\theta}{T-t}. \tag{18}$$

We note that

$$\frac{\partial Y(t,T)}{\partial r} = \bar{B}(t,T) > 0, \ t \le T \tag{19}$$

$$\frac{\partial Y(t,T)}{\partial \theta} = \bar{C}(t,T) \ge 0, \ t \le T, \tag{20}$$

where

$$\begin{aligned}\bar{B}(t,T) &= \frac{B(t,T)}{T-t}\\ \bar{C}(t,T) &= \frac{C(t,T)}{T-t}.\end{aligned}$$

Thus the yield of any maturity increases if either r or θ increases. Taking the proper limits of (18), it is straightforward to verify that

$$\begin{aligned}Y^{\infty} &= \lim_{T\to\infty} Y(t,T)\\ &= \beta\left(-\frac{\alpha}{\eta^2}+\sqrt{\frac{\alpha^2}{\eta^4}+\frac{4}{(\gamma+\kappa)\eta^2}}\right),\end{aligned}$$

i.e., the long-run yield is independent of the factors and positive. Straightforward differentiation of Y^{∞} with respect to the different parameters gives the following:

$$\begin{aligned}\frac{\partial Y^{\infty}}{\partial \kappa} &= -\frac{1}{\gamma(\gamma+\kappa)}\frac{2\beta^2}{\alpha\beta+\eta^2 Y^{\infty}} < 0\\ \frac{\partial Y^{\infty}}{\partial \sigma} &= -\frac{1}{\gamma(\gamma+\kappa)^2}\frac{4\beta^2\sigma}{\alpha\beta+\eta^2 Y^{\infty}} < 0\\ \frac{\partial Y^{\infty}}{\partial \alpha} &= -\frac{\beta Y^{\infty}}{\eta^2 Y^{\infty}+\alpha\beta} < 0\\ \frac{\partial Y^{\infty}}{\partial \beta} &= \frac{Y^{\infty}}{\beta} > 0\\ \frac{\partial Y^{\infty}}{\partial \eta} &= -\frac{2\alpha\beta}{\eta\left(\eta^2 Y^{\infty}+\alpha\beta\right)}\left(\frac{\beta}{\alpha}\frac{2}{\gamma+\kappa}-Y^{\infty}\right) < 0,\end{aligned}$$

and thus the long-run yield will decrease (increase) if κ, σ, α or η increases (β increases).

In this and the following section, several figures showing yield curves and forward rate curves are presented. The parameters used for these figures are $\kappa = 0.25, \sigma = 0.15, \alpha = 0.76, \beta = 0.023$ and $\eta = 0.035$ giving a long-run yield of 10.35%. Some of these parameter values (κ, α, and β) have been taken from BDFS [1], while others (σ and η) have been chosen, since the aim here is to generate examples of what yield curves and forward rate curves might look like in this model.[3]

[3] The parameters α and κ represent the reversion speed of the two factors, and their interpretation and values are the same as in the BDFS model [1]. The parameter β is defined somewhat differently here compared to the BDFS model, but its value has been obtained from the corresponding parameters in the BDFS model. The parameters σ and η do not have any natural corresponding interpretation in the BDFS model, but their values have been chosen here so that the (instantaneous) volatilities of the factors are about the same for the two models for typical values of the factors.

When generating examples of different yield curves and forward rate curves, we have to truncate the infinite summation in the solutions for $A(t,T)$ and $C(t,T)$ at some point. As it turns out, the summation converges very quickly, i.e., only a few terms need to be included in the summation to ensure convergence. There are two main reasons for this. First, the magnitude of the coefficients decays rapidly due to the recursive relation in (15). Second, the series expansion is evaluated at $x = e^{-\gamma(T-t)}$, which is strictly less than one for $0 \le t < T$. Thus, higher order terms become less and less significant. Performed tests for different parameter values suggest that the summation may in fact be truncated after as few as five terms, but due to the complex form of the yields, a general error estimation has not been obtained.

We now turn to examining the term structure. In Figure 1 and Figure 2, the short rate has been varied for two different values of θ. In Figure 3, the value of θ has been varied, and this effect is shown for three different values of r.

In Figures 1–3 we see that the model can generate yield curves of many interesting shapes. Apart from the "usual" increasing, decreasing and slightly humped yield curves generated by the pure short-rate models, we also have examples of initially decreasing and thereafter concavely increasing yield curves, and initially increasing and thereafter convexly decreasing yield curves. With the parameter values as above, the following main conclusions can be drawn from Figures 1–3:

- a change in r has a large effect on short-term yields, and then the effect decays rapidly although a change in r has a small effect on long-term yields up to at least 30 years
- a change in θ has practically no effect on very short-term yields, then the effect grows slowly and reaches a maximum for intermediate yields (around 5 years), after which the effect decays slowly although a change in θ has a small effect on long-term yields up to at least 30 years
- a small (large) θ initially causes the yield curve to be decreasing (increasing).

Remark 1 *The initial slope of the yield curve can be shown to be equal to half the drift of the short rate, i.e., $Y_T(t,t) = \frac{\theta - ar}{2}$. A sufficiently large (small) θ will then cause the yield curve to be initially increasing (decreasing). Conversely, if r is sufficiently large (small), the yield curve will initially be decreasing (increasing).*

From (19)–(20) it follows that the functions $\bar{B}(t,T)$ and $\bar{C}(t,T)$ determine the effect that a change in r and θ respectively will have on the yields for different maturities. This is shown in Figure 4.

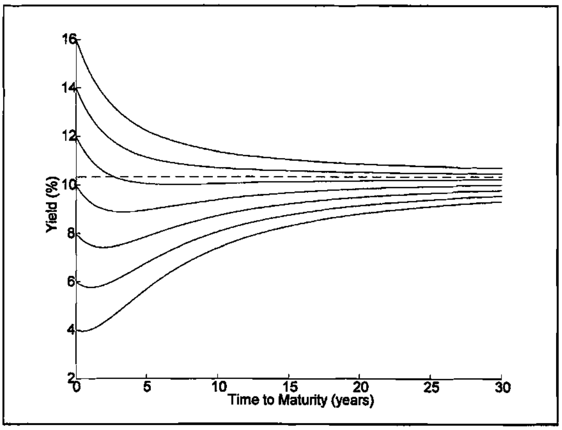

FIGURE 1. Yield Curves for Decreasing r (top-down)
Yield curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all yield curves, $\theta = 0.5\%$. The dashed line represents the long-run yield of 10.35%.

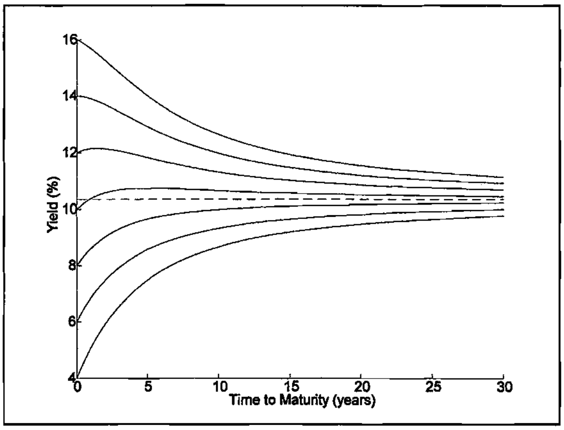

FIGURE 2. Yield Curves for Decreasing r (top-down)
Yield curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all yield curves, $\theta = 3.5\%$. The dashed line represents the long-run yield of 10.35%.

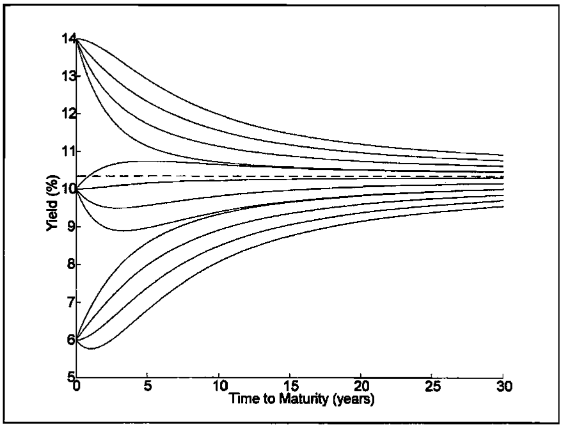

FIGURE 3. Yield Curves for Decreasing θ (top-down)
Yield curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3.5\%$ (top), $\theta = 2.5\%$, $\theta = 1.5\%$ and $\theta = 0.5\%$ (down). The dashed line represents the long-run yield of 10.35%.

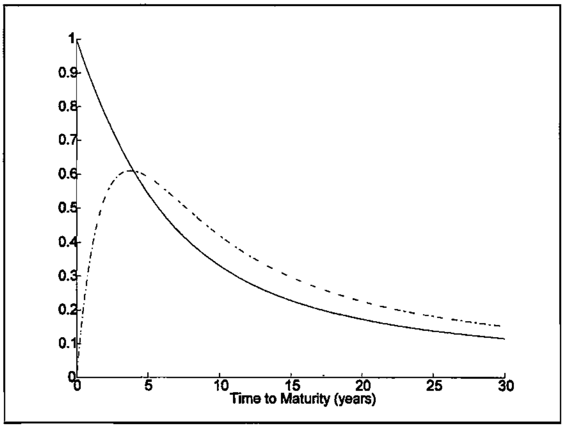

FIGURE 4. Changes in the Yield Curve with Respect to the Factors
The functions $\bar{B}(t,T)$ (solid) and $\bar{C}(t,T)$ (dash-dotted) for different maturities.

5 Forward Rates

The forward rates are defined as $f(t,T) = -\frac{\partial}{\partial T}\log P(t,T)$, and using the discount bond prices in (5), and that we have a time homogenous model, we obtain

$$f(t,T) = A_t(t,T) - B_t(t,T)r - C_t(t,T)\theta.$$

It is straightforward to verify that the long-run forward rate f^∞ is the same as the long-run yield Y^∞, i.e.,

$$\begin{aligned} f^\infty &= \lim_{T\to\infty} f(t,T) \\ &= \beta\left(-\frac{\alpha}{\eta^2} + \sqrt{\frac{\alpha^2}{\eta^4} + \frac{4}{(\gamma+\kappa)\eta^2}}\right). \end{aligned}$$

Furthermore, we have from above that

$$\begin{aligned} f_r(t,T) &= -B_t(t,T) < 0 \\ f_\theta(t,T) &= -C_t(t,T) \leq 0. \end{aligned}$$

Using (8)–(10), we obtain

$$\begin{aligned} f(t,T) &= \beta C(t,T) + \left(1 - \kappa B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T)\right) r \\ &\quad + \left(B(t,T) - \alpha C(t,T) - \frac{1}{2}\eta^2 C^2(t,T)\right)\theta. \end{aligned}$$

Using the same parameter values as in the previous section, different forward rate curves are plotted in Figures 5–7. In Figures 5–6 the short rate r is varied for two different values of θ, and in Figure 7 the value of θ is varied for three different values of r. We see examples of forward rate curves of many different and interesting shapes, similar to those of the yield curves in Figures 1–3.

Remark 2 *The initial slope of the forward rate curve can be shown to be equal to the drift of the short rate, i.e., $f_T(t,t) = \theta - ar$. A sufficiently large (small) θ will then cause the forward rate curve to be initially increasing (decreasing). Conversely, if r is sufficiently large (small), the forward rate curve will initially be decreasing (increasing).*

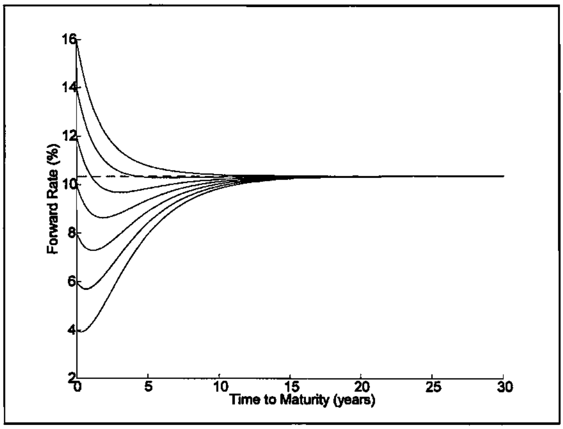

FIGURE 5. Forward Rate Curves for Decreasing r (top-down)
Forward rate curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all forward rate curves, $\theta = 0.5\%$. The dashed line represents the long-run forward rate of 10.35%.

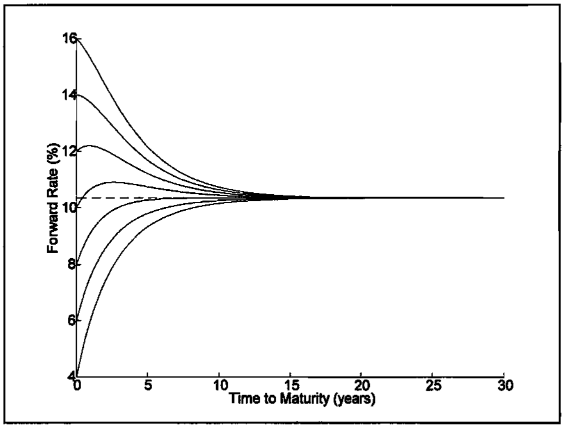

FIGURE 6. Forward Rate Curves for Decreasing r (top-down)
Forward rate curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all forward rate curves, $\theta = 3.5\%$. The dashed line represents the long-run forward rate of 10.35%.

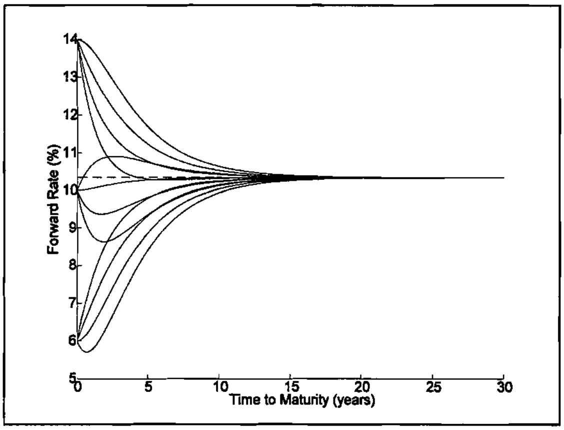

FIGURE 7. Forward Rate Curves for Decreasing θ (top-down)
Forward rate curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3.5\%$ (top), $\theta = 2.5\%$, $\theta = 1.5\%$ and $\theta = 0.5\%$ (down). The dashed line represents the long-run forward rate of 10.35%.

6 A Corresponding Two-Factor Gaussian Model

In this section, a corresponding Gaussian two-factor model is studied and compared with the model above.[4] In this essay the Gaussian model is presented as

$$\begin{aligned} dr_t &= (\theta_t - \kappa r_t)dt + \sigma dW_t^{(1)} \\ d\theta_t &= (\beta - \alpha\theta_t)dt + \eta dW_t^{(2)}, \end{aligned}$$

where $\kappa, \sigma, \alpha, \beta$ and η and are positive constants and $dW_t^{(1)} dW_t^{(2)} = \rho dt$. Thus, both factors still display mean-reverting features, the essential differences from the model above are that both factors are now driven by Gaussian noise, and that the Wiener processes are correlated. Discount bond prices are given by

$$P(t,T) = \exp\left(A(t,T) - B(t,T)r_t - C(t,T)\theta_t\right),$$

[4]See e.g. Balduzzi *et al.* [2] for a Gaussian model of this type.

where

$$
\begin{aligned}
A(t,T) &= \left(\frac{\beta}{\alpha\kappa} - \frac{1}{2\kappa^2}\left(\sigma^2 + \frac{\eta^2}{\alpha^2} + \frac{2\rho\eta\sigma}{\alpha}\right)\right)(B(t,T) - (T-t)) \\
&\quad + \left(\frac{\beta}{\alpha} - \frac{1}{\alpha+\kappa}\left(\frac{\eta^2}{2\alpha^2} + \frac{\rho\eta\sigma}{\alpha}\right)\right)C(t,T) - \frac{\eta^2}{4\alpha}C^2(t,T) \\
&\quad - \frac{1}{4\kappa}\left(\sigma^2 + \frac{\eta}{\alpha+\kappa}\left(\frac{\eta}{b} + 2\rho\sigma\right)\right)B^2(t,T) - \frac{1}{\alpha+\kappa}\left(\frac{\eta^2}{2\alpha} + \rho\eta\sigma\right)B(t,T)C(t,T) \\
B(t,T) &= \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\
C(t,T) &= \frac{\kappa(1 - e^{-\alpha(T-t)}) - \alpha\left(1 - e^{-\kappa(T-t)}\right)}{\alpha\kappa(\kappa - \alpha)}.
\end{aligned}
$$

The long-run yield and long-run forward rate can be shown to be given by

$$
Y^\infty = f^\infty = \frac{\beta}{\alpha\kappa} - \frac{1}{2\kappa^2}\left(\sigma^2 + \frac{\eta^2}{\alpha^2} + \frac{2\rho\sigma\eta}{\alpha}\right),
$$

which is not necessarily positive. The various partial derivatives are easily computed as

$$
\begin{aligned}
\frac{\partial Y^\infty}{\partial \kappa} &= -\frac{\beta}{\kappa^2\alpha} + \frac{\sigma^2\alpha^2 + \eta^2 + 2\rho\sigma\eta\alpha}{\kappa^3\alpha^2} \\
\frac{\partial Y^\infty}{\partial \sigma} &= -\frac{\sigma\alpha + \rho\eta}{\alpha\kappa^2} \\
\frac{\partial Y^\infty}{\partial \alpha} &= -\frac{\alpha\beta\kappa - \eta^2 - \alpha\rho\sigma\eta}{\kappa^2\alpha^3} \\
\frac{\partial Y^\infty}{\partial \beta} &= \frac{1}{\alpha\kappa} > 0 \\
\frac{\partial Y^\infty}{\partial \eta} &= -\frac{\eta + \rho\sigma\alpha}{\kappa^2\alpha^2},
\end{aligned}
\tag{21}
$$

where only $\frac{\partial Y^\infty}{\partial \beta}$ in (21) has a clear cut (positive) sign. Since positive long-run yields or long-run forward rates are not guaranteed, varying the parameters will sooner or later give negative yields. The question is, however, if it really matters that the model does not preclude negative values of yields and forward rates. It is, of course, very difficult to give a short, convincing answer. Instead, some yield curves and forward rate curves in the Gaussian model are presented. The parameters used are basically the same as in the previous model except for the volatility parameters. These have been chosen in comparison with other parameters in BDFS [1], and have been slightly altered to give roughly the same long-run yield as in the previous model. It should once again be stressed that the objective is to see what kind of yield curves and forward

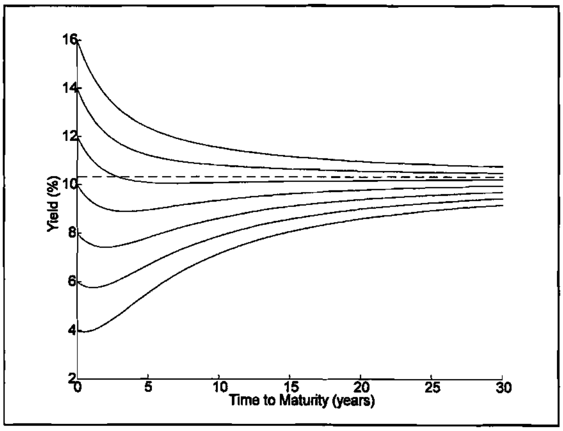

FIGURE 8. Yield Curves for Decreasing r (top-down)
Yield curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all yield curves, $\theta = 0.5\%$. The dashed line represents the long-run yield of 10.33%.

rate curves can be generated, and if the Gaussian model in this section can match the flexibility displayed by the previous model. In the figures below, $\kappa = 0.25$, $\sigma = 0.046$, $\alpha = 0.76$, $\beta = 0.023$, $\eta = 0.005$ and $\rho = -0.12$ giving a long-run yield and long-run forward rate of 10.33%.

Starting with the yield curve, the short rate is varied for two different values of θ. This is shown in Figures 8–9. In Figure 10 the factor θ is varied for three different values of r. The same procedure is repeated for the forward rates. In Figures 11–12 short rate is varied for two different values of θ. For three different values of the short rate the factor θ is varied in Figure 13.

An examination of the yield curves and forward rate curves generated by the Gaussian model in this section reveals that the Gaussian model seems to be capable of generating yield curves and forward rate curves strikingly similar to those generated by the previous model. Thus the two different models appear to be equivalent with regard to generating yield curves and forward rate curves.

7 Summary and Conclusions

In this essay, a solution is derived to discount bond prices in a two-factor model of the term structure of interest rates where the short rate and the (scaled) mean reversion level of the short rate are the factors. The solution involves a series expansion, a technique first used by Selby and Strickland [13] to simplify the implementation of the Fong and Vasicek model [8]. Several examples are given of yield curves and forward rate curves that display a greater flexibility than the corresponding curves generated by pure short-rate models.

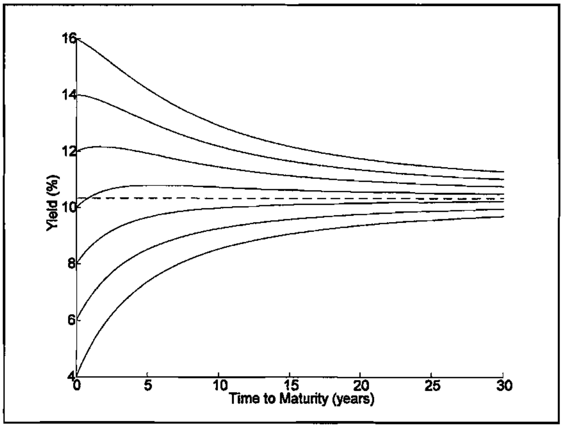

FIGURE 9. Yield Curves for Decreasing r (top-down)
Yield curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all yield curves, $\theta = 3.5\%$. The dashed line represents the long-run yield of 10.33%.

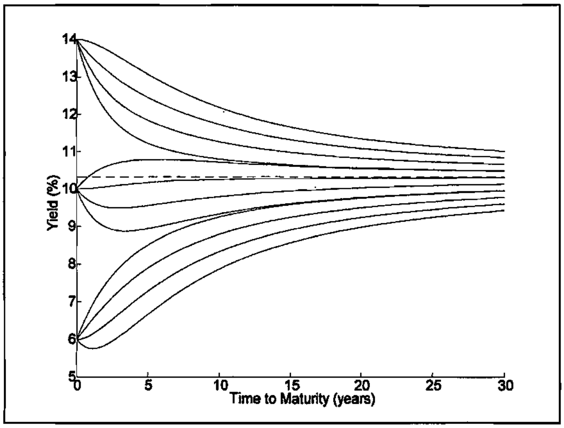

FIGURE 10. Yield Curves for Decreasing θ (top-down)
Yield curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3.5\%$ (top), $\theta = 2.5\%$, $\theta = 1.5\%$, and $\theta = 0.5\%$ (down). The dashed line represents the long-run yield of 10.33%.

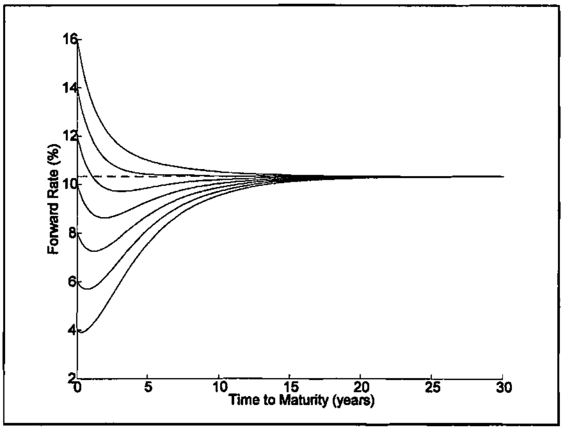

FIGURE 11. Forward Rate Curves for Decreasing r (top-down)
Forward rate curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all forward rate curves, $\theta = 0.5\%$. The dashed line represents the long-run forward rate of 10.33%.

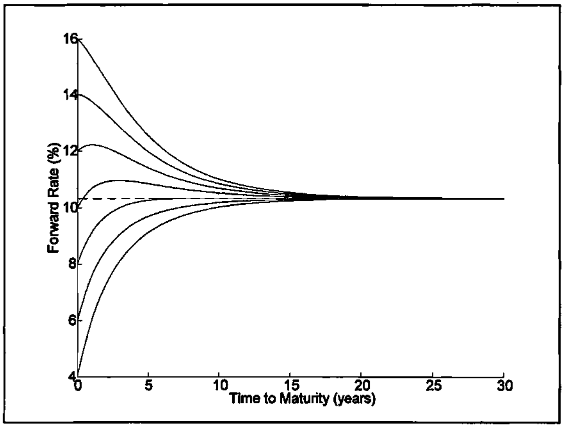

FIGURE 12. Forward Rate Curves for Decreasing r (top-down)
Forward rate curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all forward rate curves, $\theta = 3.5\%$. The dashed line represents the long-run forward rate of 10.33%.

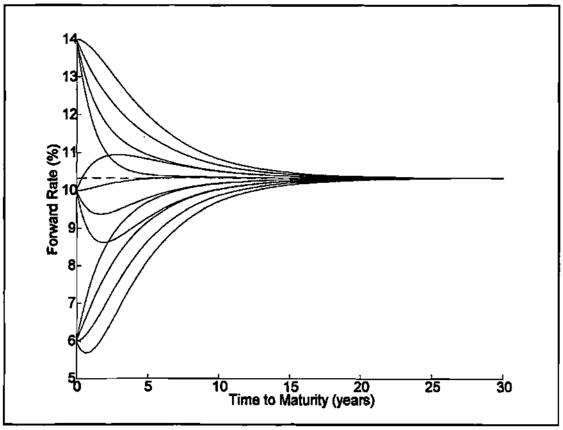

FIGURE 13. Forward Rate Curves for Decreasing θ (top-down)
Forward rate curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3.5\%$ (top), $\theta = 2.5\%$, $\theta = 1.5\%$ and $\theta = 0.5\%$ (down). The dashed line represents the long-run forward rate of 10.33%.

The model's flexibility in generating yield curves and forward rate curves of different shapes is compared with a similar Gaussian model, and no substantial difference is found between the models in this regard. One possible explanation is that in the first model presented in section 2, the volatilities of each of the factors were dependent on the level of the respective factor, while the Wiener processes were uncorrelated. In the Gaussian model, the volatility is constant for both factors, but to compensate for this, the Gaussian model allows for correlation between the Wiener processes, and thus introduces another parameter that adds flexibility to the model.

What remains to be determined in the model in section 2 are prices of other derivatives, e.g., the call option on a discount bond. Given the complexity of the PDE, it seems reasonable that other derivative prices should be calculated using Monte Carlo simulation.

Appendix A
Method of Frobenius

The following section describes a procedure known as the method of Frobenius used to obtain a solution to a certain type of linear differential equations. Coddington [5] is followed, and a brief summary of the method is presented. The aim here is to explain how the solution can be found; not to give a formal proof.

Suppose that we have a second order linear differential equation with a regular singular point at $x_0 = 0$, i.e.,

$$L(Q)(x) = x^2 Q''(x) + \tilde{a}(x) x Q'(x) + \tilde{b}(x) Q(x) = 0, \tag{22}$$

where $\tilde{a}(x)$ and $\tilde{b}(x)$ are analytic at $x_0 = 0$. Since $\tilde{a}(x)$ and $\tilde{b}(x)$ are analytic at $x_0 = 0$, they can be expressed in power series expansions

$$\begin{aligned} \tilde{a}(x) &= \sum_{k=0}^{\infty} \tilde{\alpha}_k x^k \\ \tilde{b}(x) &= \sum_{k=0}^{\infty} \tilde{\beta}_k x^k, \end{aligned}$$

which are convergent on an interval $| x |< R$ for some $R > 0$. To find a solution to (22), try an *Anzats* of the form

$$Q(x) = \sum_{k=0}^{\infty} c_k x^{k+q}, \tag{23}$$

where $c_0 \neq 0$ and q is a constant to be determined. Then,

$$\begin{aligned} Q'(x) &= \sum_{k=0}^{\infty} (k+q) c_k x^{k+q-1} \\ Q''(x) &= \sum_{k=0}^{\infty} (k+q)(k+q-1) c_k x^{k+q-2}, \end{aligned}$$

and hence

$$\begin{aligned} \tilde{b}(x) Q(x) &= x^q \sum_{k=0}^{\infty} \tilde{\beta}_k x^k \sum_{k=0}^{\infty} c_k x^k \\ &= x^q \sum_{k=0}^{\infty} \sum_{j=0}^{k} c_j \tilde{\beta}_{k-j} x^k, \end{aligned}$$

$$
\begin{aligned}
x\tilde{a}(x)Q'(x) &= x^q \sum_{k=0}^{\infty} \tilde{\alpha}_k x^k \sum_{k=0}^{\infty} c_k(k+q)x^k \\
&= x^q \sum_{k=0}^{\infty} \sum_{j=0}^{k} (j+q)c_j \tilde{\alpha}_{k-j} x^k,
\end{aligned}
$$

and

$$
x^2 Q''(x) = x^q \sum_{k=0}^{\infty} (k+q)(k+q-1)c_k x^k.
$$

Thus,

$$
\begin{aligned}
L(Q)(x) &= x^q \sum_{k=0}^{\infty} \left\{ (k+q)(k+q-1)c_k + \sum_{j=0}^{k} \left((j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right) c_j \right\} x^k \\
&= 0. \qquad (24)
\end{aligned}
$$

Relation (24) must hold for all values of x, and hence each coefficient must be identically zero. Since $c_0 \neq 0$, we have for $k = 0$

$$
p(q) = q(q-1) + q\tilde{\alpha}_0 + \tilde{\beta}_0 = 0,
$$

where $p(q)$ is called the *indicial polynomial* for (22), and the only admissible values of q are the roots of $p(q)$. Furthermore, for $k = 1, 2, \ldots,$

$$
p(q+k)c_k + \sum_{j=0}^{k-1} \left\{ (j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j = 0,
$$

or

$$
c_k = -\frac{\sum_{j=0}^{k-1} \left\{ (j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j}{p(q+k)}, \qquad (25)
$$

where c_k is well defined except at the two roots of $p(q+k)$. Now define

$$
U(x,q) = c_0 x^q + x^q \sum_{k=1}^{\infty} c_k(q) x^k. \qquad (26)
$$

If the series in (26) converges for $0 < x < R$, then clearly

$$
L(U)(x,q) = c_0 p(q) x^q,
$$

and hence if the $Q(x)$ given by (23) is a solution of (22), then q must be a root of the indicial polynomial p, and the coefficients c_k $(k \geq 1)$ are determined uniquely in terms of q and c_0 by (25), provided $p(q+k) \neq 0, k = 1, 2, \ldots$. Conversely, if q is a root of p, and if the coefficients c_k can be determined, i.e., $p(q+k) \neq 0, k = 1, 2, \ldots$, then the function Q given by $Q(x) = U(x, q)$ is a solution of (22) for any choice of c_0, provided the series in (26) converges.

In general, the nature of the solution of (22) depends on the roots of p. The only case of interest here is when the roots of p are real and distinct, and we shall assume from now on that this is the case. Suppose, therefore, that the roots of p are q_1 and q_2, such that $q_1 \neq q_2 + k$ for $k = 0, 1, 2, \ldots$. Then $p(q_1) = 0$, and no k can exist such that $p(q_1 + k) = 0$. Hence the coefficients in (25) exist for all $k = 1, 2, \ldots$, and letting $c_0 = 1$, we see that the function Q_1 given by

$$\begin{aligned} Q_1 &= U(x, q_1) \\ &= x^{q_1} \sum_{k=0}^{\infty} c_k(q_1) x^k \end{aligned}$$

is a solution of (22), since it can be shown to converge for $\mid x \mid < R$. Using the same arguments, the function Q_2 given by

$$\begin{aligned} Q_2 &= U(x, q_2) \\ &= x^{q_2} \sum_{k=0}^{\infty} c_k(q_2) x^k \end{aligned}$$

is also a solution of (22).

It can be shown that the solutions Q_1 and Q_2 are independent, and hence the most general solution of (22) on the interval $0 < \mid x \mid < R$ is given by a linear combination of the form

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2), \tag{27}$$

where K_1 and K_2 are suitable constants, and where the coefficients $c_k^{(i)}$ ($i = 1, 2$ and $k = 1, 2, \ldots$) of $U(x, q_i)$ are given by

$$c_k^{(i)} = -\frac{\sum_{j=0}^{k-1} \left\{ (j + q_i)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j^{(i)}}{p(q_i + k)},$$

where $c_0^{(1)} = c_0^{(2)} = 1$.

Remark 3 *Note that if the roots to $p(q)$ are nonnegative, then $Q(x)$ in (27) is well defined at $x = 0$.*

Appendix B
The Functions A(t,T), B(t,T) and C(t,T)

The Function B(t,T)

The ordinary differential equation and terminal condition for $B(t,T)$ is

$$\begin{aligned} B_t(t,T) - \kappa B(t,T) - \frac{1}{2}\sigma^2 B^2(t,T) + 1 &= 0 \\ B(T,T) &= 0. \end{aligned}$$

The well-known solution is

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

where $\gamma = \sqrt{\kappa^2 + 2\sigma^2}$. It is straightforward to verify that $\lim_{T\to\infty} B(t,T) = \frac{2}{\gamma+\kappa}$. Defining $\delta = \frac{\gamma-\kappa}{\gamma+\kappa}$, $B(t,T)$ can be rewritten as

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

and since $\delta e^{-\gamma(T-t)} < 1$, it follows that $B(t,T)$ can in fact be expressed in terms of a (convergent) series expansion:

$$B(t,T) = \frac{2}{\gamma+\kappa} + \frac{2\gamma}{\sigma^2}\sum_{j=1}^{\infty}(-\delta)^j e^{-\gamma j(T-t)}. \tag{28}$$

The alternative expression (28) for $B(t,T)$ will come in useful below when determining the function $C(t,T)$.

The Function C(t,T)

The ordinary differential equation and terminal condition for $C(t,T)$ is

$$\begin{aligned} C_t(t,T) - \alpha C(t,T) - \frac{1}{2}\eta^2 C^2(t,T) + B(t,T) &= 0 \qquad (29) \\ C(T,T) &= 0. \end{aligned}$$

This is a Ricatti equation, and the standard substitution to obtain $C(t,T)$ is

$$C(t,T) = -\frac{2}{\eta^2}\frac{w_t(t,T)}{w(t,T)}. \tag{30}$$

Then,

$$C_t(t,T) = -\frac{2}{\eta^2}\left(\frac{w(t,T)w_{tt}(t,T) - w_t^2(t,T)}{w^2(t,T)}\right), \tag{31}$$

and substituting (30) and (31) into (29) and rearranging, we obtain

$$w_{tt}(t,T) - \alpha w_t(t,T) - \frac{\eta^2}{2}B(t,T)w(t,T) = 0. \tag{32}$$

Now we make a change of variables according to

$$w(t,T) = x^{\varepsilon}Q(x), \tag{33}$$

where $x = e^{-\gamma(T-t)}$, and ε is a constant to be determined. We note that $\frac{dx}{dt} = \gamma x$ and hence

$$w_t(t,T) = \gamma\varepsilon x^{\varepsilon}Q(x) + \gamma x^{\varepsilon+1}Q'(x), \tag{34}$$

and

$$w_{tt}(t,T) = \gamma^2\varepsilon^2x^{\varepsilon}Q(x) + \gamma^2\left(1+2\varepsilon\right)x^{\varepsilon+1}Q'(x) + \gamma^2x^{\varepsilon+2}Q''(x). \tag{35}$$

Substituting (33), (34) and (35) into (32), and using the representation (28) with $x = e^{-\gamma(T-t)}$, we obtain after simplifications

$$x^2Q''(x) + \left(1 + 2\varepsilon - \frac{\alpha}{\gamma}\right)xQ'(x) + \left(\varepsilon^2 - \frac{\alpha}{\gamma}\varepsilon - \frac{\eta^2}{\gamma^2(\gamma+\kappa)} - \frac{\eta^2}{\gamma\sigma^2}\sum_{j=1}^{\infty}(-\delta)^j x^j\right)Q(x) = 0. \tag{36}$$

Now we choose ε to make the constant term in the coefficient of $Q(x)$ in (36) equal to zero. This is the case if

$$\varepsilon^2 - \frac{\alpha}{\gamma}\varepsilon - \frac{\eta^2}{\gamma^2(\gamma+\kappa)} = 0,$$

or

$$\varepsilon = \frac{\alpha}{2\gamma} \pm \sqrt{\left(\frac{\alpha}{2\gamma}\right)^2 + \frac{\eta^2}{\gamma^2(\gamma+\kappa)}}, \tag{37}$$

and we note that ε is real. With ε as in (37), the differential equation (36) takes the form

$$x^2 Q''(x) + \left(1 + 2\varepsilon - \frac{\alpha}{\gamma}\right) xQ'(x) - \frac{\eta^2}{\gamma\sigma^2}\sum_{j=1}^{\infty}(-\delta)^j x^j Q(x) = 0. \tag{38}$$

Now define

$$\tilde{a}(x) = \sum_{k=0}^{\infty} \tilde{\alpha}_k x^k$$

where

$$\begin{aligned} \tilde{\alpha}_0 &= 1 + 2\varepsilon - \frac{\alpha}{\gamma} \\ \tilde{\alpha}_k &= 0,\ k = 1, 2, ..., \end{aligned}$$

and

$$\tilde{b}(x) = \sum_{k=0}^{\infty} \tilde{\beta}_k x^k$$

where

$$\begin{aligned} \tilde{\beta}_0 &= 0 \\ \tilde{\beta}_k &= -\frac{\eta^2}{\gamma\sigma^2}(-\delta)^k,\ k = 1, 2, ..., \end{aligned}$$

with $\delta = \frac{\gamma-\kappa}{\gamma+\kappa}$. Note that $\tilde{a}(x)$ and $\tilde{b}(x)$ converge for $0 \le x \le 1$. Then the ODE for $Q(x)$ becomes

$$x^2 Q''(x) + \tilde{a}(x) x Q'(x) + \tilde{b}(x) Q(x) = 0, \tag{39}$$

and from Appendix A we know that the solution to (39) is given by

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2), \tag{40}$$

where K_1 and K_2 are suitable constants and $U(x, q_i)$ $(i = 1, 2)$ is defined in Appendix A. To determine the terminal condition for $Q(x)$, we have

$$\begin{aligned} C(t,T) &= -\frac{2}{\eta^2}\frac{w_t(t,T)}{w(t,T)} \\ &= -\frac{2\gamma}{\eta^2}\left(\varepsilon + \frac{xQ'(x)}{Q(x)}\right), \end{aligned}$$

where we have used (33), and ε is given by (37). The condition $C(T,T) = 0$ (remember that $x = e^{-\gamma(T-t)}$) gives

$$\frac{Q'(1)}{Q(1)} + \varepsilon = 0. \tag{41}$$

We remember that $Q(x)$ satisfies a second order differential equation, but we have at our disposal only one condition (equation (41)) to determine the constants K_1 and K_2. However, this condition will suffice since we are not interested in the function $Q(x)$ itself, but rather the quotient $Q'(x)/Q(x)$. It is therefore necessary only to determine the quotient K_2/K_1. Using (40) and (41), we then have that

$$\frac{K_1U'(1,q_1) + K_2U'(1,q_2)}{K_1U(1,q_1) + K_2U(1,q_2)} + \varepsilon = 0,$$

or

$$K^* = -\frac{U'(1,q_1) + \varepsilon U(1,q_1)}{U'(1,q_2) + \varepsilon U(1,q_2)},$$

where $K^* = K_2/K_1$. Hence the function $C(t,T)$ is given by

$$C(t,T) = -\frac{2\gamma}{\eta^2}\left(\varepsilon + \frac{e^{-\gamma(T-t)}Q'(e^{-\gamma(T-t)})}{Q(e^{-\gamma(T-t)})}\right),$$

where

$$Q(x) = K_1U(x,q_1) + K_2U(x,q_2)$$

and

$$K^* = \frac{K_2}{K_1} = -\frac{U'(1,q_1) + \varepsilon U(1,q_1)}{U'(1,q_2) + \varepsilon U(1,q_2)}.$$

Before finding the limit for $C(t,T)$, we recall that ε is given by $\varepsilon^{\pm} = \frac{1}{2}\chi \pm \frac{1}{2}\varphi$, where $\chi = \frac{\alpha}{\gamma}$ and $\varphi = \sqrt{\chi^2 + \frac{4\eta^2}{\gamma^2(\gamma+\kappa)}} > 0$. By construction, $C(t,T)$ does not depend on which ε

we choose. However, for numerical reasons, we choose $\varepsilon = \varepsilon^- = \frac{1}{2}\chi - \frac{1}{2}\varphi$, since we then have $q_2 = 1 - \tilde{\alpha}_0 = \varphi > 0$, and then $Q(x)$ is well defined for $x = 0$. We now examine the limit for $C(t,T)$:

$$\begin{aligned}\lim_{T\to\infty} C(t,T) &= -\frac{2\gamma}{\eta^2}\lim_{x\to 0}\left(\varepsilon + \frac{xQ'(x)}{Q(x)}\right)\\ &= -\frac{2\gamma\varepsilon}{\eta^2} - \frac{2\gamma}{\eta^2}\lim_{x\to 0}\frac{xQ'(x)}{Q(x)}.\end{aligned}$$

Therefore, the relevant limit to investigate is

$$\lim_{x\to 0}\frac{xQ'(x)}{Q(x)} = \lim_{x\to 0}\frac{K_1\sum_{n=0}^{\infty} c_n^{(1)} n x^n + K_2 \sum_{n=0}^{\infty} c_n^{(2)}(n+q_2)x^{n+q_2}}{K_1\sum_{n=0}^{\infty} c_n^{(1)} x^n + K_2\sum_{n=0}^{\infty} c_n^{(2)} x^{n+q_2}}.$$

Now with $\varepsilon = \varepsilon^-$, where

$$\varepsilon^- = \frac{1}{2}\chi - \frac{1}{2}\varphi \tag{42}$$

and $q_2 = \varphi > 0$, we thus have

$$\lim_{x\to 0}\frac{xQ'(x)}{Q(x)} = \frac{0}{c_0^{(1)}} = 0,$$

since we have assumed that $c_0^{(1)} = 1$. Thus

$$\begin{aligned}\lim_{T\to\infty} C(t,T) &= -\frac{2\gamma\varepsilon^-}{\eta^2}\\ &= -\frac{2\gamma}{\eta^2}\left(\frac{1}{2}\chi - \frac{1}{2}\varphi\right)\\ &= \frac{\gamma}{\eta^2}(\varphi - \chi)\\ &= \frac{\alpha}{\eta^2}\left(\sqrt{1 + \frac{4\eta^2}{\alpha^2(\gamma+\kappa)}} - 1\right).\end{aligned}$$

The Function $A(t,T)$

The ordinary differential equation and terminal condition for $A(t,T)$ is

$$\begin{aligned}A_t(t,T) &= \beta C(t,T)\\ A(T,T) &= 0.\end{aligned}$$

Integrate and use the terminal condition:

$$A(t,T) = -\beta \int_t^T C(s,T)ds.$$

We recall that $C(t,T) = -\frac{2}{\eta^2}\frac{w_t(t,T)}{w(t,T)}$, where $w(t,T) = x^{\varepsilon}Q(x)$ and $x = e^{-\gamma(T-t)}$. Then we have

$$\begin{aligned}
A(t,T) &= -\beta \int_t^T C(s,T)ds \\
&= \frac{2\beta}{\eta^2}\int_t^T \frac{w_s(s,T)}{w(s,T)}ds \\
&= \frac{2\beta}{\eta^2}\left[\log w(s,T)\right]_t^T \\
&= \frac{2\beta}{\eta^2}\log\frac{w(T,T)}{w(t,T)} \\
&= \frac{2\beta}{\eta^2}\log\frac{Q(1)}{e^{-\gamma\varepsilon(T-t)}Q(e^{-\gamma(T-t)})} \\
&= \frac{2\beta}{\eta^2}\log\frac{e^{\gamma\varepsilon(T-t)}Q(1)}{Q(e^{-\gamma(T-t)})}.
\end{aligned}$$

In summary, we obtain

$$A(t,T) = \frac{2\beta}{\eta^2}\log\frac{e^{\gamma\varepsilon(T-t)}Q(1)}{Q(e^{-\gamma(T-t)})}.$$

To find the limit for $A(t,T)$ as T tends to infinity, we begin by investigating the limit for $\frac{e^{\gamma\varepsilon(T-t)}Q(1)}{Q(e^{-\gamma(T-t)})}$, or for $\frac{x^{-\varepsilon}Q(1)}{Q(x)}$ as x tends to zero. As above, with $\varepsilon = \varepsilon^-$, we have $q_2 = \varphi > 0$ and

$$\begin{aligned}
\lim_{x\to 0}\frac{x^{-\varepsilon^-}Q(1)}{Q(x)} &= Q(1)\lim_{x\to 0}\frac{x^{-\varepsilon^-}}{K_1\sum_{n=0}^{\infty}c_n^{(1)}x^n + K_2\sum_{n=0}^{\infty}c_n^{(2)}x^{n+q_2}} \\
&= Q(1)\lim_{x\to 0}\frac{x^{-\varepsilon^-}}{K_1\sum_{n=0}^{\infty}c_n^{(1)}x^n + K_2\sum_{n=0}^{\infty}c_n^{(2)}x^{n+\varphi}} \\
&= \frac{Q(1)}{K_1 c_0^{(1)}}\lim_{x\to 0}x^{-\varepsilon^-} \\
&= 0,
\end{aligned}$$

since from (42) $\varepsilon^- = \frac{\alpha}{2\gamma} - \sqrt{\left(\frac{\alpha}{2\gamma}\right)^2 + \frac{\eta^2}{\gamma^2(\gamma+\kappa)}} < 0$. Thus, $\lim_{x\to 0} \frac{x^{-\varepsilon}Q(1)}{Q(x)} = 0$, and hence

$$\begin{aligned}
\lim_{T\to\infty} A(t,T) &= \frac{2\beta}{\eta^2} \lim_{T\to\infty} \log \frac{e^{\gamma\varepsilon(T-t)}Q(1)}{Q(e^{-\gamma(T-t)})} \\
&= \frac{2\beta}{\eta^2} \lim_{x\to 0} \log \frac{x^{-\varepsilon}Q(1)}{Q(x)} \\
&= \frac{2\beta}{\eta^2} \log \lim_{x\to 0} \frac{x^{-\varepsilon}Q(1)}{Q(x)} \\
&= -\infty.
\end{aligned}$$

This ensures that discount bond prices converge to zero as T tends to infinity.

References

[1] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram (1996a): "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income,* December, 43–53.

[2] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram (1996b): "Stochastic Mean Models of the Term Structure of Interest Rates," Working paper, Harvard University and New York University.

[3] Björk, T. and M. Hyll (2000): "On the Inversion of the Yield Curve," Working paper, Stockholm School of Economics.

[4] Chen, L. (1996): "Interest Rate Dynamics, Derivatives Pricing, and Risk Management," Springer Verlag.

[5] Coddington, E.A. (1961): "An Introduction to Ordinary Differential Equations," Prentice-Hall.

[6] Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985b): "A Theory of the Term Structure of Interest Rates," *Econometrica* **53**(2), 385–407.

[7] Duffie, D. and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance* **6**(4), 379–406.

[8] Fong, H.G. and O.A. Vasicek (1991a): "Interest Rate Volatility as a Stochastic Factor," Unpublished manuscript, Gifford Fong Associates.

[9] Fong, H.G., O.A. Vasicek, and D. Yoo (1992): "Omission Impossible," *Risk* **5**(2), 62–65.

[10] Heath, D., R. Jarrow, and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica* **60**(1), 77–105.

[11] Hyll, M. (2000d): "An Efficient Series Expansion Approach to the Balduzzi, Das, Foresi and Sundaram Model of the Term Structure of Interest Rates," Working paper, Stockholm School of Economics.

[12] Longstaff, F.A. and E.S. Schwartz (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance* **47**(4), 1259–1282.

[13] Selby, M. and C. Strickland (1995) "Computing the Fong and Vasicek Pure Discount Bond Price Formula," *The Journal of Fixed Income,* September, 78–84.

[14] Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* **5**, 177–188.

4

An Efficient Series Expansion Approach to The Balduzzi, Das, Foresi and Sundaram Model of the Term Structure of Interest Rates

ABSTRACT

This essay revisits the three-factor model of the term structure of interest rates proposed by Balduzzi, Das, Foresi and Sundaram (BDFS) where the short rate, the mean and the variance of the short rate are the factors. Using a method of series expansion, the solution to discount bond prices is provided. Moreover, the model is also extended so that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve. It is also shown that the term structure in the BDFS model allows for a change of factors, in that the original factors may be substituted for three arbitrarily chosen benchmark forward rates.

1 Introduction

From a practitioner's point of view, much is to be gained from a closed form solution to the term structure of interest rates, or any other financial derivative for that matter. Alternatives are numerical solutions to partial differential equations, discrete tree models, approximation techniques or Monte Carlo simulations, all of which are usually time-consuming and sometimes difficult to implement. As an example of the importance of an easily accessible closed form solution, we can choose to compare the Fong and Vasicek (FV) [6] and the Longstaff and Schwartz (LS) [9] models of the term structure of interest rates. Both models are two-factor models where the short rate and the volatility of the short rate are the factors. Selby and Strickland (SS) [10] argue that one of the reasons why the LS model is more widely used than the FV model is that discount bond prices in the former are obtained in closed form, while implementing discount bond prices in the latter requires more sophisticated mathematical machinery. SS show that discount bond prices can in fact be obtained in the FV model in a more direct and user-friendly manner by using the concept of a series expansion. The difference between FV and SS is in their approach to obtaining the solution to an ordinary differential equation. The solutions obtained in the two different approaches are, of course, the same, but SS express their solution as a series expansion involving only real numbers (as opposed to complex numbers), which is desirable from a computational point of view.

This essay revisits the three-factor model proposed by Balduzzi *et al.* (BDFS) [1], where the short rate, the mean and volatility of the short rate are the three factors. BDFS argue that the most natural extensions of one-factor models are those including the volatility of the short rate and the mean rate as additional factors. The reason for this is the recent evidence that volatility displays time-varying behavior, and that target rates are used as a key tool of monetary policy. The factors are assumed to follow stochastic differential equations that give rise to an affine term structure. Nevertheless, the system of ordinary differential equations that needs to be solved within the model requires a nonstandard solution method, such as expressing the solution to an ordinary differential equation in the form of a series expansion.

The main contributions of this essay are as follows:

- using a method of series expansion, prices of discount bonds are provided
- it is shown that the model allows for a change of factors (given a technical condition), i.e., the original three factors may be substituted for three (arbitrarily chosen) benchmark forward rates
- the model is extended so that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve (Since the FV model is a special case of the BDFS model, the same result is in fact shown also for that model).

The rest of the essay is organized as follows: Section 2 presents a slightly different version of the BDFS model, and the term structure is derived and solved by using a series expansion in section 3. Examples of yield curves and forward rate curves are given in sections 4 and 5, section 6 shows how the price of a discount bond can be obtained in terms of three benchmark forward rates. The model is extended in section 7, and it is shown that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve. Section 8 summarizes and concludes.

2 Model

It is assumed that the dynamics of r_t, V_t and θ_t under a martingale measure Q are given by

$$
\begin{aligned}
dr_t &= (\theta_t - \kappa r_t - \lambda V_t)dt + \sqrt{V_t}dW_t^1 && (1)\\
d\theta_t &= (\beta - \alpha\theta_t)dt + \gamma dW_t^2 && (2)\\
dV_t &= (b - aV_t)dt + \sigma\sqrt{V_t}dW_t^3, && (3)
\end{aligned}
$$

where $\kappa, \lambda, \beta, \alpha, \gamma, b, a$ and σ are constants and W_t^1, W_t^2 and W_t^3 are Wiener processes satisfying $dW_t^1 dW_t^2 = dW_t^2 dW_t^3 = 0$ and $dW_t^1 dW_t^3 = \rho dt$. In (1)–(3), r_t is the short rate, V_t is the volatility of the short rate and θ_t is a factor affecting the mean reversion level of the short rate. Note that θ_t is not the mean reversion level under Q, since we also have in the drift of the short rate a contribution from V_t. Since θ_t and V_t have mean reversion levels $\frac{\beta}{\alpha}$ and $\frac{b}{a}$ respectively, the mean reversion level for the short rate is $\frac{\beta}{\alpha\kappa} - \lambda\frac{b}{a\kappa}$. The stochastic process followed by θ_t is usually referred to as an Ornstein-Uhlenbeck process, whereas V_t follows a CIR process, or a "square root process". For fixed V_t and θ_t, the short rate r_t follows an Ornstein-Uhlenbeck process.

Fong and Vasicek extended the Vasicek model to allow for stochastic volatility, and the above model is thus an extension of the FV model in that it has a stochastic mean reversion level. Given the above dynamics of the factors, we can only expect an affine term structure if we allow correlation between W_t^1 and W_t^3. Any correlation between W_t^1 and W_t^2 or W_t^2 and W_t^3 will violate the necessary and sufficient conditions for the existence of an affine term structure[1]. It is not argued that these are the most feasible relations between the Wiener processes from an economic point of view, but they are necessary from a tractability perspective.

The main difference between our setup and that of BDFS is that they model the dynamics of their factors under an objective probability measure, where their θ_t is the true mean reversion level of the short rate. Assumptions about market prices of risk are then made to maintain the

[1] See Duffie and Kan [5].

structure of the stochastic differential equations under a risk-adjusted measure. Our specification under Q is essentially equivalent to the risk neutral formulation of BDFS.

3 The Term Structure and Properties of Discount Bond Prices

This section considers the term structure of interest rates for the model in (1)–(3), and we denote by $P(t,T)$ the price at time $t \geq 0$ of a default-free discount bond with unit payment at $T \geq t$. Using standard financial theory, we have $P(t,T) = F(t, r_t, \theta_t, V_t; T)$, where the pricing function $F(t, r, \theta, V; T)$ satisfies the so-called term structure equation

$$\begin{aligned} & F_t + (\theta - \kappa r - \lambda V)F_r + \frac{1}{2}VF_{rr} + (\beta - \alpha\theta)F_\theta + \frac{1}{2}\gamma^2 F_{\theta\theta} \\ & +(b - aV)F_V + \frac{1}{2}\sigma^2 V F_{VV} + \rho\sigma V F_{rV} = rF, \end{aligned} \tag{4}$$

with boundary condition given by $F(T, r, \theta, V; T) = 1$. Since the dynamics of the three factors satisfy the necessary and sufficient conditions for the existence of an affine term structure, the pricing function is given by

$$F(t, r, \theta, V; T) = \exp\left(A(t,T) - B(t,T)r - C(t,T)\theta - D(t,T)V\right), \tag{5}$$

where $A(t,T), B(t,T), C(t,T)$ and $D(t,T)$ are time dependent functions satisfying the terminal conditions

$$A(T,T) = B(T,T) = C(T,T) = D(T,T) = 0. \tag{6}$$

Substituting (5) into (4) we obtain

$$\begin{aligned} & A_t - \beta C + \frac{1}{2}\gamma^2 C^2 - bD + \left(-B_t + \kappa B - 1\right) r + \left(-C_t - B + \alpha C\right)\theta \\ & + \left(-D_t + \lambda B + \frac{1}{2}B^2 + aD + \frac{1}{2}\sigma^2 D^2 + \rho\sigma BD\right) V \\ = \;& 0. \end{aligned}$$

Since this relation must hold for all values of r, θ and V we have in fact a system of four ordinary differential equations:

$$\begin{aligned} A_t(t,T) &= \beta C(t,T) + bD(t,T) - \frac{1}{2}\gamma^2 C^2(t,T) & (7)\\ B_t(t,T) &= \kappa B(t,T) - 1 & (8)\\ C_t(t,T) &= \alpha C(t,T) - B(t,T) & (9)\\ D_t(t,T) &= (a + \rho\sigma B(t,T))\, D(t,T) + \frac{1}{2}\sigma^2 D^2(t,T) + \lambda B(t,T) + \frac{1}{2}B^2(t,T), & (10) \end{aligned}$$

with terminal conditions as in (6). It is not evident that the solution to (4) is well behaved and possesses reasonable economic features for all values of the parameters of the factors r_t, θ_t, and V_t. To ensure that discount bond prices are real and tend to zero as time to maturity tends to infinity, it is assumed that

$$\left(\frac{a\kappa + \rho\sigma}{2\kappa^2}\right)^2 \geq \frac{\sigma^2}{2\kappa^3}\left(\lambda + \frac{1}{2\kappa}\right), \tag{11}$$

and that

$$\frac{\beta}{\alpha\kappa} - \frac{\gamma^2}{2\alpha^2\kappa^2} > \frac{2b}{\sigma^2}\varepsilon\kappa, \tag{12}$$

where

$$\varepsilon = \frac{a\kappa + \rho\sigma}{2\kappa^2} - \sqrt{\left(\frac{a\kappa + \rho\sigma}{2\kappa^2}\right)^2 - \frac{\sigma^2}{2\kappa^3}\left(\lambda + \frac{1}{2\kappa}\right)}.$$

We then have the following proposition:

Proposition 1 *Suppose that r_t, θ_t and V_t under a martingale measure Q satisfy (1)-(3). Then discount bond prices are given by*

$$P(t,T) = \exp\left(A(t,T) - B(t,T)r_t - C(t,T)\theta_t - D(t,T)V_t\right),$$

where

$$\begin{aligned}
A(t,T) &= \left(\frac{\beta}{\alpha\kappa} - \frac{\gamma^2}{2\alpha^2\kappa^2}\right)(B(t,T)-(T-t)) + \left(\frac{\beta}{\alpha} - \frac{\gamma^2}{2\alpha^2(\alpha+\kappa)}\right) C(t,T) \\
&\quad -\frac{1}{4\alpha}\gamma^2\left(C^2(t,T) + \frac{1}{\kappa(\alpha+\kappa)}B^2(t,T) + \frac{2}{(\alpha+\kappa)}B(t,T)C(t,T)\right) \\
&\quad +\frac{2b}{\sigma^2}\log\frac{e^{\varepsilon\kappa(T-t)}Q(1)}{Q(e^{-\kappa(T-t)})} &(13)\\
B(t,T) &= \frac{1-e^{-\kappa(T-t)}}{\kappa} &(14)\\
C(t,T) &= \frac{\alpha\left(1-e^{-\kappa(T-t)}\right) - \kappa\left(1-e^{-\alpha(T-t)}\right)}{\alpha\kappa(\alpha-\kappa)} &(15)\\
D(t,T) &= -\frac{2\kappa}{\sigma^2}\left(\varepsilon + \frac{e^{-\kappa(T-t)}Q'(e^{-\kappa(T-t)})}{Q(e^{-\kappa(T-t)})}\right), &(16)
\end{aligned}$$

and

$$\varepsilon = \frac{a\kappa+\rho\sigma}{2\kappa^2} - \sqrt{\left(\frac{a\kappa+\rho\sigma}{2\kappa^2}\right)^2 - \frac{\sigma^2}{2\kappa^3}\left(\lambda+\frac{1}{2\kappa}\right)}. \tag{17}$$

The function $Q(x)$ is defined by

$$Q(x) = K_1U(x,q_1) + K_2QU(x,q_2), \tag{18}$$

where

$$\frac{K_1}{K_2} = -\frac{U'(1,q_1)+\varepsilon U(1,q_1)}{U'(1,q_2)+\varepsilon U(1,q_2)}$$

and

$$U(x,q_i) = \sum_{k=0}^{\infty} c_k^{(i)} x^{k+q_i}$$

with $q_1 = 0$ and $q_2 = 1-\tilde{\alpha}_0$. The coefficients $c_k^{(i)}$ $(i = 1,2$ and $k = 1,2,\dots)$ of $U(x,q_i)$ are given by

$$\begin{aligned}
c_1^{(i)} &= -\frac{\left(q_i\tilde{\alpha}_1 + \tilde{\beta}_1\right)c_0}{p(q_i+1)} &(19)\\
c_k^{(i)} &= -\frac{\left\{(k-1+q_i)\tilde{\alpha}_1 + \tilde{\beta}_1\right\}c_{k-1} + \tilde{\beta}_2 c_{k-2}}{p(q_i+k)} &(20)
\end{aligned}$$

$$p(q) = q(q-1) + q\tilde{\alpha}_0,$$

where

$$\begin{aligned}
\tilde{\alpha}_0 &= 1 + 2\varepsilon - \frac{a\kappa + \rho\sigma}{\kappa^2} \\
\tilde{\alpha}_1 &= \frac{\rho\sigma}{\kappa^2} \\
\tilde{\beta}_1 &= \frac{\varepsilon\rho\sigma}{\kappa^2} - \frac{\sigma^2(1+\lambda\kappa)}{2\kappa^4} \\
\tilde{\beta}_2 &= \frac{\sigma^2}{4\kappa^4},
\end{aligned}$$

and $c_0^{(1)} = c_0^{(2)} = 1.$

Proof. See Appendix A and B. ∎

Remark 1 *The series expansion, i.e., the function* $Q(x)$*, thus appears in the functions* $A(t,T)$ *and* $D(t,T)$*, which are then obtained in "quasi-analytical" forms. BDFS solve for their corresponding* $A(t,T)$ *and* $D(t,T)$ *using numerical procedures.*

Remark 2 *The condition in (11) thus ensures that* ε *is real (see (17)) and hence* $D(t,T)$ *is real as defined in (16). In turn,* $A(t,T)$ *is real by (13). Furthermore, the condition in (12) guarantees that discount bond prices tend to zero as* T *tends to infinity, and that long-run yields and forward rates are positive.*[2]

Discount bonds can be shown to be decreasing and convex in r and θ, i.e., (for $0 \le t < T$) $B(t,T) > 0$ and $C(t,T) > 0$. This implies that discount bond prices have the realistic feature that $P(t,T)$ tends to zero as $r \to \infty$ and $\theta \to \infty$. The complicated structure of $D(t,T)$ makes it somewhat difficult to analytically determine whether $D(t,T)$ is strictly positive or negative, or if the sign of $D(t,T)$ is indeterminate. Preformed tests suggest that $D(t,T)$ is in fact negative (except for very short maturities), and although not proven, this is in line with the simpler stochastic volatility model proposed by FV, where the "volatility exposure" (generally) is positive.[3] Thus, performed tests hint at $D(t,T) < 0$. This, in turn, implies that discount bond prices may not stay bounded as $V \to \infty$. One possible explanation for this is that as $V \to \infty$, the short rate might take large negative values with positive probability, and thus introduce arbitrage opportunities. The above characteristics are different from those obtained in the LS model where the signs of the partial derivatives of the bond price with respect to r and V are indeterminate.

[2] See equation (57) in Appendix B, and also equations (22) and (28).
[3] See Fong *et al.* [7].

It is also worth pointing out that the derivatives $B_t(t,T)$ and $C_t(t,T)$ are negative, but, due to the reasons mentioned above, it is difficult to determine the sign of $D_t(t,T)$. These functions play a key role in an examination of yields and forward rates.

4 Term Structure Shapes

The yields are defined as $Y(t,T) = -\frac{\log P(t,T)}{T-t}$, and the yields in the model presented above thus take the form

$$Y(t,T) = -\frac{A(t,T) - B(t,T)r - C(t,T)\theta - D(t,T)V}{T-t}. \tag{21}$$

Taking the proper limits of (21), it is straightforward to verify that

$$\lim_{T\to\infty} Y(t,T) = \frac{\beta}{\alpha\kappa} - \frac{\gamma^2}{2\alpha^2\kappa^2} - \frac{2b\kappa\varepsilon}{\sigma^2}, \tag{22}$$

i.e., the long-run yield is independent of the factors and positive due to our assumption in (12). We also note that

$$\begin{aligned} Y_r(t,T) &= \frac{B(t,T)}{T-t} > 0 \\ Y_\theta(t,T) &= \frac{C(t,T)}{T-t} > 0, \end{aligned}$$

and that

$$Y_V(t,T) = \frac{D(t,T)}{T-t},$$

where we have reason to believe that $Y_V(t,T)$ generally is negative.

This section and the following section present several figures showing yield curves and forward rate curves. The parameters used in these figures are those obtained by BDFS, namely $\kappa = 0.25, \lambda = -0.10, \alpha = 0.76, \beta = 0.023, \gamma = 0.005, a = 0.29, b = 0.0002, \sigma = 0.003$ and $\rho = -0.12$. The long-run yield is 11.43 %. It is not claimed that these values are better than any other values, or that they capture the nature of the short rate and the other factors better than any other values, but they are used to generate examples of what yield curves and forward rate curves might look like in this model.

Examining the term structure of interest rates, the first task is to motivate the title of this essay, i.e., to motivate that the above series expansion is indeed efficient in some sense. The solutions for $A(t,T)$ and $D(t,T)$ both include the infinite series expansion denoted by $Q(x)$ and defined in (18), and when implementing the model this infinite series expansion must be

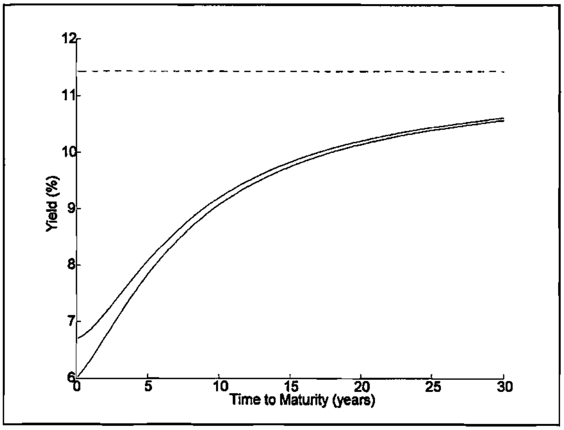

FIGURE 1. Yield Curves for Varying N
Yield curves for $N = 1$ (top) and $N = 2, 3, \ldots$(down). The dashed line represents the long-run yield of 11.43%.

truncated after, say, N terms. Loosely speaking, the series expansion is efficient if the yield curve, generated by (21), converges rapidly. There are at least two reasons to expect that the series expansion is in fact efficient. First, the magnitude of the coefficients decays rapidly due to the recursive relation in (19)–(20). As an example, with the parameters given above, the first six coefficients in each series are $c_0^{(1)} = 1, c_1^{(1)} = -7.3\ 10^{-3}, c_2^{(1)} = -3.7\ 10^{-4}, c_3^{(1)} = -8.1\ 10^{-8}, c_4^{(1)} = 1.9\ 10^{-8}, c_5^{(1)} = 2.6\ 10^{-11}$ and $c_0^{(2)} = 1, c_1^{(2)} = 3.6\ 10^{-3}, c_2^{(2)} = -8.410^{-5}, c_3^{(2)} = -3.0\ 10^{-7}, c_4^{(2)} = 2.0\ 10^{-9}, c_5^{(2)} = 7.5\ 10^{-12}$. The coefficients are recursively multiplied by small numbers, and, also, $c_k^{(i)} \sim \frac{1}{k^2}$. Second, the series expansion is evaluated at $x = e^{-\kappa(T-t)}$, which is strictly less than one for $0 \leq t < T$. Thus, higher order terms become less and less significant. In Figure 1, an example of different yield curves has been plotted for $N = 1, 2, 3, \ldots$. We note that (for these parameter values) the summation may be truncated at $N = 2$, which then motivates the efficiency of the series expansion.

We now turn to examining the term structure. In Figure 2 the short rate has been altered. In Figure 3 the value of θ has been altered; we see this effect for three different values of the short rate. Finally, in Figure 4, the volatility V of the short rate has been altered, and we see this effect for three different values of the short rate. In Figures 2–4, we see that the model can generate yield curves of many interesting shapes. Apart from the "usual" increasing, decreasing and humped yield curves, we also have examples of increasing-decreasing-increasing yield curves, and of increasing yield curves with inflection points. For the parameters chosen above, the following main conclusions can be drawn from Figures 2–4:

- a change in r has a large effect on short-term yields, and then the effect decays rapidly although a change in r has a small effect on long-term yields up to at least 30 years

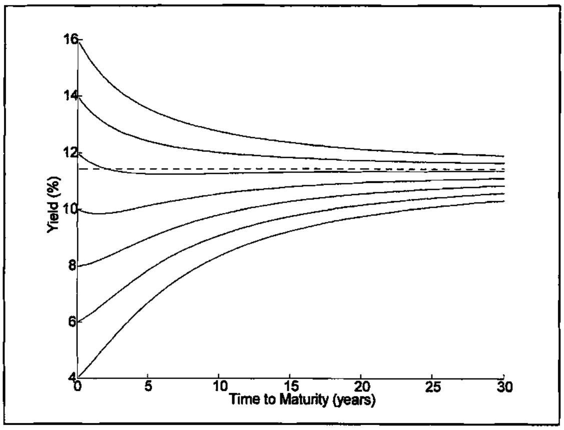

FIGURE 2. Yield Curves for Decreasing r (top-down)
Yield curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all yield curves, $\theta = 2\%$ and $V = 0.08\%$. The dashed line represents the long-run yield of 11.43%.

- a change in θ has practically no effect on very short-term yields, then the effect grows slowly and reaches a maximum for intermediate rates (around 5 years), after which the effect decays slowly although a change in θ has a small effect on long-term rates up to at least 30 years

- the effect of a change in V is practically zero for very short-term yields, and is biggest for intermediate rates (around 10 years), after which it decays slowly, displaying a significant effect on long-term maturities up to at least 30 years.

The effects that result from changes in the factors can be understood better through a careful analysis of the expression of the yields in (21). For fixed T the yield $Y(t,T)$ is a function of the three factors r, θ and V, and by making a first order Taylor expansion of the real valued function $Y(t,T,r,\theta,V)$, we obtain

$$dY(t,T,r,\theta,V) = \bar{B}(t,T)dr + \bar{C}(t,T)d\theta + \bar{D}(t,T)dV, \tag{23}$$

where

$$\bar{B}(t,T) = \frac{B(t,T)}{T-t} \tag{24}$$

$$\bar{C}(t,T) = \frac{C(t,T)}{T-t} \tag{25}$$

$$\bar{D}(t,T) = \frac{D(t,T)}{T-t}. \tag{26}$$

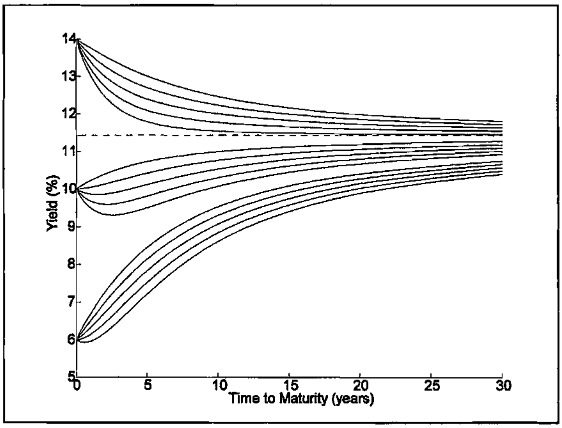

FIGURE 3. Yield Curves for Decreasing θ (top-down)
Yield curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3\%$ (top), $\theta = 2.5\%$, $\theta = 2\%$, $\theta = 1.5\%$ and $\theta = 1\%$ (down). For all yield curves, $V = 0.08\%$. The dashed line represents the long-run yield of 11.43%.

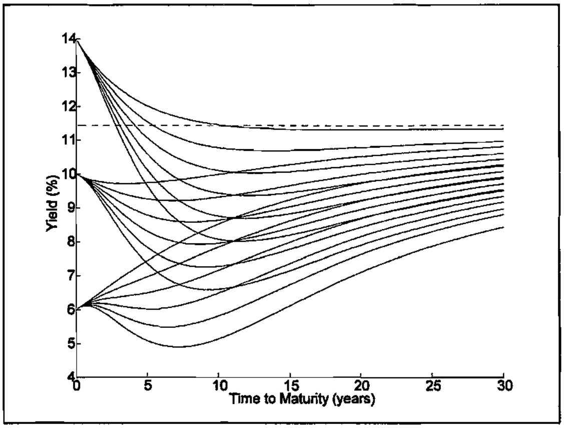

FIGURE 4. Yield Curves for Increasing V (top-down)
Yield curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $V = 0.42\%$ (top), $V = 0.84\%$, $V = 1.26\%$, $V = 1.68\%$, $V = 2.10\%$ and $V = 2.52\%$ (down). For all yield curves, $\theta = 2\%$. The dashed line represents the long-run yield of 11.43%.

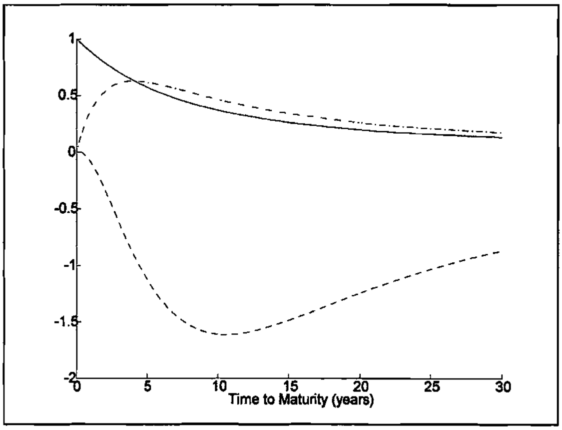

FIGURE 5. Changes in the Yield Curve with Respect to the Factors
The functions $\bar{B}(t,T)$ (solid), $\bar{C}(t,T)$ (dash-dotted) and $\bar{D}(t,T)$ (dashed) for different maturities.

The functions $\bar{B}, \bar{C}$ and $\bar{D}$ thus determine the effect that a change in r, θ or V will have on the yields for different maturities. These functions are shown in Figure 5 (for the same parameter values as above). A common way to describe the yield curve is in terms of the three components level, steepness and curvature of the yield curve. For yield curves where r is close to the long-run yield, r can easily be interpreted as the level of the yield curve. Assuming that r has the same value as the long-run yield, yield curves have been plotted for varying θ (Figure 6) and V (Figure 7). Figure 6 then suggests that θ can be interpreted as the steepness, i.e., it determines the initial slope of the yield curve. From Figure 7, it follows that V can be interpreted as the curvature, i.e., it determines the curvature of the yield curve for intermediate rates.

5 Forward Rates

The forward rates are defined as $f(t,T) = -\frac{\partial}{\partial T}\log P(t,T)$, and with discount bond prices as in (5), and using that we have a time homogenous model, we obtain

$$f(t,T) = A_t(t,T) - B_t(t,T)r - C_t(t,T)\theta - D_t(t,T)V. \tag{27}$$

It is straightforward to verify that

$$\lim_{T\to\infty} f(t,T) = \frac{\beta}{\alpha\kappa} - \frac{\gamma^2}{2\alpha^2\kappa^2} - \frac{2b\kappa\varepsilon}{\sigma^2}, \tag{28}$$

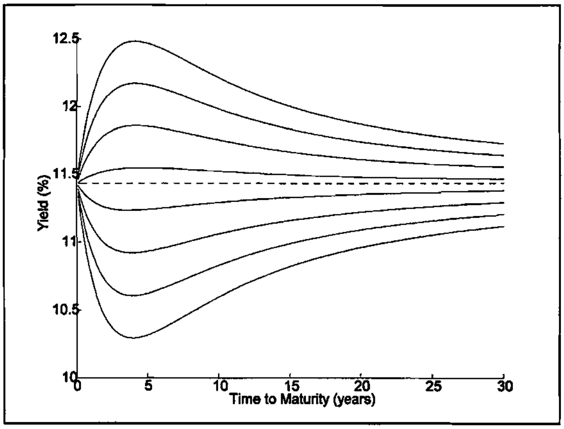

FIGURE 6. Yield Curves for Decreasing θ (top-down)
Yield curves for $\theta = 4.5\%$ (top), $\theta = 4\%$, $\theta = 3.5\%$, $\theta = 3\%$, $\theta = 2.5\%$, $\theta = 2\%$, $\theta = 1.5\%$ and $\theta = 1\%$ (down). For all yield curves, $r = 11.43\%$ and $V = 0.08\%$. The dashed line represents the long-run yield of 11.43%.

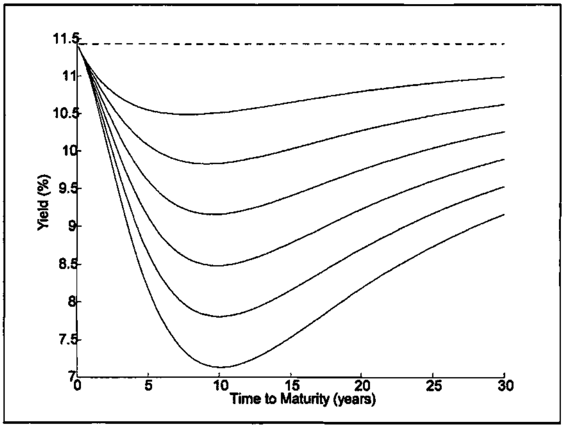

FIGURE 7. Yield Curves for Increasing V (top-down)
Yield curves for $V = 0.42\%$ (top), $V = 0.84\%$, $V = 1.26\%$, $V = 1.68\%$, $V = 2.10\%$ and $V = 2.52\%$ (down). For all yield curves, $r = 11.43\%$ and $\theta = 2\%$. The dashed line represents the long-run yield of 11.43%.

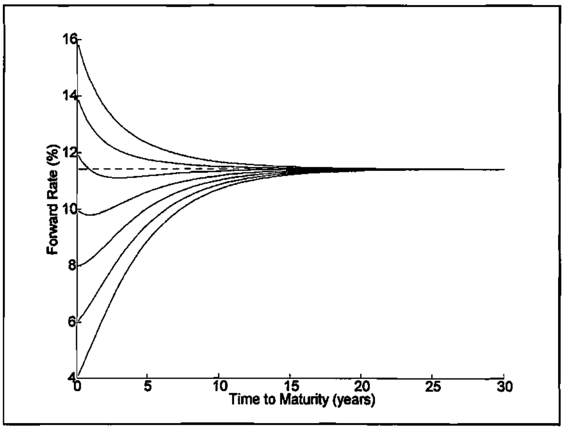

FIGURE 8. Forward Rate Curves for Decreasing r (top-down)
Forward rate curves for $r = 16\%$ (top), $r = 14\%$, $r = 12\%$, $r = 10\%$, $r = 8\%$, $r = 6\%$ and $r = 4\%$ (down). For all forward rate curves, $\theta = 2\%$ and $V = 0.08\%$. The dashed line represents the long-run forward rate of 11.43%.

i.e., the forward rates converge to the same level as the yields as $T \to \infty$. Furthermore, we have that

$$\begin{aligned} f_r(t,T) &= -B_t(t,T) < 0 \\ f_\theta(t,T) &= -C_t(t,T) < 0, \end{aligned}$$

and $f_V(t,T) = -D_t(t,T)$, but the sign of this partial derivative cannot be decided. Using (7)–(10) we obtain

$$\begin{aligned} f(t,T) &= \beta C(t,T) + bD(t,T) - \frac{1}{2}\gamma^2 C^2(t,T) \\ &\quad -(\kappa B(t,T) - 1)r - (\alpha C(t,T) - B(t,T))\theta \\ &\quad - \left((a + \rho\sigma B(t,T))\, D(t,T) + \frac{1}{2}\sigma^2 D^2(t,T) + \lambda B(t,T) + \frac{1}{2}B^2(t,T) \right) V. \end{aligned}$$

Using the same parameter values as in the previous section, different forward rate curves have been plotted in Figure 8–10. In Figure 8 the short rate has been altered. In Figure 9 the value of θ has been altered, and this effect is shown for three different values of the short rate. Finally, in Figure 10, the volatility V of the short rate has been altered, and this effect is shown for three different values of the short rate. We see examples of forward rate curves of many different shapes, similar to those of the yield curves in Figures 2–4.

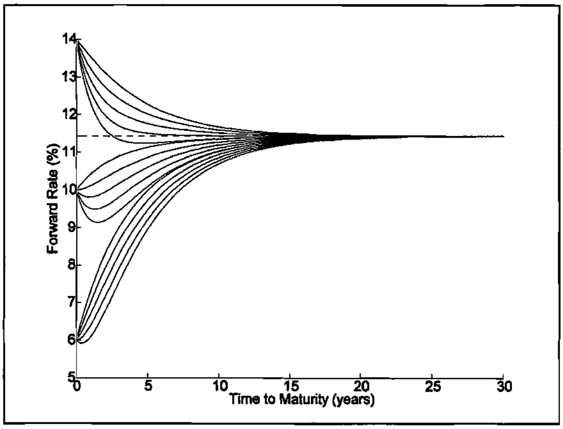

FIGURE 9. Forward Rate Curves for Decreasing θ (top-down)
Forward rate curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $\theta = 3\%$ (top), $\theta = 2.5\%$, $\theta = 2\%$, $\theta = 1.5\%$ and $\theta = 1\%$ (down). For all forward rate curves, $V = 0.08\%$. The dashed line represents the long-run forward rate of 11.43%.

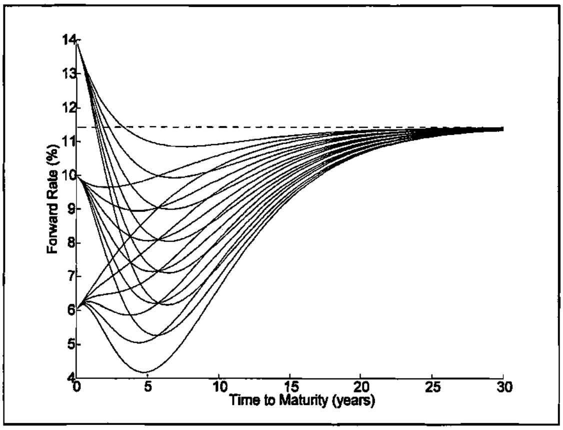

FIGURE 10. Forward Rate Curves for Increasing V (top-down)
Forward rate curves for $r = 14\%$ (top), $r = 10\%$ (center) and $r = 6\%$ (down). For each value of r, $V = 0.42\%$ (top), $V = 0.84\%$, $V = 1.26\%$, $V = 1.68\%$, $V = 2.10\%$ and $V = 2.52\%$ (down). For all forward rate curves, $\theta = 2\%$. The dashed line represents the long-run forward rate of 11.43%.

6 Bond Prices as Functions of Benchmark Forward Rates

This section demonstrates that it is possible to express discount bond prices in the model above in terms of benchmark forward rates. Let $f(t, S_i), i = 1, 2, 3$ denote benchmark forward rates with maturity at S_i. Then, from (27) we have for $i = 1, 2, 3$ that

$$f(t, S_i) = -A_T(t, S_i) + B_T(t, S_i)r + C_T(t, S_i)\theta + D_T(t, S_i)V \tag{29}$$

or, in matrix notation:

$$\begin{bmatrix} B_T(t,S_1) & C_T(t,S_1) & D_T(t,S_1) \\ B_T(t,S_2) & C_T(t,S_2) & D_T(t,S_2) \\ B_T(t,S_3) & C_T(t,S_3) & D_T(t,S_3) \end{bmatrix} \begin{bmatrix} r \\ \theta \\ V \end{bmatrix} = \begin{bmatrix} A_T(t,S_1) + f(t,S_1) \\ A_T(t,S_2) + f(t,S_2) \\ A_T(t,S_3) + f(t,S_3) \end{bmatrix}. \tag{30}$$

Provided the matrix in (30) is invertible, the factors r, θ and V may be expressed as linear combinations of the benchmark forward rates defined above:

$$X = \Lambda_X^{(1)} f(t, S_1) + \Lambda_X^{(2)} f(t, S_2) + \Lambda_X^{(3)} f(t, S_3),$$

where $X = r, \theta$ or V and $\Lambda_X^{(i)} = \Lambda_X^{(i)}(t, S_1, S_2, S_3)$ $(i = 1, 2, 3)$ is some function. If these expressions are substituted into (5), discount bond prices can be written as

$$P(t,T) = \exp\left(A(t,T) - \Omega^{(1)} f(t,S_1) - \Omega^{(2)} f(t,S_2) - \Omega^{(3)} f(t,S_3)\right),$$

where $\Omega^{(i)} = \Omega^{(i)}(t, T, S_1, S_2, S_3)$ and

$$\Omega^{(i)} = B(t,T)\Lambda_r^{(i)} + C(t,T)\Lambda_\theta^{(i)} + D(t,T)\Lambda_V^{(i)}, \qquad i = 1, 2, 3.$$

Provided that the matrix in (30) is invertible, it is thus possible to replace our original factors r, θ and V by three benchmark forward rates with arbitrary maturities. Due to the linear relation between forward rates and the factors in (29), this is to be expected. See Björk and Svensson [3] for more general findings regarding finite dimensional realizations of nonlinear forward rate models in terms of benchmark forward rates. In principal, the above program can also be carried out for benchmark yields, i.e., discount bond prices can be expressed in terms of three benchmark yields. See Duffie and Kan [5] for a discussion of this kind of variable transformation.

7 Inverting the Yield Curve

The above model can be extended so that the term structure can be fitted exactly to an arbitrary initially observed yield curve. This is sometimes referred to as inverting the yield curve. The standard procedure is to introduce a deterministic (time dependent) function $\varphi(t)$ in the drift of the short rate, and a result from Björk and Hyll allows the fitted term structure to be expressed express in terms of discount bond prices in a special case of the model above.

It is now assumed that the dynamics of the factors r_t^φ, θ_t and V_t under a martingale measure Q are given by

$$\begin{aligned} dr_t^\varphi &= (\varphi(t) + \theta_t - \kappa r_t^\varphi - \lambda V_t)dt + \sqrt{V_t} W_t^1 \\ d\theta_t &= -\alpha\theta_t dt + \gamma dW_t^2 \\ dV_t &= (b - aV_t)dt + \sigma\sqrt{V_t} dW_t^3, \end{aligned}$$

where again $\kappa, \lambda, \alpha, \gamma, b, a$ and σ are constants and W_t^1, W_t^2 and W_t^3 are Wiener processes satisfying $dW_t^1 dW_t^2 = dW_t^2 dW_t^3 = 0$ and $dW_t^1 dW_t^3 = \rho dt$. We now interpret θ_t as a stochastic disturbance term affecting the drift of the short rate satisfying $\theta_0 = 0$. As before, V_t is the variance of the short rate. We note that we still have an affine term structure, and let $P^\varphi(t,T)$ denote the price at time t of a discount bond with unit payment at T.

Let $P^0(t,T)$ denote the price at time t of a discount bond in the special case when $\beta = 0$ in the model in section 2–3, and let the corresponding forward rates are denoted by $f^0(t,T)$. Furthermore, for all $t \geq 0$, $P^*(0,t)$ and $f^*(0,t)$ are the initially observed yield curve and the corresponding forward rate curve. We then have the following proposition:[4]

Proposition 2 *The theoretical term structure implied by the model in this section can be fitted exactly to an arbitrary initially observed yield curve. A perfect initial fit is obtained by choosing*

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + \kappa f^*(0,t) - \kappa f^0(0,t),$$

and discount bonds are then given by

$$P^\varphi(t,T) = \frac{P^*(0,T)}{P^*(0,t)} \frac{P^0(0,t)}{P^0(0,T)} P^0(t,T) \exp\left(\left(f^*(0,t) - f^0(0,t)\right) B(t,T)\right). \tag{31}$$

Proof. Denote by r_s^0 the short rate in the special case of the model in section 2 when $\beta = 0$. Note that (see [2] for details) $r_s^\varphi = r_s^0 + \chi(t,s)$, $s \geq t$, where $\chi(t,s) = \int_t^s e^{-\kappa(s-u)}\varphi(u)du$.

[4] For a more general pricing formula, see Björk and Hyll [2].

Discount bond prices are given by

$$P^{\varphi}(t,T) = E^{Q}_{t,r,\theta,V}\left[e^{-\int_t^T r_t^{\varphi} du}\right],$$

and thus

$$P^{\varphi}(t,T) = e^{-\int_t^T \chi(t,u)du} E^{Q}_{t,r,\theta,V}\left[e^{-\int_t^T r_u^0 du}\right], \tag{32}$$

where the short hand notation used is $E^{Q}_{t,r,\theta,V}[\cdot] = E^{Q}[\cdot \mid r_t = r, \theta_t = \theta, V_t = V]$. Hence the expectation in (32) can be interpreted as a discount bond price in the special case when $\beta = 0$ in the model in section 2, and we thus have

$$P^{\varphi}(t,T) = e^{-\int_t^T \chi(t,u)du} P^0(t,T). \tag{33}$$

For $t = 0$ we get

$$P^{\varphi}(0,T) = e^{-\int_0^T \chi(0,u)du} P^0(0,T). \tag{34}$$

Now take the logarithm of (34), differentiate with respect to T and use the definitions of the forward rates to obtain

$$f^{\varphi}(0,T) = \chi(0,T) + f^0(0,T). \tag{35}$$

We now want to choose φ such that $f^{\varphi}(0,T) = f^*(0,T)$ for all T. We then have

$$f^*(0,T) - f^0(0,T) = \int_0^T e^{-\kappa(T-u)}\varphi(u)du. \tag{36}$$

Differentiating (36) with respect to T, we get after some simplifications

$$\varphi(t) = f_T^*(0,t) - f_T^0(0,t) + \kappa f^*(0,t) - \kappa f^0(0,t). \tag{37}$$

Finally, we obtain

$$\begin{aligned} -\int_t^T \chi(t,s)ds &= -\int_t^T \int_t^s e^{-\kappa(s-u)}\varphi(u)duds \\ &= \log \frac{P^*(0,T)}{P^*(0,t)} \frac{P^0(0,t)}{P^0(0,T)} + \left(f^*(0,t) - f^0(0,t)\right) B(t,T), \end{aligned} \tag{38}$$

which, with (33), proves (31). ■

8 Summary and Conclusions

This essay presents a solution to discount bond prices in the three-factor model suggested by Balduzzi *et al.* [1] by using a series expansion. This technique was first used by Selby and Strickland [10] to simplify the implementation of the Fong and Vasicek model [6]. The BDFS model is also extended so that the theoretical term structure can be fitted exactly to an arbitrary initially observed yield curve.

Furthermore, it is shown that the BDFS model can be modified so that discount bond prices can be obtained in terms of three (arbitrarily chosen) benchmark forward rates.

Several examples are shown of term structures and forward rate curves that display great flexibility and a variety of realistic features.

What remains to be determined are prices of other derivatives, e.g., the call option on a discount bond. Given the complexity of the PDE, it seems reasonable that other derivative prices should be calculated using Monte Carlo simulation.

Appendix A
Method of Frobenius

The following section describes a procedure known as the method of Frobenius used to obtain a solution to a certain type of linear differential equations. Coddington [4] is followed, and a brief summary of the method is presented. The aim here is to explain how the solution can be found; not to give a formal proof.

Suppose that we have a second order linear differential equation with a regular singular point at $x_0 = 0$, i.e.,

$$L(Q)(x) = x^2 Q''(x) + \tilde{a}(x) x Q'(x) + \tilde{b}(x) Q(x) = 0, \tag{39}$$

where $\tilde{a}(x)$ and $\tilde{b}(x)$ are analytic at $x_0 = 0$. Since $\tilde{a}(x)$ and $\tilde{b}(x)$ are analytic at $x_0 = 0$, they can be expressed in power series expansions

$$\begin{aligned} \tilde{a}(x) &= \sum_{k=0}^{\infty} \tilde{\alpha}_k x^k \\ \tilde{b}(x) &= \sum_{k=0}^{\infty} \tilde{\beta}_k x^k, \end{aligned}$$

which are convergent on an interval $| x |< R$ for some $R > 0$. To find a solution to (39), try an *Anzats* of the form

$$Q(x) = \sum_{k=0}^{\infty} c_k x^{k+q}, \tag{40}$$

where $c_0 \neq 0$ and q is a constant to be determined. Then

$$\begin{aligned} Q'(x) &= \sum_{k=0}^{\infty} (k+q) c_k x^{k+q-1} \\ Q''(x) &= \sum_{k=0}^{\infty} (k+q)(k+q-1) c_k x^{k+q-2}, \end{aligned}$$

and hence

$$\begin{aligned} \tilde{b}(x) Q(x) &= x^q \sum_{k=0}^{\infty} \tilde{\beta}_k x^k \sum_{k=0}^{\infty} c_k x^k \\ &= x^q \sum_{k=0}^{\infty} \sum_{j=0}^{k} c_j \tilde{\beta}_{k-j} x^k, \end{aligned}$$

$$\begin{aligned} x\tilde{a}(x)Q'(x) &= x^q \sum_{k=0}^{\infty} \tilde{\alpha}_k x^k \sum_{k=0}^{\infty} c_k (k+q) x^k \\ &= x^q \sum_{k=0}^{\infty} \sum_{j=0}^{k} (j+q) c_j \tilde{\alpha}_{k-j} x^k, \end{aligned}$$

and

$$x^2 Q''(x) = x^q \sum_{k=0}^{\infty} (k+q)(k+q-1) c_k x^k.$$

Thus

$$\begin{aligned} L(Q)(x) &= x^q \sum_{k=0}^{\infty} \left\{ (k+q)(k+q-1)c_k + \sum_{j=0}^{k} \left((j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right) c_j \right\} x^k \\ &= 0. \end{aligned} \tag{41}$$

Relation (41) must hold for all values of x, and hence each coefficient must be identically zero. Since $c_0 \neq 0$, we have for $k = 0$

$$p(q) = q(q-1) + q\tilde{\alpha}_0 + \tilde{\beta}_0 = 0,$$

where $p(q)$ is called the *indicial polynomial* for (39), and the only admissible values of q are the roots of $p(q)$. Furthermore, for $k = 1, 2, \ldots,$

$$p(q+k)c_k + \sum_{j=0}^{k-1} \left\{ (j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j = 0,$$

or

$$c_k = -\frac{\sum_{j=0}^{k-1} \left\{ (j+q)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j}{p(q+k)}, \tag{42}$$

where c_k is well defined except at the two roots of $p(q+k)$. Now define

$$U(x,q) = c_0 x^q + x^q \sum_{k=1}^{\infty} c_k(q) x^k. \tag{43}$$

If the series in (43) converges for $0 < x < R$, then clearly

$$L(U)(x,q) = c_0 p(q) x^q,$$

and hence if the $Q(x)$ given by (40) is a solution of (39), then q must be a root of the indicial polynomial p, and the coefficients c_k $(k \geq 1)$ are determined uniquely in terms of q and c_0 by (42), provided $p(q+k) \neq 0, k = 1, 2, \ldots.$ Conversely, if q is a root of p, and if the coefficients c_k can be determined, i.e., $p(q+k) \neq 0, k = 1, 2, \ldots,$ then the function Q given by $Q(x) = U(x, q)$ is a solution of (39) for any choice of c_0, provided the series in (43) converges.

In general, the nature of the solution of (39) depends on the roots of p. The only case of interest here is when the roots of p are real and distinct, and we shall assume from now on that this is the case. Suppose, therefore, that the roots of p are q_1 and q_2, such that $q_1 \neq q_2 + k$ for $k = 0, 1, 2, \ldots.$ Then $p(q_1) = 0$, and no k can exist such that $p(q_1 + k) = 0$. Hence the coefficients in (42) exist for all $k = 1, 2, \ldots,$ and letting $c_0 = 1$ we see that the function Q_1 given by

$$\begin{aligned} Q_1 &= U(x, q_1) \\ &= x^{q_1} \sum_{k=0}^{\infty} c_k(q_1) x^k \end{aligned}$$

is a solution of (39), since it can be shown that it converges for $\mid x \mid < R$. Using the same arguments, the function Q_2 given by

$$\begin{aligned} Q_2 &= U(x, q_2) \\ &= x^{q_2} \sum_{k=0}^{\infty} c_k(q_2) x^k \end{aligned}$$

is also a solution of (39).

It can be shown that the solutions Q_1 and Q_2 are independent, and hence the most general solution of (39) on the interval $0 <\mid x \mid< R$ is given by a linear combination of the form

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2), \tag{44}$$

where K_1 and K_2 are suitable constants, and where the coefficients $c_k^{(i)}$ $(i = 1, 2$ and $k = 1, 2, \ldots)$ of $U(x, q_i)$ are given by

$$c_k^{(i)} = -\frac{\sum_{j=0}^{k-1} \left\{ (j + q_i)\tilde{\alpha}_{k-j} + \tilde{\beta}_{k-j} \right\} c_j^{(i)}}{p(q_i + k)},$$

where $c_0^{(1)} = c_0^{(2)} = 1$.

Remark 3 *Note that if the roots to $p(q)$ are nonnegative, then $Q(x)$ defined in (44) is well defined at $x = 0$.*

Appendix B
The Functions A(t,T), B(t,T), C(t,T) and D(t,T)

The Functions $B(t,T)$ and $C(t,T)$

The ordinary differential equation and terminal condition for $B(t,T)$ is

$$\begin{aligned} B_t(t,T) &= \kappa B(t,T) - 1 \\ B(T,T) &= 0. \end{aligned}$$

The solution is

$$B(t,T) = \frac{1 - e^{-\kappa(T-t)}}{\kappa},$$

and it is easily verified that $B(t,T) \geq 0$, $B_t(t,T) \leq 0$ and that $\lim_{T\to\infty} B(t,T) = \frac{1}{\kappa}$.

The ordinary differential equation and terminal condition for $C(t,T)$ is

$$\begin{aligned} C_t(t,T) &= \alpha C(t,T) - B(t,T) \\ C(T,T) &= 0. \end{aligned}$$

Multiplying with the integrating factor $e^{-\alpha t}$, we have

$$\frac{\partial}{\partial t}\left(e^{-\alpha t} C(t,T)\right) = -e^{-\alpha t} B(t,T).$$

Integrating from t to T, we obtain

$$C(t,T) = \frac{\alpha\left(1 - e^{-\kappa(T-t)}\right) - \kappa\left(1 - e^{-\alpha(T-t)}\right)}{\alpha\kappa(\alpha - \kappa)}.$$

Now

$$C_T(t,T) = \frac{e^{-\kappa(T-t)} - e^{-\alpha(T-t)}}{\alpha - \kappa} \geq 0,$$

and hence $C(t,T) \geq 0$. Furthermore, $\lim_{T\to\infty} C(t,T) = \frac{1}{\alpha\kappa}$.

The Function $D(t,T)$

The ordinary differential equation and terminal condition for $D(t,T)$ is

$$\begin{aligned} D_t(t,T) &= (a+\rho\sigma B(t,T))\,D(t,T)+\frac{1}{2}\sigma^2 D^2(t,T)+\lambda B(t,T)+\frac{1}{2}B^2(t,T) \\ D(T,T) &= 0. \end{aligned} \tag{45}$$

This is a Ricatti equation, and the standard substitution is

$$D(t,T) = -\frac{2}{\sigma^2}\frac{w_t(t,T)}{w(t,T)}. \tag{46}$$

Then

$$D_t = -\frac{2}{\sigma^2}\left(\frac{ww_{tt}-w_t^2}{w^2}\right), \tag{47}$$

and substituting (46) and (47) into (45), we obtain

$$w_{tt} - (a+\rho\sigma B)\,w_t + \frac{\sigma^2}{2}\left(\lambda B + \frac{1}{2}B^2\right) w = 0. \tag{48}$$

Now we make a change of variables according to

$$w(t,T) = x^{\varepsilon}Q(x), \tag{49}$$

where $x = e^{-\kappa(T-t)}$ and ε is some constant to be determined. We note that $\frac{dx}{dt} = \kappa x$ and hence

$$w_t = \varepsilon\kappa x^{\varepsilon}Q(x) + \kappa x^{\varepsilon+1}Q'(x), \tag{50}$$

and

$$w_{tt} = \varepsilon^2\kappa^2 x^{\varepsilon}Q(x) + \kappa^2(2\varepsilon+1)\,x^{\varepsilon+1}Q'(x) + \kappa^2 x^{\varepsilon+2}Q''(x). \tag{51}$$

Substituting (49), (50) and (51) into (48) and simplifying, we obtain

$$\begin{aligned} &\kappa^2x^2Q''(x) + \left(\kappa^2(1+2\varepsilon) - a\kappa - \rho\sigma + \rho\sigma x\right)xQ'(x) \\ &+ \left(\varepsilon^2\kappa^2 - \varepsilon(a\kappa+\rho\sigma) + \frac{\sigma^2}{2\kappa}\left(\lambda + \frac{1}{2\kappa}\right) + \left(\varepsilon\rho\sigma - \frac{\lambda\sigma^2}{2\kappa} - \frac{\sigma^2}{2\kappa^2}\right)x + \frac{\sigma^2}{4\kappa^2}x^2\right)Q(x) \\ =\ & 0. \end{aligned} \tag{52}$$

Now we choose ε to make the constant term in the coefficient of $Q(x)$ equal to zero. This is the case if

$$\varepsilon^2\kappa^2 - \varepsilon\left(a\kappa+\rho\sigma\right) + \frac{\sigma^2}{2\kappa}\left(\lambda+\frac{1}{2\kappa}\right) = 0,$$

or

$$\varepsilon = \frac{a\kappa+\rho\sigma}{2\kappa^2} \pm \sqrt{\left(\frac{a\kappa+\rho\sigma}{2\kappa^2}\right)^2 - \frac{\sigma^2}{2\kappa^3}\left(\lambda+\frac{1}{2\kappa}\right)}, \tag{53}$$

and it is assumed that $\left(\frac{a\kappa+\rho\sigma}{2\kappa^2}\right)^2 - \frac{\sigma^2}{2\kappa^3}\left(\lambda+\frac{1}{2\kappa}\right) \geq 0$ so that ε is real. With ε as in (53), the differential equation (52) takes the form

$$x^2Q''(x) + \left(1+2\varepsilon-\frac{a\kappa+\rho\sigma}{\kappa^2}+\frac{\rho\sigma}{\kappa^2}x\right)xQ'(x) + \left(\left(\frac{\varepsilon\rho\sigma}{\kappa^2}-\frac{\sigma^2\left(1+\lambda\kappa\right)}{2\kappa^4}\right)x+\frac{\sigma^2}{4\kappa^4}x^2\right)Q(x) = 0.$$

Now define

$$\tilde{a}(x) = \sum_{k=0}^{\infty}\tilde{\alpha}_k x^k,$$

where

$$\begin{aligned}
\tilde{\alpha}_0 &= 1+2\varepsilon-\frac{a\kappa+\rho\sigma}{\kappa^2}\\
\tilde{\alpha}_1 &= \frac{\rho\sigma}{\kappa^2}\\
\tilde{\alpha}_k &= 0,\ k=2,3,...,
\end{aligned}$$

and

$$\tilde{b}(x) = \sum_{k=0}^{\infty}\tilde{\beta}_k x^k,$$

where

$$\begin{aligned}
\tilde{\beta}_0 &= 0\\
\tilde{\beta}_1 &= \frac{\varepsilon\rho\sigma}{\kappa^2}-\frac{\sigma^2\left(1+\lambda\kappa\right)}{2\kappa^4}\\
\tilde{\beta}_2 &= \frac{\sigma^2}{4\kappa^4}\\
\tilde{\beta}_k &= 0,\ k=3,4,....
\end{aligned}$$

Note that $\tilde{a}(x)$ and $\tilde{b}(x)$ are convergent for $0 \leq x \leq 1$. Then the ODE for $Q(x)$ becomes

$$x^2 Q''(x) + \tilde{a}(x) x Q'(x) + \tilde{b}(x) Q(x) = 0, \tag{54}$$

and from Appendix A we know that the solution to (54) is given by

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2) \tag{55}$$

where K_1 and K_2 are suitable constants and where $U(x, q_i)$ $(i = 1, 2)$ is defined in Appendix A. To determine the terminal condition for $Q(x)$, we use that

$$\begin{aligned} D(t,T) &= -\frac{2}{\sigma^2} \frac{w_t(t,T)}{w(t,T)} \\ &= -\frac{2\kappa}{\sigma^2} \frac{\varepsilon Q(x) + x Q'(x)}{Q(x)} \\ &= -\frac{2\kappa}{\sigma^2} \left(\varepsilon + \frac{x Q'(x)}{Q(x)} \right). \end{aligned}$$

Since $D(T,T) = 0$, we get (remember that $x = e^{-\kappa(T-t)}$)

$$\frac{Q'(1)}{Q(1)} + \varepsilon = 0. \tag{56}$$

Using (55) and (56), we have that

$$\frac{K_1 U'(1, q_1) + K_2 U'(1, q_2)}{K_1 U(1, q_1) + K_2 U(1, q_2)} + \varepsilon = 0,$$

or

$$\begin{aligned} K^* &= \frac{K_2}{K_1} \\ &= -\frac{U'(1, q_1) + \varepsilon U(1, q_1)}{U'(1, q_2) + \varepsilon U(1, q_2)}. \end{aligned}$$

Hence the function $D(t,T)$ is given by

$$D(t,T) = -\frac{2\kappa}{\sigma^2} \left(\varepsilon + \frac{e^{-\kappa(T-t)} Q'(e^{-\kappa(T-t)})}{Q(e^{-\kappa(T-t)})} \right),$$

where

$$Q(x) = K_1 U(x, q_1) + K_2 U(x, q_2)$$

and

$$K^* = \frac{K_2}{K_1} = -\frac{U'(1,q_1)+\varepsilon U(1,q_1)}{U'(1,q_2)+\varepsilon U(1,q_2)}.$$

Before finding the limit for $D(t,T)$, we recall that ε is given by $\varepsilon^{\pm} = \frac{1}{2}\chi \pm \frac{1}{2}\varphi$, where $\chi = \frac{a\kappa+\rho\sigma}{\kappa^2}$ and $\varphi = \sqrt{\chi^2 - \frac{2\sigma^2}{\kappa^3}\left(\lambda+\frac{1}{2\kappa}\right)} > 0$. By construction, $D(t,T)$ does not depend on which ε we choose. For numerical reasons, $\varepsilon = \varepsilon^- = \frac{1}{2}\chi - \frac{1}{2}\varphi$ is chosen, since then we have $q_2 = \varphi > 0$ as it is assumed that $\left(\frac{a\kappa+\rho\sigma}{2\kappa^2}\right)^2 - \frac{\sigma^2}{2\kappa^3}\left(\lambda+\frac{1}{2\kappa}\right) \geq 0$. It then follows that $Q(x)$ is well defined for $x=0$. We now examine the limit for $D(t,T)$, and we have

$$\begin{aligned}\lim_{T\to\infty} D(t,T) &= \lim_{x\to 0} -\frac{2\kappa}{\sigma^2}\left(\varepsilon + \frac{xQ'(x)}{Q(x)}\right)\\ &= -\frac{2\kappa\varepsilon}{\sigma^2} - \frac{2\kappa}{\sigma^2}\lim_{x\to 0}\frac{xQ'(x)}{Q(x)},\end{aligned}$$

so the relevant limit to investigate is

$$\lim_{x\to 0}\frac{xQ'(x)}{Q(x)} = \lim_{x\to 0}\frac{\sum_{n=0}^{\infty} c_n^{(1)} n x^n + K^* \sum_{n=0}^{\infty} c_n^{(2)}(n+q_2)x^{n+q_2}}{\sum_{n=0}^{\infty} c_n^{(1)} x^n + K^* \sum_{n=0}^{\infty} c_n^{(2)} x^{n+q_2}}.$$

Now $\varepsilon = \varepsilon^-$ and $q_2 = \varphi > 0$, and hence

$$\begin{aligned}\lim_{x\to 0}\frac{xQ'(x)}{Q(x)} &= \frac{0}{c_0^{(1)}}\\ &= 0,\end{aligned}$$

since it is assumed that $c_0^{(1)} = 1$. Thus

$$\lim_{T\to\infty} D(t,T) = -\frac{2\kappa\varepsilon^-}{\sigma^2}.$$

The Function $A(t,T)$

The ordinary differential equation and terminal condition for $A(t,T)$ is

$$\begin{aligned}A_t(t,T) &= \beta C(t,T) + bD(t,T) - \frac{1}{2}\gamma^2 C^2(t,T)\\ A(T,T) &= 0.\end{aligned}$$

Integrate and use the terminal condition:

$$A(t,T) = -\beta \int_t^T C(s,T)ds + \frac{1}{2}\gamma^2 \int_t^T C^2(s,T)ds - b \int_t^T D(s,T)ds.$$

Routine integration gives

$$\int_t^T C(s,T)ds = \frac{T-t-B(t,T)}{\alpha\kappa} - \frac{1}{\alpha}C(t,T)$$

and

$$\begin{aligned}\int_t^T C^2(s,T)ds &= -\frac{1}{2\alpha}\left(C^2(t,T) + \frac{1}{(\alpha+\kappa)}\frac{1}{\kappa}B^2(t,T) + \frac{2}{(\alpha+\kappa)}B(t,T)C(t,T)\right) \\ &\quad + \frac{1}{\alpha^2\kappa^2}(T-t) - \frac{1}{\alpha^2\kappa^2}B(t,T) - \frac{1}{\alpha^2(\alpha+\kappa)}C(t,T).\end{aligned}$$

To evaluate $\int_t^T D(s,T)ds$, we recall that

$$D(t,T) = -\frac{2}{\sigma^2}\frac{w_t(t,T)}{w(t,T)},$$

where $w(t) = x^\varepsilon Q(x)$ and $x = e^{-\kappa(T-t)}$. Then we get

$$\begin{aligned}\int_t^T D(s,T)ds &= -\frac{2}{\sigma^2}\int_t^T \frac{w_s(s)}{w(s)}ds \\ &= -\frac{2}{\sigma^2}\left[\log w(s)\right]_t^T \\ &= -\frac{2}{\sigma^2}\log\frac{w(T)}{w(t)} \\ &= -\frac{2}{\sigma^2}\log\frac{Q(1)}{x^\varepsilon Q(x)} \\ &= \frac{2}{\sigma^2}\log\frac{x^\varepsilon Q(x)}{Q(1)} \\ &= \frac{2}{\sigma^2}\log\frac{e^{-\varepsilon\kappa(T-t)}Q(e^{-\kappa(T-t)})}{Q(1)}.\end{aligned}$$

In summary, we obtain

$$
\begin{aligned}
A(t,T) &= \left(\frac{\beta}{\alpha\kappa}-\frac{\gamma^2}{2\alpha^2\kappa^2}\right)(B(t,T)-(T-t))+\left(\frac{\beta}{\alpha}-\frac{\gamma^2}{2\alpha^2(\alpha+\kappa)}\right)C(t,T)\\
&\quad -\frac{1}{4\alpha}\gamma^2\left(C^2(t,T)+\frac{1}{\kappa(\alpha+\kappa)}B^2(t,T)+\frac{2}{(\alpha+\kappa)}B(t,T)C(t,T)\right)\\
&\quad +\frac{2b}{\sigma^2}\log\frac{e^{\varepsilon\kappa(T-t)}Q(1)}{Q(e^{-\kappa(T-t)})}.
\end{aligned}
$$

Now we turn to finding the limit $\lim_{T\to\infty} A(t,T)$, and we note that since the corresponding limits for $B(t,T)$ and $C(t,T)$ exist, we need to consider the limit of the function $A^*(t,T)$ defined by

$$
\begin{aligned}
A^*(t,T) &= \frac{2b}{\sigma^2}\log\frac{e^{\varepsilon\kappa(T-t)}Q(1)}{Q(e^{-\kappa(T-t)})}-\left(\frac{\beta}{\alpha\kappa}-\frac{\gamma^2}{2\alpha^2\kappa^2}\right)(T-t)\\
&= \frac{2b}{\sigma^2}\log e^{\varepsilon\kappa(T-t)}+\frac{2b}{\sigma^2}\log\frac{Q(1)}{Q(e^{-\kappa(T-t)})}-\left(\frac{\beta}{\alpha\kappa}-\frac{\gamma^2}{2\alpha^2\kappa^2}\right)(T-t)\\
&= \frac{2b}{\sigma^2}\log\frac{Q(1)}{Q(e^{-\kappa(T-t)})}-\left(\left(\frac{\beta}{\alpha\kappa}-\frac{\gamma^2}{2\alpha^2\kappa^2}\right)-\frac{2b}{\sigma^2}\varepsilon\kappa\right)(T-t). \qquad (57)
\end{aligned}
$$

Since $Q(0)$ is well defined, we see that $A^*(t,T)$, and hence $A(t,T)$, tends to $-\infty$ if

$$
\frac{\beta}{\alpha\kappa}-\frac{\gamma^2}{2\alpha^2\kappa^2}>\frac{2b}{\sigma^2}\varepsilon\kappa,
$$

and this relation is satisfied by our assumption in (12). This assumption thus guarantees that discount bond prices tend to zero as T tends to infinity.

References

[1] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram (1996): "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income,* December, 43–53.

[2] Björk, T. and M. Hyll (2000): "On the Inversion of the Yield Curve," Working paper, Stockholm School of Economics.

[3] Björk, T. and L. Svensson (1999): "On the Existence of Finite Dimensional Realizations for Nonlinear Forward Rate Models," Working paper, Stockholm School of Economics.

[4] Coddington, E.A. (1961): "An Introduction to Ordinary Differential Equations," Prentice-Hall.

[5] Duffie, D. and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance* **6**(4), 379–406.

[6] Fong, H.G. and O.A. Vasicek (1991): "Interest Rate Volatility as a Stochastic Factor," Unpublished manuscript, Gifford Fong Associates.

[7] Fong, H.G., O.A. Vasicek, and D. Yoo (1992): "Omission Impossible," *Risk* **5**(2), 62–65.

[8] Heath, D., R. Jarrow, and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica* **60**(1), 77–105.

[9] Longstaff, F.A. and E.S. Schwartz (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *The Journal of Finance* **47**(4), 1259–1282.

[10] Selby, M. and C. Strickland (1995) "Computing the Fong and Vasicek Pure Discount Bond Price Formula," *The Journal of Fixed Income,* September, 78–84.

5

Quasi Arbitrage-Free Discount Bond Prices in the Cox, Ingersoll and Ross Model

ABSTRACT
In a stochastic short-rate model of the term structure of interest rates, the price of a discount bond with maturity T is given by the risk neutral expectation of a contingent claim with unit payment at time T. By using the Feynman-Kac representation, discount bond prices can also be obtained as the solution to a certain partial differential equation (PDE), usually referred to as the term structure equation. Regularity conditions on the solution to the PDE ensure that there is a unique solution identical to the one obtained from calculating the risk neutral expectation. This essay considers a less regular solution to the term structure equation in the Cox, Ingersoll and Ross (CIR) model. Under additional conditions, this solution remains non-negative and less than or equal to one for all maturities and positive values of the short rate, and is therefore referred to as a "quasi arbitrage-free term structure". Furthermore, the shape of the yield curves and forward rate curves are explored, and find that the quasi arbitrage-free term structure is much more flexible and displays more complex and realistic features than the classic CIR model does.

1 Introduction

From standard financial theory, given purely technical conditions, there exists a martingale measure Q such that the price at time $t \geq 0$ of a contingent claim with payoff X at $T \geq t$ is given by the expression

$$E^Q\left[e^{-\int_t^T r_u du} X \mid \mathcal{F}_t\right]. \tag{1}$$

This essay examines the term structure of interest rates, and the main asset under consideration is therefore the discount bond, with payoff function $X = 1$ at T. The price at time t of a discount bond with unit payment at T is denoted by $P(t,T)$. Assuming that the dynamics of the short rate under Q are given by

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t, \tag{2}$$

and given certain technical conditions on the functions $\mu(t,r)$ and $\sigma(t,r)$, the price at time t of a discount bond with unit payment at T is given by $P(t,T) = F(t, r_t; T)$, where $F(t,r;T)$ solves the so-called term structure equation[1]

$$F_t(t,r;T) + \mu(t,r)F_r(t,r;T) + \frac{1}{2}\sigma^2(t,r)F_{rr}(t,r;T) = rF(t,r;T) \tag{3}$$

$$F(T,r;T) = 1. \tag{4}$$

From now on, the T-variable in F will often be omitted. A direct computation of the expectation in (1) is often difficult, and therefore the PDE (3) must be solved analytically or using numerical methods. However, without further conditions on F, a solution to the PDE (3) satisfying (4) is generally not unique.[2] Thus, apart from the true arbitrage-free price given by (1), one would expect the existence of a variety of "wild" or less regular solutions. The question is whether any of these less regular solutions have some or any economic interpretation at all.

The purpose of this essay is to give an example of a solution to the term structure equation that does not correspond to the true arbitrage-free price obtained by (1), but that nevertheless shares at least some important features with an arbitrage-free bond market. In fact, in this essay the short rate is assumed to satisfy the stochastic differential equation proposed by Cox, Ingersoll, and Ross (CIR) [1], and a solution to the term structure equation is given different from the one obtained by CIR themselves. This less regular solution of the term structure equation is not compatible with an arbitrage-free bond market, and hence we may expect arbitrage opportunity to appear. Nevertheless, it possesses some important economic features, e.g., it tends to zero for $0 \leq t < T$ as $r \to \infty$. Moreover, if the short rate is positive, we will see

[1] See e.g. Duffie [3].

[2] See e.g. Fritz [4].

that with additional conditions, the less regular solution will be non-negative and less than or equal to one for all t, T such that $0 \leq t < T$, and can thus be interpreted as a term structure.

The rest of this essay is organized as follows: Section 2 defines the concepts of perturbation function and quasi arbitrage-free term structure, and a quasi arbitrage-free term structure is derived for the case when the short rate follows the dynamics proposed by CIR. Section 3 presents a proposition that ensures that the quasi arbitrage-free discount bond prices are non-negative and less than or equal to one for all maturities if r is positive. Various examples of the shapes of the quasi arbitrage-free term structure and forward rate curves are then given in section 4 and section 5. Section 6 summarizes and concludes.

2 Model

This section opens with the following definition:

Definition 1 *The function $\chi(t,T,r)$ is called a* ***perturbation function*** *if it satisfies the PDE (3), and, for all positive r, satisfies the additional conditions*

$$\lim_{T\to t}\chi(t,T,r) = 0 \tag{5}$$

$$\lim_{T\to\infty}\chi(t,T,r) = 0. \tag{6}$$

For simplicity, the r-variable in χ will often be omitted.

Let $P(t,T)$ denote the arbitrage-free price of a discount bond, i.e., the solution to (3) that corresponds to (1). With this definition of the perturbation function, the function $P^*(t,T) = P(t,T)+\chi(t,T,r)$ satisfies the PDE (3) with the correct terminal condition at T. Furthermore, as T tends to infinity or t, $\chi(t,T)$ by definition tends to zero. We can therefore expect $P^*(t,T)$ to be similar to the arbitrage-free term structure $P(t,T)$ for long and short maturities. Consequently, we expect the biggest differences between $P^*(t,T)$ and the arbitrage-free term structure $P(t,T)$ to appear for intermediate maturities.

A first natural question to ask, given the dynamics of the short rate in (2), is whether a function $\chi(t,T)$ exists as defined above. It cannot be assumed to exist for all (or any) dynamics of the short rate. The purpose of this essay is not to investigate under what conditions the function $\chi(t,T)$ may or may not exist, so therefore an example of where it exists is given. As an example, it assumed that the short rate under a martingale measure satisfy the stochastic differential equation

$$dr_t = a(b-r_t)dt + \sigma\sqrt{r_t}dW_t, \tag{7}$$

which is the short-rate process suggested by CIR [1]. The PDE (3) thus takes the form

$$F_t(t,r) + a(b-r)F_r(t,r) + \frac{1}{2}\sigma^2 r F_{rr}(t,r) = rF(t,r). \tag{8}$$

We then have the following lemma:

Lemma 1 *Define $C(t,T)$ and $D(t,T)$ by*

$$\begin{aligned} C(t,T) &= \frac{2ab}{\sigma^2}\log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1}\right) && (9)\\ D(t,T) &= \frac{2}{a+\gamma} + \frac{2\gamma}{\sigma^2\left(e^{\gamma(T-t)}-1\right)}, && (10) \end{aligned}$$

where

$$\gamma = \sqrt{a^2+2\sigma^2}. \tag{11}$$

For given fixed $r > 0$, define

$$\chi(t,T) = \alpha \exp\left(C(t,T) - D(t,T)r\right),$$

where α is a constant. Then $\chi(t,T)$ is a perturbation function.

Proof. Inserting $\chi(t,T)$ in the PDE (8), we find that $C(t,T)$ and $D(t,T)$ satisfy the system

$$\begin{aligned} C_t(t,T) &= abD(t,T) && (12)\\ D_t(t,T) &= aD(t,T) + \frac{1}{2}\sigma^2 D^2(t,T) - 1. && (13) \end{aligned}$$

It is straightforward to verify that $C(t,T)$ and $D(t,T)$ defined by (9)–(10) satisfy the system (12)–(13). Furthermore, it follows from Lemma 4 below that $\chi(t,T)$ satisfies $\lim_{T\to t}\chi(t,T) = 0$ and $\lim_{T\to\infty}\chi(t,T) = 0$ for all positive r.[3] ■

The true arbitrage-free price in the CIR model is given by

$$P(t,T) = \exp\left(A(t,T) - B(t,T)r\right),$$

[3] A complete formal solution to the system (12)–(13) is given in Appendix A.

where $A(t,T)$ and $B(t,T)$ satisfy the system

$$\begin{aligned} A_t(t,T) &= abB(t,T) && (14) \\ B_t(t,T) &= aB(t,T) + \frac{1}{2}\sigma^2 B^2(t,T) - 1, && (15) \end{aligned}$$

and terminal conditions $A(T,T) = B(T,T) = 0$. $A(t,T)$ and $B(t,T)$ are given by

$$\begin{aligned} A(t,T) &= \frac{2ab}{\sigma^2} \log \left(\frac{2\gamma e^{\frac{a+\gamma}{2}(T-t)}}{(a+\gamma)\left(e^{\gamma(T-t)} - 1\right) + 2\gamma} \right) && (16) \\ B(t,T) &= \frac{2\left(e^{\gamma(T-t)} - 1\right)}{(a+\gamma)\left(e^{\gamma(T-t)} - 1\right) + 2\gamma}, && (17) \end{aligned}$$

where γ is given in (11), and thus, $P^*(t,T)$ can in the CIR model be expressed as

$$P^*(t,T) = \exp\left(A(t,T) - B(t,T)r\right) + \alpha \exp\left(C(t,T) - D(t,T)r\right). \quad (18)$$

We now have the following practical lemmas:

Lemma 2 *The following hold.*

$$\begin{aligned} A(t,T) - C(t,T) &= \frac{2ab}{\sigma^2} \log\left(\gamma B(t,T)\right) && (19) \\ B(t,T) - D(t,T) &= \frac{2}{\sigma^2} \frac{B_t(t,T)}{B(t,T)}. && (20) \end{aligned}$$

Proof. Relation (19) is easily verified from the definitions of $A(t,T)$, $B(t,T)$ and $C(t,T)$. Differentiate (19) with respect to t and use (12) and (14) to obtain (20). ∎

Lemma 3 *The functions $A(t,T)$, $B(t,T)$, $C(t,T)$ and $D(t,T)$ satisfy*

$$\begin{aligned} \lim_{T\to\infty} A(t,T) &= -\infty && (21) \\ \lim_{T\to\infty} B(t,T) &= \frac{2}{a+\gamma} && (22) \\ \lim_{T\to t} C(t,T) &= \infty && (23) \\ \lim_{T\to\infty} C(t,T) &= -\infty && (24) \\ \lim_{T\to t} D(t,T) &= \infty && (25) \\ \lim_{T\to\infty} D(t,T) &= \frac{2}{a+\gamma} && (26) \end{aligned}$$

Furthermore, we have

$$B_T(t,T) = \frac{4\gamma e^{\gamma(T-t)}}{\left((a+\gamma)\left(e^{\gamma(T-t)}-1\right)+2\gamma\right)^2} > 0, \qquad 0 \le t \le T < \infty. \tag{27}$$

Proof. An application of L'Hôpital's rule gives (21), and (22) follows directly from the definition of $B(t,T)$. The limit in (23) follows directly from the definition of $C(t,T)$, and using L'Hôpital's rule, we have

$$\begin{aligned}
\lim_{T\to\infty} \frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1} &= \frac{a+\gamma}{2\gamma} \lim_{T\to\infty} \frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}} \\
&= \frac{a+\gamma}{2\gamma} \lim_{T\to\infty} e^{\frac{a-\gamma}{2}(T-t)} \\
&= 0,
\end{aligned} \tag{28}$$

since $\gamma > a$. Hence, (24) follows from (28), since $\lim_{x\to 0+} \log x = -\infty$. The limits in (25) and (26) are easily verified using the definition of $D(t,T)$. Finally, (27) follows directly from the definition of $B(t,T)$. ∎

Lemma 4 *For any constant $\beta > 0$, we have*

$$\lim_{T\to t} \left(C(t,T) - \beta D(t,T)\right) = -\infty \tag{29}$$

$$\lim_{T\to\infty} \left(C(t,T) - \beta D(t,T)\right) = -\infty. \tag{30}$$

Proof. Proving (29), we have, using the definitions of $C(t,T)$ and $D(t,T)$, that

$$\begin{aligned}
&\lim_{T\to t} \left(C(t,T) - \beta D(t,T)\right) \\
&= \lim_{T\to t} \left\{ \frac{2ab}{\sigma^2} \log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1} \right) - \beta\left(\frac{2}{\gamma+a} + \frac{2\gamma}{\sigma^2\left(e^{\gamma(T-t)}-1\right)} \right) \right\} \\
&= \lim_{T\to t} \left\{ \frac{2ab}{\sigma^2} \log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1} \right) + \frac{2ab}{\sigma^2}\left(-\frac{\beta(\gamma-a)}{2ab} - \frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)} \right) \right\} \\
&= \frac{2ab}{\sigma^2} \lim_{T\to t} \left\{ \log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1} \right) + \log\exp\left(-\frac{\beta(\gamma-a)}{2ab} - \frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)} \right) \right\} \\
&= \frac{2ab}{\sigma^2} \lim_{T\to t} \log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1} \exp\left(-\frac{\beta(\gamma-a)}{2ab} - \frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)} \right) \right) \\
&= \frac{2ab}{\sigma^2} \lim_{T\to t} \log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)} e^{-\frac{\beta(\gamma-a)}{2ab}}}{e^{\gamma(T-t)}-1} \exp\left(-\frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)} \right) \right).
\end{aligned}$$

Then

$$
\begin{aligned}
&\lim_{T\to t}\left(C(t,T)-\beta D(t,T)\right)\\
=\;& \frac{2ab}{\sigma^2}\log\left(e^{-\frac{\beta(\gamma-a)}{2ab}}\right)+\frac{2ab}{\sigma^2}\lim_{T\to t}\log\left(e^{\frac{a+\gamma}{2}(T-t)}\right)+\frac{2ab}{\sigma^2}\lim_{T\to t}\log\left(\frac{\exp\left(-\frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)}\right)}{e^{\gamma(T-t)}-1}\right)\\
=\;& \frac{2ab}{\sigma^2}\log\left(e^{-\frac{\beta(\gamma-a)}{2ab}}\right)+0+\frac{2ab}{\sigma^2}\lim_{T\to t}\log\left(\frac{\exp\left(-\frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)}\right)}{e^{\gamma(T-t)}-1}\right).
\end{aligned}
$$

To find the last limit, write

$$
\begin{aligned}
\frac{2ab}{\sigma^2}\lim_{T\to t}\log\left(\frac{\exp\left(-\frac{\gamma\beta}{ab(e^{\gamma(T-t)}-1)}\right)}{e^{\gamma(T-t)}-1}\right) &= \frac{2ab}{\sigma^2}\lim_{x\to 0}\log\left(\frac{\exp\left(-\frac{\gamma\beta}{abx}\right)}{x}\right)\\
&= \frac{2ab}{\sigma^2}\lim_{y\to\infty}\log\left(y\exp\left(-\frac{\gamma\beta}{ab}y\right)\right).
\end{aligned}
$$

Providing a, b, γ and β are all positive, we may use the well-known limit $\lim_{y\to\infty} y\exp\left(-\frac{\gamma\beta}{ab}y\right)=0$, and hence

$$
\frac{2ab}{\sigma^2}\lim_{y\to\infty}\log\left(y\exp\left(-\frac{\gamma\beta}{ab}y\right)\right)=-\infty,
$$

since $\lim_{x\to 0^+}\log x=-\infty$. Thus, we have shown that

$$
\lim_{T\to t}\left(C(t,T)-D(t,T)\beta\right)=-\infty.
$$

Finally, relations (24) and (26) prove (30). ■

3 The Term Structure

From (18) we see that the parameter α determines the degree deviation from the arbitrage-free prices, since the special case $\alpha=0$ corresponds to the arbitrage-free CIR prices. It is also clear from (18) that $P^*(t,T)$ becomes negative for sufficiently small α (α may be negative), and greater than one for sufficiently large α. Natural conditions to impose on $P^*(t,T)$ are that it remains non-negative and less than or equal to one for all maturities. In other words, we will only consider values of α such that for a given fixed $r > 0$, we have for all $T \geq t \geq 0$ that

$$0 \leq \exp\left(A(t,T) - B(t,T)r\right) + \alpha \exp\left(C(t,T) - D(t,T)r\right) \leq 1, \tag{31}$$

or, equivalently,

$$\psi(t,T) \leq \alpha \leq \varphi(t,T),$$

where

$$\begin{aligned} \psi(t,T) &= -\frac{\exp\left(A(t,T) - B(t,T)r\right)}{\exp\left(C(t,T) - D(t,T)r\right)} \\ \varphi(t,T) &= \frac{1 - \exp\left(A(t,T) - B(t,T)r\right)}{\exp\left(C(t,T) - D(t,T)r\right)}. \end{aligned}$$

It is clear that (31) holds for $\alpha = 0$, which corresponds to the arbitrage-free CIR prices, and one wonders if any $\alpha \neq 0$ may exist for which (31) holds for a given fixed $r > 0$ and all $T \geq t \geq 0$. Before this question is answered, the following results, concerning the upper bound for $\psi(t,T)$ and the lower bound for $\varphi(t,T)$, are presented:

Proposition 1 *The function $\psi(t,T)$ satisfies*

$$\lim_{T\to t} \psi(t,T) = -\infty \tag{32}$$

$$\lim_{T\to\infty} \psi(t,T) = -\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}}. \tag{33}$$

Furthermore, let x be the number of roots to the equation

$$y^2 - \frac{2ab}{\sigma^2 r}y + \frac{2}{\sigma^2} = 0 \tag{34}$$

in the interval $0 < y < \frac{2}{a+\gamma}$, and define

$$\psi^\infty = -\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}}. \tag{35}$$

We then have one of the following three scenarios:

1. $x = 0$. *For $T \geq t \geq 0$, we then have $\sup_T \psi(t,T) = \psi^\infty$.*

2. $x = 1$. *For* $T \geq t \geq 0$, *we then have* $\sup_T \psi(t,T) = \psi^{\max}$, *where*

$$\psi^{\max} = -\exp\left(\frac{2ab}{\sigma^2}\left(1+\log\left(\gamma y^-\right)\right) - 2r\left(\frac{a}{\sigma^2}+y^-\right)\right), \tag{36}$$

and y^- *is the root to (34).*

3. $x = 2$. *For* $T \geq t \geq 0$, *we then have* $\sup_T \psi(t,T) = \max\left[\psi^\infty, \psi^{\max}\right]$, *where* ψ^∞ *and* $\psi^{\max}$ *are defined in (35) and (36), with* y^- *being the smallest root to (34).*

Proof. Proving (32), we have

$$\begin{aligned}\lim_{T\to t}\psi(t,T) &= -\lim_{T\to t}\frac{\exp\left(A(t,T)-B(t,T)r\right)}{\exp\left(C(t,T)-D(t,T)r\right)}\\ &= -\infty,\end{aligned}$$

since

$$\lim_{T\to t}\exp\left(A(t,T)-B(t,T)r\right) = 1$$

and, from (29) (with $\beta = r > 0$), we have

$$\lim_{T\to t}\exp\left(C(t,T)-D(t,T)r\right) = 0.$$

To prove (33), we have, using (19) and (20), that

$$\begin{aligned}\lim_{T\to\infty}\psi(t,T) &= -\lim_{T\to\infty}\exp\left(\frac{2ab}{\sigma^2}\log\left(\gamma B(t,T)\right) - \frac{2r}{\sigma^2}\frac{B_t(t,T)}{B(t,T)}\right)\\ &= -\exp\left(\frac{2ab}{\sigma^2}\log\frac{2\gamma}{a+\gamma}\right)\\ &= -\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}},\end{aligned}$$

where we have used (22). Moreover, we have that

$$\psi(t,T) = -\exp\left(\frac{2ab}{\sigma^2}\log\left(\gamma B(t,T)\right) - \frac{2r}{\sigma^2}\frac{B_t(t,T)}{B(t,T)}\right),$$

and a first order condition for local extreme values is $\psi_T(t,T) = 0$, giving

$$\frac{2ab}{\sigma^2}\frac{B_T(t,T)}{B(t,T)} = \frac{2r}{\sigma^2}\frac{\partial}{\partial T}\left(\frac{B_t(t,T)}{B(t,T)}\right). \tag{37}$$

Now $B(t,T)$ satisfies the ODE (15), and hence the first order condition (37) simplifies to

$$B^2(t,T) - \frac{2ab}{\sigma^2 r}B(t,T) + \frac{2}{\sigma^2} = 0. \tag{38}$$

This is a quadratic equation in the variable $B(t,T)$, with roots given by

$$B^{\pm} = \frac{ab}{\sigma^2 r} \pm \sqrt{\left(\frac{ab}{\sigma^2 r}\right)^2 - \frac{2}{\sigma^2}},$$

and we note that $B^{\pm} > 0$. Since $B(t,T)$ is non-negative and increasing in T, it follows from (22) that we must have $0 < B^{\pm} < \frac{2}{a+\gamma}$, and we note that there is only one value $T = T^{\pm}$ such that $B(t,T^{\pm}) = B^{\pm}$. Hence if $T^{\pm}$ is such that $0 < B^{\pm} < \frac{2}{a+\gamma}$, then $\psi(t,T)$ has a local extreme value at $T^{\pm}$. Now let x be the number of roots to the equation $y^2 - \frac{2ab}{\sigma^2 r}y + \frac{2}{\sigma^2} = 0$ in the interval $0 < y < \frac{2}{a+\gamma}$. Then, since $\psi(t,T)$ is continuous and satisfies (32)–(33), it follows that we have one of the three possible scenarios:

1. $x = 0$. Then $\psi(t,T)$ does not have local maximum, and hence for $T \geq t \geq 0$, we have

$$\begin{aligned} \psi^{\infty} &= \sup_T \psi(t,T) \\ &= -\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}}. \end{aligned}$$

2. $x = 1$. Then $\psi(t,T)$ has a local maximum, and this occurs at $B = B^-$, and then for $T \geq t \geq 0$, we have

$$\begin{aligned} \psi^{\max} &= \sup_T \psi(t,T) \\ &= -\exp\left(\frac{2ab}{\sigma^2}\log\left(\gamma B^-\right) + \frac{2r}{\sigma^2}\left(\frac{1}{B^-} - a - \frac{1}{2}\sigma^2 B^-\right)\right) \\ &= -\exp\left(\frac{2ab}{\sigma^2}\left(1 + \log\left(\gamma B^-\right)\right) - 2r\left(\frac{a}{\sigma^2} + B^-\right)\right), \end{aligned}$$

where we have used (15) and (38), and we note that if $x = 1$, we also have that $\sup_T \psi(t,T) = \psi^{\max} > \psi^{\infty}$.

3. $x = 2$. Then $\psi(t,T)$ has a local maximum at $B = B^-$, and a local minimum at $B = B^+$, and hence for $T \geq t \geq 0$, we have

$$\sup_T \psi(t,T) = \max\left[\psi^{\max}, \psi^{\infty}\right],$$

and the proof is complete. ■

Proposition 2 *The function* $\varphi(t,T)$ *satisfies*

$$\lim_{T\to t}\varphi(t,T) = \infty \tag{39}$$
$$\lim_{T\to\infty}\varphi(t,T) = \infty. \tag{40}$$

Furthermore, for $T \geq t \geq 0$ *we have* $\inf_T \varphi(t,T) = \varphi^{\min} > 0$.

Proof. Proving (39), we first note that from (29) there exists $t^* > t$ such that $C(t,T) - \beta D(t,T) < 0$ whenever $T < t^*$. In particular, for $\beta = \frac{r}{2}$, we thus have

$$e^{-C(t,T)} > e^{-\frac{r}{2}D(t,T)}, \tag{41}$$

for $T < t^*$. Using L'Hôpital's rule, we then get

$$\begin{aligned}\lim_{T\to t}\varphi(t,T) &= \lim_{T\to t}\frac{1-e^{A(t,T)-B(t,T)r}}{e^{C(t,T)-D(t,T)r}}\\ &= -\lim_{T\to t}\frac{(A_T(t,T)-B_T(t,T)r)\,e^{A(t,T)-B(t,T)r}}{(C_T(t,T)-D_T(t,T)r)\,e^{C(t,T)-D(t,T)r}}\\ &= -L_1 * L_2,\end{aligned}$$

where

$$\begin{aligned}L_1 &= \lim_{T\to t}(A_T(t,T)-B_T(t,T)r)\,e^{A(t,T)-B(t,T)r}\\ L_2 &= \lim_{T\to t}\frac{e^{-C(t,T)+D(t,T)r}}{(C_T(t,T)-D_T(t,T)r)}.\end{aligned}$$

For the first limit, L_1, we have

$$\begin{aligned}L_1 &= \lim_{T\to t}(A_T(t,T)-B_T(t,T)r)\,e^{A(t,T)-B(t,T)r}\\ &= -\lim_{T\to t}(abB(t,T)+B_T(t,T)r)\,e^{A(t,T)-B(t,T)r}\\ &= -r,\end{aligned}$$

and for the second limit, L_2, we note that

$$\begin{aligned} L_2 &= \lim_{T\to t} \frac{e^{-C(t,T)+D(t,T)r}}{(C_T(t,T) - D_T(t,T)r)} \\ &= \lim_{T\to t} \frac{e^{-C(t,T)+D(t,T)r}}{\left(-abD(t,T) - \left(1 - aD(t,T) - \frac{1}{2}\sigma^2 D^2(t,T)\right) r\right)} \\ &\geq \lim_{T\to t} \frac{e^{-\frac{r}{2}D(t,T)+D(t,T)r}}{\left(-abD(t,T) - \left(1 - aD(t,T) - \frac{1}{2}\sigma^2 D^2(t,T)\right) r\right)} \\ &= \lim_{T\to t} \frac{e^{\frac{r}{2}D(t,T)}}{\left(-r + a(r-b)D(t,T) + \frac{1}{2}\sigma^2 r D^2(t,T)\right)}, \end{aligned}$$

where we have used (41). Thus, from (25), it follows that

$$L_2 \geq \lim_{y\to\infty} \frac{e^{\frac{r}{2}y}}{-r + a\,(r-b)\,y + \frac{1}{2}\sigma^2 r y^2},$$

and using L'Hôpital's rule repeatedly we obtain

$$\begin{aligned} L_2 &\geq \lim_{y\to\infty} \frac{e^{\frac{r}{2}y}}{-r + a\,(r-b)\,y + \frac{1}{2}\sigma^2 r y^2} \\ &= \frac{r}{2} \lim_{y\to\infty} \frac{e^{\frac{r}{2}y}}{a\,(r-b) + \sigma^2 r y} \\ &= \frac{r}{4\sigma^2} \lim_{y\to\infty} e^{\frac{r}{2}y} \\ &= \infty. \end{aligned}$$

Hence, we have proved that

$$\begin{aligned} \lim_{T\to t} \varphi(t,T) &= -L_1 * L_2 \\ &= \infty. \end{aligned}$$

To prove (40), we have

$$\begin{aligned} \lim_{T\to\infty} \varphi(t,T) &= \lim_{T\to\infty} \frac{1 - e^{A(t,T)-B(t,T)r}}{e^{C(t,T)-D(t,T)r}} \\ &= \infty, \end{aligned}$$

since $\lim_{T\to\infty} \exp\left(A(t,T) - B(t,T)r\right) = 0$ and $\lim_{T\to\infty} \exp\left(C(t,T) - D(t,T)r\right) = 0$. Furthermore, we note that $1 - \exp\left(A(t,T) - B(t,T)r\right)$ is positive for all T except at $T = t$, where it is equal to zero. However, as T tends to t, we have shown above that $\lim_{T\to t} \varphi(t,T) = \infty$. Then since $\varphi(t,T)$

is continuous and satisfies $\lim_{T\to\infty} \varphi(t,T) = \infty$, it follows that $\varphi(t,T)$ has a global minimum $\varphi^{\min}$ which is positive. ■

The main result in this essay is the following:

Proposition 3 *Suppose that for a given fixed $r > 0$ $P^*(t,T)$ are given by*

$$P^*(t,T) = \exp\left(A(t,T) - B(t,T)r\right) + \alpha \exp\left(C(t,T) - D(t,T)r\right),$$

where α is a constant and $A(t,T), B(t,T), C(t,T)$ and $D(t,T)$ are given by (16),(17),(9) and (10), respectively. Then, if

$$\sup_T \psi(t,T) < \alpha < \inf_T \varphi(t,T),$$

where $\sup_T \psi(t,T)$ is determined in Proposition 1, and $\inf_T \varphi(t,T)$ is shown to be positive in Proposition 2, we have that

$$0 \leq P^*(t,T) \leq 1, \qquad 0 \leq t \leq T. \tag{42}$$

Proof. This follows from the definitions of $\psi(t,T)$ and $\varphi(t,T)$, and from Proposition 1 and Proposition 2. ■

The following definition is now made:

Definition 2 *$P^*(t,T)$ is called a* ***Quasi arbitrage-free term structure*** *if α is such that $P^*(t,T)$ satisfies (42).*

We now move on to explore the shapes of the quasi arbitrage-free term structure.

4 Term Structure Shapes

It is generally agreed that one of the major shortcomings of one-factor models is their inability to generate flexible and realistic shapes of the yield curve. Standard models such as the models by Vasicek [5] and Cox, Ingersoll and Ross [1] display simple yield curves that are essentially increasing, decreasing or slightly humped. In sharp contrast to this, yield curves generated by the quasi arbitrage-free term structure in (18) display various realistic shapes and interesting nonlinear phenomena. From the definition of the quasi arbitrage-free bond prices in (18), we expect, as argued above, that the biggest difference between the yields generated by the quasi arbitrage-free bond prices and those generated by the arbitrage-free CIR model to appear for intermediate rates. In practice, the effect of using the quasi arbitrage-free bond prices can essentially be interpreted as introducing a hump on the otherwise simple CIR yield curve. Dependent on the magnitude and sign of α, this hump may be large or small, or inverted.

We define the quasi arbitrage-free yields by

$$Y^*(t,T) = -\frac{1}{T-t}\log P^*(t,T), \tag{43}$$

where $P^*(t,T)$ is defined in (18), and the following lemma shows some of their desirable properties.

Lemma 5 *The quasi arbitrage-free yields satisfy*

$$\lim_{T\to t} Y^*(t,T) = r \tag{44}$$

$$\lim_{T\to\infty} Y^*(t,T) = \frac{2ab}{a+\gamma}. \tag{45}$$

Proof. See Appendix B. ■

Remark 1 *The limits obtained for the quasi arbitrage-free yields in Lemma 5 are thus the same as the limits for the yields in the arbitrage-free CIR model.*

The parameter α has been varied in Figure 1 and Figure 2. In both figures, the special case $\alpha = 0$ representing the simple CIR yield curve is the dash-dotted curve. We see that it is possible to generate yield curves with humps or inverted humps, for both increasing and decreasing yield curves, and that the yield curves are thus more complex, showing more realistic features. We note that a negative value for α tends to introduce a hump, whereas a positive α tends to create an inverted hump. This can be understood from (43), since it follows that

$$\frac{\partial Y^*(t,T)}{\partial \alpha} = -\frac{1}{T-t}\frac{e^{C(t,T)-D(t,T)r}}{P^*(t,T)} < 0, \tag{46}$$

i.e., starting from the dash-dotted CIR yield curve, a positive (negative) change in α will then push the yield curve downwards (upwards). It is worth pointing out that the quasi arbitrage-free yield curves more or less coincide with the simple CIR yield curve for short-term maturities, and that the difference from the CIR yield curve first becomes visible for maturities greater then about 4 years. For long-term maturities, the yield curves converge to the same value given in (45).

The short rate has been varied in Figure 3 and Figure 4. In Figure 3, $\alpha = 0.01$ and in Figure 4, $\alpha = -0.01$. Again, we see that a negative value for α tends to introduce a hump, whereas a positive α tends to create an inverted hump. We also note again that the effect is most significant for intermediate maturities, and that the effect is larger the smaller the value of the short rate. It is also worth pointing out that (at least for these parameter values) the yield curves intersect each other for $\alpha = -0.01$.

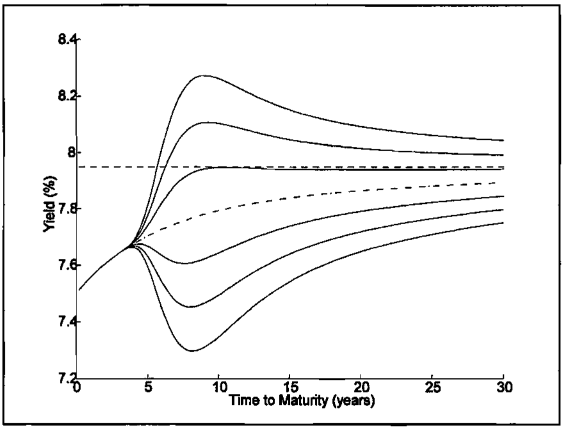

FIGURE 1. Yield Curves for Increasing α (top-down)
Yield curves for $\alpha = -0.06$ (top), $-0.04, ..., 0.04, 0.06$ (down). Parameters used: $a = 0.25$, $b = 0.08$, $\sigma^2 = 0.0008$ and $r = 7.5\%$. The dashed line represents the long-run yield of 7.95%, and the dash-dotted curve represents the corresponding CIR yield curve ($\alpha = 0$).

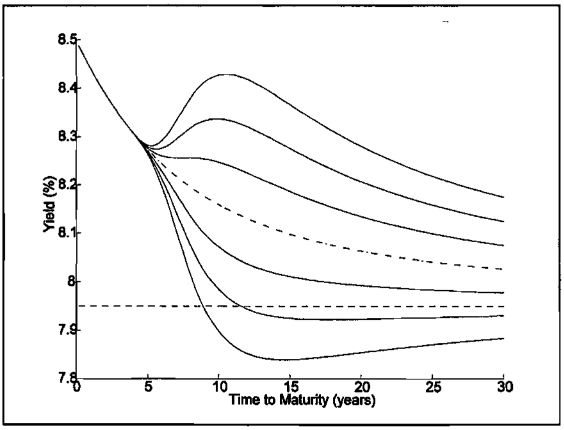

FIGURE 2. Yield Curves for Increasing α (top-down)
Yield curves for $\alpha = -0.06$ (top), $-0.04, ..., 0.04, 0.06$ (down). Parameters used: $a = 0.25$, $b = 0.08$, $\sigma^2 = 0.0008$ and $r = 8.5\%$. The dashed line represents the long-run yield of 7.95%, and the dash-dotted curve represents the corresponding CIR yield curve ($\alpha = 0$).

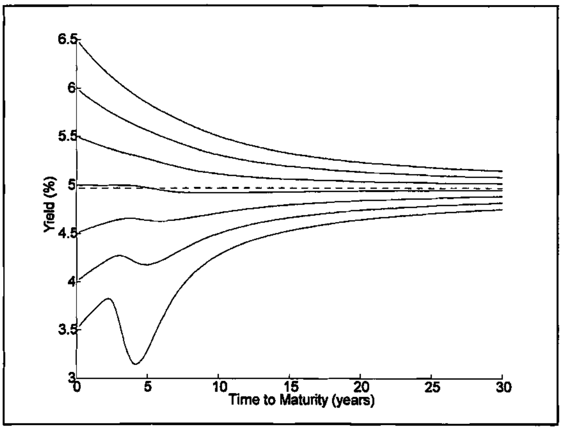

FIGURE 3. Yield curves for Decreasing r and $\alpha = 0.01$ (top-down)
Yield curves for $r = 6.5\%$ (top), 6%, ..., 4% and 3.5% (down). Parameters used: $a = 0.25$, $b = 0.05$, $\sigma^2 = 0.0008$ and $\alpha = 0.01$. The dashed line represents the long-run yield of 4.97%.

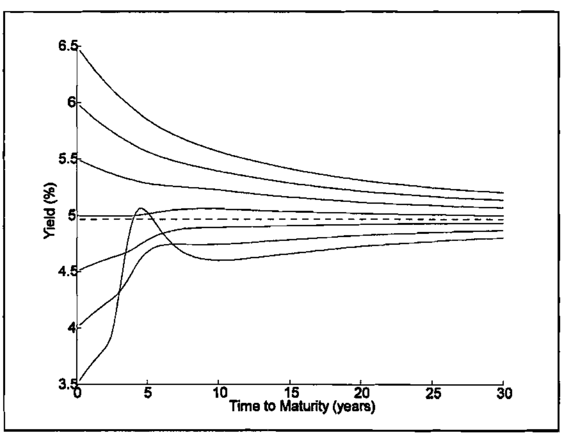

FIGURE 4. Yield Curves for Decreasing r and $\alpha = -0.01$ (top-down)
Yield curves for $r = 6.5\%$ (top), 6%, ..., 4% and 3.5% (down). Parameters used: $a = 0.25$, $b = 0.05$, $\sigma^2 = 0.0008$ and $\alpha = -0.01$. The dashed line represents the long-run yield of 4.97%.

5 Forward Rates

The forward rates are defined as $f^*(t,T) = -\frac{P_T^*(t,T)}{P^*(t,T)}$, and using the definition of $P^*(t,T)$ from (18) we obtain

$$\begin{aligned} f^*(t,T) &= \frac{r + a(b-r)B(t,T) - \frac{1}{2}r\sigma^2 B^2(t,T)}{1 + \alpha e^{-A+Br}e^{C-Dr}} \\ &\quad + \alpha \frac{r + a(b-r)D(t,T) - \frac{1}{2}r\sigma^2 D^2(t,T)}{1 + \alpha e^{-A+Br}e^{C-Dr}} e^{C-A-(D-B)r}. \end{aligned}$$

Without any restrictions on α, the forward rates may become negative at some point. From (43) it follows that

$$f^*(t,T) = Y^*(t,T) + (T-t)Y_T^*(t,T), \tag{47}$$

and thus negative forward rates may appear where the corresponding yield curve has a large negative slope.

It is easily verified that the limits for $f^*(t,T)$ are the same for those obtained by the yields in (44)–(45), i.e.,

$$\begin{aligned} \lim_{T\to t} f^*(t,T) &= r \\ \lim_{T\to\infty} f^*(t,T) &= \frac{2ab}{a+\gamma}. \end{aligned} \tag{48}$$

In Figure 5 and Figure 6, the parameter α has been varied. In both figures, the special case $\alpha = 0$ representing the simple CIR forward rate curve is the dash-dotted curve. As was the case with the yield curves, we see that the forward rate curves can also exhibit humps or inverted humps, for both increasing and decreasing forward rate curves. We note that a negative value for α tends to introduce a hump, whereas a positive α tends to create an inverted hump. Combining (43), (46) and (47), we have

$$\frac{\partial f^*(t,T)}{\partial \alpha} = \left(r - f^*(t,T) + a(b-r)D(t,T) - \frac{1}{2}\sigma^2 r D^2(t,T)\right) \frac{e^{C(t,T)-D(t,T)r}}{P^*(t,T)},$$

so the effect on the forward rate curve from a change in α depends on if the forward rate is above or below its starting point (r), the difference between b and r, and the factor $\frac{1}{2}\sigma^2 r D^2(t,T)$. We also note that the forward rate curves coincide with the simple CIR yield curve for short-term maturities, and that the difference from the CIR forward rate curve first becomes visible for maturities greater than about 4 years. For long-term maturities, the forward rates converge to the same value as the long-run yield.

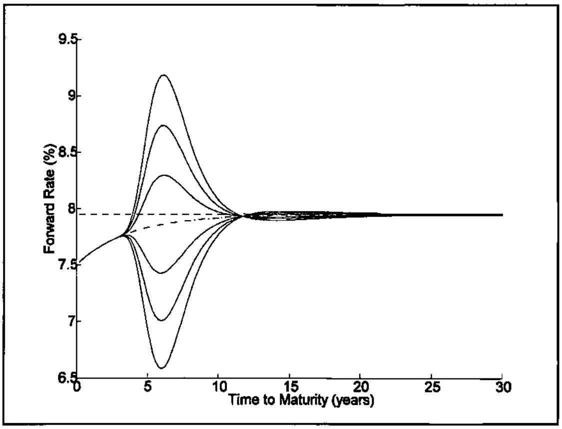

FIGURE 5. Forward Rate Curves for Increasing α (top-down)
Forward rate curves for $\alpha = -0.06$ (top), -0.04, ..., 0.04 and 0.06 (down). Parameters used: $a = 0.25$, $b = 0.08$, $\sigma^2 = 0.0008$, $r = 7.5\%$. The dashed line represents the long-run forward rate of 7.95%, and the dash-dotted curve represents the corresponding CIR forward rate curve ($\alpha = 0$).

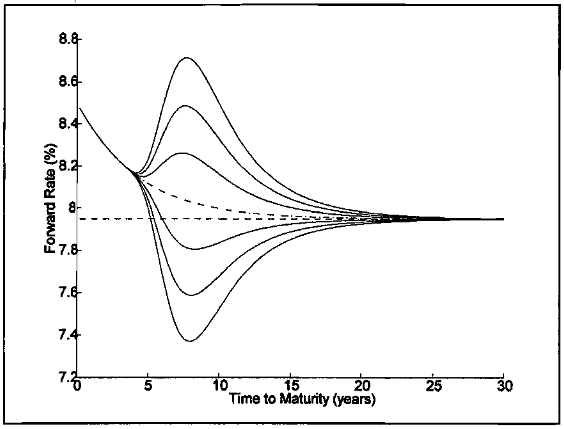

FIGURE 6. Forward Rate Curves for Increasing α (top-down)
Forward rate curves for $\alpha = -0.06$ (top), -0.04, ..., 0.04 and 0.06 (down). Parameters used: $a = 0.25$, $b = 0.08$, $\sigma^2 = 0.0008$ and $r = 8.5\%$. The dashed line represents the long-run forward rate of 7.95%, and the dash-dotted curve represents the corresponding CIR forward rate curve ($\alpha = 0$).

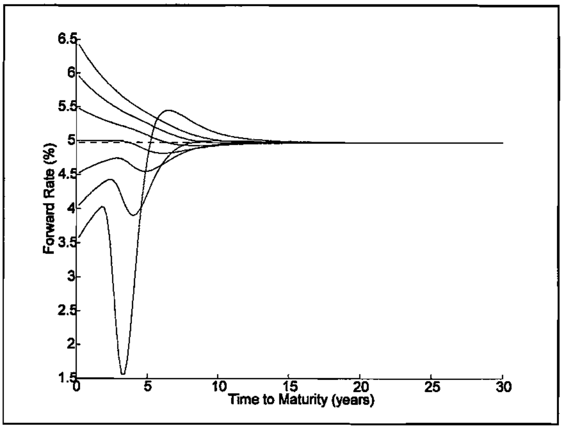

FIGURE 7. Forward Rate Curves for Decreasing r and $\alpha = 0.01$ (top-down)
Forward rate curves for $r = 6.5\%$ (top), 6%, ..., 4% and 3.5% (down). Parameters used: $a = 0.25$, $b = 0.05$, $\sigma^2 = 0.0008$, $\alpha = 0.01$. The dashed line represents the long-run forward rate of 4.97%.

In Figure 7 and Figure 8, the short rate has been varied. In Figure 7, $\alpha = 0.01$ and in Figure 8, $\alpha = -0.01$. Again, we see that a negative value for α tends to introduce a hump, whereas a positive α tends to create an inverted hump. We also note again that the effect is most significant for intermediate maturities, and that the effect is larger the smaller the value of the short rate. It is also worth pointing out that (at least for these parameter values) the forward rate curves intersect each other.

6 Summary and Conclusions

Natural boundary conditions to impose on the solution to (8) when the short rate follows the dynamics in (7), i.e., the CIR model, are that bond prices tend to zero as r tends to infinity and remain finite when $r = 0$[4]. Essentially, if a solution to (8) is bounded for all $0 \leq t \leq T$ and non-negative r, then this is the solution obtained from (1)[5]. These conditions are fulfilled by the CIR discount bond prices. For the quasi arbitrage-free prices, it is easily shown that they tend to zero for $0 \leq t < T$ as r tends to infinity, since $B(t,T)$ is non-negative and $D(t,T)$ is strictly positive. However, it can be verified that $\lim_{T \to t} \lim_{r \to 0} P^*(t,T) = \infty$. Thus, the quasi arbitrage-free bond prices do not remain finite for $r = 0$ as $T \to t$. It is a natural economic condition that bond prices stay bounded when $r = 0$. Nevertheless, it is worth pointing out that the quasi arbitrage-free prices $P^*(t,T)$ are bounded for all *positive* r for all $0 \leq t \leq T$. Thus, for all

[4]See Dewynne *et al.* [2].
[5]See Øksendal [6].

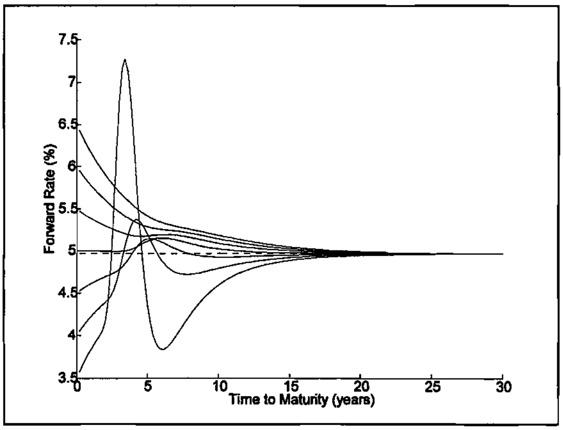

FIGURE 8. Forward Rate Curves for Decreasing r and $\alpha = -0.01$ (top-down)
Forward rate curves for $r = 6.5\%$ (top), 6%, ..., 4% and 3.5% (down). Parameters used: $a = 0.25$, $b = 0.05$, $\sigma^2 = 0.0008$ and $\alpha = -0.01$. The dashed line represents the long-run forward rate of 4.97%.

practical purposes, since we may always expect a positive short rate, $P^*(t,T)$ stays bounded for all $0 \leq t \leq T$.

Given that the short rate follows the stochastic differential equation in (7), then the "true" arbitrage-free discount bond prices in this model are those obtained by CIR, i.e., $P(t,T) = \exp(A(t,T) - B(t,T)r)$, where $A(t,T)$ and $B(t,T)$ are given by (16)–(17). We can therefore expect arbitrage opportunities to appear if we use any other expression for discount bond prices; in particular there will be arbitrage opportunities if we use the discount bond prices $P^*(t,T)$ given in (18). On the other hand, assuming that the short rate can only take on positive values, we can regard $P^*(t,T)$ as a term structure that has much in common with the CIR term structure, while displaying more flexible and realistic shapes. In Proposition 3, it was shown that for given fixed $r > 0$, there exist a strictly negative lower bound and a strictly positive upper bound for α to ensure that $0 \leq P^*(t,T) \leq 1$ for all $0 \leq t \leq T$.

Finally, it is worth pointing out that the general perturbation function in Definition 1, given by Lemma 1 in the CIR model, can be added to the price function of any derivative in the CIR model. For instance, let $c(t,T)$ denote the price of a European call option in the CIR model. Then we can define a similar quasi arbitrage-free call option price by $c^*(t,T) = c(t,T) + \alpha\chi(t,T)$, where $\chi(t,T)$ is the perturbation function in the CIR model given in Lemma 1. It then remains to be determined whether it is possible to find bounds on α so that option prices remain nonnegative.

Appendix A
The Functions C(t,T) and D(t,T)

The Function $D(t,T)$

The Ricatti equation (13) can be transformed to a second order ordinary differential equation by making the substitution

$$D(t,T) = -\frac{2}{\sigma^2}\frac{w'(t)}{w(t)}. \tag{49}$$

The ODE obtained for $w(t)$ is then

$$w''(t) - aw'(t) - \frac{\sigma^2}{2}w(t) = 0,$$

with a general solution given by

$$w(t) = K_1 e^{\frac{a+\gamma}{2}t} + K_2 e^{\frac{a-\gamma}{2}t}, \tag{50}$$

where K_1 and K_2 are arbitrary constants and γ is as in (11). Thus we have

$$\begin{aligned} D(t,T) &= -\frac{2}{\sigma^2}\frac{w'(t)}{w(t)} \\ &= -\frac{1}{\sigma^2}\frac{(a+\gamma)e^{\gamma t} + (a-\gamma)\frac{K_2}{K_1}}{e^{\gamma t} + \frac{K_2}{K_1}}. \end{aligned}$$

Since $\chi(t,T)$ is defined to be time-homogenous, and satisfies (5)-(6), we choose $\frac{K_2}{K_1} = -e^{\gamma T}$. Then

$$D(t,T) = \frac{2}{\gamma + a} + \frac{2\gamma}{\sigma^2\left(e^{\gamma(T-t)} - 1\right)}.$$

See Lemma 3 for the limits $\lim\limits_{T\to t} D(t,T)$ and $\lim\limits_{T\to\infty} D(t,T)$.

The Function $C(t,T)$

The differential equation for $C(t,T)$ is

$$C_t(t,T) = abD(t,T),$$

and using (49) we obtain

$$C_t(t,T) = -\frac{2ab}{\sigma^2}\frac{w'(t)}{w(t)}.$$

Thus

$$C(t,T) = g(T) - \frac{2ab}{\sigma^2}\log w(t),$$

where $g(T)$ is an arbitrary function of T and $w(t)$ is given by (50). We get

$$\begin{aligned}
C(t,T) &= g(T) - \frac{2ab}{\sigma^2}\log\left(K_1 e^{\frac{a+\gamma}{2}t} + K_2 e^{\frac{a-\gamma}{2}t}\right)\\
&= g(T) - \frac{2ab}{\sigma^2}\log\left\{K_1 e^{\frac{a+\gamma}{2}t}\left(1+\frac{K_2}{K_1}e^{-\gamma t}\right)\right\}\\
&= g(T) - \frac{2ab}{\sigma^2}\log\left\{e^{\frac{a+\gamma}{2}t}\right\} - \frac{2ab}{\sigma^2}\log\left\{K_1\left(1+\frac{K_2}{K_1}e^{-\gamma t}\right)\right\}\\
&= g(T) - \frac{2ab}{\sigma^2}\frac{a+\gamma}{2}t - \frac{2ab}{\sigma^2}\log\left\{-K_1\left(e^{\gamma(T-t)}-1\right)\right\},
\end{aligned}$$

where we have used that $\frac{K_2}{K_1} = -e^{\gamma T}$. Since we require $C(t,T)$ to be time-homogenous, we choose $g(T) = \frac{2ab}{\sigma^2}\frac{a+\gamma}{2}T$. Then,

$$\begin{aligned}
C(t,T) &= \frac{2ab}{\sigma^2}\frac{a+\gamma}{2}(T-t) - \frac{2ab}{\sigma^2}\log\left\{-K_1\left(e^{\gamma(T-t)}-1\right)\right\}\\
&= \frac{2ab}{\sigma^2}\log e^{\frac{a+\gamma}{2}(T-t)} - \frac{2ab}{\sigma^2}\log\left\{-K_1\left(e^{\gamma(T-t)}-1\right)\right\}\\
&= \frac{2ab}{\sigma^2}\log\left(-K_1\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1}\right),
\end{aligned}$$

and we note that K_1 must be some arbitrary *negative* constant. Without loss of generality, we may assume that $K_1 = -1$. This constant will be absorbed in α. Thus we obtain

$$C(t,T) = \frac{2ab}{\sigma^2}\log\left(\frac{e^{\frac{a+\gamma}{2}(T-t)}}{e^{\gamma(T-t)}-1}\right).$$

See Lemma 3 for the limits $\lim\limits_{T\to t} C(t,T)$ and $\lim\limits_{T\to\infty} C(t,T)$.

Appendix B
Proof of Lemma 5

Throughout the proof, it is assumed that $r > 0$. Using L'Hôpital's rule, we have for $\tau = \infty$ or $\tau = t$ that

$$\begin{aligned}
\lim_{T\to\tau} Y^*(t,T) &= -\lim_{T\to\tau} \frac{1}{T-t} \log P^*(t,T) \\
&= -\lim_{T\to\tau} \frac{P_T^*(t,T)}{P^*(t,T)} \\
&= -\lim_{T\to\tau} \frac{(A_T(t,T) - B_T(t,T)r)\, e^{A-Br} + \alpha\,(C_T(t,T) - D_T(t,T)r)\, e^{C-Dr}}{e^{A-Br} + \alpha e^{C-Dr}} \\
&= \lim_{T\to\tau} \frac{abB(t,T) + B_T(t,T)r}{1 + \alpha e^{C-Dr} e^{-A+Br}} + \alpha \lim_{T\to\tau} \frac{abD(t,T) + D_T(t,T)r}{e^{A-Br} + \alpha e^{C-Dr}} e^{C-Dr}. \qquad (51)
\end{aligned}$$

For $\tau = t$, the first limit in (51) is given by

$$\lim_{T\to t} \frac{abB(t,T) + B_T(t,T)r}{1 + \alpha e^{C-Dr} e^{-A+Br}} = r, \qquad (52)$$

where we have used (22) and (29). Rewriting the second limit in (51) as

$$\lim_{T\to t} \frac{abD(t,T) + D_T(t,T)r}{e^{A-Br} + \alpha e^{C-Dr}} e^{C-Dr} = L_1 * L_2 * L_3,$$

where

$$\begin{aligned}
L_1 &= \lim_{T\to t} \frac{1}{e^{A-Br} + \alpha e^{C-Dr}} \\
L_2 &= \lim_{T\to t} e^{C-D\frac{r}{2}} \\
L_3 &= \lim_{T\to t} (abD(t,T) + D_T(t,T)r)\, e^{-D\frac{r}{2}}.
\end{aligned}$$

Using (29), and the definitions of $A(t,T)$ and $B(t,T)$, we have $L_1 = 1$ and $L_2 = 0$. Finally, using (13) and (25), we obtain

$$\begin{aligned}
L_3 &= \lim_{T\to t} (abD(t,T) + D_T(t,T)r)\, e^{-D\frac{r}{2}} \\
&= \lim_{T\to t} \left(abD(t,T) + \left(1 - aD(t,T) - \frac{1}{2}\sigma^2 D^2(t,T)\right) r \right) e^{-D\frac{r}{2}} \\
&= \lim_{y\to\infty} \left(r + a\,(b-r)\,y - \frac{1}{2} r\sigma^2 y^2 \right) e^{-y\frac{r}{2}} \\
&= 0,
\end{aligned}$$

and thus the second limit in (51) is

$$\lim_{T\to t} \frac{abD(t,T)+D_T(t,T)r}{e^{A-Br}+\alpha e^{C-Dr}} e^{C-Dr} = 0. \tag{53}$$

Hence, (52) and (53) prove (44).

For $\tau = \infty$, we have for the first and second limit in (51) that

$$\lim_{T\to\infty} \frac{abB(t,T)+B_T(t,T)r}{1+\alpha e^{C-A}e^{(B-D)r}} = \frac{2ab}{a+\gamma} \frac{1}{1+\alpha\left(\frac{2\gamma}{a+\gamma}\right)^{-\frac{2ab}{\sigma^2}}},$$

and

$$\lim_{T\to\infty} \frac{abD(t,T)+D_T(t,T)r}{e^{A-C-(B-D)r}+\alpha} = \frac{2ab}{a+\gamma} \frac{1}{\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}}+\alpha},$$

respectively, where we have used (22), (19), (20), and (25). Thus,

$$\begin{aligned}\lim_{T\to\infty} Y^*(t,T) &= \frac{2ab}{a+\gamma}\left(\frac{1}{1+\alpha\left(\frac{2\gamma}{a+\gamma}\right)^{-\frac{2ab}{\sigma^2}}} + \frac{\alpha}{\left(\frac{2\gamma}{a+\gamma}\right)^{\frac{2ab}{\sigma^2}}+\alpha}\right)\\ &= \frac{2ab}{a+\gamma},\end{aligned}$$

which then proves (45).

References

[1] Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica* **53**(2), 385–407.

[2] Dewynne, J., S. Howison, and P. Wilmott (1993): "Option Pricing," Oxford Financial Press.

[3] Duffie, D. (1996): "Dynamic Asset Pricing," (2nd edn.), Princeton University Press.

[4] Fritz, J. (1982): "Partial Differential Equations," Springer Verlag.

[5] Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* **5**, 177–188.

[6] Øksendal, B. (1995): "Stochastic Differential Equations," (4th edn.), Springer Verlag.

EFI
The Economic Research Institute

Reports since 1995
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title

2000

Berg-Suurwee, U., Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.
Bergkvist, L., Advertising Effectiveness Measurement: Intermediate Constructs and Measures.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Koncernchefen och ägarna.
Bornefalk, A., Essays on Social Conflict and Reform.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform.
Edman, J., Information Use and Decision Making in Groups.
Ericson, M., Strategi, kalkyl, känsla.
Gunnarsson, J., Wahlund, R., Flink, H., Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.
Hellman, N., Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.
Håkansson, P., Beyond Private Label – The Strategic View on Distributor Own Brands.
I huvudet på kunden. Söderlund, M., (red).
Karlsson Stider, A., Familjen och firman.
Ljunggren, U., Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.
Stenström, E., Konstiga företag.
Styrning av team och processer – Teoretiska perspektiv och fallstudier. Bengtsson, L., Lind, J., Samuelson, L.A., (red).
Sweet, S., Industrial Change Towards Environmental Sustainability - The Case of Replacing Chloroflouorocarbons.

1999

Adler, N., Managing Complex Product Development.
Allgulin, M., Supervision and Monetary Incentives.
Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Eliasson, A-C., Smooth Transitions in Macroeconomic Relationships.
Flink, H., Gunnarsson, J., Wahlund, R., Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Helgesson, C-F., Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
Japanese Production Management in Sunrise or Sunset. Karlsson, C., (red).

Jönsson, B., Jönsson, L., Kobelt, G., Modelling Disease Progression And the Effect of Treatment in Secondary Progressive MS. Research Report.
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdelsnämndsreformen – en kartläggning.
Ljunggren, U., En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – Med tillämpning i Stockholms stads grundskolor.
Lundbergh, S., Modelling Economic High-Frequency Time Series.
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Mölleryd, B.G., Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
Nilsson, K., Ledtider för ledningsinformation.
Osynlig Företagsledning. Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red).
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting.
Sandström, M., Evaluating the Benefits and Effectiveness of Public Policy.
Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.
Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
Söderström, U., Monetary Policy under Uncertainty.
Werr, A., The Language of Change The Roles of Methods in the Work of Management Consultants.
Wijkström, F., Svenskt organisationsliv – Framväxten av en ideell sektor.

1998

Andersson, M., On Testing and Forecasting in Fractionally Integrated Time Series Models.
Berg-Suurwee, U., Styrning av kultur- och fritidsförvaltning innan stadsdelsnämndsreformen.
Berg-Suurwee, U., Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdelsnämndsreformen.
Bergström, F., Essays on the Political Economy of Industrial Policy.
Bild, M., Valuation of Takeovers.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
Eriksson-Skoog, G., The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
Gredenhoff, M., Bootstrap Inference in Time Series Econometrics.
Ioannidis, D., I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
Johansson, S., Savings Investment, and Economic Reforms in Developing Countries.
Levin, J., Essays in Company Valuation.
Ljunggren, U., Styrning av grundskolan i Stockholms stad innan stadsdelsnämndsreformen.
Mattsson, S., Från stat till marknad – effekter på nätverksrelationer vid en bolagiseringsreform.
Nyberg, A., Innovation in Distribution Channels – An Evolutionary Approach.
Olsson, P., Studies in Company Valuation.
Reneby, J., Pricing Corporate Securities.

Roszbach, K., Essays on Banking Credit and Interest Rates.
Runsten, M., The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.
Segendorff, B., Essays on Bargaining and Delegation.
Sjöberg, L., Bagozzi, R., Ingvar, D.H., Will and Economic Behavior.
Sjögren, A., Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.
Studier i kostnadsintäktsanalys. Jennergren, P., (red)
Söderholm, J., Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.
Thorburn, K., Cash Auction Bankruptcy and Corporate Restructuring
Wijkström, F., Different Faces of Civil Society.
Zethraeus, N., Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.
Andersson, B., Essays on the Swedish Electricity Market.
Berggren, N., Essays in Constitutional Economics.
Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
Charpentier, C., Budgeteringens roller, aktörer och effekter. En studie av budget-processerna i en offentlig organisation.
De Geer, H., Silfverberg, G., Citizens'Trust and Authorities' Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.
Friberg, R., Prices, Profits and Exchange Rates.
Från optionsprissättning till konkurslagstiftning. Bergström, C., Björk, T., (red)
Hagerud, G.E., A New Non-Linear GARCH Model.
Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.
He, C., Statistical Properties of Garch Processes.
Holmgren, M., Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.
Jennergren, P., Tutorial on the McKinsey Model for Valuation of Companies.
Lagerlöf, J., Essays on Political Economy, Information, and Welfare.
Lange, F., Wahlund, R., Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.
Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.
Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.
Sjöberg, L., Ramsberg, J., En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.
Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.
Sällström, S., On the Dynamics of Price Quality.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.

Zhang Gang., Chinese Rural Enterprises Between Plan and Market.
Åhlström, P., Sequences in the Profess of Adopting Lean Production.
Åkesson, G., Företagsledning i strategiskt vakuum. Om aktörer och förändrings-processer.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.

1996

Advancing your Business. People and Information Systems in Concert. Lundeberg, M., Sundgren, B (red).
Att föra verksamheten framåt. Människor och informationssystem i samverkan. red. Lundeberg, M., Sundgren, B.
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.
Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.
Nielsen, S., Omkostningskalkulation for avancerede produktions-omgivelser - en sammenligning af stokastiske og deterministiske omkost-ningskalkulationsmodeller.
Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civil-ekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.
Ulfsdotter, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.

Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.

Örtendahl, M., Health and Time - A Problem of Discounting.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. Jonung, L (red).
Gunnarsson, J., Wahlund, R., Hushållens finansiella strategier. En explorativ studie.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Statisfaction in a Store Performance Framework. Research Report.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.