Essays on Interest Rate Theory

Mikael Elhouar

Essays on Interest Rate Theory

Mikael Elhouar

EFI THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the Degree of Doctor of Philosophy, Ph.D
Stockholm School of Economics 2008.

ISBN 978-91-7258-761-8

Keywords:
Interest rate models, Heath-Jarrow-Morton, Regime-Switching, LIBOR market model, swap market model, state space models, Markovian realizations, expectations hypotheses.

Printed by:
Elanders, Vällingby 2008

Distributed by:
EFI, The Economic Research Institute
Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden
www.hhs.se/efi

To Chloé and Raphael

"I say, that's rather a difficult question. Won't you give the answer yourself?"
"No, because it's worthless unless you yourself discover it."
Somerset Maugham, "Of Human Bondage."

EFI THE ECONOMIC RESEARCH INSTITUTE

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution that works independently of economic, political and sectional interests. It conducts theoretical and empirical research in the 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. At EFI, the researchers select their projects based on the need for theoretical or practical development of a research domain, on their methodological interests, and on the generality of a problem.

Research Organization

The research activities at the Institute are organized into 21 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	*Centre Director:*
Management and Organisation (A)	Sven-Erik Sjöstrand
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr (acting)
Centre for Innovation and Operations Management (T)	Pär Åhlström
Centre for Media and Economic Psychology (P)	Guje Sevón
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson (acting)
Marketing, Distribution and Industrial Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Centre for Business and Economic History (BEH)	Håkan Lindgren
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Magnus Johannesson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Paul Segerstrom
Economic Statistics (ES)	Anders Westlund
Law (RV)	Johnny Herre
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address

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

Acknowledgments

First and foremost I would like to thank my supervisor Tomas Björk. It has been a privilege to have him as a supervisor. He has been very generous with his time, shared his knowledge and ideas and provided the necessary encouragement to complete this project. In addition to being an excellent supervisor, he has also been a good friend. I will surely miss his resolute knockings on my office door.

Former and present fellow Ph.D. students have made my working days (and evenings) at SSE a pleasant and memorable experience. I would especially like to thank the "math-finance gang", Agatha Murgoci, Irina Slinko, Linus Kaisajuntti, Magnus Blix, Mia Hinnerich, Raquel Gaspar, Sven Skallsjö, my former office mate Raana Farooqi Lind, and of course all the others namely Andriy Bodnaruk, Anete Pajuste, Damian Migueles Chazarreta, Daniel Sunesson, Håkan Blomdahl, Johan Söderström, Krister Ahlersten, Linus Siming, Maria Strömquist, Otto Elmgart, Per Östberg, Reimo Juks, Roland Nilsson, and Stefan Engström.

It has been a pleasure to have Anneli Sandbladh, Eva Eklund and Marita Rosing as office neighbours. Without you the department would be a dull place. Thank you for always giving prompt assistance with administrative and practical matters.

I would like to thank the current head of department, Mike Burkart, as well as the former one, Peter Englund, for providing me with a good working environment.

The financial support from the Jan Wallander and Tom Hedelius foundation made the work with this thesis possible. I am grateful for having been given this opportunity.

I am grateful to Sven-Olof Fridolfsson for encouraging me to apply to the Ph.D. program. I was a sound advice.

The work with this thesis has taken up a great number of evenings and week-ends during the last year. I am very grateful to my wife, Elisabeth, for her understanding and support.

Finally, I would like to thank my mother for her support and lifelong encouragement.

Dōmo! [Bow].

Stockholm, July 2008
Mikael Elhouar

Contents

Introduction

This thesis is composed of three papers dealing with different areas of interest rates theory. This introduction will situate the papers in the broader context of mathematical models of interest rates and highlight the common themes.

The first mathematical models of interest rates were based on the specification of a stochastic differential equation (SDE) for the dynamics of the instantaneous short rate process r. These models, commonly referred to as *short rate models*, made their appearance some years after the now famous model of Black and Scholes (1973) for the pricing of European stock options. By assuming that the stock price follows a geometric Brownian motion (GBM), Black and Scholes were able to derive a closed form formula for the price of a European put and call. It was therefore natural to use the same modelling idea to obtain prices for options on bonds.

A pioneering model in this area is the model by Vasiček (1977), where the instantaneous short rate is assumed to follow an Orstein-Uhlenbeck process. This is a mean reverting Gaussian process which is easier to handle analytically than a GBM, particularly when it comes to pricing bonds. From an economic standpoint it also seems better suited for capturing the time series properties of interest rates. The major contribution in this model however does not lie in the modelling of the short rate itself but in the solution it provides to the issue that interest rates are not tradable assets.

One weakness of the Vasiček (1977) model is that a Gaussian process allows the short rate to become negative. Dothan (1978) therefore proposed to use a GBM,thereby guarantying positive interest rates. Cox et al. (1985) proposed a squared Gaussian process, which also provides positive interest rates. In order to improve the quality of the calibration of the models to real world data, short rate models with time-dependent parameters were introduced such as those by Ho and Lee (1986) and Hull and White (1990).

One criticism of the short-rate models is that a single state variable drives the entire term structure of interest rates. Models with two or more state variables were therefore a natural next step in the modelling effort. This may also have been a response to the demands of practitioners for increased flexibility. One also finds a parallel trend in the modelling of stock prices with the development of stochastic volatility models. Famous examples of two-factor interest models are the models of Longstaff and Schwartz (1992), Hull and White (1994), and Duffie and Kan (1996).

In **Paper 3** of this thesis, a Gaussian two-factor model from Brigo and Mercurio (2006) is used to investigate, from a theoretical perspective, the various expectations hypotheses concerning interest rates that have been extensively studied in the empirical literature. Similar studies have been conducted previously using the short rate models of Vasiček (1977), in Campbell (1986), and Cox et al. (1985), in Frachot and Lesne (1993).

After the short rate models and the multiple factor models a new strand in the literature emerged with the introduction of the forward rate models with the paper by Heath et al. (1992). Their idea is to specify the dynamics of the forward rate curve, which is an infinite dimensional object, with the help of an SDE. In practice this mean choosing a particular functional form for the drift and volatility of the SDE. But since there is a one-to-one correspondence between forward rates and bond prices, a model of this type in fact specifies the dynamics of infinitely many assets. There is therefore a major risk that arbitrage opportunities are introduced in the model. It turns out that the condition of absence of arbitrage places very strong restriction on the model, namely that the drift of the SDE is entirely determined by the volatility (the so called Heath-Jarrow-Morton (HJM) drift condition).

Modelling in the HJM setting therefore amounts to choosing a functional form for the volatility. As an example if we set the volatility to a constant we obtain the Ho and Lee (1986) (one-factor) short-rate model. Choosing a particular exponential function produces the Hull and White (1990) short rate model. The problem however is that other functional forms may produce a non-Markovian and therefore infinite dimensional stochastic system. As we are modelling an infinite dimensional object, obtaining a finite dimensional model should be the exception rather than the rule.

At the end of the 1990's a number of papers appeared addressing the question of how the volatility should be specified in order to produce a HJM model possessing a finite-dimensional realization (FDR): Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997), Inui and Kijima (1998), Chiarella and Kwon (2001). These papers provided different examples of volatilities producing FDRs. Björk and Svensson (2001), building on previous work by Björk and Christensen (1999), were the first to produce both necessary and sufficient conditions on the volatility to guarantee the existence of a FDR. This work, which was at first only treated deterministic volatilities, was later extended in Björk et al. (2004) to also include stochastic volatilities driven by a diffusion.

Paper 1 in this thesis uses the differential geometric framework provided by Björk and Christensen (1999) and Björk and Svensson (2001) to study the FDR of HJM models with a volatility driven by a Markov chain, a so called regime-switching volatility. Both necessary and sufficient conditions on the volatility that guarantee the existence of a FDR are provided. Then some commonly used deterministic volatilities, known to provide a one-dimensional realization, are made stochastic and their FDR studied.

Despite the increased complexity of the interest models practitioners still use a formula based on an extension of the Black (1976) model when pricing caps and swaps. This pricing formula is based on the famous Black and Scholes (1973) model for the pricing of European stock options. The derivation of this formula is however not satisfactory from a theoretical point of view. At one point of the proof the interest rate is assumed to be constant and at another it is assumed to be stochastic.

Another issue with short rate models and forward rate models is that the underlying variables are infinitesimal quantities which are not observable. This gives rise to a lot of calibration problems when the model are implemented. This may be one reason for the persistent use by market practitioners of the theoretically unfounded formulas based on Black (1976).

In the end of the 1990's models of a new kind, so called market models, were introduced to provide a solution to the mentioned shortcomings of the interest rate models of the previous decades. In the LIBOR market model, introduced in the work of Miltersen et al. (1997) and Brace et al. (1997), the LIBOR rates are used as underlying stochastic state variables. By assuming the rates to be log-normally distributed, a formula of the Black (1976) type for the pricing of caplets is derived which this time is put on firm theoretical ground. Jamshidian (1997) introduced the swap market model, where the underlying variable is the swap rate. By assuming the forward swap rates to be log-normally distributed a Black (1976) type formula for swaptions is derived.

The major shortcoming of the market models is that they tend to have a large dimension. When pricing an interest rate derivative in this framework it is necessary to use one state variable (LIBOR or swap rate) for every cash flow. The pricing of a contract with 10 years to maturity and quarterly payments, which is not uncommon on the market, will require a 39-dimensional market model.

Paper 2 in this thesis tries to address the dimension problem of the market models with the help of the geometrical framework of Björk and Svensson (2001). Although this framework was originally developed to study FDRs of infinite dimensional models it can also be used in a finite dimensional setting. The markets models are viewed as stochastic systems and the framework is used to examine if there exist a realization of the these systems with a smaller dimension.

The mathematical modelling of interest rates has of course produced other important types of models which have not been mentioned in this introduction. There are two important omissions. Firstly the models by Flesaker and Hughston (1996) and Rogers (1994) which link bond pricing with stochastic potential theory and which guarantee positive interest rates. Secondly there is the Markov functional models introduced in Hunt et al. (2000).

Bibliography

Bhar, Ramaprasad, and Carl Chiarella, 1997, Transformation of Heath-Jarrow-Morton models to Markovian systems, *The European Journal of Finance* 3, 1–26.

Björk, Tomas, and Bent Jesper Christensen, 1999, Interest rate dynamics and consistent forward rate curves, *Mathematical Finance* 9, 323–348.

Björk, Tomas, Camilla Landén, and Lars Svensson, 2004, Finite dimensional Markovian realizations for stochastic volatility forward rate models, *Proceedings of the Royal Society, Series A* 460, 53–83.

Björk, Tomas, and Lars Svensson, 2001, On the existence of finite-dimensional realizations for nonlinear forward rate models, *Mathematical Finance* 11, 205–243.

Black, Fisher, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3, 167–179.

Black, Fisher, and Myron S. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Brace, Alan, Dariuz Gatarek, and Marek Musiela, 1997, The market model of interest rate dynamics, *Mathematical Finance* 7, 127–154.

Brigo, Damiano, and Fabio Mercurio, 2006, *Interest Rate Models – Theory and Practice* (Springer).

Campbell, John Y., 1986, A defence of traditional hypotheses about the term structure of interest rates, *Journal of Finance* 41, 183–193.

Chiarella, Carl, and Oh Kang Kwon, 2001, Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model, *Finance and Stochastics* 5, 237–257.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Dothan, L. Uri, 1978, On the term structure of interest rates, *Journal of financial economics* 6, 59–69.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Flesaker, Bjørn, and Lane Hughston, 1996, Positive interest, *Risk Magazine* 9, 46–49.

Frachot, Antoine, and J.P. Lesne, 1993, Expectations hypothesis and stochastic volatilities, notes d'études et de recherche 24, Banque de France.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates, *Econometrica* 60, 77–105.

Ho, Thomas, and Sang-Bin Lee, 1986, Term structure movements and pricing interest rate contingent claims, *Journal of Finance* 41, 1011–1029.

Hull, John, and Alan White, 1990, Pricing interest-rate derivative securities, *Review of Financial Studies* 3, 573–592.

Hull, John, and Alan White, 1994, Numerical procedures for implementing term structure models ii: Two-factor models, *The Journal of Derivatives* 2, 37–48.

Hunt, P. J., J. E. Kennedy, and Antoon Pelsser, 2000, Markov-functional interest rate models, *Finance and Stochastics* 4, 391–408.

Inui, Koji, and Masaaki Kijima, 1998, A Markovian framework in multi-factor Heath-Jarrow-Morton models, *Journal of Financial and Quantitative Analysis* 33, 423–440.

Jamshidian, Farshid, 1997, Libor and swap market models and measures, *Finance and Stochastics* 1, 293–330.

Longstaff, Francis A., and E. S. Schwartz, 1992, A two-factor interest rate model and contingent claims valuation, *The Journal of Fixed Income* 3, 16–23.

Miltersen, Kristian R., Klaus Sandmann, and Dieter Sondermann, 1997, Closed form solutions for term structure derivatives with log-normal interest rates, *Journal of Finance* 52, 409–430.

Ritchken, Peter, and L. Sankarasubramanian, 1995, Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance* 5, 55–72.

Rogers, L.C.G., 1994, The potential approach to the term structure of interest rates, *Mathematical Finance* 7, 157–176.

Vasiček, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 177–188.

Summaries

Paper 1: Finite-dimensional Realizations of Regime-Switching HJM Models

Another version of this paper has appeared in
Applied Mathematical Finance, Vol.15, No. 4, 331–354, August 2008.

This paper studies Heath-Jarrow-Morton-type models with regime-switching stochastic volatility. In this setting the forward rate volatility is allowed to depend on the current forward rate curve as well as on a continuous time Markov chain y with finitely many states. Employing the framework developed by Björk and Svensson we find necessary and sufficient conditions on the volatility guaranteeing the representation of the forward rate process by a finite dimensional Markovian state space model. These conditions allow us to investigate regime-switching generalizations of some well-known models such as those by Ho-Lee, Hull-White, and Cox-Ingersoll-Ross.

Paper 2: Minimal Realizations of Market Models

Joint work with Tomas Björk and Damir Filipović

Using the differential geometric framework of Björk and Svensson we study realizations of the LIBOR market model and the swap market model of interest rates. The volatilities are allowed to be both constant and time-dependent. We show that there does not exist realizations of a smaller dimension than that of the original model neither for the LIBOR market model nor for the swap market model.

Paper 3: Expectations Hypotheses in a Gaussian Two-Factor Model

Using an additive Gaussian two-factor model of the instantaneous short rate we study the expectations hypotheses of interest rates as defined by Cox-Ingersoll-Ross. The expectations hypothesis are studied both under the risk-neutral and the objective probability measure. The Girsanov kernel for the measure transformation is allowed to depend on the state variables. We show that, when the Girsanov kernel is stochastic, all the expectations hypotheses fail to hold under the objective probability measure, even if we allow for an additive premium as suggested by Campbell. Frachot-Lesne have stated that it is necessary to use a model with stochastic volatility to obtain this result. Our findings contradict their statement.

Paper 1

Paper 1

Finite-dimensional Realizations of Regime-Switching HJM Models

Another version of this paper has appeared in
Applied Mathematical Finance, Vol. 15, No. 4, 331–354, August 2008.

1.1 Introduction

The modelling of the term structure of interest rates in the Heath-Jarrow-Morton (HJM) framework (Heath, Jarrow, and Morton, 1992), usually begins with the specification of the volatility term in the stochastic differential equation (SDE) for the forward rate.[1] By assuming absence of arbitrage it can be shown that the drift term is completely determined by the volatility under the risk neutral measure (the HJM drift condition). Thus, to obtain a specific model of the term structure in this framework one only has to specify the volatility.

The object that is modelled, the forward rate curve, is an infinite dimensional object, so the SDE that we are specifying is in fact infinite dimensional. From this perspective there is no reason why the solution to this SDE can be realized by a finite-dimensional state space model. It can of course happen. To take a concrete example, if we set the volatility

[1] For an overview of the HJM framework see, for example, chapter 23 in Björk (2004).

to be a constant, then one can show that this gives rise to a one dimensional model equivalent to the Ho and Lee (1986) model for the short rate fitted to the initial term structure. For more complicated volatility functions however we are not *a priori* certain to obtain a model that has a finite-dimensional representation. In fact, considering the inherent infinite dimensionality of the problem, this should be the exception rather than the rule.

In the end of the 1990s, numerous papers address the question of how one should to specify the volatility in order to guaranty that the model produced had a finite-dimensional realization (henceforth FDR): Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997), Inui and Kijima (1998), and Chiarella and Kwon (2001).[2] These paper provide examples of volatilities known to yield a FDR. Björk and Svensson (2001), building on earlier results by Björk and Christensen (1999), were the first to provide both necessary and sufficient conditions for the existence of a FDR. To complement this pure existence result, Björk and Landén (2002) address the question of the actual construction of the realization.

The present paper uses the geometric framework developed by Björk and coauthors to investigate HJM models with regime-shifting stochastic volatility. Here the volatility is a function of a continuous time Markov chain with a finite number of states and of the current forward rate curve. The main purpose is to provide necessary and sufficient conditions on the volatility to guaranty a FDR.

From a modeling perspective the use of a Markov chain could be motivated by the need to make the model more flexible without making it intractable. There could also be economic reasons why an interest rate model driven by a Markov chain is attractive. Business cycle expansions and contraction can potentially have a first-order impact on expectations concerning inflation, monetary policy, and nominal interest rates. Such regime shifts could have a major impact on the whole term structure of interest rates.

There is a large number of empirical studies where one tries to fit the dynamics of bond yields (short rates in particular) using discrete-time econometric models with regime-switching: Hamilton (1988), Garcia and Perron (1996), Gray (1996), and Ang and Bekaert (2002), to name a few. These studies suggest that regime-switching models capture historical data better than single regime models. Some authors propose so-called dynamic term structure models in continuous time; Naik and Lee (1997), Bansal and Zhou (2002), Dai, Singleton, and Yang (2007). When estimated, these models provide additional empirical support for the regime-switching models. Moreover, the inferred regimes appear to be associated with indicators

[2]See Björk (2003) for an overview.

of business cycles. While containing some theoretical results, these papers are, however, mostly concerned with econometric issues.

Previous attempts to introduce regime shifts in the parameters of theoretical interest models are quite recent and have been mainly concerned with short rate models. Hansen and Poulsen (2000) extend the Vasiček (1977) model by allowing for shifts in the local mean of the short rate between a high and a low state. Numerical methods are then used to compute bond prices. Landén (2000) considers a short rate model where both the drift and volatility are driven by a Wiener process and a marked point process. For the special case when the underlying process is a Markov chain with only two states, a closed form expression for the bond price is derived.

Valchev (2004) seems to be the first and, to our knowledge, the only author to introduce regime-shifting coefficients in a HJM type of model. He first provides a semi-martingale representation of the Markov chain modulated volatility. Then he considers some explicit representations of the volatility, and derives the corresponding short rate dynamics. The paper is restricted to Gaussian models, i.e. models where the volatility does not depend on the current forward rate curve. The question of the existence of FDRs is not discussed and no realizations are constructed.

The paper by Björk, Landén, and Svensson (2004) is related to the present one in the sense that it treats FDRs of stochastic volatility models within the same geometric framework. The difference is that the authors consider a volatility driven by a diffusion, and that this is essential for the methods employed. In the present case with Markov chain modulated volatility the geometric picture is different and therefore needs to be tackled by other means.

The rest of the paper is organized as follows. Section 1.2 provides the theoretical foundations needed in the rest of the paper. In particular the FDR problem is formalized and a very general theorem on the FDR of regime-switching volatility models is proved. After these quite abstract results, more concrete matters are discussed. Section 1.3 treats the case of deterministic volatility in detail. This allow us to obtain concrete results concerning the possible choices of volatilities and their FDRs. The regime-switching generalizations of Ho and Lee (1986) model and of the Hull and White (1990) model are then presented and their FDR derived. Section 1.4 treats the special case of separable volatility. Possible regime-switching models of this class are discussed and finally the example of the generalized Cox, Ingersoll, and Ross (1985) model is analyzed.

1.2 Model

As in Heath, Jarrow, and Morton (1992) we consider a bond market modelled on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, Q)$ where Q is the risk-neutral martingale measure. The market is assumed to be default and arbitrage free. The space carries a one dimensional Wiener process W, and a m-dimensional continuous-time Markov chain $\{y_t\}$, $t \geq 0$, the finite state space of which is the set $S = \{e_1, \ldots, e_m\}$ of unit vectors on $\mathbb{R}^m$.[3] We restrict ourselves to a scalar driving Wiener process to simplify notation by avoiding to differentiate between indices related to Wiener processes and states of the Markov chain. The restriction has no bearing on the results.

The price at time t of a zero coupon bond maturing at $t + x$ is denoted $p(t, x)$. Notice the Musiela parametrization,[4] which uses the time *to* maturity x instead of the usual time *of* maturity T. The instantaneous forward rate is defined as

$$r_t(x) := -\frac{\partial}{\partial x} \log p(t, x),$$

i.e. the price at time t of a zero-coupon bond maturing at $t + x > t$ is given by $p(t, x) = \exp\{-\int_0^x r_t(s)\, ds\}$.

The modelling will focus on the forward rate curve, $x \mapsto r_t(x)$, rather than on the individual forward rate $r_t(x)$. To stress this fact the forward rate curve is denoted r_t. The forward rate process $\{r_t\}$ is thus viewed a stochastic process taking values in the Hilbert space $\mathcal{H}$ of forward rate curves.[5]

The present paper studies a HJM model of the forward rate where the volatility σ is modulated by the continuous time Markov chain y. We take as given a smooth volatility function σ of the form

$$\sigma : \mathcal{H} \times S \times \mathbb{R}_+ \to \mathbb{R},$$

that is $\sigma(r, y, x)$ is a functional of the infinite-dimensional r-variable, the continuous time Markov chain y taking values in the finite set S, and the positive real variable x.

Definition 1.1. *The forward rate is given by the following stochastic dif-*

[3] The notation used for important variables is listed at the end of the article.

[4] See Musiela (1993). The use of the Musiela parametrization is motivated by the emphasis on the forward rate curve $x \mapsto r_t(x)$ as an infinite dimensional object in the framework of Björk and Svensson (2001).

[5] For details concerning the choice of space, see Björk (2003) and Filipović and Teichmann (2003).

ferential equation

$$\begin{cases} dr_t(x) &= \mu_0(r_t, y_t, x)\,dt + \sigma(r_t, y_t, x)dW(t), \\ r_0(x) &= r_0^o(x), \end{cases} \tag{1.1}$$

where μ_0 *is given by the well known HJM drift condition emanating from the assumption of absence of arbitrage:*

$$\mu_0(r_t, y_t, x) = \frac{\partial}{\partial x} r_t(x) + \sigma(r_t, y_t, x) \int_0^x \sigma(r_t, y_t, s)\,ds.$$

As we will be using results from differential geometry, for instance the Frobenius theorem, it is convenient to formulate the model of the forward rate in terms of Stratonovich integrals instead of the usual Itō integrals. This allows us to use the theorems directly without having to translate them beforehand into the Itō setting.

The Stratonovich formulation of the model is given by the following stochastic differential equation:

$$\begin{cases} dr_t(x) &= \mu(r_t, y_t, x)\,dt + \sigma(r_t, y_t, x) \circ dW(t), \\ r_0(x) &= r_0^o(x), \end{cases} \tag{1.2}$$

where $\circ$ indicates a Stratonovich differential. The HJM drift condition now gives

$$\begin{aligned} \mu(r_t, y_t, x) = \frac{\partial}{\partial x} r_t(x) &+ \sigma(r_t, y_t, x) \int_0^x \sigma(r_t, y_t, s)\,ds, \\ &- \frac{1}{2} \sigma_r'(r_t, y_t, x)[\sigma(r_t, y_t, x)]. \end{aligned} \tag{1.3}$$

Here $\sigma_r'(r_t, y_t)[\sigma(r_t, y_t)]$ denotes the Fréchet derivative $\sigma_r'(r_t, y_t)$ operating on $\sigma(r_t, y_t)$. Setting

$$\mathbb{F} := \frac{\partial}{\partial x}, \tag{1.4}$$

$$\Phi(r_t, y_t, x) := \sigma(r_t, y_t, x) \int_0^x \sigma(r_t, y_t, s)\,ds, \tag{1.5}$$

$$\mathcal{C}(r_t, y_t, x) := \sigma_r'(r_t, y_t, x)[\sigma(r_t, y_t, x)], \tag{1.6}$$

we can write the drift more compactly as

$$\mu(r_t, y_t, x) = \mathbb{F} r_t(x) + \Phi(r_t, y_t, x) - \frac{1}{2} \mathcal{C}(r_t, y_t, x). \tag{1.7}$$

Henceforth we will use this Stratonovich formulation of the model.

Remark 1.1. *When the HJM drift condition, as expressed in* (1.7)*, is formulated using the usual parametrization and Itō integrals it only consists of the term* Φ*. The present setting yields two more terms. The term* $\mathbb{F}r_t$ *comes from the use of the Musiela parametrization. The term,* $-1/2\,\mathcal{C}$*, comes from the use of Stratonovich integrals.*

Remark 1.2. *To simplify notation we define*

$$\begin{aligned}\mu_i &:= \mu(r_t, e_i, x),\\ \sigma_i &:= \sigma(r_t, e_i, x),\\ \Phi_i &:= \Phi(r_t, e_i, x).\end{aligned}$$

where $i \in \{1, 2, \ldots, n\}$*.*

Remark 1.3. *We will use the following notation when treating linear vector spaces. If* v_1 *and* v_2 *are vectors we will denote the space spanned by* v_1 *and* v_2 *by*

$$\langle v_1, v_2\rangle.$$

The Lie algebra generated by two vector fields v_1 *and* v_2 *will be denoted*

$$\langle v_1, v_2\rangle_{LA}.$$

1.2.1 Problem Formulation

The main question to be answered in the present paper is when does the regime-switching volatility model in (1.2) admits a finite-dimensional realization (FDR). To be more specific we present the following definition.

Definition 1.2 (FDR)**.** *The SDE in* (1.2) *is said to have a* finite dimensional realization *if there exist smooth* d*-dimensional vector fields* a *and* b*, initial points* $z_0 \in \mathbb{R}^d$ *and* $y_0 \in I$*, and a (smooth) mapping* $G : \mathbb{R}^d \to \mathcal{H}$ *such that* r_t *has a representation of the following form*

$$r_t(x) = G(z_t, x), \tag{1.8}$$

$$dz_t = a(z_t, y_t)\, dt + b(z_t, y_t) \circ dW_t, \tag{1.9}$$

$$z_0 = z^0. \tag{1.10}$$

Remark 1.4. *Notice that, in the definition above, the Markov chain* y *only appears in the dynamics of the state variable* z_t *and not in the mapping* G*. Setting* $r_t = G(z_t, y_t)$ *would mean that* r_t *is allowed to jump with the Markov chain, but this would be in conflict with the representation of dynamics of* r_t *in* (1.2)*, where* y_t *only affect the coefficients but no jump term is present.*

Remark 1.5. *The model is formulated under the risk neutral probability measure Q in* (1.2). *But we could as well have formulated the model under the objective probability measure P. If the equality in* (1.8) *holds Q almost surely it also holds P almost surely as P and Q are equivalent martingale measures. Formulating the model directly under Q has the advantage of leaving the concept of market price of risk, which has no impact on our results, outside of the modelling.*

The framework developed in Björk and Christensen (1999) and Björk and Svensson (2001) allows us to formulate the following theorem which answer the question concerning the existence of a finite-dimensional realizations.

Theorem 1.1. *The forward rate model in* (1.2) *admits a finite-dimensional realization if and only if the Lie algebra $\mathscr{L}$ generated by $\langle \mu_i, \sigma_i : i \in I \rangle$ denoted $\langle \mu_i, \sigma_i : i \in I \rangle_{LA}$ is finite dimensional at r_0. Here μ_i and σ_i are viewed as vector fields on $\mathcal{H}$.*

Proof. This proof follows the proofs of Theorem 4.1 in Björk and Christensen (1999) and Theorem 3.2 in Björk and Svensson (2001). It turns out that the introduction of the Markov chain y does not alter the original theorem. Assume that the model admits a FDR. Differentiation of (1.8) with the help of the Itō formula and use of (1.9) gives

$$dr_t = G'_z(z_t)a(z_t, y_t)\,dt + G'_z(z_t)b(z_t, y_t) \circ dW_t.$$

Comparing this with (1.2) produces

$$\begin{aligned}\mu(r_t, y_t) &= G'_z(z_t)a(z_t, y_t),\\ \sigma(r_t, y_t) &= G'_z(z_t)b(z_t, y_t).\end{aligned}$$

The forward curve manifold $\mathcal{G} = \operatorname{Im} G = \{G(z, \cdot) \in \mathcal{H}; z \in \mathbb{R}^d\}$ is by assumption finite dimensional. The tangent space of the manifold $\mathcal{G}$ at the point z, denoted $T_{\mathcal{G}}(z)$, is therefore also finite dimensional. We observe that $\operatorname{Im} G'_z(z) = T_{\mathcal{G}}(z)$, so for every $z \in \mathbb{R}^d$ and every $y \in S = \{e_1, e_2, \dots e_m\}$ we have

$$\begin{aligned}\mu(r, y) &\in T_{\mathcal{G}}(z),\\ \sigma(r, y) &\in T_{\mathcal{G}}(z),\end{aligned}$$

where $r = G(z)$. By standard theory[6] we know that the Lie bracket of two vector fields tangential to a manifold is also tangential to the manifold, so we also have $[\mu_i(r), \sigma_j(r)] \in T_{\mathcal{G}}(z)$ for all $i, j \in I$. We can therefore conclude that the Lie algebra is finite dimensional.

[6] See Warner (1971) or Olver (1986), Lemma 1.37.

Assume that the Lie algebra is finite dimensional with, say, dimension n. Frobenius' theorem supplies a n-dimensional (smooth) embedded submanifold $\mathcal{G}$ with a local parametrization provided by a mapping $G : \mathbb{R}^n \to \mathcal{H}$, i.e. $\mathcal{G} = \operatorname{Im} G = \{G(z,\cdot) \in \mathcal{H}, z \in \mathbb{R}^d\}$, so that for $r = G(z)$, $\mu_i(r) \in T_{\mathcal{G}}(z)$ and $\sigma_i(r) \in T_{\mathcal{G}}(z)$ for all $i \in I$. We first want to find a factor process z_t with dynamics as in (1.9) and, second, show that the mapping G actually gives us r_t as $r_t = G(z_t)$. Recall that $T_{\mathcal{G}}(z) = \operatorname{Im} G'_z(z)$, so $\mu_i(r), \sigma_i(r) \in \operatorname{Im} G'_z(z)$. We can therefore find, for each $i \in I$, $a_i(z)$ and $b_i(z)$ such that $\mu_i(r) = G'_z(z)a_i(z)$, and $\sigma_i(r) = G'_z(z)b_i(z)$, $r = G(z)$, or using more compact notation

$$\begin{aligned} \mu(r,y) &= G'_z(z)a(z,y), \\ \sigma(r,y) &= G'_z(z)b(z,y). \end{aligned} \tag{1.11}$$

The injectivity of G', following from Frobenius' theorem, guarantees the uniqueness of $a_i(z)$ and $b_i(z)$. As $G'_z(z)$ has a smooth left inverse $\big(G'_z(z)\big)^{-1}$, we can solve for $a(z,y)$ and $b(z,y)$ and obtain

$$\begin{aligned} a(z,y) &= \big(G'_z(z)\big)^{-1}\mu(r,y), \\ b(z,y) &= \big(G'_z(z)\big)^{-1}\sigma(r,y), \end{aligned}$$

where $r = G(z)$. The smoothness of G and of the inverse $\big(G'_z(\zeta)\big)^{-1}$ gives smoothness of $a(z,\cdot)$ and $b(z,\cdot)$ and therefore local Lipschitz continuity, yielding a unique strong solution to the SDE (1.9) for arbitrary initial condition $z_0 \in \mathbb{R}^d$ and $y_0 \in S$. This completes the quest for the factor process z_t. To show that $r_t = G(z_t)$, define $q_t = G(z_t)$. Differentiation gives

$$dq_t = G'_z(z_t)a(z_t,y_t)\,dt + G'_z(z_t)b(z_t,y_t)\,dt.$$

Using (1.11) we see that this equation is identical to (1.2). Given an arbitrary $r_0 \in \mathcal{G}$, since $\mathcal{G} = \operatorname{Im} G(z)$ and $G(z)$ is injective there is a unique z_0 such that $r_0 = G(z_0)$. If we now set $q_0 = z_0$, by uniqueness of the strong solution to (1.2) we obtain that $r_t = q_t = G(z_t)$. □

Remark 1.6. *The finiteness of state space of the Markov chain is not needed to prove the theorem. When proving necessity the finite dimensionality of the Lie algebra follows from the finite dimensionality of the manifold $\mathcal{G}$, not from the number of states in the Markov chain. When proving sufficiency, we assume the Lie algebra to be finite dimensional, so we implicitly assume that the only finite many μ_i and σ_i are linearly independent, regardless of the number of states in the Markov chain.*

1.2.2 Constructing realizations

Having settled the question of existence of a finite-dimensional realization, the next natural step is to construct such a realization. We will need the following definition to describe the parametrization of forward curve manifold.

Definition 1.3. *Let f be a smooth vector field on the space $\mathcal{H}$, and let y be an arbitrary point on $\mathcal{H}$. The solution to the following ODE*

$$\frac{dy_t}{dt} = f(y_t), \qquad y_0 = y,$$

will be denoted

$$y_t = e^{tf} y.$$

We can now describe the manifold using Theorem 4.2 of Björk and Svensson (2001) which we reproduce here for convenience.

Theorem 1.2. *Assume that the Lie algebra $\langle \mu, \sigma \rangle_{LA}$ is spanned by the smooth vector fields $f_1, \dots, f_d$. Then, for the initial point r^0, all forward rate curves produced by the model will belong to the induced tangential manifold $\mathcal{G}$, which can be parametrized as $\mathcal{G} = \mathrm{Im}[G]$, where*

$$G(z_1, \dots, z_d) = e^{z_d f_d} \cdots e^{z_1 f_1} r^0,$$

and where the operator $e^{z_i f_i}$ is defined as above.

Remark 1.7. *Notice that the preceding theorem shows that the dimension of the realization is equal to the size of the Lie algebra.*

To construct the finite-dimensional realization we follow the procedure outlined in Björk and Landén (2002). Given a volatility $\sigma(r, y, t)$ for which the Lie algebra $\mathscr{L}$ is finite dimensional and given an initial forward rate curve r^0 we perform the following steps:

(i) Select a finite number of vector fields $f_1, \dots, f_d$ spanning the Lie algebra $\mathscr{L}$.

(ii) Determine the manifold $\mathcal{G}$ by using Theorem 1.2 above.

(iii) Set $r_t = G(z_t)$. Make the following *Ansatz* for the dynamics of the state variable z

$$dz_t = a(z_t, y_t)\, dt + b(z_t, y_t) \circ dW_t.$$

Using the Stratonovich version of the Itō formula yields

$$dr_t = dG(z_t) = G'(z_t)\, a(z_t, y_t)\, dt + G'(z_t)\, b(z_t, y_t) \circ dW_t. \tag{1.12}$$

Identifying the coefficients of the drift and diffusion vector fields of the r process we obtain

$$\begin{aligned} G'(z_t)a(z_t, y_t) &= \mu(G(z_t), y_t) \\ G'(z_t)b(z_t, y_t) &= \sigma(G(z_t), y_t). \end{aligned} \tag{1.13}$$

These equations can now be used to solve for the vector fields a and b.

Remark 1.8. *It can sometimes be very burdensome to find the* minimal *set of vector fields spanning the Lie algebra. In those cases it is more practical to work with a slightly larger set of vector fields. Occasionally this can cause the equations in* (1.13) *to have multiple solutions. This is of no consequence since it's enough to find one solution, and any solution will be adequate for our purpose.*

1.3 Deterministic volatility

We can now use the abstract results above to derive concrete results for particular model choices. In this section we consider the simple case where the volatility does not depend on r. We first derive the Lie algebra, then we determine the invariant manifold and find the state variable dynamics. Finally, we study regime-switching generalizations of the Ho-Lee and Hull-White models.

1.3.1 The Lie algebra

We assume that the volatility is of the form

$$\sigma(r, x, y) = \sigma(x, y), \tag{1.14}$$

i.e. the volatility does not depend on r. When the Markov chain is in state e_i the volatility $\sigma_i(x)$ is a constant vector field on $\mathcal{H}$.

Using the notation in (1.7) for the drift μ, the generators of the Lie algebra will be the following vector fields (one for each state of the Markov chain y):

$$\begin{aligned} \sigma(r, x, y) &= \sigma_i, \qquad\qquad i \in I, \\ \mu(r, x, y) &= \mathbb{F}r + \Phi_i. \end{aligned}$$

Because the σ_i:s are constant as functions of r the Fréchet derivatives $\sigma_i' = 0$ and therefore $\mathcal{C} = \sigma_r'[\sigma] = 0$ in (1.7).

To compute Lie brackets we need the Fréchet derivative of μ_i which is $\mu_i' = \mathbb{F}$ as $\mathbb{F}$ is a linear operator (and in the space $\mathcal{H}$ also a bounded[7] operator).

Taking Lie brackets we obtain

$$\begin{aligned}
[\mu_i, \mu_j] &= \mathbb{F}\big(\mathbb{F}r + \Phi_j\big) - \mathbb{F}\big(\mathbb{F}r + \Phi_i\big) = \mathbb{F}\big(\Phi_j - \Phi_i\big), \\
[\mu_i, \sigma_j] &= \mathbb{F}\sigma_j, \\
[\sigma_i, \sigma_j] &= 0,
\end{aligned}$$

where $i, j \in I$. Now the only remaining non-constant vector field is μ_i, so computing the remaining Lie brackets is easy and yields the following system of generators for the Lie algebra:

$$\mathscr{L} = \langle \mathbb{F}r + \Phi_i, \mathbb{F}^k\sigma_i, \mathbb{F}^\ell(\Phi_i - \Phi_j); i, j \in I; k, \ell = 0, 1, \ldots \rangle_{LA}. \tag{1.15}$$

As we are interested in a FDR of the model, we want to know when the Lie algebra $\mathscr{L}$ is finite dimensional (remember Remark 1.7). Now for $\mathscr{L}$ to be finite dimensional, first the space $\langle F^k\sigma_i; k = 0, 1, \ldots \rangle$ must be finite-dimensional, for every fixed $i \in I$. But for this to occurs, the vector fields $\mathbb{F}^k\sigma_i$ must satisfy a linear relation of the form

$$\mathbb{F}^{n_i}\sigma_i = \sum_{k=0}^{n_i-1} c_k^i \mathbb{F}^k \sigma_i, \tag{1.16}$$

where $c_k^i \in \mathbb{R}$. We recall that $\mathbb{F} := \partial/\partial x$, so that (1.16) is in fact an ODE with constant coefficients in the variable σ_i. By general ODE theory, σ_i is a solution to (1.16) if and only if σ_i is a quasi-exponential function.[8]

Definition 1.4 (QE function)**.** *A quasi-exponential function is a function of the form*

$$f(x) = \sum_i p_i(x)e^{\lambda_i x} + \sum_j e^{\alpha_j x}\big[p_j(x)\cos(\omega_j x) + q_j(x)\sin(\omega_j x)\big],$$

where $\lambda_i, \alpha_j, \omega_j$ are real numbers and p_i, p_j, q_j are polynomials with real coefficients.

The preceding argument shows that a necessary condition for the model to posses a FDR is that σ_i is QE. It turns out that this condition is also sufficient. To show this we will use the following Lemma which lists some well known the properties of QE functions.

[7]The boundedness of $\mathbb{F}$ is the reason underlying the choice of space $\mathcal{H}$. For details, consult Björk (2003) and Filipović and Teichmann (2003).

[8]See, for example, Schwartz (1992), chapter IV, §4.

Lemma 1.1. *Quasi-exponential functions have the following properties:*

(i) A function is QE if and only if it is a component of the solution of a vector-valued linear ODE with constant coefficients.

(ii) A function is QE if and only if it can be written as $f(x) = ce^{Ax}b$.

(iii) If f is QE, then f' is QE.

(iv) If f is QE, then its primitive function is QE.

(v) If f and g are QE, then fg is QE.

We now have the following result.

Proposition 1.1. *Assume that the volatility is deterministic i.e. of the form*

$$\sigma(r, y, x) = \sigma(y, x).$$

Then the model in (1.2) *possesses a FDR if and only if $\sigma(x, y)$ is a quasi-exponential function in the x variable.*

Proof. Having proved necessity, we assume that σ_i is QE for every $i \in I$. Then by (i) we obtain that $\langle \mathbb{F}^k \sigma_i; k = 0, 1, \ldots \rangle$ is finite dimensional. The function $\Phi_i = \sigma \int_0^x \sigma^\top \, ds$ is QE by (v) and (iv) since it is composed of a QE multiplied by the integral of a QE. In addition $\Phi_i - \Phi_j$ is QE as linear combinations of QE functions are QE. We can therefore also conclude that $\langle \mathbb{F}^{\ell+1}(\Phi_i - \Phi_j); \ell = 0, 1, \ldots \rangle$ is finite dimensional. □

Knowing that the volatility σ is quasi-exponential in the variable x allows us to give a simpler system of generators for the Lie algebra $\mathscr{L}$ than the one provided in (1.15). This will facilitate the calculation of the dimension of the Lie algebra in Proposition 1.3 and the construction of realization Section 1.3.3. We have the following result.

Proposition 1.2. *Assume that the volatility $\sigma_i(x)$ has the following form*

$$\sigma_i(x) = \sum_{j=1}^{N_i} p_{i,j}(x) \, e^{-\alpha_{i,j} x},$$

where $p_{i,j}(x)$ is a polynomial of degree $n_{i,j}$, $\alpha_{i,j} \in \mathbb{R}$, $\alpha_{i,j} \neq 0$ and $N_i \in \mathbb{N}$. Then

$$\langle \mu_i, \sigma_i \rangle_{LA} = \langle \mathbb{F}r, \mathbb{F}^k \sigma_i, \mathbb{F}^\ell \Phi_i; i \in I; k = 0, \ldots, \mathfrak{n}_1^i; \ell = 0, \ldots, \mathfrak{n}_2^i \rangle_{LA}, \quad (1.17)$$

where

$$\mathfrak{n}_1^i = \sum_{j=1}^{N_j} n_j^i + N_i - 1,$$

$$\mathfrak{n}_2^i = \sum_{\substack{(j,k)\\ j\le k}} (n_j^i + n_k^i) + \frac{1}{2}N_i(N_i+1) - 1.$$

We will prove this proposition in four steps, each step giving rise to a lemma. The first lemma shows that the linear space $\langle \mathbb{F}^k\sigma_i; k=0,\dots,\rangle$ is spanned by monomials $x^j e^{\alpha_k x}$. Although this lemma is not used explicitly in the proof of the proposition above, it is used to prove the next three lemmas. Before stating the second lemma we will substitute the vector field Φ_i occurring in (1.15) by the vector field $\tilde{\Phi}_i := \Phi_i - K\sigma_i$, where K is a specific constant. The second lemma then shows that the linear space $\langle \mathbb{F}^\ell(\tilde{\Phi}_i - \tilde{\Phi}_j); i,j\in I; \ell=0,\dots,\rangle$ is spanned by monomials $x^k e^{-(\alpha_{i,j}+\alpha_{i,k})x}$ thereby allowing us to disentangle the vector fields $\mathbb{F}(\tilde{\Phi}_i - \tilde{\Phi}_j)$ into $\mathbb{F}\tilde{\Phi}_i$ and $\mathbb{F}\tilde{\Phi}_j$. The third lemma shows that we can also disentangle the vector field $\mathbb{F}r + \tilde{\Phi}_i$ into $\mathbb{F}r$ and $\tilde{\Phi}_i$. Finally, in Proposition 1.3 we calculate the size of the Lie algebra $\mathscr{L}$, which enable us to specify $\mathfrak{n}_1^i$ and $\mathfrak{n}_2^i$ mentioned in the Proposition 1.2.

To avoid cluttering the notation with (unnecessary) double indices we will temporarily remove the dependence on y from the expressions. The first lemma used in the proof of Proposition 1.2 is as follows.

Lemma 1.2. *Assume that the volatility $\sigma(x)$ has the following form*

$$\sigma(x) = \sum_{j=1}^{N} p_j(x) e^{-\alpha_j x},$$

where $p_j(x)$ is a polynomial of degree n_j, $\alpha_j \in \mathbb{R}$, $\alpha_j \neq 0$ and $N \in \mathbb{N}$. Then

$$\begin{aligned}&\langle \mathbb{F}^k\sigma(x); k=0,1,\dots\rangle \\ &\quad= \langle \mathbb{F}^k\big(x^{n_j}e^{-\alpha_j x}\big); k=0,1,\dots; j=1,\dots,N\rangle.\end{aligned} \tag{1.18}$$

Proof. We start by showing that

$$\langle \mathbb{F}^k\sigma(x); k=0,1,\dots\rangle = \langle x^i\, e^{-\alpha_j x}; i=0,\dots,n_j; j=1,\dots,N\rangle. \tag{1.19}$$

The equality (1.18) then follows from the trivial fact that

$$\begin{aligned}&\langle \mathbb{F}^k\big(x^{n_j}\, e^{-\alpha_j x}\big); k=0,1,\dots; j=1,\dots N\rangle \\ &\quad= \langle x^i\, e^{-\alpha_j x}; j=1,\dots,N; i=0,\dots,n_j\rangle.\end{aligned}$$

To prove (1.19) we first have to show that every element on the left-hand side can be written as a linear combination of the elements on the right-hand side. But this is trivial since we know that σ is quasi-exponential. Conversely, we have to show that every monomial $x^i e^{-\alpha_j x}$ on the right-hand side of (1.19) can be written as a linear combination of the vector fields $\mathbb{F}^k\sigma$ on the left-hand side. To this end we will use the operator $(\mathbb{F}+\alpha_j)$ which lowers the degree of the polynomial $p_j(x)$. To make things as transparent as possible, we start by assuming that $N=1$, i.e. that $\sigma = p(x)e^{-\alpha x}$, where $p(x)$ is a polynomial of degree n. We then have $(\mathbb{F}+\alpha)\sigma = p'(x)e^{-\alpha x} + p(x)(-\alpha)e^{-\alpha x} + \alpha p(x)e^{-\alpha x} = p'(x)e^{-\alpha x}$ where $p'(x) = dp(x)/dx$.

Applying powers of the operator $\mathbb{F}+\alpha_1$ yields

$$\begin{aligned}
(\mathbb{F}+\alpha_1)^0\sigma(x) &= p_n(x)e^{-\alpha x}\\
(\mathbb{F}+\alpha_1)^1\sigma(x) &= p_{n-1}(x)e^{-\alpha x}\\
&\vdots\\
(\mathbb{F}+\alpha_1)^n\sigma(x) &= p_0(x)e^{-\alpha x},
\end{aligned}$$

where $p_i(x)$ denotes a polynomial of degree i. Having at our disposal one polynomial of each degree allow us to perform Gaussian elimination and thereby express each monomial $x^i e^{-\alpha x}$ as a linear combination of different powers of the operator $\mathbb{F}+\alpha$. It only remains to show that every power of the operator $\mathbb{F}+\alpha$ can itself be expressed as a linear combination of the vector fields $\mathbb{F}^k\sigma$. A simple application of the binomial theorem shows that

$$(\mathbb{F}+\alpha)^k\sigma(x) = \sum_{i=0}^{n}\binom{n}{i}\alpha^{n-i}\mathbb{F}^i\sigma(x),$$

and we are done with the case $N=1$.

Next we have to show that we can obtain the same result when $N>1$. To keep the notation simple we now treat the case $N=2$. To show that the result also holds for higher values of N is trivial. So now we assume that $\sigma = q_1(x)e^{-\alpha_1 x} + q_2(x)e^{-\alpha_2 x}$, where the $q_i(x)$ denotes a polynomial of degree n_i. Now note that the operator $\mathbb{F}+\alpha_1$ only lowers the degree of the polynomial in front of $e^{-\alpha_1 x}$ leaving the degree of the polynomial in front of $e^{-\alpha_2 x}$ unchanged, as can be seen in the following calculation:

$$(\mathbb{F}+\alpha_1)\sigma(x) = q_1'(x)e^{-\alpha_1 x} + [q_2'(x) + (\alpha_1-\alpha_2)q_2(x)]e^{-\alpha_2 x}.$$

This means that we can recover the monomials $x^i e^{-\alpha_1 x}$ by first annihilating the polynomial $q_2(x)$ from σ as $(\mathbb{F}+\alpha_2)^{n_2+1}\sigma = \pi_1(x)e^{-\alpha_1 x}$, where $\pi_1(x)$ is a polynomial of degree n_1. Then we just apply the same procedure as in the $N = 1$ case, by operating with powers of the operator $\mathbb{F} + \alpha_1$ thereby recovering all monomials $x^i e^{-\alpha_1 x}$. We can then apply the same procedure to recover all monomials $x^i e^{-\alpha_2 x}$ by first annihilating $q_1(x)$ with the help of $(\mathbb{F} + \alpha_1)^{n_1}$. □

Before stating the next lemma we will replace the vector fields Φ_i in the generator set of $\mathscr{L}$ in (1.15) with simpler vector fields $\tilde{\Phi}_i$. This will facilitate the analysis of the generator set $\langle \mathbb{F}^\ell(\Phi_i - \Phi_j); i,j \in I; \ell = 0,1,\ldots\rangle$ and allows us to make use of Lemma 1.2 in the proofs. Let us for a moment put aside the dependence on y and study $\Phi(x)$. To this end we partially integrate the integral of $\sigma(x)$ in the expression for $\Phi(x)$ as follows

$$\begin{aligned}\Phi(x) &= \Big(\sum_{i=1}^{N} p_i(x) e^{-\alpha_i s}\Big)\Big(\int_0^x \sum_{j=1}^{N} p_j(s) e^{-\alpha_j s} ds\Big)\\ &= \Big(\sum_{i=1}^{N} p_i(x) e^{-\alpha_i s}\Big)\Big(\sum_{j=1}^{N} q_j(x) e^{-\alpha_j x} + K\Big),\end{aligned}$$

where

$$\begin{aligned}q_j(x) &= -\left(\frac{p_j(x)}{\alpha_j} + \frac{\mathbb{F}p_j(x)}{\alpha_j^2} + \ldots \frac{\mathbb{F}^n p_j(x)}{\alpha_j^{n+1}}\right),\\ K &= \sum_{j=1}^{N}\Big(\sum_{i=0}^{n_j} \frac{1}{\alpha_j^{i+1}} \mathbb{F}^i p_j(0)\Big).\end{aligned}$$

By expanding the product we obtain

$$\Phi(x) = \sum_{(i,j)} p_i(x) q_j(x) e^{-(\alpha_i + \alpha_j)x} + K\,\sigma(x). \tag{1.20}$$

We note that the last term of this expression for Φ involves σ and therefore monomials of the form $x^i e^{-\alpha_j x}$. We now define

$$\begin{aligned}\tilde{\Phi}(x) &= \Phi(x) - K\sigma(x)\\ &= \sum_{(i,j)} p_i(x) q_j(x) e^{-(\alpha_i+\alpha_j)x}.\end{aligned} \tag{1.21}$$

We can now replace every occurrence of Φ_i in (1.15) with $\tilde{\Phi}_i$. This can be shown by simple use of the relation $\tilde{\Phi}_i = \Phi_i + K_i\sigma_i$. The vector fields $\mathbb{F}r + \tilde{\Phi}_i = (\mathbb{F}r + \Phi_i) - K_i\sigma_i$, so $\tilde{\Phi}_i \in \mathscr{L}$. We can therefore add them to the set of generators of $\mathscr{L}$. Next as $\tilde{\Phi}_i - \tilde{\Phi}_j = \Phi_i - \Phi_j - K_i\sigma_i + K_j\sigma_j$ we see that $\mathbb{F}^k(\tilde{\Phi}_i - \tilde{\Phi}_j) = \mathbb{F}^k(\Phi_i - \Phi_j) - K_i\mathbb{F}^k\sigma_i - K_j\mathbb{F}^k\sigma_j$, so $\mathbb{F}^k(\tilde{\Phi}_i - \tilde{\Phi}_j) \in \mathscr{L}$. We can therefore also add $\mathbb{F}^k(\tilde{\Phi}_i - \tilde{\Phi}_j)$ to the set of generators of $\mathscr{L}$.

The next lemma treats the generators $\mathbb{F}^\ell(\tilde{\Phi}_i - \tilde{\Phi}_j)$. We are therefore forced to reintroduce the index for the Markov chain y in the notation.

Lemma 1.3. *Assume that the volatility σ_i has the following form*

$$\sigma_i(x) = \sum_{j=1}^{N_i} p_{i,j}(x)e^{-\alpha_{i,j}x},$$

where $i \in I$, $p_{i,j}$ is a polynomial of degree $n_{i,j}$, $\alpha_{i,j} \in \mathbb{R}$, $\alpha_{i,j} \neq 0$ and $N \in \mathbb{N}$. Then

$$\begin{aligned}&\langle \mathbb{F}^\ell(\tilde{\Phi}_i - \tilde{\Phi}_j); i,j \in I; \ell = 0,1,\ldots\rangle \\ &\quad = \langle \mathbb{F}^\ell\tilde{\Phi}_i; i \in I; \ell = 0,1,\ldots,n_{i,j}+n_{i,k}\rangle \\ &\quad = \langle x^\ell e^{-(\alpha_i^j+\alpha_i^k)x}; i \in I; j,k = 0,1,\ldots; \ell = 0,\ldots,n_{i,j}+n_{i,k}\rangle.\end{aligned}$$

Proof. As $\tilde{\Phi}_i(x)$ has exactly the same functional form as $\sigma(x)$ we can apply Lemma 1.2 on $\tilde{\Phi}_i(x)$ and obtain

$$\begin{aligned}&\langle \mathbb{F}^\ell\tilde{\Phi}_i(x); i \in I; \ell = 0,1,\ldots\rangle \\ &\quad = \langle \mathbb{F}^\ell(x^{n_j+n_k}e^{-(\alpha_j+\alpha_k)x}); i \in I; j,k = 1,\ldots,N_i; \ell = 0,1,\ldots\rangle,\end{aligned}$$

which proves the second equality.

To prove the first equality we just note that $\tilde{\Phi}_i - \tilde{\Phi}_j$ also has the functional form of σ so Lemma 1.2 can be applied again. □

The next lemma disentangles the generators $\mathbb{F}r + \tilde{\Phi}_i$ into $\mathbb{F}r$ and $\tilde{\Phi}_i$. This turns out to be be important for the study of the invariant manifold in the next section, as it separates the effect of the vector field $\mathbb{F}r$, making calculations a lot easier and producing a neater parametrization.

Lemma 1.4. *Assume that the volatility σ_i has the form*

$$\sigma_i(x) = \sum_{j=1}^{N_i} p_{i,j}(x)e^{-\alpha_{i,j}x},$$

where $i \in I$, $p_{i,j}$ is a polynomial of degree $n_{i,j}$, $\alpha_{i,j} \in \mathbb{R}$, $\alpha_{i,j} \neq 0$ and $N \in \mathbb{N}$. Then

$$\langle \mathbb{F}r + \tilde{\Phi}_i; i \in I\rangle = \langle \mathbb{F}r, \tilde{\Phi}_i; i \in I\rangle.$$

Proof. Lemma 1.3 shows that $\tilde{\Phi}_i \in \mathscr{L}$. Therefore we obtain that $\mu_i - \tilde{\Phi}_i = \mathbb{F}r \in \mathscr{L}$, meaning that we can add $\mathbb{F}r$ to the set of generators, thereby disentangling the generator $\mathbb{F}r + \tilde{\Phi}_i$. □

The next proposition gives us an upper bound for the dimension of the Lie algebra for a given finite-dimensional deterministic volatility model.

Proposition 1.3. *The dimension of the Lie algebra in* (1.15) *is* at most *given by*

$$\dim(\mathscr{L}) = 1 + \sum_{i=1}^{m} \Big(\sum_{j} n_j^i + \sum_{\substack{(j,k) \\ j \le k}} (n_j^i + n_k^i) + \frac{1}{2} N_i (N_i + 3) \Big),$$

where $j, k = 1, \dots, N_i$.

Proof. For a given state j of the Markov chain, $\langle \mathbb{F}^k \sigma \rangle$ is spanned by at most $\sum_{j=1}^{N} (n_j + 1) = \sum_{j=1}^{N} n_j + N$ monomials of the form $x^k \, e^{-\alpha_j x}$. The space $\langle \mathbb{F}^k \tilde{\Phi} \rangle$ is spanned by at most $\sum_{(j \le k)} (n_j + n_k + 1) = \sum_{(j \le k)} (n_j + n_k) + N(N+1)/2$ monomials of the form $x^\ell e^{-(\alpha_j + \alpha_k)}$. Finally, we sum over all states of the Markov chain and add the vector field $\mathbb{F}r$. As it can happen that some combinations of indices give rise to the same function $e^{-\alpha x}$ or that some coefficients in the polynomials are zero, the calculation only gives an upper bound on the dimension. □

By an argument analogous to the one used in the preceding proof it is easy to see what $\mathfrak{n}_1^i$ and $\mathfrak{n}_2^i$ to use Proposition 1.2 namely

$$\mathfrak{n}_1^i = \sum_{j=1}^{N_i} n_j^i + N_i - 1,$$

$$\mathfrak{n}_2^i = \sum_{\substack{(j,k) \\ j \le k}} (n_j^i + n_k^i) + \frac{1}{2} N_i (N_i + 1) - 1.$$

Example 1.1 (A simple regime-switching model)**.** *In a simple regime-switching model with deterministic volatility, the Markov chain has two states and the volatility has the form* $e^{-\alpha_i x}$*, where* $\alpha_i > 0$*. In terms of the above proposition this means that* $m = 2$, $N_1 = N_2 = 1$, $n_1^1 = n_1^2 = 0$, *so that the size of the realization is* $\dim(\mathscr{L}) = 5$.

Remark 1.9. *It follows from* (1.21) *that we can use* Φ_i *as a generator instead of* $\tilde{\Phi}_i$ *in* (1.17)*. But, in the next section, when computing the invariant manifold and constructing realizations we will use* $\tilde{\Phi}_i$ *as we know that it satisfies a particular ODE.*

1.3.2 The Invariant manifold

Having derived the Lie algebra we can now parametrize the manifold in which the forward rate curve evolves. This is done by applying step (ii) in the procedure outlined in Section 1.2.2. The parametrization will be used in the next section when the state variable dynamics of the FDR are derived.

First we have to select which vector fields to use to span the Lie algebra. As the preceding section demonstrates, there are a lot of possible choices. It turns out that, for the task at hand, the most important choice is that the vector field $\mu_i = \mathbb{F}r + \Phi_i$ is broken up into $\mathbb{F}r$ and Φ_i. This is because we are dealing with deterministic volatility, so the only non-constant vector field, in any parametrization, will be one involving the term $\mathbb{F}r$. Here we use the following generator set for the Lie algebra $\mathscr{L}$:

$$\langle \mathbb{F}r, \mathbb{F}^k\sigma_i, \mathbb{F}^\ell\tilde{\Phi}_i; i \in I; k = 0, 1, \ldots, \mathfrak{n}_1^i; \ell = 0, 1, \ldots, \mathfrak{n}_2^i\rangle_{LA}.$$

To find a parametrization of the invariant manifold we apply Theorem 1.2. In practice this amounts to calculating the operator e^{tf_0}, $e^{tf_{1,k}^i}$, and $e^{tf_{2,\ell}^i}$ where

$$\begin{aligned} f_0 &= \mathbb{F}r, \\ f_{1,k}^i &= \mathbb{F}^k\sigma_i, \qquad k = 0, 1, \ldots, \mathfrak{n}_1^i, \\ f_{2,\ell}^i &= \mathbb{F}^\ell\tilde{\Phi}_i, \qquad \ell = 0, 1, \ldots, \mathfrak{n}_2^i, \end{aligned}$$

and then applying them in turn to the initial forward rate curve r_0.

Recall that according to Definition 1.3, to calculate the operator e^{tf_0} we need to solve the equation

$$\frac{dr}{dt} = \mathbb{F}r,$$

which is a linear equation with solution

$$r_t = e^{t\mathbb{F}}r_0.$$

As $\mathbb{F}$ is the generator of the semi-group of left translations[9], given the initial forward rate curve r_0, the solution can be written as

$$(e^{t\mathbb{F}}r_0)(x) = r_0(x+t). \tag{1.22}$$

The vector fields $f_{1,k}^i$ and $f_{2,\ell}$ are constant so the solution to the corresponding ODE's are easily given by

$$\begin{aligned} (e^{tf_{1,k}^i}r_0)(x) &= r_0(x) + \mathbb{F}^k\sigma_i(x)\,t, \\ (e^{tf_{2,k}^i}r_0)(x) &= r_0(x) + \mathbb{F}^k\tilde{\Phi}_i(x)\,t. \end{aligned}$$

[9]For details see, for example, Section 2.10 in Engel and Nagel (2000).

Applying all the operators in turn on the initial forward rate produces the following parametrization of the invariant manifold.

Proposition 1.4. *The invariant manifold generated by the initial forward rate curve r_0 is parametrized as*

$$G(\mathbf{z}) = r_0(x+z_0) + \sum_{i=1}^{m}\Big(\sum_{k=0}^{\mathfrak{n}_1^i}\mathbb{F}^k\sigma_i(x)z_{1,k}^i + \sum_{\ell=0}^{\mathfrak{n}_2^i}\mathbb{F}^\ell\tilde{\Phi}_i(x)z_{2,\ell}^i\Big), \qquad (1.23)$$

where $\mathbf{z} = (z_0, z_{1,k}^i, z_{2,\ell}^i; i \in I; k,l = 0,1,\ldots)$.

Remark 1.10. *In the previous proposition we can replace the vector fields $\mathbb{F}^k\sigma_i$ and $F^\ell\tilde{\Phi}_i$ by the monomials of the quasi-exponential functions that span the same space. This provides us with the following parametrization of the invariant manifold:*

$$G(\mathbf{z}) = r_0(x+z_0) + \sum_{i=1}^{m}\Big(\sum_{j=1}^{N_i}\sum_{k=0}^{n_{i,j}} x^k e^{-\alpha_{i,j}x}\cdot z_{i,j,k}^1 + \sum_{(j,k)}\sum_{\ell=0}^{n_{i,j}+n_{i,k}} x^\ell e^{-(\alpha_{i,j}+\alpha_{i,k})x}\cdot z_{i,j,k,\ell}^2\Big).$$

1.3.3 State variable dynamics

Having computed the manifold G we can now proceed to step (iii) in the procedure outlined in section 1.2.2 and solve for the state variable dynamics.

As $\langle\mathbb{F}^k\sigma_i, k = 0,1,\ldots\rangle$ and $\langle\mathbb{F}^\ell\tilde{\Phi}_i, \ell = 0,1,\ldots\rangle$ are finite-dimensional linear spaces of dimension $\mathfrak{n}_1^i + 1$ and $\mathfrak{n}_2^i + 2$, respectively. Therefore we know that σ_i and $\tilde{\Phi}_i$ must satisfy linear relations of the form

$$\begin{aligned}\mathbb{F}^{\mathfrak{n}_1^i+1}\sigma_i &= \sum_{k=0}^{\mathfrak{n}_1^i} c_{i,k}^1\mathbb{F}^k\sigma_i,\\ \mathbb{F}^{\mathfrak{n}_2^i+1}\tilde{\Phi}_i &= \sum_{\ell=0}^{\mathfrak{n}_2^i} c_{i,\ell}^2\mathbb{F}^\ell\tilde{\Phi}_i.\end{aligned} \qquad (1.24)$$

We start by computing the Fréchet derivative of G. For $i \in I$ let $\mathbf{h}_1^i := (h_{1,0}^i,\ldots,h_{1,\mathfrak{n}_1^i}^i)$ and $\mathbf{h}_2^i := (h_{2,0}^i,\ldots,h_{2,\mathfrak{n}_2^i}^i)$. Then

$$G'(\mathbf{z})(h_0, h_1^i, h_2^i; i\in I) = \mathbb{F}r_0(x+z_0)\,h_0 + \sum_{i=1}^{m}\Big(\sum_{k=0}^{\mathfrak{n}_1^i}\mathbb{F}^k\sigma_i(x)h_{1,k}^i + \sum_{\ell=0}^{\mathfrak{n}_2^i}\mathbb{F}^\ell\tilde{\Phi}_i(x)h_{2,\ell}^i\Big).$$

Next, to solve the equation $G'(\mathbf{z}) \cdot \mathbf{a}(\mathbf{z}, y_t) = \mu(r_t, y_t)$ for $\mathbf{a}(\mathbf{z}, y_t)$ we need to compute $\mu(r_t, y_t) = \mathbb{F}r_t + \tilde{\Phi}(x, y_t) = \mathbb{F}G(\mathbf{z}) + \tilde{\Phi}(x, y_t)$. Taking the derivative of (1.23) we obtain

$$
\begin{aligned}
\mu(r_t, y_t) &= \mathbb{F}r_0(x+z_0) + \sum_{i=1}^{m} \Big(\sum_{k=0}^{\mathfrak{n}_1^i} \mathbb{F}^{k+1}\sigma_i(x) z_{1,k}^i + \sum_{\ell=0}^{\mathfrak{n}_2^i} \mathbb{F}^{\ell+1}\tilde{\Phi}_i(x) z_{2,\ell}^i \Big) \\
&\quad + \tilde{\Phi}(x, y_t) \\
&= \lceil \text{using } (1.24) \rceil \\
&= \mathbb{F}r_0(x+z_0) + \sum_{i=1}^{m} \Big(\sum_{k=0}^{\mathfrak{n}_1^i - 1} \mathbb{F}^{k+1}\sigma_i(x) z_{1,k}^i + \sum_{k=0}^{\mathfrak{n}_1^i} c_{i,k}^1 \mathbb{F}^k \sigma_i(x) z_{1,\mathfrak{n}_1^i}^i \\
&\quad + \sum_{\ell=0}^{\mathfrak{n}_2^i - 1} \mathbb{F}^{\ell+1}\tilde{\Phi}_i(x) z_{2,\ell}^i + \sum_{\ell=0}^{\mathfrak{n}_2^i} c_{i,\ell}^2 \mathbb{F}^\ell \tilde{\Phi}_i(x) z_{2,\mathfrak{n}_2^i}^i \Big) + \tilde{\Phi}(x, y_t) \\
&= \lceil \text{after change of indices} \rceil \\
&= \mathbb{F}r_0(x+z_0) + \sum_{i=1}^{m} \Big(\sum_{k=1}^{\mathfrak{n}_1^i} \mathbb{F}^{k}\sigma_i(x) z_{1,k-1}^i + \sum_{k=0}^{\mathfrak{n}_1^i} c_{i,k}^1 \mathbb{F}^k \sigma_i(x) z_{1,\mathfrak{n}_1^i}^i \\
&\quad + \sum_{\ell=1}^{\mathfrak{n}_2^i} \mathbb{F}^{\ell}\tilde{\Phi}_i(x) z_{2,\ell-1}^i + \sum_{\ell=0}^{\mathfrak{n}_2^i} c_{i,\ell}^2 \mathbb{F}^\ell \tilde{\Phi}_i(x) z_{2,\mathfrak{n}_2^i}^i \Big) + \tilde{\Phi}(x, y_t).
\end{aligned}
$$

By setting $G'(\mathbf{z})\, \mathbf{a}(\mathbf{z}, y_t) - \mu(r_t, y_t) = 0$ and rearranging, we obtain

$$
\begin{aligned}
&\mathbb{F}r_0(x+z_0)[a_0 - 1] + \sum_{i=1}^{m} \Big(\sum_{k=1}^{\mathfrak{n}_1^i} \mathbb{F}^k \sigma_i(x) [a_{1,k}^i - z_{1,k-1}^i - c_{1,k}^i z_{1,\mathfrak{n}_1^i}] \\
&\quad + \sigma_i(x)[a_{1,0}^i - c_{1,0}^i z_{1,\mathfrak{n}_1^i}] + \sum_{\ell=1}^{\mathfrak{n}_2^i} \mathbb{F}^\ell \tilde{\Phi}_i(x) [a_{2,\ell}^i - z_{2,\ell-1}^i - c_{2,\ell}^i z_{2,\mathfrak{n}_2^i}] \\
&\quad + \tilde{\Phi}_i [a_{2,0}^i - c_{2,0} z_{2,\mathfrak{n}_2^i}^i - \mathbb{I}_t^i] \Big) = 0.
\end{aligned}
$$

Here $\mathbb{I}_t^i$ is the indicator function such that $\mathbb{I}_t^i = 1$ if $y_t = e_i$ and $\mathbb{I}_t^i = 0$ otherwise. Now as the preceding equation must hold for all x we can solve for a_0, $a^i_{1,k}$, and $a^i_{2,\ell}$ and obtain

$$\begin{aligned} a_0 &= 1, \\ a^i_{1,0} &= c^i_{1,0} z^i_{1,\mathfrak{n}^i_1}, \\ a_{1,k} &= z^i_{1,k-1} + c^i_{1,k} z^i_{1,\mathfrak{n}^i_1}, \qquad k = 1, \ldots, \mathfrak{n}^i_1, \\ a^i_{2,0} &= c^i_{2,0} z^i_{2,\mathfrak{n}^i_2} + \mathbb{I}^i_t, \\ a^i_{2,\ell} &= z^i_{2,\ell-1} + c^i_{2,\ell} z_{2,\mathfrak{n}^i_2}, \qquad \ell = 1, \ldots, \mathfrak{n}^i_2. \end{aligned}$$

Next we solve the equation $G'(\mathbf{z}) \cdot \mathbf{b}(\mathbf{z}, y_t) = \sigma(y_t)$ for $\mathbf{b}(\mathbf{z}, y_t)$

$$\mathbb{F}r_0(x+z_0)\, b_0 + \sum_{i=1}^{m} \Big(\sum_{k=1}^{\mathfrak{n}^i_1} \mathbb{F}^k \sigma(x) b^i_{1,k} + \sigma_i(x)[b^i_{1,0} - \mathbb{I}^i_t] + \sum_{\ell=0}^{\mathfrak{n}^i_2} \mathbb{F}^\ell \tilde{\Phi}_i(x) b^i_{2,\ell} \Big) = 0.$$

The solution is

$$\begin{aligned} b_0 &= 0, \\ b^i_{1,0} &= \mathbb{I}^i_t, \\ b^i_{1,k} &= 0, \qquad k = 1, \ldots, \mathfrak{n}^i_1, \\ b^i_{2,\ell} &= 0, \qquad \ell = 0, \ldots, \mathfrak{n}^i_2. \end{aligned}$$

Proposition 1.5. *Given a deterministic volatility as in* (1.14) *of the quasi-exponential form and an initial forward rate curve* r_0*, the forward rate model in* (1.2) *has a finite-dimensional realization in the sense of Definition 1.2 where the dynamics of the state variable* $\mathbf{z}$ *are given by*[10]

$$\begin{aligned} dz_0 &= dt, \\ dz^i_{1,0} &= c^i_{1,0} z^i_{1,\mathfrak{n}^i_1}\, dt + \mathbb{I}^i_t\, dW_t, \\ dz^i_{1,k} &= (z^i_{1,k-1} + c^i_{1,k} z^i_{1,\mathfrak{n}^i_1})\, dt, \qquad k = 1, \ldots, \mathfrak{n}^i_1, \\ dz^i_{2,0} &= (c^i_{2,0} z^i_{2,\mathfrak{n}^i_2} + \mathbb{I}^i_t)\, dt, \\ dz^i_{2,\ell} &= (z^i_{2,\ell-1} + c^i_{2,\ell} z^i_{2,\mathfrak{n}^i_2})\, dt, \qquad \ell = 1, \ldots, \mathfrak{n}^i_2. \end{aligned}$$

1.3.4 Example: Ho-Lee

The simplest possible case of regime-switching deterministic volatility one can think of is when the volatility jumps between constant states, i.e.

[10] Recall that when the volatility is deterministic the Itō and Stratonovich dynamics are the same.

$\sigma(r, y, x) = \sigma(y_t)$. In terms of Proposition 1.2 this is a degenerate case since it implies that $\alpha_{i,j} = 0$. To derive the invariant manifold and the dynamics of the state variables we therefore cannot directly apply Proposition 1.4 and Proposition 1.5. Instead we will obtain these results from scratch, thereby providing a further illustration of simplicity of the use of the Lie algebraic machinery.

The Lie algebra $\mathscr{L}$ is generated by the vector fields

$$\begin{aligned} \sigma_i(r, x) &= \sigma_i, && i \in I, \\ \mu_i(r, x) &= \mathbb{F}r + \Phi_i = \mathbb{F}r + \sigma_i^2 x. \end{aligned}$$

First we note that all σ_i are multiples of the same unit vector field $e(x) \equiv 1 \in \mathscr{L}$. We also note that, for $i, j \in I$, $\mu_i - \mu_j = (\sigma_i^2 - \sigma_j^2)x \in \mathscr{L}$. Therefore we can add e and x to the set of generators and after simplification we are left with the following three vector fields

$$\begin{aligned} f_0 &= \mathbb{F}r, \\ f_1 &= e, \\ f_2 &= x. \end{aligned}$$

Remark 1.11. *The number of states of the Markov chain has no effect on the structure of the Lie algebra, and consequently neither on the invariant manifold nor on the state variable dynamics.*

To obtain the invariant manifold we apply Theorem 1.2 by first solving for the operators e^{tf_0}, e^{tf_1}, and e^{tf_2}, and then letting them act on the initial forward rate curve r_0. As in (1.22), using that $\mathbb{F}$ is the generator of the semi-group of left translation, we obtain

$$(e^{z_0\mathbb{F}} r_0)(x) = r_0(x + z_0).$$

The vector fields e and x being constant we obtain

$$\begin{aligned} (e^{z_1 f_1} r_0)(x) &= r_0(x) + z_1, \\ (e^{z_2 f_2} r_0)(x) &= r_0(x) + x\, z_2. \end{aligned}$$

Applying the operators in turn produces the following parametrization of the invariant manifold:

$$\begin{aligned} G(z_0, z_1, z_2) &= e^{z_2 f_2} e^{z_1 f_1} e^{z_0 f_0} r_0(x) \\ &= r_0(x + z_0) + z_1 + x z_2. \end{aligned}$$

The dynamics of the state variables z_0, z_1, and z_3 are obtained by comparing the coefficients of the SDE in (1.2) with those obtained in (1.12).

For the drift we obtain the equation $G'(\mathbf{z})\,\mathbf{a} = \mu$. Here

$$\begin{aligned} G'(\mathbf{z})\,\mathbf{a} &= (\mathbb{F}r_0(x+z_0), e, x) \cdot (a_1, a_2, a_3) \\ &= \mathbb{F}r_0(x+z_0)a_1 + a_2 + a_3 x, \end{aligned}$$

and

$$\begin{aligned} \mu &= \mathbb{F}r + \Phi = \mathbb{F}G(\mathbf{z}) + \Phi \\ &= \mathbb{F}r_0(x+z_0) + z_2 + \sigma^2(y)x. \end{aligned}$$

Setting $G'(\mathbf{z})\mathbf{a} - \mu = 0$ yields

$$[a_1 - 1]\mathbb{F}r_0(x+z_0) + [a_3 - \sigma^2(y)]x + a_2 - z_2 = 0.$$

As this equation must hold for all x we obtain

$$\begin{aligned} a_1 &= 1, \\ a_2 &= z_2, \\ a_3 &= \sigma^2(y). \end{aligned}$$

For the volatility we set $G'(\mathbf{z})\mathbf{b} - \sigma(y) = 0$ and obtain

$$\mathbb{F}r_0(x+z_0)b_1 + xb_3 + b_2 - \sigma(y) = 0.$$

Again, as this equation must hold for all x, we obtain

$$\begin{aligned} b_1 &= 0, \\ b_2 &= \sigma(y), \\ b_3 &= 0. \end{aligned}$$

We can now sum up our findings in the following proposition.

Proposition 1.6 (Ho-Lee example)**.** *Assume that the volatility is given by $\sigma(r_t, y_t, x) = \sigma(y_t)$. Then the invariant manifold generated by the initial forward rate curve is given by*

$$G(z_0, z_1, z_2) = r_0(x+z_0) + z_1 + xz_2,$$

where the state variables have the following dynamics

$$\begin{aligned} dz_0 &= dt, \\ dz_1 &= z_2\,dt + \sigma(y_t)dW_t, \\ dz_2 &= \sigma^2(y_t)dt. \end{aligned}$$

1.3.5 Example: Hull-White

As our second example we look at the Hull and White (1990) extension of the Vasiček (1977) model where the volatility takes the form $\sigma(r,x) = \beta e^{-\alpha x}$. If we let both parameters α and β be driven by the Markov chain y we obtain the following volatility:

$$\sigma(y,x) = \beta(y)e^{-\alpha(y)x}.$$

To calculate the invariant manifold we apply Proposition 1.4 and obtain:

$$G(z_0, z^i_{1,0}, z^i_{2,0}) = r_0(x+z_0) + \sum_{i=1}^{m} \Big(\sigma_i(x) z^i_{1,0} + \tilde{\Phi}_i(x) z^i_{2,0}\Big).$$

The volatility solves the following ODE

$$\mathbb{F}\sigma_i(x) = -\alpha_i\,\sigma_i(x),$$

meaning that $c^i_{1,0} = -\alpha_i$. Next

$$\Phi_i = \frac{\beta_i^2}{\alpha_i}(e^{-\alpha_i x} - e^{-2\alpha_i x}),$$

therefore

$$\tilde{\Phi}_i = \Phi_i - \frac{\beta_i}{\alpha_i}\sigma_i = -\frac{\beta_i^2}{\alpha_i} e^{-2\alpha_i x},$$

and $\tilde{\Phi}$ solves the ODE

$$\mathbb{F}\tilde{\Phi}_i(x) = -2\alpha_i \tilde{\Phi}_i(x)$$

giving $c^i_{2,0} = -2\alpha_i$. By Proposition 1.5 the state variables $\mathbf{z}$ will have the following dynamics:

$$\begin{aligned} dz_0 &= dt, \\ dz^i_{1,0} &= -\alpha_i z^i_{1,\mathfrak{n}^i_1}\,dt + \mathbb{I}^i_t\,dW_t, \\ dz^i_{2,0} &= (-2\alpha_i z^i_{2,0} + \mathbb{I}^i_t)\,dt. \end{aligned}$$

1.4 Separable volatility

After deterministic volatility, at the next (tractable) level of complexity we consider separable volatility, i.e. volatility of the form

$$\sigma(r,x) = \varphi(r)\lambda(x).$$

Here the vector field σ has constant direction $\lambda \in \mathcal{H}$ but has varying length determined by the smooth scalar field φ. When introducing regime-switching in this type of volatility we can make either φ or λ, or both, dependent on the Markov chain y. We will consider these three cases in turn.

1.4.1 Stochastic length

Assume that the volatility is given by

$$\sigma(r, x, y) = \varphi(r, y)\lambda(x).$$

The Lie algebra is generated by the vector fields (recall equation (1.7))

$$\mu_i = \mathbb{F}r + \Phi_i - \frac{1}{2}\mathcal{C}_i = \mathbb{F}r + \underbrace{\varphi_i^2}_{\phi_i} \underbrace{\lambda(x)\int_0^x \lambda(s)\,ds}_{\Lambda} - \frac{1}{2}\varphi_i'[\lambda]\varphi_i\lambda,$$

$$\sigma_i = \varphi_i\lambda.$$

To simplify notation we define

$$\Lambda(x) := \lambda(x)\int_0^x \lambda(s)\,ds,$$
$$\phi(r, y) := \varphi^2(r, y).$$

To avoid treating trivial cases we make the following assumption.

Assumption 1.1. *Assume that $\varphi(r, y) \neq 0$ for all $r \in \mathcal{H}$.*

We observe that all the generators σ are multiples of the single vector field λ. Next we can remove suitable multiples of λ from all generators μ_i. After this operation we are left with the following vector fields

$$g_1^i := \mathbb{F}r + \phi_i(r)\Lambda, \qquad i \in I,$$
$$g_2 := \lambda.$$

As $g_1^i - g_1^j = (\phi_i - \phi_j)\Lambda$ and $\phi_i - \phi_j \neq 0$, we can add Λ to the set of generators which after simplification consists of

$$f_1 := \mathbb{F}r,$$
$$f_2 := \lambda,$$
$$f_3 := \Lambda.$$

Notice that all the generators are now independent of the Markov chain y. As both f_2 and f_3 are constant, taking Lie brackets is easy and we obtain the following system of generators for the Lie algebra:

$$\langle \mathbb{F}r, \mathbb{F}^k\lambda, \mathbb{F}^k\Lambda; k = 0, 1, \ldots\rangle_{LA}.$$

Remark 1.12. *We obtain the same Lie algebra as when $\varphi(r)$ does not depend on y and we make the assumption that $\phi''[\lambda;\lambda] \neq 0$. This case is studied by (Björk and Svensson, 2001, p.228)*

We can now apply the same reasoning as in section 1.3.1. The space $\langle \mathbb{F}^k \lambda \rangle_{LA}$ is finite dimensional if and only if λ is a QE function of x. If λ is QE, then Λ will also be QE according to point (iv) and (v) of Lemma 1.1, as it is an integral of λ multiplied by λ. Consequently, $\langle \mathbb{F}^k \Lambda \rangle_{LA}$ also spans a finite-dimensional space, and we have proved the following proposition.

Proposition 1.7. *Assume that the volatility has the form*

$$\sigma(r, y, x) = \varphi(r, y)\lambda(x).$$

Assume furthermore that Assumption 1.1 is satisfied. Then the model in (1.2) *possesses a FDR if and only if $\lambda(x)$ is a quasi-exponential function in the variable x. The smooth vector field $\varphi(r, y)$ can be chosen freely.*

Remark 1.13. *As the Lie algebra has the same structure as in Proposition 1.2, with vector fields of the quasi-exponential form we obtain the same result concerning the representation by monomials and dimension as in section 1.3.1.*

1.4.2 Stochastic direction

Now assume that the volatility is given by

$$\sigma(r, x, y) = \varphi(r)\lambda(x, y).$$

Using the same notation as in the previous section, the Lie algebra is generated by the vector fields

$$\mu_i = \mathbb{F}r + \phi\Lambda_i - \frac{1}{2}\varphi'[\lambda_i]\varphi\lambda_i,$$
$$\sigma_i = \varphi\lambda_i.$$

Also here we have to make an assumption concerning φ.

Assumption 1.2. *Assume that $\varphi(r) \neq 0$ for all $r \in \mathcal{H}$.*

This assumption allows us to remove suitable multiples of λ_i from μ_i. After this operation we are left with the following generators:

$$g_1^i = \mathbb{F}r + \phi\Lambda_i, \qquad i \in I,$$
$$g_2^i = \lambda_i.$$

To achieve greater simplification we take Lie brackets between the vector fields

$$[g_1^i, g_2^j] = \mathbb{F}\lambda_j + \phi'[\lambda_j]\Lambda_i.$$

Taking one bracket produces

$$[[g_1^i, g_2^j], g_2^k] = \phi''[\lambda_j, \lambda_k]\Lambda_i.$$

To proceed we make the following assumption.

Assumption 1.3. *Assume that $\phi''[\lambda_j, \lambda_k] \neq 0$ for all $r \in \mathcal{H}$ for at least two states j and k of the Markov chain y.*

This assumption allows us to remove suitable multiples of Λ_i from g_1^i. After these simplifications we are left with the following generators:

$$f_0 = \mathbb{F}r,$$
$$f_1 = \lambda_i,$$
$$f_2 = \Lambda_i,$$

so the set of generators of the Lie algebra now becomes

$$\langle \mathbb{F}r, \mathbb{F}^k\lambda_i, \mathbb{F}^k\Lambda_i; i \in I; k = 0, 1, \ldots \rangle_{LA}.$$

Applying the same reasoning as in the previous section we obtain the following result.

Proposition 1.8. *Assume that the volatility has the form*

$$\sigma(r, y, x) = \varphi(r)\lambda(x, y).$$

Suppose furthermore that Assumption 1.2 and 1.3 are satisfied. Then the model in (1.2) *possesses a FDR if and only if $\lambda(x, y)$ is a quasi-exponential function in the variable x. The smooth vector field $\varphi(r)$ can be chosen freely.*

What if Assumption 1.3 is not satisfied? This is the case in the example treated in the next section. There, as we will see, $\phi''(r)[\lambda_i, \lambda_j] = 0$ for $i, j \in I$ and $\lambda_i(x)$ is not a QE function.

1.4.3 Example: Cox-Ingersoll-Ross

To remove the possibility of negative interest rates present in Gaussian models, Cox, Ingersoll, and Ross (1985) (CIR) introduced the following process to model the short rate $R = r(0)$:

$$dR_t = a(b - R_t)\,dt + \rho\sqrt{R_t}\,dW_t.$$

This model has an equivalent HJM formulation where the volatility takes the form

$$\sigma(r,x) = \sqrt{r(0)}\,\lambda(x,\rho,a), \tag{1.25}$$

where

$$\lambda(x,\rho,a) = \rho\frac{\partial}{\partial x}\left(\frac{2(e^{\gamma x}-1)}{(\gamma+a)(e^{\gamma x}-1)+2\gamma}\right),$$

and where

$$\gamma = \sqrt{a^2+2\rho^2}.$$

A natural question to ask now is if the parameters a and ρ can be made stochastic, i.e. functions of the Markov chain y. We note that the first order Fréchet derivative is $\phi'(r)[\lambda_i] = \lambda_i(0)$, i.e. it amounts to a point evaluation. Next $\phi''(r)[\lambda_i,\lambda_j] = 0$, so we see that we cannot use the results of the previous section.

Remark 1.14. *This example was also treated in Björk, Landén, and Svensson (2004), but there a and ρ were driven by a diffusion and it was proved that an FDR did not exist. However, in their proof the diffusion assumption was crucial, to their result is not applicable here.*

To simplify the treatment we will assume that the Markov chain only has two states, the generalization to more states being trivial. We assume that the parameters a and ρ are functions of the Markov chain, taking values a_i and ρ_i, $i = 1,2$. The Lie algebra is spanned by the following vector fields

$$\begin{aligned} g_1 &= \mathbb{F}r + \phi(r)\Lambda_1,\\ g_2 &= \Lambda_1 - \Lambda_2,\\ g_3 &= \lambda_1,\\ g_4 &= \lambda_2. \end{aligned}$$

To make the relation between λ_i and Λ_i clear we will express these vector fields in terms of

$$L_i = \int_0^x \lambda_i(s)\,ds, \qquad i = 1,2.$$

As $\mathbb{F}L_i = \lambda_i$ and $\mathbb{F}(L_i^2) = 2L_i\mathbb{F}L_i = 2\Lambda_i$, we obtain[11]

$$\begin{aligned} g_1 &= \mathbb{F}r + \frac{1}{2}\phi(r)\mathbb{F}L_1^2,\\ g_2 &= \mathbb{F}(L_1^2 - L_2^2),\\ g_3 &= \mathbb{F}L_1,\\ g_4 &= \mathbb{F}L_2. \end{aligned}$$

[11] The constant 1/2 in g_2 has been simplified away.

We note that every vector field except the first is constant.

As the CIR model admits an affine term structure[12] it can be shown that L_i must satisfy the following Riccati equation:

$$\mathbb{F}L_i(x) + a_i L_i(x) + \frac{\rho_i}{2} L_i^2 = \rho_i.$$

Taking the derivative of this equation with respect to x and solving for $\mathbb{F}L_i^2$ yields

$$\mathbb{F}L_i^2 = -\frac{2}{\rho_i}\Big(\mathbb{F}^2 L_i + a_i \mathbb{F}L_i\Big).$$

By substituting this expression into g_1 and g_2 we obtain

$$\begin{aligned} g_1 &= \mathbb{F}r - \frac{1}{\rho_1}\phi(r)(\mathbb{F}^2 L_1 + a_i \mathbb{F}L_1) \\ g_2 &= -\frac{2}{\rho_1}\Big(\mathbb{F}^2(L_1 - L_2) + a_i \mathbb{F}(L_1 - L_2)\Big). \end{aligned}$$

By removing suitable multiples of g_3 and g_4 from these expressions we end up with the following system of generators for the Lie algebra:

$$\begin{aligned} h_1 &= \mathbb{F}r - \frac{\phi(r)}{\rho_1}\mathbb{F}^2 L_1, \\ h_2 &= \mathbb{F}^2(L_1 - L_2), \\ h_3 &= \mathbb{F}L_1, \\ h_4 &= \mathbb{F}L_2. \end{aligned}$$

Now we can start to take Lie brackets and we obtain

$$\begin{aligned} A_1 &:= [h_1, h_2] = \mathbb{F}^3(L_1 - L_2) + \alpha_1 \mathbb{F}^2 L_1, \\ B_1 &:= [h_1, h_3] = \mathbb{F}^2 L_1 + \beta_1 \mathbb{F}^2 L_1, \\ C_1 &:= [h_1, h_4] = \mathbb{F}^2 L_2 + \gamma_1 \mathbb{F}^2 L_1, \end{aligned}$$

where

$$\alpha_1 := -\frac{1}{\rho_1}\phi'(r)[h_2], \qquad \beta_1 := -\frac{1}{\rho_1}\phi'(r)[h_3], \qquad \gamma_1 := -\frac{1}{\rho_1}\phi'(r)[h_4].$$

A small calculation reveals that $\phi'(r)[h_3] = \phi'(r)[\lambda_1] = \lambda_1(0) = \rho_1$ so that $\beta_1 = -1$ and therefore the bracket $B_1 = 0$. Further brackets will therefore only involve A_1 and C_1.

[12] See Björk (2004), chapter 22.

Continuing taking bracket for $n \geq 2$ we obtain

$$A_n := [h_1, A_{n-1}] = \mathbb{F}^{n+2}(L_1 - L_2) + \sum_{i=1}^{n} \alpha_i \, \mathbb{F}^{n+2-i} L_1,$$

$$C_n := [h_1, C_{n-1}] = \mathbb{F}^{n+1} L_2 + \sum_{i=1}^{n} \gamma_i \, \mathbb{F}^{n+2-i} L_1,$$

where we define $\alpha_i := -(1/\rho)\phi'(r)[A_{i-1}]$ and $\gamma_i := -(1/\rho)\phi'(r)[C_{i-1}]$, for $i > 1$.

Now, aiming at a contradiction we assume that the Lie algebra is finite dimensional. Then there exists a N such that C_N can be written as a linear combination of $C_1, \ldots, C_{N-1}$ and therefore there are constants $c_1, c_2, \ldots, c_N$ such that

$$c_1 C_1 + c_2 C_2 + \ldots + c_N C_N = 0.$$

If we substitute in the expressions for $C_1, C_2, \ldots, C_N$ we obtain

$$\sum_{i=1}^{N} c_i F^{i+1} L_2 + \sum_{i=1}^{N} \eta_i \mathbb{F}^{i+1} L_1 = 0, \tag{1.26}$$

where the scalars η_i are linear combinations of the scalars c_i. By observing how parameters a_i and ρ_i enter into exponential terms in both the numerator and denominator of L_i,

$$L_i = \rho_i \frac{2(e^{\gamma_i x} - 1)}{(\gamma_i + a_i)(e^{\gamma_i x} - 1) + 2\gamma_i},$$

it is clear that there cannot be any linear relation between the derivatives of L_1 and L_2. This means that both the sums in (1.26) must be equal to zero, i.e

$$\sum_{i=1}^{N} c_i F^{i+1} L_2 = 0,$$

$$\sum_{i=1}^{N} \eta_i \mathbb{F}^{i+1} L_1 = 0.$$

Recognizing that the two preceding equations are linear ODE's with constant coefficient in the variables L_1 and L_2, respectively, we know that L_1 and L_2 should be quasi-exponential functions, which they are not. We therefore have a contradiction. The assumption that the Lie algebra is finite dimensional is therefore false. We can therefore conclude that, a CIR model with parameters modulated by a Markov chain does not posses a finite-dimensional realization.

1.4.4 Stochastic length and direction

In Section 1.4.1 the length φ of the volatility $\sigma = \varphi\lambda$ was made stochastic. We showed that this did only have a minor impact on the structure of the Lie algebra. The same Lie algebra as in the case without a Markov chain emerged but without having to assume that $\varphi''[\lambda;\lambda] \neq 0$. It turns out that we obtain a similar result when we now let both length and direction depend on y,

$$\varphi(r, y, x) = \varphi(r, y)\lambda(x, y).$$

We apply the same reasoning as in Section 1.4.2, and we make Assumption 1.1 as well as the following assumption.

Assumption 1.4. *Assume that $\phi_i''[\lambda_j, \lambda_k] \neq 0$ for all $r \in \mathcal{H}$ for at least two states j and k of the Markov chain y.*

We simplify the generators and end up with the following system

$$\langle \mathbb{F}r, \mathbb{F}^k\lambda_i, \mathbb{F}^k\Lambda_i; i \in I; k = 0, 1, \ldots\rangle_{LA}. \tag{1.27}$$

As expected we obtain the following result.

Proposition 1.9. *Assume that the volatility has the following form*

$$\sigma(r, y, x) = \varphi(r, y)\lambda(x, y).$$

Suppose furthermore that Assumption 1.2 and Assumption 1.4 are satisfied. Then the model in (1.2) possesses a FDR if and only if $\lambda(x, y)$ is a quasi-exponential function in the variable x. The smooth vector field $\varphi(r, y)$ can be chosen freely.

It turns out that we can replace Assumption 1.4 in the previous proposition by a weaker assumption. To see how, we start by simplifying the generators of the Lie algebra in the same way as in Section 1.4.2. This gives the following generators:

$$\begin{aligned} g_1^i &= \mathbb{F}r + \phi_i\Lambda_i, \\ g_2^i &= \lambda_i. \end{aligned}$$

Now we observe that

$$\begin{aligned} g_1^i - g_1^j &= \phi_i\Lambda_i - \phi_j\Lambda_j, \\ [g_1^i, \lambda_k] - [g_1^j, \lambda_k] &= \phi_i'[\lambda_k]\Lambda_i - \phi_j'[\lambda_k]\Lambda_j. \end{aligned}$$

We can now introduce the following assumption.

Assumption 1.5. *Assume that* $\phi_j \phi_i'[\lambda_k] - \phi_i \phi_j'[\lambda_k] \neq 0$ *for all* $i, j \in i$ *and for some* k.

This allows us to solve for Λ_i and Λ_j and obtain the system of generators in (1.27).

The following example shows that Assumption 1.5 is weaker than Assumption 1.4.

Example 1.2. *In a two-state model, setting* $\varphi_1 = \sqrt{a_1 + br(0)}$, $\varphi_2 = \sqrt{a_2 + br(0)}$, *where* $a_1, a_2, b \in \mathbb{R}$, *satisfies Assumption 1.5, but not Assumption 1.4.*

Remark 1.15. *In the CIR model studied in Section 1.4.3 only the direction is stochastic. If we alter the model by introducing stochastic length, by using the recipe in the previous example, we immediately can conclude that the model does not posses a FDR since* λ *is not QE.*

Notation

m : Number of states in the Markov chain.

$I = \{1, 2, \ldots, m\}$ Set to index the states of the Markov chain.

d number of vector fields spanning the Lie algebra i.e. dimension of realization.

N number of exponential terms in σ

n_j degree of the polynomial in front of the exponential term $e^{-\alpha_j x}$.

Bibliography

Ang, Andrew, and Geert Bekaert, 2002, Regime switching in interest rates, *Journal of Business and Economic Statistics* 20, 163–182.

Bansal, Ravi, and Hao Zhou, 2002, Term structure of interest rates with regime shifts, *The Journal of Finance* 57, 1997–2043.

Bhar, Ramaprasad, and Carl Chiarella, 1997, Transformation of Heath-Jarrow-Morton models to Markovian systems, *The European Journal of Finance* 3, 1–26.

Björk, Tomas, 2003, On the geometry of interest rate models, in *Paris-Princeton Lectures on Mathematical Finance*, volume 1847 of *Springer Lecture Notes in Mathematics*, 133–215 (Springer).

Björk, Tomas, 2004, *Arbitrage Theory in Continuous Time* (Oxford University Press), 2nd edition.

Björk, Tomas, and Bent Jesper Christensen, 1999, Interest rate dynamics and consistent forward rate curves, *Mathematical Finance* 9, 323–348.

Björk, Tomas, and Camilla Landén, 2002, On the construction of finite dimensional realizations for nonlinear forward rate models, *Finance and Stochastics* 6, 303–331.

Björk, Tomas, Camilla Landén, and Lars Svensson, 2004, Finite dimensional Markovian realizations for stochastic volatility forward rate models, *Proceedings of the Royal Society, Series A* 460, 53–83.

Björk, Tomas, and Lars Svensson, 2001, On the existence of finite-dimensional realizations for nonlinear forward rate models, *Mathematical Finance* 11, 205–243.

Chiarella, Carl, and Oh Kang Kwon, 2001, Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model, *Finance and Stochastics* 5, 237–257.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2007, Regime shifts in a dynamic term structure model of u.s. treasury bond yields, *The Review of Finantial Studies* 20, 1669–1706.

Engel, Klaus-Jochen, and Rainer Nagel, 2000, *One-Parameter Semigroups for Linear Evolution Equations*, GTM (Springer).

Filipović, Damir, and Josef Teichmann, 2003, Existence of invariant manifolds for stochastic equations in infinite dimension, *Journal of Functional Analysis* 197, 398–432.

Garcia, René, and Pierre Perron, 1996, An analysis of the real interest rate under regime shifts, *The Review of Economics and Statistics* 78, 111–125.

Gray, Stephen F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27–62.

Hamilton, James D., 1988, Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates, *Journal of Economic Dynamics and Control* 12, 385–423.

Hansen, Asbjørn, and Rolf Poulsen, 2000, A simple regime switching term structure model, *Finance and Stochastics* 4, 409–429.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates, *Econometrica* 60, 77–105.

Ho, Thomas, and Sang-Bin Lee, 1986, Term structure movements and pricing interest rate contingent claims, *Journal of Finance* 41, 1011–1029.

Hull, John, and Alan White, 1990, Pricing interest-rate derivative securities, *Review of Financial Studies* 3, 573–592.

Inui, Koji, and Masaaki Kijima, 1998, A Markovian framework in multifactor Heath-Jarrow-Morton models, *Journal of Financial and Quantitative Analysis* 33, 423–440.

Landén, Camilla, 2000, Bond pricing in a hidden markov model of the short rate, *Finance and Stochastics* 4, 371–389.

Musiela, Marek, 1993, Stochastic PDE:s and the term structure models, Journées Internationales de Finance, IGR-AFFI, la Baule, France.

Naik, Vasant, and Moon Hoe Lee, 1997, Yield curve dynamics with discrete shifts in economic regimes: Theory and estimation, working paper, Faculty of Commerce, University of British Columbia.

Olver, Peter J., 1986, *Applications of Lie Groups to Differential Equations* (Springer), second edition.

Ritchken, Peter, and L. Sankarasubramanian, 1995, Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance* 5, 55–72.

Schwartz, Laurent, 1992, *Analyse II: Calcul Différentiel et Équations Différentielles* (Hermann).

Valchev, Stoyan, 2004, Stochastic volatility Gaussian Heath-Jarrow-Morton models, *Applied Mathematical Finance* 11, 347–368.

Vasiček, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 177–188.

Warner, Frank W., 1971, *Foundations of Differentiable Manifolds and Lie Groups* (Springer).

Paper 2

Paper 2

Minimal Realizations of Market Models

Joint work with Tomas Björk and Damir Filipović

2.1 Introduction

Since their introduction in the end of the 1990s, the LIBOR market model and the swap market model have been very popular among practitioners.[1] This popularity stems from their applicability.

Firstly, the models allows the use of real discrete market rates as inputs, the LIBOR rate in one case and the swap rate in the other. This contrast with traditional interest rate models where instantaneous quantities are used. The first models developed to price and hedge interest rate derivative securities used the instantaneous short rate as the underlying stochastic factor. This is for example the case in the seminal model by Vasiček (1977) or in the short rate models that followed such as those by Cox, Ingersoll, and Ross (1985), Ho and Lee (1986), and Hull and White (1990). Later on, Heath, Jarrow, and Morton (1992) used the instantaneous forward rate as the underlying variable, thereby initiating a second strand in the literature. Then followed models such as those by Ritchken and Sankarasubramanian (1995) or Mercurio and Moraleda (2000). The main advantage of using instantaneous quantities is of course that they are easy to handle mathematically. Unfortunately these entities are theoretical abstraction that cannot be observed in real life. This causes problems for market participants when they try to fit the parameters of the models to real world data.

[1] See for example (Brigo and Mercurio, 2006).

Another issue with traditional interest rate models concerns the wide use in the two largest interest rate option markets of the world, the cap market and the swaptions market, of pricing formulas based on an extension of Black (1976) model, this in spite off the existence of the more sophisticated models mentioned in the previous paragraph. The pricing formula used is a simple alteration of the Black and Scholes (1973) formula for the pricing of stock options. This formula clearly lacks theoretical foundation since in its derivation, at some point, it is assumed that the interest is constant, while in another it is assumed to be stochastic.

In order to overcome these shortcomings a new kind of interest rate models was introduced in the end of the nineties. The LIBOR market model was introduced in 1997 in the work of Miltersen, Sandmann, and Sondermann (1997) and Brace, Gatarek, and Musiela (1997). The swap market models was introduced by Jamshidian (1997). Models of this kind are commonly referred to as *market models*. Here the real market interest rates, the LIBOR rates or swap rates are used as underlying stochastic variables. By assuming that the variables are log-normally distributed, the pricing formulas for caps and swaptions, based on Black (1976) are recovered and as result, a sound theoretical foundation for their use is obtained.

Despite their benefits, market models have one important shortcoming namely their dimension. The models have the same dimension as the number of rates that are being modelled. The pricing of a derivative security having cash flows at N different points in time requires a model of size N. To take a concrete example, a ten year swap with quarterly payments necessitates a 39-dimensional model. It would therefore be convenient if it was possible to reduce the dimension of the model in some way.

The object of the present paper is to investigate if it is possible to reduce the dimension of the market models by analysing them from a geometric point of view. To this end we will use the framework developed by Björk and Svensson (2001) to study finite dimensional realizations of forward rate models. It turns out that the minimal dimension of the markets models is equal to the size of the Lie algebra generated by the drift and volatility of SDE defining the models.

The rest of the paper is structured as follows. Section 2.2 provides the necessary background on the LIBOR market model and the swap market model. First the LIBOR market model is presented in Section 2.2.1. In particular the Black-76 pricing formula and the dynamics of the LIBOR rates under the terminal measure are derived. Then, in Section 2.2.2, after a brief presentation of the swap market model the relation between the swap market model and the LIBOR market is studied. Section 2.3 states the problem of finding minimal realization of market models formally. In

Section 2.4 minimal realizations of models with constant volatility are treated. Finally, in Section 2.5 we tackle the slightly more general case of realizations of time-dependent volatility models.

2.2 Market Models

2.2.1 The LIBOR market model

All standard mathematical assumptions relevant in the context of contingent claim pricing are made.[2] All processes are defined on a filtered probability space $(\Omega, \{\mathcal{F}_t; t \geq 0\}, \mathbb{Q})$. The basis is assumed to carry a Wiener process W and the filtration is assumed to be the internal one generated by W.[3]

As is conventional, we assume the existence of a default free bond for every maturity. The bond market is assumed to be free of arbitrage. Let $p(t, T)$ denote the price, at time $t \geq 0$, of a zero coupon bond maturing at time $T \geq t$. We will make use of a given fixed set of maturities $T_0, T_1, \ldots, T_N$ denote the tenor by $\alpha_i = T_i - T_{i-1}$, where $i = 1, \ldots, N$. On the real market tenors of 3 or 5 months are common.

In the model we consider, money placed into the money-market account will earn the LIBOR rate. The model will first be used to price caps, and in particular caplets, that are functions of the *forward* LIBOR rate which is defined as follows:[4]

Definition 2.1 (LIBOR forward rate)**.** *The* LIBOR *forward rate, contracted at time t, for the period* $[T_{i-1}, T_i]$ *is defined as*

$$L_i(t) = -\frac{1}{\alpha_i} \frac{p_i(t) - p_{i-1}(t)}{p_i(t)}, \tag{2.1}$$

where $i = 1, \ldots, N$*.*

The price of a cap

A cap is a contract composed of a portfolio of call options on the LIBOR rate called caplets. The purpose of the contract is to provide an insurance against LIBOR rates rising above the level R, the cap rate. The payoff of caplet number i, denoted X_i, is given by

$$X_i = \alpha_i \max[L_i(T_{i-1}) - R, 0], \tag{2.2}$$

[2] See for example Musiela and Rutkowski (1997).

[3] Concerning the choice of a one-dimensional Wiener process see Remark 2.1.

[4] More advanced derivative securities can of course also be priced within the same framework.

where $i = 1, \ldots, N$. Notice that $L_i(T_{i-1})$ is known at time T_{i-1} since it is the LIBOR spot rate at time T_{i-1} for the period $\alpha_i = [T_{i-1}, T_i]$, but the amount is not payed out until time T_i. The payoff is very natural if we consider the hedging of interest payments on a loan with variable LIBOR interest rate. The rate for the period $[T_{i-1}, T_i]$ is then settled at time T_{i-1} and payed at time T_i. A cap with resettlement dates $T_0, T_1, \ldots, T_N$ entitles the holder of the contract to the amount X_i at time T_i for $i = 1, \ldots, N$.

Having put the necessary notation in place we now go on to price a caplet. Recalling (2.2) and using standard risk-neutral pricing we know that the price, at time t, of caplet number i, with payoff X_i, is given by

$$c_i(t) = \alpha_i \, \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^{T_i} r_s ds} \max[L_i(T_{i-1}) - R, 0] | \mathcal{F}_t],$$

where r_s denotes the short rate and the expectation is taken under the risk-neutral measure $\mathbb{Q}$, i.e. with the bank account as numeraire. In the present context it is far more convenient to use $p(t, T_i)$ as a numeraire and price the claim under the T_i-forward measure.[5] The price can then be written as

$$c_i(t) = \alpha_i p(t, T_i) \, \mathbb{E}^{T_i}[\max[L_i(T_{i-1}) - R, 0] | \mathcal{F}_t]. \tag{2.3}$$

This removes the exponential term from the pricing expression, so we do not even have to assume the existence of the short rate. Next we observe that

$$\alpha_i L_i(T_{i-1}) = -\frac{p_i(T_{i-1}) - p_{i-1}(T_{i-1})}{p_i(T_{i-1})} = \frac{p_{i-1}(T_{i-1})}{p_i(T_{i-1})} - 1$$

is a martingale under the T_i-forward measure, henceforth denoted $\mathbb{Q}^i$, on the interval $[0, T_{i-1}]$. This follows from the fact that a constant, trivially, is a martingale under any measure and that $p_i(T_{i-1})$ is divided by the numeraire $p_i(T_i)$. Now we arrive at the main idea of the LIBOR market modelling approach. To obtain a pricing formula of the Black-76 type we assume that L_i is log-normally distributed under the T_i forward measure, i.e. of the form

$$dL_i(t) = L_i(t)\sigma_i(t) dW^i(t), \qquad i = 1, \ldots, N, \tag{2.4}$$

where W^i denotes the $\mathbb{Q}^i$-Wiener process. The expectation in (2.3) can now be computed explicitly yielding a Black-76 pricing formula of the form

$$c_i(t) = \alpha_i p(t, T_i) \big[L_i(t) \, \mathrm{N}(d_1) - R \, \mathrm{N}(d_2) \big],$$

[5] For a review of the change of numeraire technique, see Björk (2004, Chap. 24) or Brigo and Mercurio (2006, Chap. 2).

where

$$\begin{aligned} d_1 &= \frac{1}{\Sigma_i(t,T_{i-1})}\Big[\log\Big(\frac{L_i(t)}{R}\Big) + \frac{1}{2}\Sigma_i^2(t,T_{i-1})\Big], \\ d_2 &= d_1 - \Sigma_i(t,T_{i-1}), \\ \Sigma_i^2(t,T) &= \int_t^T \|\sigma_i(s)\|^2 ds. \end{aligned}$$

With an explicit formula, calibration of the model to cap prices becomes trivial, at least as long as we assume the volatility to be time-independent. With σ_i constant, we obtain $\Sigma_i^2(t,T) = \sigma_i(T-t)$ and therefore only one unknown quantity in the pricing formula. Given caplet prices we can compute the implied volatility. This is actually the quantity quoted on the market. It can then be used to calibrate the model to existing cap prices and afterwards pricing other derivative instruments on the market.

Dynamics under the terminal measure

In the derivation above, we have assumed that each forward LIBOR rate L_i is log-normally distributed under its own measure $\mathbb{Q}^i$. In order to price an instrument that depends on several rates we need to model all LIBOR rates under a common measure. The canonical way to do this is to use the terminal measure, i.e. the measure $\mathbb{Q}^N$ as a common measure.

Remark 2.1. *Here we assume that the* LIBOR *rates are driven by the same Wiener process* W*. We could of course use a vector Wiener process instead but this would only render the task of finding a small realization harder. This will become apparent ones we start calculating the Lie algebra generated by the drift and volatility of the model.*

Assuming that we have specified the deterministic volatilities $\sigma_1, \ldots, \sigma_N$, the dynamics of the LIBOR rates will be of the following form under the measure $\mathbb{Q}^N$:

$$dL_i(t) = \mu_i(t, L_i(t))L_i(t)dt + \sigma_i(t)L_i(t)dW^N(t), \qquad i = 1, \ldots, N.$$

We now need to find $\mu_1, \ldots, \mu_N$ so that all the rates $L_i(t)$ will have the dynamics given by (2.4) under their own measure $\mathbb{Q}^i$. The tool needed for this endeavour is the Girsanov Theorem.

The Radon-Nikodym derivative needed to change measure from $\mathbb{Q}^i$ to $\mathbb{Q}^j$, here denoted ρ_i^j, is given by

$$\rho_i^j(t) = \frac{d\mathbb{Q}^j}{d\mathbb{Q}^i} = \frac{p_i(0)}{p_j(0)}\frac{p_j(t)}{p_i(t)}, \qquad 0 \le t \le T.$$

In our particular case we need

$$\rho_i^{i-1}(t) = \frac{d\mathbb{Q}^{i-1}}{d\mathbb{Q}^i} = \frac{p_i(0)}{p_{i-1}(0)} \frac{p_{i-1}(t)}{p_i(t)} = a_i \frac{p_{i-1}(t)}{p_i(t)}, \qquad 0 \le t \le T,$$

where we set $a_i := p_i(0)/p_{i-1}(0)$. To rewrite this last equation we use that, according to (2.1),

$$\frac{p_{i-1}(t)}{p_i(t)} = 1 + \alpha_i(t) L_i(t), \tag{2.5}$$

so

$$\rho_i^{i-1}(t) = a_i[1 + \alpha_i L_i(t)]. \tag{2.6}$$

By taking the differential we obtain the $\mathbb{Q}^i$-dynamics of ρ_i^{i-1}:

$$d\rho_i^{i-1}(t) = a_i \alpha_i dL_i(t).$$

Using (2.4) we rewrite the dynamics as

$$d\rho_i^{i-1}(t) = a_i \alpha_i L_i(t) \sigma_i(t) dW^i(t).$$

Next we use (2.6) to obtain

$$\begin{aligned} d\rho_i^{i-1}(t) &= a_i \rho_i^{i-1} \frac{1}{\rho_i^{i-1}} \alpha_i L_i(t) \sigma_i(t) dW^i(t) \\ &= a_i \rho_i^{i-1} \underbrace{\frac{1}{a_i} \frac{1}{1 + \alpha_i(t) L_i(t)}}_{1/\rho_i^{i-1}} \alpha_i L_i(t) \sigma_i(t) dW^i(t). \end{aligned}$$

After simplification we obtain

$$d\rho_i^{i-1}(t) = \rho_i^{i-1} \frac{\alpha_i L_i(t)}{1 + \alpha_i(t).L_i(t)} dW^i(t).$$

We can now identify the Girsanov kernel of ρ_i^{i-1} to be

$$\frac{\alpha_i L_i(t)}{1 + \alpha_i L_i(t)} \sigma_i(t).$$

By the Girsanov theorem we therefore know that

$$dW^i(t) = \frac{\alpha_i L_i(t)}{1 + \alpha_i(t) L_i(t)} \sigma_i(t) dt + dW^{i-1}(t). \tag{2.7}$$

By inserting this into the expression for the process $L_{N-1}(t)$ we obtain

$$\begin{aligned} dL_{N-1}(t) &= \sigma_{N-1}(t)L_{N-1}(t)dW^{N-1}(t) \\ &= \sigma_{N-1}(t)L_{N-1}(t)\left(dW^N(t) - \frac{\alpha_N L_N(t)}{1+\alpha_N L_N(t)}\sigma_N(t)dt\right) \\ &= -\frac{\alpha_N L_N(t)}{1+\alpha_N L_N(t)}\sigma_N(t)\sigma_{N-1}(t)L_{N-1}(t)dt \\ &\quad + \sigma_{N-1}(t)L_{N-1}(t)dW^N(t). \end{aligned}$$

Using (2.7) repeatedly it is easy to show, using induction, the following well known result.

Proposition 2.1. *Under the terminal measure $\mathbb{Q}^N$ the* LIBOR *rates have the following dynamics:*

$$dL_i(t) = -\left(\sum_{k=i+1}^{N} \frac{\alpha_k L_k(t)}{1+\alpha_k L_k(t)}\sigma_k(t)\right)\sigma_i(t)L_i(t)dt + \sigma_i(t)L_i(t)dW^N,$$

for $i = 1, \ldots, N$ and where the convention $\sum_{N+1}^{N}(\ldots) = 0$ is used.

To facilitate the algebraic treatment of the model that will take place in Section 2.4 and 2.5, we will now renumber the variables and the parameters of the model, and write the dynamics in logarithmic form.

We set $X_1 := L_N$, $X_2 := L_{N-1}$, $\ldots$, $X_N := L_1$ and by a slight abuse of notation we also renumber the parameters in the following way: $\sigma_i = \sigma_{N-i+1}$ and $\alpha_i = \alpha_{N-i+1}$ for $i = 1, \ldots N$. With this notation the model now reads

$$dX_i(t) = -\Big(\sum_{k=1}^{i-1} \frac{\alpha_k X_k(t)}{1+\alpha_k X_k(t)}\sigma_k(t)\Big)\sigma_i(t)X_i(t)dt + \sigma_i(t)X_i(t)dW^N,$$

for $i = 1, \ldots, N$.

To get the model in logarithmic form we set $Y_i = \log X_i$, for $i = 1, \ldots, N$. By a simple application of the itō formula we obtain

$$dY_i(t) = -\left(\sum_{k=1}^{i-1} \frac{\alpha_k e^{Y_i(t)}}{1+\alpha_k e^{Y_i(t)}}\sigma_k(t) + \frac{1}{2}\sigma_i(t)\right)\sigma_i(t)dt + \sigma_i(t)dW^N, \quad (2.8)$$

where $i = 1, \ldots, N$. These are the dynamics that will be used in the subsequent treatment of the LIBOR market model.

2.2.2 The swap market model

The swap market model is based on the same ideas as the LIBOR market model. By modelling the forward swap rate as a GBM, and using a suitable numeraire, a closed-form formula for swaptions prices of the Black-76 type is obtained. In this section, we will content ourselves with a presentation of the variables of the model and the relationship to the LIBOR market model. For details the refer the reader to the expositions by Björk (2004, Chap. 25) or Brigo and Mercurio (2006, Chap. 6).

A swap involves the exchange of a set of floating payments for a set of fixed payment. In a receiver swap with tenor $T_N - T_n$, the holder of the swap will, at the dates $T_{n+1}, T_{n+2}, \ldots T_N$, receive a fixed payment and pay a floating payment. This kind of swap is referred to as a $T_n \times (T_N - T_n)$ swap. The corresponding forward swap rate is defined as follows:

Definition 2.2 (Swap rate)**.** *A forward swap rate $R_n^N(t)$ of a $T_n \times (T_N - T_n)$ swap is*

$$R_n^N(t) = \frac{p_n(t) - p_N(t)}{\sum_{i=n+1}^N \alpha_i p_i(t)}, \tag{2.9}$$

where $\alpha_i = T_i - T_{i-1}$.

The denominator in (2.9) is an important quantity in it's own right so we make the following definition.

Definition 2.3 (Accrual factor)**.** *The process $S_n^k(t)$, $n < K$, given by*

$$S_n^k(t) = \sum_{i=n+1}^{k} \alpha_i p_i(t),$$

is called the accrual factor.

The accrual factor is in fact the value at time t of a traded quantity, namely the value of a buy-and-hold portfolio composed, for each i, of α_i units of zero-coupon bond maturing at time T_i. Being the value of a traded asset, the accrual factor can be used as a numeraire. We will denote the martingale measure for the numeraire S_n^k by $\mathbb{Q}_n^k$.

The basic idea of the swap market model is to model the swap rate R_n^N as a GBM and to use the accrual factor as a numeraire. This allows swaptions to be priced with a Black-76 type formula.

Definition 2.4 (Swap market model)**.** *Let T_0, T_1, T_n be resettlement dates and $\mathcal{N} = \{(n,k) : n,k \in \mathbb{N}, 0 \leq n < k \leq N\}$. A swap market model with volatilities $\sigma_{n,k}$ is specified by assuming that the swap rates have the dynamics*

$$dR_n^k(t) = R_n^k(t)\sigma_{n,k}(t)dW_n^k(t),$$

where $W_n^k(t)$ is a Wiener process, possibly multidimensional under the measure $\mathbb{Q}_n^k$.

As with the LIBOR market model, it is possible to express all the swap rates under a common measure. We will not pursue this issue here but will instead concentrate on the relation between the swap market model and the LIBOR market model. For our present purpose of studying the (minimal) dimension of market models it turns out, as will become clear in the next section, that it is sufficient to study the LIBOR market model. The swap rate being a function of the LIBOR rate, any conclusion we can draw concerning the minimal dimension of one model also applies to the other model. We have chosen to concentrate on the LIBOR market model because the calculations turn out to be more succinct.

We will now express the swap rate as a nonlinear function of the LIBOR rate. Using (2.1) we have

$$\frac{p_i(t)}{p_{i-1}(t)} = \frac{1}{1+\alpha_i L_i(t)}.$$

This enable us to write

$$\begin{aligned}
\frac{p_N(t)}{p_n(t)} &= \frac{p_N(t)}{p_{N-1}(t)} \cdot \frac{p_{N-1}(t)}{p_{N-2}(t)} \cdots \frac{p_{n+1}(t)}{p_n(t)} \\
&= \frac{1}{1+\alpha_N L_N(t)} \cdot \frac{1}{1+\alpha_{N-1} L_{N-1}(t)} \cdots \frac{1}{1+\alpha_{n+1} L_{n+1}(t)} \qquad (2.10)\\
&= \prod_{i=n+1}^{N} \frac{1}{1+\alpha_i L_i(t)}.
\end{aligned}$$

The swap rate, given in (2.9), can be rewritten as

$$\begin{aligned}
R_n^N(t) &= \frac{p_n(t) - p_N(t)}{\sum_{i=n+1}^{N} \alpha_i p_i(t)} = p_n(t)\left(1 - \frac{p_N(t)}{p_n(t)}\right)\frac{1}{\sum_{i=n+1}^{N} \alpha_i p_i(t)} \\
&= \left(1 - \frac{p_N(t)}{p_n(t)}\right)\frac{1}{\sum_{i=n+1}^{N} \alpha_i p_i(t)/p_n(t)}.
\end{aligned} \qquad (2.11)$$

By inserting (2.10) into (2.11) we obtain

$$R_n^N(t) = \frac{1 - \prod_{i=n+1}^{N} \frac{1}{1+\alpha_i L_i(t)}}{\sum_{i=n+1}^{N} \alpha_i \prod_{j=n+1}^{i} \frac{1}{1+\alpha_j L_j(t)}}. \qquad (2.12)$$

Remark 2.2. *When the LIBOR rates are modelled as log-normal stochastic variables under their own measure it is clear from the preceding expression* *that we should not expect the swap rates to be log-normal variables under the measure $\mathbb{Q}_n^N$. This highlights the classic problem of the market model approach, namely that the* LIBOR *market model and the swap market model are incompatible. Surprisingly however, simulations show that the swap rates are in fact not far from being log-normal.*[6]

2.3 Problem formulation

One major drawback of the market models is that the dimension of the Markov system they are made up of is the same as the number of LIBOR or swap rates. The pricing of a derivative security having cash flows at N different points in time, requires a system of size N. It would therefore be very convenient if it was possible the transform the model in such a way as to reduce its dimension. In this section we will see how this could in principle be possible.

A market model $\mathcal{M}$ is formulated as a system of stochastic differential equations (SDE's). Each equation models the behaviour of a single LIBOR or swap rate. In the definitions that follow we will focus on the LIBOR market model. The same definitions can of course be expressed in terms of swap rates.

Instead of viewing the model $\mathcal{M}$ as a system consisting of N equations we can view it as a single equation specifying the dynamics of an N-dimensional object. In vector form the model presented in Proposition 2.1 becomes

$$(\mathcal{M}) \qquad dL(t) = \mu(t, L_t)dt + \Sigma(t, L_t)dW_t, \tag{2.13}$$

where

$$dL(t) = \begin{bmatrix} dL_1(t) \\ dY_2(t) \\ \vdots \\ dL_N(t) \end{bmatrix}, \qquad \Sigma(t, L_t) = \begin{bmatrix} \sigma_1(t)L_1(t) \\ \sigma_2(t)L_2(t) \\ \vdots \\ \sigma_N(t)L_N(t) \end{bmatrix},$$

$$\mu(t, L_t) = \begin{bmatrix} -\sigma_1(t)\sum_{k=2}^{N} \frac{\alpha_k L_k(t)}{1+\alpha_k L_k(t)}\sigma_k(t) \\ -\sigma_2(t)\sum_{k=3}^{N} \frac{\alpha_k L_k(t)}{1+\alpha_k L_k(t)}\sigma_k(t) \\ \vdots \\ -\sigma_{N-1}(t)\frac{\alpha_N L_N(t)}{1+\alpha_1 L_N(t)}\sigma_N(t) \\ 0 \end{bmatrix}.$$

[6] See Brigo and Mercurio (2006, Chap. 8).

The solution to (2.16), which can be shown to exist, is an N-dimensional stochastic process $L(t)$.[7]

We now make the following definitions:

Definition 2.5 (Realization)**.** *An m-dimensional realization of the SDE* (2.13) *is a point $z_o \in \mathbb{R}^m$, an open set $\mathcal{Z}$ of $\mathbb{R}^m$, smooth functions a and b, and a map $G : \mathbb{R}^m \to \mathbb{R}^N$ such that $L(t)$ has the local representation*

$$L(t) = G(z_t) \qquad P.\ a.s.$$

where z_t is the strong solution of the m-dimensional SDE

$$\begin{aligned} dz(t) &= a(z_t)dt + b(z_t)dW_t \\ z(0) &= z_0, \end{aligned} \tag{2.14}$$

and where the Wiener process W in (2.14) *is the same as in* (2.13)*. The process z is called the state process.*

Remark 2.3. *Note that the dimension of the realization is defined pointwise, in a neighborhood of the point z_0.*

Main Problem 2.1. *The objective is to find a a representation with smaller dimension than that of the SDE of $L(t)$ in* (2.13)*, i.e we want $m < N$. Otherwise we could just choose G to be the identity map and set $z(t) = L(t)$.*

Ultimately we are searching for the minimal realization.

Definition 2.6 (Minimal Realization)**.** *The dimension of a realization is defined as the dimension of the state process z. A realization is minimal if there is no other realization of smaller dimension.*

The object of the present paper is to find the dimension of the minimal realization of the LIBOR market model and of the swap market model.

Instead of doing a separate investigation of the two market models, it in fact enough to study the realization problem for the LIBOR market. From equation (2.12) in Section 2.2.2 we know that the swap rate $R_n^N(t)$ is a nonlinear function H of the LIBOR rate, given by

$$H(L_t) = \frac{1 - \prod_{i=n+1}^{N} \frac{1}{1+\alpha_i L_i(t)}}{\sum_{i=n+1}^{N} \alpha_i \prod_{j=n+1}^{i} \frac{1}{1+\alpha_j L_j(t)}}. \tag{2.15}$$

Finding the minimal realization for the swap market, given by Definition 2.4, amounts to finding a mapping $\tilde{G} : \mathbb{R}^m \to \mathbb{R}^N$, $m < N$, and a

[7] See, for example, Björk (2004, Chap. 25) or Hunt and Kennedy (2000, Chap. 18).

state process z so that

$$R_n^N(t) = \tilde{G}(z_t)$$
$$dz_t = a(z_t)dt + b(z_t)dW_t.$$

Assuming we have already found a minimal realization for the corresponding LIBOR market model, given by a map G and a m-dimensional state process z, the minimal realization for the swap market model will simply be given by the composition of H and G as follows:

$$\tilde{G}(z_t) = H \circ G(z_t).$$

We have thereby showed the following

Proposition 2.2. *Every minimal realization of the swap market model can be translated into a minimal realization of the corresponding* LIBOR *market model.*

In order to find the size of the minimal realization we will employ the framework of Björk and Svensson (2001). This framework was developed to study finite dimensional realizations of infinite dimensional forward rate models in the Heath, Jarrow, and Morton (1992) setting. It can however without modification be employed to study realizations of finite dimensional models.

We will not present the Björk-Svensson framework here but refer the reader to the original article Björk and Svensson (2001) or to the overview in Björk (2003).

For our purpose the only result needed to conduct our investigation is the following which we provide without proof.

Proposition 2.3. *The dimension of the minimal realization of the* LIBOR *market model is equal to the dimension of the Lie algebra, evaluated pointwise, generated by μ and σ in* (2.13).

2.4 Constant volatility

In this section we will study the Lie algebra generated by the vector fields μ and Σ in the logarithmic formulation of the LIBOR market model.

Writing (2.8) in vector we obtain.

$$dY(t) = \mu(t, Y_t)dt + \Sigma(t, Y_t)dW_t, \tag{2.16}$$

where

$$dY(t) = \begin{bmatrix} dY_1(t) \\ dY_2(t) \\ \vdots \\ dY_N(t) \end{bmatrix}, \qquad \Sigma(t, Y_t) = \begin{bmatrix} \sigma_1(t) \\ \sigma_2(t) \\ \vdots \\ \sigma_N(t) \end{bmatrix},$$

$$\mu(t, Y_t) = \begin{bmatrix} -\frac{1}{2}\sigma_1^2(t) \\ -\left(\frac{\alpha_1 e^{Y_1(t)}}{1+\alpha_1 e^{Y_1(t)}}\sigma_1(t) + \frac{1}{2}\sigma_2(t)\right)\sigma_2(t) \\ \vdots \\ -\left(\sum_{k=1}^{N-1} \frac{\alpha_k e^{Y_i(t)}}{1+\alpha_k e^{Y_i(t)}}\sigma_k(t) + \frac{1}{2}\sigma_N(t)\right)\sigma_N(t) \end{bmatrix}.$$

In the present section we assume that the volatility is constant, the time-dependent case being treated in the next section. Hence, we set the components of Σ to be constants $\sigma_1, \dots, \sigma_N$. Next, for simplification, we multiply the vector μ by -2. We end up with the following vector fields:

$$f_0(y_1, \dots, y_{N-1}) = \begin{bmatrix} \sigma_1^2 \\ \left(\frac{2\alpha_1 e^{y_1}}{1+\alpha_1 e^{y_1}}\sigma_1 + \sigma_2\right)\sigma_2 \\ \vdots \\ \left(\sum_{k=1}^{N-1} \frac{2\alpha_k e^{y_k}}{1+\alpha_k e^{y_k}}\sigma_k + \sigma_N\right)\sigma_N \end{bmatrix} \qquad \sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_N \end{bmatrix}.$$

We set

$$G_k := \frac{2\alpha_k e^{y_k}}{1+\alpha_k e^{y_k}}\sigma_k,$$

and rewrite f_0 as

$$f_0 = \begin{bmatrix} \sigma_1^2 \\ (G_1 + \sigma_2)\sigma_2 \\ (G_1 + G_2 + \sigma_3)\sigma_3 \\ \vdots \\ (\sum_{k=1}^{N-1} G_k + \sigma_N)\sigma_N \end{bmatrix}.$$

With the vector fields in the appropriate form we can start computing Lie brackets, the first bracket being

$$[f_0, \sigma] = f_0'[\sigma] - \sigma'[f_0].$$

Here $f_0'(y)[\sigma]$ denotes the Jacobian matrix of $f_0(y)$ operating on σ. As σ is constant, trivially $\sigma' = 0$. The Jacobian matrix of f_0 is

$$f_0' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \sigma_2 G_1' & 0 & 0 & \dots & 0 \\ \sigma_3 G_1' & \sigma_3 G_2' & 0 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \sigma_N G_1' & \sigma_N G_2' & \dots & \sigma_N G_{N-1}' & 0 \end{bmatrix},$$

where

$$G_k' = \frac{\partial G_k}{\partial y_k}.$$

Taking brackets we get

$$f_1 := [f_0, \sigma] = f_0'\sigma = \begin{bmatrix} 0 \\ \sigma_2(\sigma_1 G_1') \\ \sigma_3(\sigma_1 G_1' + \sigma_2 G_2') \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1} \sigma_k G_k') \end{bmatrix}.$$

We now continue to take brackets of f_1 with σ. The Jacobian of f_1 is given by

$$f_1' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \sigma_2\sigma_1 G_1'' & 0 & 0 & \dots & 0 \\ \sigma_3\sigma_1 G_1'' & \sigma_3\sigma_2 G_2'' & 0 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \sigma_N\sigma_1 G_1'' & \sigma_N\sigma_2 G_2'' & \dots & \sigma_N\sigma_{N-1} G_{N-1}'' & 0 \end{bmatrix}.$$

Taking the bracket of f_1 with σ produces

$$f_2 := [f_1, \sigma] = f_1'\sigma = \begin{bmatrix} 0 \\ \sigma_2(\sigma_1^2 G_1'') \\ \sigma_3(\sigma_1^2 G_1'' + \sigma_2^2 G_2'') \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1} \sigma_k^2 G_k'') \end{bmatrix}.$$

Next the Jacobian of f_2 is given by

$$f_1' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \sigma_2\sigma_1^2 G_1^{(3)} & 0 & 0 & \dots & 0 \\ \sigma_3\sigma_1^2 G_1^{(3)} & \sigma_3\sigma_2^2 G_2^{(3)} & 0 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \sigma_N\sigma_1^2 G_1^{(3)} & \sigma_N\sigma_2^2 G_2^{(3)} & \dots & \sigma_N\sigma_{N-1}^2 G_{N-1}^{(3)} & 0 \end{bmatrix}.$$

Taking the bracket of f_2 with σ yields

$$f_3 := [f_2, \sigma] = f_2'\sigma = \begin{bmatrix} 0 \\ \sigma_2(\sigma_1^3 G_1^{(3)}) \\ \sigma_3(\sigma_1^3 G_1^{(3)} + \sigma_2^3 G_2^{(3)}) \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1} \sigma_k^3 G_k^{(3)}) \end{bmatrix}.$$

By continuing to take brackets in the same fashion, m times in total, we obtain

$$f_m := [f_1, [f_2, [\dots, [f_{m-1}, \sigma]\dots]]] = \begin{bmatrix} 0 \\ \sigma_2(\sigma_1^m G_1^{(m)}) \\ \sigma_3(\sigma_1^m G_1^{(m)} + \sigma_2^m G_2^{(m)}) \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1} \sigma_k^m G_k^{(m)}) \end{bmatrix}.$$

Given a point $y^N = (y_1, y_2, \dots, y_N)$, we want to investigate whether the N vector fields σ, f_0, f_1, $\dots$, f_{N-2} are linearly independent in a neighborhood $B_N(y^N) = \{\xi \in \mathbb{R}^N : |\xi - y^N| < \epsilon\}$ of y^N. This amounts to studying the ranks of the following matrix:

$$\begin{bmatrix} \vdots & \vdots & \vdots & & \vdots \\ \sigma & f_0 & f_1 & \dots & f_{N-2} \\ \vdots & \vdots & \vdots & & \vdots \end{bmatrix},$$

which in full details reads

$$\begin{bmatrix} \sigma_1 & \sigma_1^2 & 0 & \dots & 0 \\ \sigma_2 & \sigma_2(G_1 + \sigma_2) & \sigma_2\sigma_1 G_1' & \dots & \sigma_2\sigma_1^{N-2} G_1^{(N-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma_N & \sigma_N(\sum_{k=1}^{N-1} G_k + \sigma_N) & \sigma_N \sum_{k=1}^{N-1} \sigma_k G_k' & \dots & \sigma_N \sum_{k=1}^{N-1} \sigma_k^{N-2} G_k^{(N-2)} \end{bmatrix}.$$

In order to facilitate the study of the rank of the matrix, we perform some Gaussian elimination. First, we divide row number i with σ_i for $1 \leq i \leq N$. This yields

$$\begin{bmatrix} 1 & \sigma_1 & 0 & \dots & 0 \\ 1 & G_1 + \sigma_2 & \sigma_1 G_1' & \dots & \sigma_1^{N-2} G_1^{(N-2)} \\ 1 & G_1 + G_2 + \sigma_3 & \sigma_1 G_1' + \sigma_2 G_2' & \dots & \sigma_1^{N-2} G_1^{(N)} + \sigma_2^{(N-2)} G_2^{(N-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \sum_{k=1}^{N-1} G_k + \sigma_N & \sum_{k=1}^{N-1} \sigma_k G_k' & \dots & \sum_{k=1}^{N-1} \sigma_k^{N-2} G_k^{(N-2)} \end{bmatrix}.$$

Next, we subtract the first row from the remaining rows. This yields

$$\begin{bmatrix} 1 & \sigma_1 & 0 & \dots & 0 \\ 0 & G_1 + \sigma_2 - \sigma_1 & \sigma_1 G_1' & \dots & \sigma_1^{N-2} G_1^{(N-2)} \\ 0 & G_1 + G_2 + \sigma_3 - \sigma_1 & \sigma_1 G_1' + \sigma_2 G_2' & \dots & \sigma_1^{N-2} G_1^{(N)} + \sigma_2^{(N-2)} G_2^{(N-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \sum_{k=1}^{N-1} G_k + \sigma_N - \sigma_1 & \sum_{k=1}^{N-1} \sigma_k G_k' & \dots & \sum_{k=1}^{N-1} \sigma_k^{N-2} G_k^{(N-2)} \end{bmatrix}.$$

Then we subtract row i from row $i+1$ when $2 \leq i \leq N-1$ and obtain

$$\begin{bmatrix} 1 & \sigma_1 & 0 & \dots & 0 \\ 0 & G_1 + \sigma_2 - \sigma_1 & \sigma_1 G_1' & \dots & \sigma_1^{N-2} G_1^{(N-2)} \\ 0 & G_2 + \sigma_3 - \sigma_2 & \sigma_2 G_2' & \dots & \sigma_2^{(N-2)} G_2^{(N-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & G_{N-1} + \sigma_N - \sigma_{N-1} & \sigma_{N-1} G_{N-1}' & \dots & \sigma_{N-1}^{N-2} G_{N-1}^{(N-2)} \end{bmatrix}.$$

Finally, we subtract the σ_1 times the first column from the second column to obtain

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & G_1 + \sigma_2 - \sigma_1 & \sigma_1 G_1' & \dots & \sigma_1^{N-2} G_1^{(N-2)} \\ 0 & G_2 + \sigma_3 - \sigma_2 & \sigma_2 G_2' & \dots & \sigma_2^{(N-2)} G_2^{(N-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & G_{N-1} + \sigma_N - \sigma_{N-1} & \sigma_{N-1} G_{N-1}' & \dots & \sigma_{N-1}^{N-2} G_{N-1}^{(N-2)} \end{bmatrix}$$

At this point a change in notation is convenient. We set $n := N-1$ and

$y^n = (y_1, y_2, \ldots, y_n)$ and obtain the following $n \times n$ matrix.

$$H_n(y^n) = \begin{bmatrix} G_1 + \sigma_2 - \sigma_1 & \sigma_1 G_1' & \ldots & \sigma_1^{n-1} G_1^{(n-1)} \\ G_2 + \sigma_3 - \sigma_2 & \sigma_2 G_2' & \ldots & \sigma_2^{(n-1)} G_2^{(n-1)} \\ \vdots & \vdots & & \vdots \\ G_n + \sigma_{n+1} - \sigma_n & \sigma_n G_n' & \ldots & \sigma_n^{n-1} G_n^{(n-1)} \end{bmatrix}. \tag{2.17}$$

Our goal now is to show that the matrix $H_n(y^n)$ has full rank. In the next proposition, we will in fact prove a slightly more general result. This has the benefit of making the proof more transparent, but more importantly, will enable us to use the result in the next section where we study the time-dependent volatility case.

Define the matrix $F_n(y)$ in the following way

$$F_n(y) = \begin{bmatrix} f_1^1(y_1) & f_1^2(y_1) & \ldots & f_1^{n-1}(y_1) & f_1^n(y_1) \\ f_2^1(y_2) & f_2^2(y_2) & \ldots & f_2^{n-1}(y_2) & f_2^n(y_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f_n^1(y_n) & f_n^2(y_n) & \ldots & f_n^{n-1}(y_n) & f_n^n(y_n) \end{bmatrix}. \tag{2.18}$$

We will also consider the submatrix of F_n given by

$$F_{n-1}(y_1, \ldots, y_{n-1}) = \begin{bmatrix} f_1^1(y_1) & f_1^2(y_1) & \ldots & f_1^{n-1}(y_1) \\ f_2^1(y_2) & f_2^2(y_2) & \ldots & f_2^{n-1}(y_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n-1}^1(y_{n-1}) & f_{n-1}^2(y_{n-1}) & \ldots & f_{n-1}^{n-1}(y_{n-1}) \end{bmatrix}.$$

We make the following assumption:

Assumption 2.1.

(i) the functions $f_i^j(y_i)$, $i, j = 1, \ldots n$, are continuous functions on the real line $\mathbb{R}$.

(ii) In each row i of the matrix $F_n(y)$, the functions $f_i^1, f_i^2, \ldots, f_i^n$ are linearly independent over the real field.

Given a point $y_0 \in \mathbb{R}^n$, we study the matrix $F_n(y)$ in a neighborhood $B_n(y_0)$ of the point $y_0 \in \mathbb{R}^n$. Let $B_n(y_0) := \{x \in \mathbb{R}^n : \|x - y_0\| < \epsilon\}$ where the real $\epsilon > 0$. The submatrix F_{n-1} is studied on the projection of the neighborhood $B_n(y_0)$ on $\mathbb{R}^{n-1}$, denoted $B_{n-1}(y_0)$, i.e.

$$B_{n-1}(y) := \operatorname{proj}_{\mathbb{R}^{n-1}} B_n(y).$$

Note that $B_{n-1}(y_0)$ is a function of $y_1, \ldots, y_{n-1}$ only. To avoid any confusion in this respect, elements of B_n will be denoted x and elements of B_{n-1} will be denoted by z.

To prove that the matrix $F_n(y)$ has full rank, we will use the following result.

Lemma 2.1. *If* $\det F_n(x) \equiv 0$ *for all* $x \in B_n(y)$ *then* $\det F_{n-1}(z) \equiv 0$ *for all* $z \in B_{n-1}(y)$.

Proof. The statement will be proved by contradiction. We therefore assume that $\det F_n(x) \equiv 0$ for all $x \in B_n(y)$ and that there exists a $z \in B_{n-1}(y)$ such that $\det F_{n-1}(z) \equiv 0$. Another way to express the later is by letting $C_{n-1} := \{z \in \mathbb{R}^{n-1} : \det H_{n-1}(z) \neq 0\}$ and assuming that $B_{n-1} \cap C_{n-1} \neq \emptyset$. Note that as the determinant is a continuous function, the set C_{n-1} is an open set being the inverse image of an open set.

For all z in the non-empty and open set $B_{n-1} \cap C_{n-1}$, the column vectors of the matrix $F_{n-1}(z)$

$$v_1^{n-1}(z) = \begin{bmatrix} f_1^1(z_1) \\ f_2^1(z_2) \\ \vdots \\ f_{n-1}^1(z_{n-1}) \end{bmatrix}, v_2^{n-1}(z) = \begin{bmatrix} f_1^2(z_1) \\ f_2^2(z_2) \\ \vdots \\ f_{n-1}^2(z_{n-1}) \end{bmatrix}, \ldots,$$

$$v_{n-1}^{n-1}(z) = \begin{bmatrix} f_1^{n-1}(z_1) \\ f_2^{n-1}(z_2) \\ \vdots \\ f_{n-1}^{n-1}(z_{n-1}) \end{bmatrix},$$

are linearly independent as $z \in C_{n-1}$. Adding up to $n-1$ independent vectors, they consequently span $\mathbb{R}^{n-1}$. In particular the vector

$$v_{n-1}^n(z) = \begin{bmatrix} f_1^n(z_1) \\ f_2^n(z_2) \\ \vdots \\ f_{n-1}^n(z_{n-1}) \end{bmatrix}$$

lies in their span. Therefore there exists unique functions $h_1(z), \ldots, h_{n-1}(z)$ such that

$$v_{n-1}^n(z) = \sum_{i=1}^{n-1} h_i(z) v_i^{n-1}. \tag{2.19}$$

If we interpret this equation component-wise, for row i, $1 \leq i \leq n-1$, we obtain

$$f_i^n(z_i) = h_1(z) f_i^1(z_i) + h_2(z) f_i^2(z_i) + \ldots + h_{n-1}(z) f_i^{n-1}(z_i). \tag{2.20}$$

Next, we consider the column vectors of the matrix $F_n(x)$ for all x in B_n. We can identify the first $n-1$ component of the vector x with z, i.e.

$$x = (x_1, x_2, \ldots, x_{n-1}, x_n) = (z_1, z_2, \ldots, z_{n-1}, x_n).$$

as $z \in B_{n-1}$. Therefore we have $v_i^{n-1}(z) = v_i^{n-1}(x)$ for $i = 1, \ldots n-1$.

Since the vectors $v_1^{n-1}, v_2^{n-1}, \ldots, v_{n-1}^{n-1}$ are linearly independent it follows that also the vectors

$$v_1^n(x) = \begin{bmatrix} f_1^1(x_1) \\ f_2^1(x_2) \\ \vdots \\ f_{n-1}^1(x_{n-1}) \\ f_n^1(x_n) \end{bmatrix}, v_2^n(x) = \begin{bmatrix} f_1^2(x_1) \\ f_2^2(x_2) \\ \vdots \\ f_{n-1}^2(x_{n-1}) \\ f_n^2(x_n) \end{bmatrix}, \ldots,$$

$$v_{n-1}^n(x) = \begin{bmatrix} f_1^{n-1}(x_1) \\ f_2^{n-1}(x_2) \\ \vdots \\ f_{n-1}^{n-1}(x_{n-1}) \\ f_n^{n-1}(x_n) \end{bmatrix},$$

are linearly independent. Knowing that $\det F_n = 0$ for all $x \in B_n$, the vector v_n^n must consequently lie in their span. Therefore there exists unique functions $k_1(x), k_2(x), \ldots, k_{n-1}(x)$ such that

$$v_n^n(x) = \sum_{i=1}^{n-1} k_i(x) v_i^n(x). \tag{2.21}$$

We will now focus on the first $n-1$ components of (2.21). Component-wise for row i, where $1 \leq i \leq n-1$, we have

$$f_i^n(x_i) = k_1(x) f_i^1(x_i) + k_2(x) f_i^2(x_i) + \ldots + k_{n-1}(x) f_i^{n-1}(x_i).$$

By observing that $x_i = z_i$ for $1 \leq i \leq n-1$ we can rewrite this equation as

$$f_i^n(z_i) = k_1(x) f_i^1(z_i) + k_2(x) f_i^2(z_i) + \ldots + k_{n-1}(x) f_i^{n-1}(z_i). \tag{2.22}$$

By comparing (2.20) and (2.22) and by using the uniqueness of $k_i(x)$ and $h_i(z)$ respectively, we can conclude that $k_i(x) = h_i(z)$, $1 \le i \le n-1$. This means that the functions $k_i(x)$ we have solved for, are in fact only functions of $x_1, \ldots, x_{n-1}$. The last component of (2.21) can therefore be written

$$f_n^n(x_n) = \sum_{i=1}^{n} h_i(z) f_n^n(x_n). \tag{2.23}$$

The important fact here is that the functions $h_i(z)$ do not depend on x_n. This allows us to finalize our proof. We have reached a contradiction since we initially assumed that the functions $f_n(x_n)$ are linearly independent over the real field. The assumption that $B_{n-1} \cap C_{n-1} \neq \emptyset$ must therefore be wrong. We can conclude that $\det F_{n-1} \equiv 0$ for all $z \in B_{n-1}(y)$. □

We can now immediately apply Lemma 2.1 and prove the linear independence of the columns of the matrix $F_n(x)$.

Proposition 2.4. *Let $y_0 \in \mathbb{R}^n$ be given. Set $B_n(y_0) = \{x \in \mathbb{R}^n : \|x - y_0\| < \epsilon\}$, $\epsilon > 0$, and $B_{n-i} = \mathrm{proj}_{\mathbb{R}^{n-i}} B_n$. Under Assumption 2.1 the matrix $F_n(x)$ defined in (2.18) has full rank for all $x \in B_n(y_0)$.*

Proof. Assume on the contrary that $\det F_n(x) \equiv 0$ on B_n. Application of Lemma 2.1 gives $\det F_{n-1}(z) \equiv 0$ for all $z \in B_{n-1}$. By repeated use of Lemma 2.1 we finally get, after $n-1$ steps, that $\det F_1(\xi) \equiv 0$ for $\xi \in B_1$ or more precisely that $f_1^1(\xi) \equiv 0$. This is in conflict with Assumption 2.1, where we state that the functions f_i^j where linearly independent. We thus have a contradiction. The assumption that $\det F_n(x) \equiv 0$ on B_n must therefore be false, and we are done. □

Returning to the LIBOR market model we apply the previous theorem on the matrix $H_n(x)$ in (2.17).

Proposition 2.5. *The matrix $H_n(x)$ has full rank in a for all $x \in B_n(y_0)$.*

Proof. We simply observe that the functions G_i, which we recall are given by

$$G_i = \frac{2\alpha_i e^{y_i}}{1 + \alpha_i e^{y_i}} \sigma_i, \qquad 1 \le i \le n,$$

as well as their derivatives $G_j^{(m)}$, $1 \le m \le n-1$, are linearly independent over $\mathbb{R}$. The matrix $H_n(x^n)$ has therefore the same form as $F_n(y)$. The statement therefore follows from application of Proposition 2.4. □

We can now conclude our examination of the dimension of market models in the constant volatility case.

Proposition 2.6. *The N-dimensional* LIBOR *market model with constant volatility and the corresponding swap market model do not posses a realization of smaller dimension, in the sense of Definition 2.5.*

2.5 Time-dependent volatility

Next we tackle the slightly more general case of time-dependent volatility. The Lie algebra becomes a little more involved but as we will see we can use the same ideas as in the previous section. The relevant vector fields are now

$$\hat{f}_0(y_1,\ldots,y_{N-1},t) = \begin{bmatrix} \sigma_1^2(t) \\ \left(\frac{2\alpha_1 e^{y_1}}{1+\alpha_1 e^{y_1}}\sigma_1(t)+\sigma_2(t)\right)\sigma_2(t) \\ \vdots \\ \left(\sum_{k=1}^{N-1}\frac{2\alpha_k e^{y_k}}{1+\alpha_k e^{y_k}}\sigma_k(t)+\sigma_N(t)\right)\sigma_N(t) \\ 1 \end{bmatrix},$$

$$\hat{\sigma} = \begin{bmatrix} \sigma_1(t) \\ \sigma_2(t) \\ \vdots \\ \sigma_N(t) \\ 0 \end{bmatrix}.$$

We set

$$G_k := H_k\sigma_k(t) \quad \text{where } H_k := \frac{2\alpha_k e^{y_k}}{1+\alpha_k e^{y_k}},$$

and rewrite $\hat{f}_0$ as

$$\hat{f}_0 = \begin{bmatrix} \sigma_1^2(t) \\ (G_1(t)+\sigma_2(t))\sigma_2(t) \\ (G_1(t)+G_2(t)+\sigma_3(t))\sigma_3(t) \\ \vdots \\ (\sum_{k=1}^{N-1} G_k(t)+\sigma_N(t))\sigma_N(t) \\ 1 \end{bmatrix}.$$

In order to save the space and make the matrices more readable we will from now on not specify the time dependence of $\sigma_i(t)$ and $G_i^{(m)}(t)$ in the matrices.

The first N colums of the Jacobian matrix of $\hat{f}_0$ are

$$\hat{f}_0' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \sigma_2 G_1' & 0 & 0 & \dots & 0 \\ \sigma_3 G_1' & \sigma_3 G_2' & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_N G_1' & \sigma_N G_2' & \sigma_N G_3' & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

and the last column of the Jacobian matrix of $\hat{f}_0$ is

$$\begin{bmatrix} 2\sigma_1 \partial_t \sigma_1 \\ \partial_t \sigma_2 (G_1 + \sigma_2) + \sigma_2 \partial_t (G_1 + \sigma_2) \\ \partial_t \sigma_3 (G_1 + G_2 + \sigma_3) + \sigma_3 \partial_t (G_1 G_2 + \sigma_3) \\ \vdots \\ \partial_t \sigma_N (\sum_{k=1}^{N-1} G_k + \sigma_N) + \sigma_N \partial_t (\sum_{k=1}^{N-1} G_k + \sigma_N) \\ 0 \end{bmatrix}$$

where

$$G_k' = \frac{\partial G_k}{\partial y_k} \qquad \text{and} \qquad \partial_t \sigma_i = \frac{\partial \sigma_i}{\partial t}$$

The Jacobian matrix of $\hat{\sigma}(t)$

$$\hat{\sigma}' = \begin{bmatrix} 0 & \dots & 0 & \partial_t \sigma_1 \\ 0 & \dots & 0 & \partial_t \sigma_2 \\ \dots & & & \\ 0 & \dots & 0 & \partial_t \sigma_N \\ 0 & \dots & 0 & 0 \end{bmatrix}.$$

Taking brackets we get

$$\hat{f}_1 := [\hat{f}_0, \hat{\sigma}] = \hat{f}_0' \hat{\sigma} - \sigma' \hat{f}_0 = \begin{bmatrix} -\partial_t \sigma_1 \\ \sigma_2 (\sigma_1 G_1') - \partial_t \sigma_2 \\ \sigma_3 (\sigma_1 G_1' + \sigma_2 G_2') - \partial_t \sigma_3 \\ \vdots \\ \sigma_N (\sum_{k=1}^{N-1} \sigma_k G_k') - \partial_t \sigma_N \\ 0 \end{bmatrix}.$$

The next step is to compute the Jacobian matrix of f_1:

$$\hat{f}_1' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_1(t) \\ \sigma_2\sigma_1 G_1'' & 0 & 0 & \dots & 0 & a_2(t) \\ \sigma_3\sigma_1 G_1'' & \sigma_3\sigma_2 G_2'' & 0 & \dots & 0 & a_3(t) \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ \sigma_N\sigma_1 G_1'' & \sigma_N\sigma_2 G_2'' & \dots & \sigma_N\sigma_{N-1} G_{N-1}'' & 0 & a_N(t) \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Taking the bracket of $\hat{f}_1$ with $\hat{\sigma}$ produces

$$\hat{f}_2 := [\hat{f}_1, \hat{\sigma}] = \hat{f}_1'\hat{\sigma} - \hat{\sigma}'\hat{f}_1 = \begin{bmatrix} -\partial_t\sigma_1 \\ \sigma_2(\sigma_1^2 G_1'') - \partial_t\sigma_2 \\ \sigma_3(\sigma_1^2 G_1'' + \sigma_2^2 G_2'') - \partial_t\sigma_3 \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1}\sigma_k^2 G_k'') - \partial_t\sigma_N \\ 0 \end{bmatrix}.$$

The next bracket is given by

$$\hat{f}_3 := [\hat{f}_2, \sigma] = \hat{f}_2'\sigma - \hat{\sigma}'\hat{f}_2 = \begin{bmatrix} -\partial_t\sigma_1 \\ \sigma_2(\sigma_1^3 G_1^{(3)}) - \partial_t\sigma_2 \\ \sigma_3(\sigma_1^3 G_1^{(3)} + \sigma_2^3 G_2^{(3)}) - \partial_t\sigma_3 \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1}\sigma_k^3 G_k^{(3)}) - \partial_t\sigma_N \\ 0 \end{bmatrix}.$$

After taking brackets m times we obtain

$$\hat{f}_m := [f_1, [f_2, [\dots, [f_{m-1}, \sigma]\dots]]] = \begin{bmatrix} -\partial_t\sigma_1 \\ \sigma_2(\sigma_1^m G_1^{(m)}) - \partial_t\sigma_2 \\ \sigma_3(\sigma_1^m G_1^{(m)} + \sigma_2^m G_2^{(m)}) - \partial_t\sigma_3 \\ \vdots \\ \sigma_N(\sum_{k=1}^{N-1}\sigma_k^m G_k^{(m)}) - \partial_t\sigma_N \\ 0 \end{bmatrix}.$$

In the same way as in the previous section, we collect all the vector fields in a matrix:

$$\begin{bmatrix} \vdots & \vdots & \vdots & & \vdots \\ \hat{\sigma} & \hat{f}_0 & \hat{f}_1 & \dots & \hat{f}_{N-2} \\ \vdots & \vdots & \vdots & & \vdots \end{bmatrix}.$$

The first N column are

$$\begin{bmatrix} \sigma_1 & \sigma_1^2 & -\partial_t \sigma_1 & \dots \\ \sigma_2 & \sigma_2(G_1 + \sigma_2) & \sigma_2 \sigma_1 G_1' - \partial_t \sigma_2 & \dots \\ \sigma_3 & \sigma_3(G_1 + G_2 + \sigma_3) & \sigma_3(\sigma_1 G_1' + \sigma_2 G_2') - \partial_t \sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ \sigma_N & \sigma_N \sum_{k=1}^{N-1} G_k + \sigma_N & \sigma_N \sum_{k=1}^{N-1} \sigma_k G_k' - \partial_t \sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and the last column is

$$\begin{bmatrix} -\partial_t \sigma_1 \\ \sigma_2 \sigma_1^{N-1} G_1^{(N-1)} - \partial_t \sigma_2 \\ \sigma_3(\sigma_1^{N-1} G_1^{(N-1)} + \sigma_2^{(N-1)} G_2^{(N-1)}) - \partial_t \sigma_3 \\ \vdots \\ \sigma_N \sum_{k=1}^{N-1} \sigma_k^{N-1} G_k^{(N-1)} - \partial_t \sigma_N \\ 0 \end{bmatrix}.$$

The next step is to perform Gaussian elimination. By subtracting suitable multiples of the last row we can remove all other entries in the second column of the matrix and obtain for the first N columns

$$\begin{bmatrix} \sigma_1 & 0 & -\partial_t \sigma_1 & \dots \\ \sigma_2 & 0 & \sigma_2 \sigma_1 G_1' - \partial_t \sigma_2 & \dots \\ \sigma_3 & 0 & \sigma_3(\sigma_1 G_1' + \sigma_2 G_2') - \partial_t \sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ \sigma_N & 0 & \sigma_N \sum_{k=1}^{N-1} \sigma_k G_k' - \partial_t \sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and for the last column we obtain

$$\begin{bmatrix} -\partial_t\sigma_1 \\ \sigma_2\sigma_1^{N-1}G_1^{(N-1)} - \partial_t\sigma_2 \\ \sigma_3(\sigma_1^{N-1}G_1^{(N-1)} + \sigma_2^{(N-1)}G_2^{(N-1)}) - \partial_t\sigma_3 \\ \vdots \\ \sigma_N \sum_{k=1}^{N-1} \sigma_k^{N-1}G_k^{(N-1)} - \partial_t\sigma_N \\ 0 \end{bmatrix}$$

Next we divide row number i with σ_i for $1 \leq i \leq N$. For the first N column we obtain

$$\begin{bmatrix} 1 & 0 & -\partial_t\sigma_1/\sigma_1 & \dots \\ 1 & 0 & \sigma_1 G_1' - \partial_t\sigma_2/\sigma_2 & \dots \\ 1 & 0 & \sigma_1 G_1' + \sigma_2 G_2' - \partial_t\sigma_3/\sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ 1 & 0 & \sum_{k=1}^{N-1} \sigma_k G_k' - \partial_t\sigma_N/\sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and for the last column we obtain

$$\begin{bmatrix} -\partial_t\sigma_1/\sigma_1 \\ \sigma_1^{N-1}G_1^{(N-1)} - \partial_t\sigma_2/\sigma_2 \\ \sigma_1^{N-1}G_1^{(N-1)} + \sigma_2^{(N-1)}G_2^{(N-1)} - \partial_t\sigma_3/\sigma_3 \\ \vdots \\ \sum_{k=1}^{N-1} \sigma_k^{N-1}G_k^{(N-1)} - \partial_t\sigma_N/\sigma_N \\ 0 \end{bmatrix},$$

Next we subtract the first row from row number 2 to N thereby obtaining, for the first N columns

$$\begin{bmatrix} 1 & 0 & -\partial_t\sigma_1/\sigma_1 & \dots \\ 0 & 0 & \sigma_1 G_1' - \partial_t\sigma_2/\sigma_2 & \dots \\ 0 & 0 & \sigma_1 G_1' + \sigma_2 G_2' - \partial_t\sigma_3/\sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \sum_{k=1}^{N-1} \sigma_k G_k' - \partial_t\sigma_N/\sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and for the last column

$$\begin{bmatrix} -\partial_t\sigma_1/\sigma_1 \\ \sigma_1^{N-1}G_1^{(N-1)} - \partial_t\sigma_2/\sigma_2 \\ \sigma_1^{N-1}G_1^{(N-1)} + \sigma_2^{(N-1)}G_2^{(N-1)} - \partial_t\sigma_3/\sigma_3 \\ \vdots \\ \sum_{k=1}^{N-1}\sigma_k^{N-1}G_k^{(N-1)} - \partial_t\sigma_N/\sigma_N \\ 0 \end{bmatrix}.$$

Then we subtract the first column from the column 3 to $N+1$ to produce, for the first N columns

$$\begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & \sigma_1 G_1' - \partial_t\sigma_2/\sigma_2 & \dots \\ 0 & 0 & \sigma_1 G_1' + \sigma_2 G_2' - \partial_t\sigma_3/\sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \sum_{k=1}^{N-1}\sigma_k G_k' - \partial_t\sigma_N/\sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and for the last column

$$\begin{bmatrix} 0 \\ \sigma_1^{N-1}G_1^{(N-1)} - \partial_t\sigma_2/\sigma_2 \\ \sigma_1^{N-1}G_1^{(N-1)} + \sigma_2^{(N-1)}G_2^{(N-1)} - \partial_t\sigma_3/\sigma_3 \\ \vdots \\ \sum_{k=1}^{N-1}\sigma_k^{N-1}G_k^{(N-1)} - \partial_t\sigma_N/\sigma_N \\ 0 \end{bmatrix}.$$

By subtracting row i from row $i+1$ when $2 \le i \le N-1$ we obtain, for the first N columns

$$\begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & \sigma_1 G_1' - \partial_t\sigma_2/\sigma_2 & \dots \\ 0 & 0 & \sigma_2 G_2' + \partial_t\sigma_2/\sigma_2 - \partial_t\sigma_3/\sigma_3 & \dots \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \sigma_{N-1}G_{N-1}' + \partial_t\sigma_{N-1}/\sigma_{N-1} - \partial_t\sigma_N/\sigma_N & \dots \\ 0 & 1 & 0 & \dots \end{bmatrix},$$

and for the last column

$$\begin{bmatrix} 0 \\ \sigma_1^{N-1} G_1^{(N-1)} - \partial_t \sigma_2/\sigma_2 \\ \sigma_2^{(N-1)} G_2^{(N-1)} + \partial_t \sigma_2/\sigma_2 - \partial_t \sigma_3/\sigma_3 \\ \vdots \\ \sigma_{N-1}^{N-1} G_{N-1}^{(N-1)} + \partial_t \sigma_{N-1}/\sigma_{N-1} - \partial_t \sigma_N/\sigma_N \\ 0 \end{bmatrix}.$$

We set

$$\begin{aligned} n &:= N-1, \\ y^n &:= (y_1, y_2, \dots, y_n), \\ g_i^j(y_i) &:= \sigma_i^j G_i^{(j)}, \\ \phi_1(t) &:= -\partial_t \sigma_2/\sigma_2, \\ \phi_i(t) &:= \partial_t \sigma_i/\sigma_i - \partial_t, \sigma_{i+1}/\sigma_{i+1}, \qquad \text{for } i = 2, \dots, n-1, \end{aligned}$$

and finally obtain

$$H_n(x^n, t) = \begin{bmatrix} g_1^1 + \phi_1(t) & g_1^2 + \phi_1(t) & \dots & g_1^{n-1} + \phi_1(t) & g_1^n + \phi_1(t) \\ g_2^1 + \phi_2(t) & g_2^2 + \phi_2(t) & \dots & g_2^{n-1} + \phi_2(t) & g_2^n + \phi_2(t) \\ \vdots & \vdots & & \vdots & \vdots \\ g_n^1 + \phi_n(t) & g_n^2 + \phi_n(t) & \dots & g_n^{n-1} + \phi_n(t) & g_2^n + \phi_n(t) \end{bmatrix}. \tag{2.24}$$

Note that the vector

$$\begin{bmatrix} \phi_1(t) \\ \phi_2(t) \\ \vdots \\ \phi_n(t) \end{bmatrix}$$

is added to every column of the matrix. To study the linear independence of the columns of the matrix we can remove this vector leaving us with the following matrix

$$\begin{bmatrix} g_1^1 & g_1^2 & \dots & g_1^{n-1} & g_1^n \\ g_2^1 & g_2^2 & \dots & g_2^{n-1} & g_2^n \\ \vdots & \vdots & & \vdots & \vdots \\ g_n^1 & g_n^2 & \dots & g_n^{n-1} & g_2^n \end{bmatrix}. \tag{2.25}$$

This matrix corresponds to (2.18) as the functions g_i^j on row i of the matrix are linearly independent. We can apply Proposition 2.4 again and obtain

Proposition 2.7. *The matrix* $H_n(x^n, t)$ *in* (2.24) *has full rank in a neighborhood* B_n *of the point* (y^n, t).

We can now bring our examination of the dimension of market models in the time-dependent volatility case to an end.

Proposition 2.8. *The* N*-dimensional* LIBOR *market model with time-dependent volatility and the corresponding swap market model do not posses a realization of smaller dimension, in the sense of Definition 2.5.*

2.6 Conclusion

It is not possible to reduce the dimension of the LIBOR market model by finding a realization of smaller dimension in the sense of Definition 2.5. Neither in the constant volality case nor in the time-dependent case.

The results concerning the LIBOR market model are also valid for the swap market model. This follows from the fact that the swap rate can be expressed as a (nonlinear) function of the LIBOR rate. The problem of finding a realization of small dimension for the swap market model can therefore be reformulated into a realization problem for LIBOR rates.

Bibliography

Björk, Tomas, 2003, On the geometry of interest rate models, in *Paris-Princeton Lectures on Mathematical Finance*, volume 1847 of *Springer Lecture Notes in Mathematics*, 133–215 (Springer).

Björk, Tomas, 2004, *Arbitrage Theory in Continuous Time* (Oxford University Press), 2nd edition.

Björk, Tomas, and Lars Svensson, 2001, On the existence of finite-dimensional realizations for nonlinear forward rate models, *Mathematical Finance* 11, 205–243.

Black, Fisher, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3, 167–179.

Black, Fisher, and Myron S. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Brace, Alan, Dariuz Gatarek, and Marek Musiela, 1997, The market model of interest rate dynamics, *Mathematical Finance* 7, 127–154.

Brigo, Damiano, and Fabio Mercurio, 2006, *Interest Rate Models – Theory and Practice* (Springer).

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates, *Econometrica* 60, 77–105.

Ho, Thomas, and Sang-Bin Lee, 1986, Term structure movements and pricing interest rate contingent claims, *Journal of Finance* 41, 1011–1029.

Hull, John, and Alan White, 1990, Pricing interest-rate derivative securities, *Review of Financial Studies* 3, 573–592.

Hunt, P. J., and J. E. Kennedy, 2000, *Financial Derivatives in Theory and Practice* (John Wiley and sons, ltd).

Jamshidian, Farshid, 1997, Libor and swap market models and measures, *Finance and Stochastics* 1, 293–330.

Mercurio, Fabio, and Juan Moraleda, 2000, An analytically tractable interest rate model with humped volatility, *European Journal of Operational Research* 120, 205–214.

Miltersen, Kristian R., Klaus Sandmann, and Dieter Sondermann, 1997, Closed form solutions for term structure derivatives with log-normal interest rates, *Journal of Finance* 52, 409–430.

Musiela, Marek, and Marek Rutkowski, 1997, *Martingale Methods in Financial Modelling* (Springer).

Ritchken, Peter, and L. Sankarasubramanian, 1995, Volatility structures of forward rates and the dynamics of the term structure, *Mathematical Finance* 5, 55–72.

Vasiček, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 177–188.

Paper 3

Paper 3

Expectations Hypotheses in a Gaussian Two-Factor Model

3.1 Introduction

Expectations hypotheses have a long tradition in the economics literature that can be traced back to the work of Fisher (1896), Hicks (1939) and Lutz (1940). They occur both in interest rate theory and international economics.

The common claim of these hypotheses is that a forward quantity, such as the forward rate, is an unbiased predictor of some spot quantity such as the future spot rate.

There is an extensive empirical literature dealing with these hypotheses and it is still an active area of research among econometricians. The purely theoretical work is however comparatively scarce. The interest in these hypotheses is understandable since it would we be very valuable to obtain information about future values of economic variables using only the current ones.

The first thorough theoretical work on expectations hypotheses for interest rates, using the modern theory of arbitrage pricing, was done by Cox et al. (1981). They start by providing precise definitions of the hypotheses, they proceed to show that these hypotheses are in general incompatible and that only one of them is consistent with a general equilibrium model. Subsequent empirical work by Fama (1984), Fama and Bliss (1987), Fama (1990), Campbell (1986), Campbell and Shiller (1991), to name just a few, strongly reject expectations hypotheses in this form. It is generally argued that it is necessary to add a term premium, constant or time dependent, to the equations for them to make sense.

The first article to address the CIR critique of the expectations hypotheses from a theoretical perspective is Campbell (1986). Using a Vasiček (1977) model he shows that the expectations hypotheses are valid if a constant premium is added to the equations.

Frachot and Lesne (1993) argue that the conclusion of Campbell (1986) is wrong because he uses a Gaussian model. They instead suggest the use of a Cox et al. (1985) model, where the short rate follows a squared Gaussian process. As this model exhibits stochastic volatility the expectation hypotheses have to be corrected in a different way. Adding a term premium to the value in the classical expectation hypothesis is not enough. The value should also be multiplied by a term premium giving rise to an affine relation between values in the classical expectations.

In his Ph.D. dissertation Jadoui (1998), in joint work with Frachot, also considers the expectations hypotheses outlined in Cox et al. (1981), but instead uses a quadratic Gaussian model. He argues that the relation between the values in the classical expectations hypotheses are in the form of a second-degree polynomial.

Together, the previous studies seem to suggest that the relation between the values in the classical expectations hypothesis, say $\mathbb{E}_t[r_T]$ and $f(t,T)$, is determined by the form of the function relating the values to the state variable. If the values are affine functions of the state variable, say $\mathbb{E}_t[r_T] = \alpha + \beta x$ and $f(t,T) = \gamma + \delta x$, where x is the state variable, then in turn the relationship between the values is affine: $\mathbb{E}_t[r_T] = \eta f(t,T) + \mu$ where $\eta = \beta/\delta$ and $\mu = \alpha - \beta\gamma/\delta$. If the values are determined by a second order polynomial in the state variable, a (unique) relation between the variables in the form of a second order polynomial can similarly be derived. A natural question to ask here is what happens when the values are functions of two state variables, say $\mathbb{E}_t[r_T] = \alpha x + \beta y$ and $f(t,T) = \gamma x + \delta y$. Unless the vectors (α, β) and (γ, δ) are parallel we cannot express $\mathbb{E}_t[r_T]$ as a function of $f(t,T)$ without coefficients that depend on the state variables.

In this paper we study the expectations hypotheses in an additive Gaussian two-factor model of the short rate specified under the risk-neutral measure $\mathbb{Q}$. This model has a constant volatility. Then an affine state dependent Girsanov kernel is used to formulate the model under the objective probability measure $\mathbb{P}$. We then proceed to examine the different expectations hypotheses defined as in Cox et al. (1981). We show that these expectations hypotheses fail to hold if we allow the Girsanov kernel to be state dependent despite that the model has a deterministic volatility.

The paper is structured as follows. Section 3.2 establishes the notation and presents the classification of the expectations hypotheses by Cox et al. (1981). Section 3.3 provides a brief theoretical overview of the expectations hypotheses using the change of numeraire framework and the Heath-Jarrow-Morton approach. In Section 3.4 the additive two-factor model of Brigo and Mercurio (2006) is presented and then used to investigate the various expectations hypotheses outlined in Section 3.2. The relationship to the work of Campbell (1986), Frachot and Lesne (1993) and Jadoui (1998) is then discussed.

3.2 Formulation of Expectations Hypotheses

The economy is modelled on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ denotes the objective probability measure. We assume the market to be free of arbitrage. This is equivalent to assuming the existence at least one martingale measure $\mathbb{Q}$ making the discounted asset prices $\mathcal{F}$-martingales.[1]

3.2.1 Notation

The time t *price* of a zero coupon bond with time to maturity $T - t$ is denoted $P(t, T)$. The bond is assumed to deliver one Euro at maturity T so $P(T, T) = 1$. The *continuously compounded spot rate (or yield)* $Y(t, T)$ for the period $[t, T]$ is defined as

$$Y(t, T) := -\frac{\log P(t, T)}{T - t}.$$

The *continuously compounded forward rate* $F_t(S, T)$ contracted at time t is defined as

$$F_t(S, T) := -\frac{\log P(t, T) - \log P(t, S)}{T - S}.$$

The *instantaneous forward rate* with maturity T contracted at time t is defined as

$$f(t, T) := -\frac{\partial \log P(t, T)}{\partial T}.$$

The *instantaneous short rate* at time t is defined as

$$r(t) := f(t, t).$$

[1] See e.g. Björk (2004) chapter 10.

3.2.2 CIR classification of EH hypotheses

We will adopt the terminology of Cox, Ingersoll, and Ross (1981) when we refer to expectations hypotheses. The measure under which the expectations are taken is intentionally left unspecified.

Local Expectation Hypothesis (L-EH)

According to this hypothesis the expected instantaneous return of holding a zero coupon bond is equal to the short rate. By standard arbitrage theory we know that this hypothesis is always true under the risk neutral measure $\mathbb{Q}$.

$$\mathbb{E}\left[\lim_{h\to 0, h>0} \frac{P(t+h,T)-P(t,T)}{P(t,T)}\middle| \mathcal{F}_t\right] = \mathbb{E}\left[\frac{dP(t,T)}{P(t,T)}\middle| \mathcal{F}_t\right] = r(t)\,dt. \tag{3.1}$$

Return To Maturity Expectation Hypothesis (RTM-EH)

Under this hypothesis the return of holding a zero coupon bond until maturity is equal to the expected return from rolling over a series of single-period bonds

$$\frac{1}{P(t,T)} = \mathbb{E}\left[\exp\left(\int_t^T r(s)\,ds\right)\middle| \mathcal{F}_t\right]. \tag{3.2}$$

Unbiased Expectation Hypothesis (U-EH)

This hypothesis says that the forward rate is an unbiased estimate of the future short rate

$$f(t,T) = \mathbb{E}\left[r(T)|\, \mathcal{F}_t\right]. \tag{3.3}$$

Yield To Maturity Expectation Hypothesis (YTM-EH)

By integrating (3.3) with respect to T we get

$$Y(t,T) = \frac{1}{T-t}\,\mathbb{E}\left[\int_t^T r(s)\,ds\middle| \mathcal{F}_t\right]. \tag{3.4}$$

3.3 Theoretical Overview

Before studying the expectations hypotheses in a particular model it is informative to see what kind of results we can obtain in a general setting and what kind of assumption we need to make in able to draw any conclusions.

3.3.1 The change of numeraire approach

The change of numeraire technique was popularized in the paper by Geman et al. (1995).[2]. We will use it here to treat the U-EH.

We start by deriving a classical result namely that U-EH is only valid under the T-forward measure. Under the T-forward measure, the price of a T-claim $\mathcal{X}$ is given by $\Pi(t) = p(t,T)\,\mathbb{E}^T\left[\mathcal{X}\middle|\,\mathcal{F}_t\right]$. Setting $\mathcal{X} = r(T)$, we obtain

$$\Pi(t) = p(t,T)\,\mathbb{E}^T\left[r(T)\middle|\,\mathcal{F}_t\right].$$

On the other hand, using ordinary risk neutral valuation under the martingale measure $\mathbb{Q}$ we obtain

$$\Pi(t) = \mathbb{E}^{\mathbb{Q}}\left[r(T)e^{-\int_t^T r(s)ds}\middle|\,\mathcal{F}_t\right].$$

By equating these expressions for the price and using that $p(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds}\middle|\,\mathcal{F}_t\right]$ and that the forward rate is defined by $f(t,T) := -\partial \log p(t,T)/\partial T$, we obtain

$$\begin{aligned}
\mathbb{E}^T\left[r(T)\middle|\,\mathcal{F}_t\right] &= \frac{1}{p(t,T)}\,\mathbb{E}^{\mathbb{Q}}\left[r(T)e^{-\int_t^T r(s)ds}\middle|\,\mathcal{F}_t\right] \\
&= -\frac{1}{p(t,T)}\,\mathbb{E}^{\mathbb{Q}}\left[\frac{\partial}{\partial T}e^{-\int_t^T r(s)ds}\middle|\,\mathcal{F}_t\right] \\
&= -\frac{1}{p(t,T)}\frac{\partial}{\partial T}\,\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds}\middle|\,\mathcal{F}_t\right] \\
&= -\frac{1}{p(t,T)}\frac{\partial p(t,T)}{\partial T} = -\frac{\partial \log p(t,T)}{\partial T} \\
&= f(t,T).
\end{aligned}$$

In the framework of Platen and Heath (2006, Chap. 10) the growth optimal portfolio (GOP), with value $N(t)$ at time t, is the numeraire under the objective probability measure.

[2]For an overview see also Björk (2004, Chap. 24) or Brigo and Mercurio (2006, Chap. 2)

Starting from the objective probability we obtain

$$
\begin{aligned}
\mathbb{E}^P[r(T)|\mathcal{F}_t] &= \mathbb{E}^T\left[r(T)\frac{p(t,T)}{p(T,T)}\frac{N(T)}{N(t)}|\mathcal{F}_t\right] \\
&= \frac{p(t,T)}{N(t)}\,\mathbb{E}^T[r(T)N(T)|\mathcal{F}_t] \\
&= \frac{p(t,T)}{N(t)}\Big\{\,\mathbb{E}^T[N(T)|\mathcal{F}_t]\,\mathbb{E}^T[r(T)|\mathcal{F}_t] + \mathrm{Cov}[r(T),N(T)]\Big\} \\
&= \frac{p(t,T)}{N(t)}\Big\{\,\mathbb{E}^T[N(T)|\mathcal{F}_t]f(t,T) + \mathrm{Cov}[r(T),N(T)]\Big\} \\
&= \frac{p(t,T)}{N(t)}\Big\{\frac{N(t)}{p(t,T)}f(t,T) + \mathrm{Cov}[r(T),N(T)]\Big\} \\
&= f(t,T) + \frac{p(t,T)}{N(t)}\,\mathrm{Cov}[r(T),N(T)]\Big\}
\end{aligned}
$$

The forward rate rate is thus a biased estimate of the future spot rate. The size of the premium as well as its sign depends on the covariance between the short rate and the GOP. By making some concrete assumptions about the dynamics of the GOP it should be possible to deduce interesting properties about this bias.

3.3.2 The HJM approach

By the assumption of absence of arbitrage, the Heath-Jarrow-Morton condition gives the following dynamics for the forward rate under the risk-neutral measure $\mathbb{Q}$

$$
\begin{aligned}
f(\tau,T) = f(t,T) &+ \int_t^\tau \sigma(s,T)[\int_s^T \sigma(s,\xi)^\top d\xi]ds \\
&+ \int_t^\tau \sigma(s,T)dW_s^{\mathbb{Q}},
\end{aligned}
\tag{3.5}
$$

where $^\top$ denotes transposition and where W is a d-dimensional Wiener process. The short rate $r(T)$ is then given by

$$
\begin{aligned}
r(T) = f(T,T) = f(t,T) &+ \int_t^T \sigma(s,T)[\int_s^T \sigma(s,\xi)^\top d\xi]ds \\
&+ \int_t^T \sigma(s,T)dW_s^{\mathbb{Q}}.
\end{aligned}
$$

We immediately observe the kind of relation that prevails between the short rate and the forward rate. We already now that the U-EH is only valid under the T-forward measure. Here, we can derive the functional form of the premium under the risk-neutral measure ones we have specified a volatility. To get an idea of the sign of the premium and its properties with respect to time thus need to make further assumptions.

If we now assume that the volatility σ is deterministic we immediately obtain

$$\mathbb{E}^{\mathbb{Q}}\left[r(T) | \mathcal{F}_t \right] = f(t,T) + \int_t^T \sigma(s,T))[\int_s^T \sigma(s,\xi)^\top d\xi]ds.$$

We let $\mathbb{P}$ denote objective probability measure. By absence of arbitrage there exists a vector $\lambda(t) = (\lambda_1(t), \ldots, \lambda_d(t))^\top$ (the Girsanov kernel) so that $dW^{\mathbb{P}} = dW^{\mathbb{Q}} + \lambda_t dt$. This gives

$$\begin{aligned} r(T) = f(t,T) &+ \int_t^T \sigma(s,T)[\int_s^T \sigma(s,\xi)^\top d\xi - \varphi(s)]ds \\ &+ \int_t^T \sigma(s,T) dW_s^{\mathbb{P}}. \end{aligned}$$

The change of probability measure from $\mathbb{Q}$ to $\mathbb{P}$ gives rise to a new premium, that will depend on the product of the volatility and the Girsanov kernel. If we assume the Girsanov kernel to be deterministic we immediately obtain

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left[r(T) | \mathcal{F}_t \right] = f(t,T) &+ \int_t^T \sigma(s,T)) \int_s^T \sigma(s,\xi)^\top d\xi ds \\ &- \int_t^T \sigma(s,T)\varphi(s) ds. \end{aligned}$$

Let us now provide two examples showing how the HJM framework can be used to get analytical expressions for the biases. We only compute the bias under the risk neutral measure. To obtain the bias under the objective probability measure we also have to make an assumption concerning the functional form of the Girsanov kernel.

Example: Hoo-Lee

Assume that $d = 1$ and that σ is constant. Then $\int_s^T \sigma(s,\xi)d\xi = (T-s)\sigma$ and therefore

$$\mathbb{E}^{\mathbb{Q}}\left[r(T) | \mathcal{F}_t \right] = f(t,T) + \int_t^T \sigma^2 (T-s)\, ds = f(t,T) + \frac{\sigma^2}{2}(T-t)^2.$$

Under the risk neutral measure the forward rate underestimates the future spot rate. We also see that the bias is a quadratically increasing function of the time to maturity

Example:Hull-White(Generalized Vasiček)

Assume $d = 1$ and $\sigma(s, T) = \sigma e^{-\gamma(T-s)}$. Then

$$\int_s^T \sigma e^{-\gamma(\xi - s)}\, d\xi = \frac{\sigma}{\gamma}(1 - e^{-\gamma(T-s)}),$$

and therefore

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[r(T)|\, \mathcal{F}_t\right] &= f(t,T) + \frac{\sigma^2}{\gamma}\int_t^T e^{-\gamma(T-s)}(1 - e^{-\gamma(T-s)})ds \\ &= f(t,T) + \frac{\sigma^2}{2\gamma^2}(1 - e^{-\gamma(T-t)})^2\end{aligned}$$

Also here the forward rate underestimates the future spot rate. The bias is also an increasing function of the time to maturity.

According to the theory of Björk and Svensson (2001) all interesting volatility functions are of the quasi-exponential form, because they produce finite dimensional realizations. To obtain more general result in this HJM setting, we could therefore try study the bias for this type of function.

3.4 The Additive Two-Factor Gaussian Model (G2++)

We will study expectations hypothesis in an additive two-factor Gaussian model, which goes under the name G2++ in the terminology of Brigo and Mercurio (2006). This model can, by a suitable change of variables, be transformed into the Hull and White (1994) model where the state variables are the short rate and the mean reversion level.[3] As the model remains Gaussian this will have not implications on our results.

The model

The instantaneous short rate is assumed to have the form

$$\begin{aligned} r(t) &= x_1(t) + x_2(t) + \psi(t), \\ r(0) &= r_0, \end{aligned} \tag{3.6}$$

where the state variables $x_1(t)$ and $x_2(t)$ have the following dynamics under the risk-neutral measure $\mathbb{Q}$:

$$\begin{aligned} dx_1(t) &= -b_1 x_1(t)dt + \sigma_1 dW_1^{\mathbb{Q}}(t), \qquad x_1(0) = 0 \\ dx_1(t) &= -b_2 x_2(t)dt + \sigma_2 dW_2^{\mathbb{Q}}(t), \qquad x_2(0) = 0. \end{aligned} \tag{3.7}$$

[3]See Brigo and Mercurio (2006, p.159).

Here (W_1, W_2) is a two-dimensional Wiener process with instantaneous correlation ρ such that

$$\rho dt = dW_1(t)dW_2(t).$$

Here $r_0, b_1, b_2, \sigma_1, \sigma_2$ are positive constants. The deterministic function $\psi(t)$ will be used to fit the model to the current yield curve. We assume that $\psi(0) = r_0$.

The assumption of absence of arbitrage is characterized by the existence of a square integrable process φ, the Girsanov kernel, such that

$$dW^{\mathbb{Q}} = dW^{\mathbb{P}} - \varphi(t)dt, \tag{3.8}$$

Here we assume this Girsanov to have the following form

$$\varphi(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix} = \begin{pmatrix} \gamma_1 + \delta_1 x_1(t) \\ \gamma_2 + \delta_2 x_2(t) \end{pmatrix}.$$

We note that $\varphi(t)$ depends on the state variables. This will turn out to be important for the treatment of the expectations hypotheses.

Using (3.8), we rewrite the dynamics of the state variables under the objective probability measure $\mathbb{P}$ as

$$\begin{aligned} dx_1(t) &= \Big[-\sigma_1\gamma_1 - (b_1 + \sigma_1\delta_1)x_1(t) \Big] dt + \sigma_1 dW_1^{\mathbb{P}}, \\ dx_2(t) &= \Big[-\sigma_2\gamma_2 - (b_2 + \sigma_2\delta_2)x_2(t) \Big] dt + \sigma_2 dW_2^{\mathbb{P}}. \end{aligned} \tag{3.9}$$

Setting

$$\begin{aligned} \alpha_1 &:= -\sigma_1\gamma_1, \qquad & \beta_1 &:= b_1 + \sigma_1\delta_1, \\ \alpha_2 &:= -\sigma_2\gamma_2, \qquad & \beta_2 &:= b_2 + \sigma_2\delta_2, \end{aligned}$$

we obtain

$$\begin{aligned} dx_1(t) &= \big(\alpha_1 - \beta_1 x_1(t)\big)dt + \sigma_1 dW_1^{\mathbb{P}}, \\ dx_2(t) &= \big(\alpha_2 - \beta_2 x_2(t)\big)dt + \sigma_2 dW_2^{\mathbb{P}}. \end{aligned} \tag{3.10}$$

By (3.6) and a simple integration of (3.10) from t to T we obtain the following representation of the instantaneous short rate

$$\begin{aligned} r(T) = \psi(T) &+ x_1(t)e^{-\beta_1(T-t)} + \frac{\alpha_1}{\beta_1}(e^{-\beta_1(T-t)} - 1) \\ &+ x_2(t)e^{-\beta_2(T-t)} + \frac{\alpha_2}{\beta_2}(e^{-\beta_2(T-t)} - 1) \\ &+ \sigma_1 \int_t^T e^{-\beta_1(T-s)} dW_1(s) + \sigma_2 \int_t^T e^{-\beta_2(T-s)} dW_2(s). \end{aligned}$$

From this expression it is clear that, conditional on $\mathcal{F}_t$, $r(T)$ is normally distributed. The expected value and variance are

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}[r(T)|\mathcal{F}_t] &= \psi(T) + \frac{\alpha_1}{\beta_1}(e^{-\beta_1(T-t)} - 1) + \frac{\alpha_2}{\beta_2}(e^{-\beta_2(T-t)} - 1) \\ &\quad + x_1(t)e^{-\beta_1(T-t)} + x_2(t)e^{-\beta_2(T-t)}, \\ \mathrm{Var}^{\mathbb{P}}[r(T)|\mathcal{F}_t] &= \frac{\sigma_1^2}{2\beta_1}\left[1 - e^{-2\beta_1(T-t)}\right] + \frac{\sigma_2^2}{2\beta_2}\left[1 - e^{-2\beta_2(T-t)}\right] \\ &\quad + 2\rho\frac{\sigma_1\sigma_2}{\beta_1+\beta_2}\left[1 - e^{-(\beta_1+\beta_2)(T-t)}\right].\end{aligned} \tag{3.11}$$

3.4.1 The price of a zero-coupon bond

By risk-neutral valuation the price at time t of a zero-coupon bond maturing at time T is given by

$$p(t,T) = \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r(s)ds}|\mathcal{F}_t].$$

In order to compute this expectation, we note that the integral

$$Z(t,T) = \int_t^T r(s)ds = \int_t^T \psi(s)ds + \int_t^T x_1(s)ds + \int_t^T x_2(s)ds,$$

is, conditional to the sigma field $\mathcal{F}_t$, a normally distributed random variable. If we denote the (conditional) mean by $m(t,T)$ and variance by $v(t,T)$, we know from the moment generating function of Z that

$$\mathbb{E}^{\mathbb{Q}}[e^{Z(t,T)}] = e^{m(t,T)+\frac{1}{2}v(t,T)}.$$

Using (3.7) we obtain

$$\begin{aligned}m(t,T) &= \mathbb{E}^{\mathbb{Q}}[Z(t,T)] = \int_t^T \psi(s)ds \\ &\quad + \frac{1}{b_1}(1 - e^{-b_1(T-t)})x_1(t) + \frac{1}{b_2}(1 - e^{-b_2(T-t)})x_2(t). \\ v(t,T) &= \frac{\sigma_1^2}{b_1^3}\left[b_1(T-t) - \frac{3}{2} + 2e^{-b_1(T-t)} - \frac{1}{2}e^{-2b_1(T-t)}\right] \\ &\quad + \frac{\sigma_2^2}{b_2^3}\left[b_2(T-t) - \frac{3}{2} + 2e^{-b_2(T-t)} - \frac{1}{2}e^{-2b_2(T-t)}\right] \\ &\quad + 2\rho\frac{\sigma_1\sigma_2}{b_1 b_2}\Big[T - t + \frac{1}{b_1}(e^{b_1(T-t)} - 1) + \frac{1}{b_2}(e^{b_2(T-t)} - 1) \\ &\quad - \frac{1}{b_1+b_2}(e^{-(b_1+b_2)(T-t)} - 1)\Big]\end{aligned} \tag{3.12}$$

By using these expressions for the mean and variance we obtain the following.

Proposition 3.1. *The price at time t of a zero-coupon bond with maturity T is*

$$\begin{aligned} P(t,T) &= \exp\Big\{ -m(t,T) + \frac{1}{2}v(t,T)\Big\} \\ &= \exp\Big\{ -\int_t^T \psi(s)ds + \frac{1}{b_1}(e^{-b_1(T-t)}-1)x_1(t) \\ &\quad + \frac{1}{b_2}(e^{-b_2(T-t)}-1)x_2(t) + \frac{1}{2}v(t,T)\Big\}. \end{aligned} \tag{3.13}$$

If we assume that the observed yield curve on the market today (at time $t=0$) is a smooth function $T \mapsto P^\star(0,T)$, we can use it to fit the model to current prices. We will use the function ψ for this task. According to (3.13) we have

$$P^\star(0,T) = \exp\Big\{ -\int_0^T \psi(s)ds + \frac{1}{2}v(0,T)\Big\}.$$

We can therefore write

$$\exp\Big\{ -\int_t^T \psi(s)ds\Big\} = \frac{P^\star(0,T)}{P^\star(0,t)}\exp\Big\{ -\frac{1}{2}[v(0,T)-v(0,t)]\Big\},$$

and obtain the following corollary.

Corollary 3.1. *The price at time t of a zero coupon bond with maturity T is*

$$P(t,T) = \exp\Big\{A_0(t,T) - A_1(t,T)x_1(t) - A_2(t,T)x_2(t)\Big\},$$

where

$$\begin{aligned} A_0(t,T) &= \log\frac{P^\star(0,T)}{P^\star(0,t)} + \frac{1}{2}[v(t,T)+v(0,t)-v(0,T)] \\ A_1(t,T) &= \frac{1}{b_1}(1-e^{-b_1(T-t)}) \\ A_2(t,T) &= \frac{1}{b_2}(1-e^{-b_2(T-t)}) \end{aligned}$$

and where $P^\star(0,t)$ and $P^\star(0,T)$ are the current prices observed on the market of zero-coupon bonds with maturity t and T.

Now that the we have solved for the bond price and made assumptions concerning the change of measure between $\mathbb{Q}$ and $\mathbb{P}$ we are ready to investigate under what conditions, if any, the expectations hypotheses hold. We will follow the CIR classification discussed in Section 3.2.2.

3.4.2 Expectations hypotheses in G2++

Local Expectation Hypothesis

We know that this hypothesis is by construction satisfied under the risk neutral measure i.e.

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{dP(t,T)}{P(t,T)}\middle|\mathcal{F}_t\right] = r(t)$$

To study this hypothesis under the objective probability measure $\mathbb{P}$ we apply the itō formula to the price equation in (3.13), using the state variable dynamics in (3.10). If we then compare the resulting expression with the corresponding expression when using the state variable dynamics under $\mathbb{Q}$ we obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\frac{dP(t,T)}{P(t,T)}\middle|\mathcal{F}_t\right] = r(t) - A_1(t,T)\sigma_1(\underbrace{\gamma_1 + \delta_1 x_1(t)}_{\varphi_1(t)}) \\ - A_2(t,T)\sigma_2(\underbrace{\gamma_2 + \delta_2 x_2(t)}_{\varphi_2(t)}).\end{aligned}$$

By assuming that the Girsanov kernel is state dependent, i.e $\delta_i \neq 0$ we obtain a stochastic premium.

As the functions A_i are positive, for the premium to be positive, we have to assume that the Girsanov kernels are negative. We note that the functions A_i are increasing in the maturity yielding higher expected returns for long maturities.

Return to Maturity Expectation Hypothesis

To study this expectation hypothesis we need to compute the expectation

$$\mathbb{E}^{\mathbb{P}}[e^{\int_t^T r(s)ds}|\mathcal{F}_t].$$

We use the same technique as in the derivation of the price of a zero-coupon bond in Section 3.4.1. The integral

$$Z(t,T) = \int_t^T r(s)ds = \int_t^T \psi(s)ds + \int_t^T x_1(s)ds + \int_t^T x_2(s)ds,$$

is a normally distributed stochastic variable, conditional on $\mathcal{F}_t$. If we denote the mean $M(t,T)$ and the variance $v(t,T)$, we thus have

$$\mathbb{E}^{\mathbb{P}}[e^{Z(t,T)}|\mathcal{F}_t] = e^{M(t,T)+\frac{1}{2}v(t,T)}.$$

The conditional mean of $Z(t,T)$ is

$$\begin{aligned} M(t,T) &= \mathbb{E}^{\mathbb{P}}[\int_t^T r(s)|\mathcal{F}_t] \\ &= \int_t^T \psi(s)ds + (\frac{\alpha_1}{\beta_1} + \frac{\alpha_2}{\beta_2})(T-t) \\ &+ \frac{1}{\beta_1}\Big(x_1(t) - \frac{\alpha_1}{\beta_1}\Big)(1-e^{-\beta_1(T-t)}) \\ &+ \frac{1}{\beta_2}\Big(x_2(t) - \frac{\alpha_2}{\beta_2}\Big)(1-e^{-\beta_2(T-t)}) \end{aligned} \tag{3.14}$$

The variance $v(t,T)$ is the same as the one we calculated previously when solving for the price of a zero-coupon bond in (3.12). Comparing the expression for $M(t,T)$ in (3.14) and $m(t,T)$ in (3.12) we immediately see that the return to maturity expectation hypothesis cannot hold even if the Girsanov kernel is not state dependent i.e. if $\delta_1 = \delta_2 = 0$ so that $\beta_1 = b_1$ and $\beta_2 = b_2$. In the present setting we instead obtain

$$\mathbb{E}^{\mathbb{P}}[e^{\int_t^T r(s)ds}|\mathcal{F}_t] = \frac{1}{p(t,T)}e^{M(t,T)-m(t,T)+v(t,T)}.$$

The expected value is thus equal to the inverse of the price times a state-dependent stochastic premium. We also note that the hypothesis does not even hold under the measure $\mathbb{Q}$. If we set the Girsanov kernel to zero, thereby obtaining $m(t,T) = M(t,T)$, we still have to multiply the inverse of the price by $e^{v(t,T)}$ to obtain equality in the equation.

Yield to Maturity Expectation Hypothesis

The yield to maturity expectation hypothesis is the most interesting to study from a theoretical perspective since almost all empirical studies are based on it.

Here we need to compare $(T-t)Y(t,T)$ with $\mathbb{E}^{\mathbb{P}}[\int_t^T r(s)ds|\mathcal{F}_t]$.

From the definition of $Y(t,T)$ and Corollary 3.1 we obtain

$$(T-t)Y(t,T) = -A_0(t,T) + A_1(t,T)x_1(t) + A_2(t,T)x_2(t).$$

where

$$\begin{aligned}
-A_0(t,T) &= \int_t^T \psi(s)ds \\
&\quad - \frac{\sigma_1^2}{2b_1^2}\Big[(T-t) - \frac{3}{2b_1} + \frac{2}{b_1}e^{-b_1(T-t)} - \frac{1}{2b_1}e^{-2b_1(T-t)}\Big] \\
&\quad - \frac{\sigma_1^2}{2b_2^2}\Big[(T-t) - \frac{3}{2b_2} + \frac{2}{b_2}e^{-b_2(T-t)} - \frac{1}{2b_2}e^{-2b_2(T-t)}\Big] \\
&\quad - \rho\frac{\sigma_1\sigma_2}{b_1b_2}\Big[T-t+\frac{1}{b_1}(e^{b_1(T-t)}-1) + \frac{1}{b_2}(e^{b_2(T-t)}-1) \\
&\qquad - \frac{1}{b_1+b_2}(e^{-(b_1+b_2)(T-t)}-1)\Big] \\
A_1(t,T) &= \frac{1}{b_1}(1-e^{-b_1(T-t)}) \\
A_2(t,T) &= \frac{1}{b_2}(1-e^{-b_2(T-t)})
\end{aligned}$$

To compute the expected value we use the Fubini theorem to obtain

$$\mathbb{E}^{\mathbb{P}}\Big[\int_t^T r(s)ds|\mathcal{F}_t\Big] = \int_t^T \mathbb{E}^{\mathbb{P}}[r(s)|\mathcal{F}_t]ds.$$

Integration of the expectation in (3.11) gives the desired result.

$$\mathbb{E}^{\mathbb{P}}\Big[\int_t^T r(s)ds|\mathcal{F}_t\Big] = B_0(t,T) + B_1(t,T)x_1(t) + B_2(t,T)x_2(t)$$

where

$$\begin{aligned}
B_0(t,T) &= \int_t^T \psi(s)ds + \frac{\alpha_1}{\beta_1}\Big[\frac{1}{\beta_1}\Big(1-e^{-\beta_1(T-t)}\Big) - (T-t)\Big] \\
&\quad + \frac{\alpha_2}{\beta_2}\Big[\frac{1}{\beta_2}\Big(1-e^{-\beta_2(T-t)}\Big) - (T-t)\Big] \\
B_1(t,T) &= \frac{1}{\beta_1}\Big(1-e^{-\beta_1(T-t)}\Big) \\
B_2(t,T) &= \frac{1}{\beta_2}\Big(1-e^{-\beta_2(T-t)}\Big)
\end{aligned}$$

By removing the state dependence in the Girsanov kernel we can obtain $A_1 = B_1$ and $A_2 = B_2$ but we will still have $-A_0 \neq B_0$. We will thus have to add a deterministic term for the equation to be valid. We recall that the yield to maturity expectation hypothesis is just an integrated version of the unbiased expectation hypothesis and that the latter is only valid under the T-forward measure.

This theoretical result could help explain the empirical failure of the yield-to-maturity hypothesis common in the literature. If we run a regression of $\int_t^T r(s)ds$ on $Y(t,T)$ we are thus not only expected to find that the intercept is different from zero but also that the model is mispecified since the error term is correlated with the right hand side variables.

Frachot and Lesne (1993) study the expectations hypotheses using a Cox et al. (1985) model for the interest rate of the form

$$dr(t) = \big(\phi(t) - \lambda r(t)\big)dt + \sqrt{\alpha r(t) + \beta}dW^{\mathbb{Q}}(t).$$

We note that this model collapses to a Vasiček (1977) model when $\alpha = 0$. To compute expectations under the objective probability measure Frachot-Lesne use a Girsanov kernel of the form $\varphi(t) = q\sqrt{\alpha r(t) + \beta}$. This is the usual way to model the change of measure in this type of model.[4] They then obtain an affine relation between the variables in the YTM-H of the form $Y(t,T) = a + b\,\mathbb{E}_t^{\mathbb{P}}[\int_t^T r(s)ds]$. This is compared to the result of Campbell (1986) who obtains $b = 1$. Frachot and Lesne (1993) argue that Campbells result is due to the use of a Gaussian interest model (α set to zero). They therefore advocate that a stochastic volatility model should be used when studying the expectations hypothesis. There is however a problem with their argument. As the same parameter α occurs both in the volatility and in the Girsanov kernel it is impossible to know if it is the stochastic volatility or the stochastic Girsanov kernel that drives their result. Our analysis in a Gaussian model clearly indicates that it is the stochastic Girsanov kernel, and not the stochastic volatility which is behind the affine relation.

Our analysis in a two-factor model also highlights another important issue with Frachot-Lesnes idea. In a CIR model the quantities occurring in the expectations hypotheses are affine functions of the state variables. This allows us to express one of the quantities as an affine function of the other in a unique way. When Jadoui (1998) uses a quadratic Gaussian model of the short-rate he finds that there is a quadratic relationship between the quantities. But as in the affine case, if the quantities can be expressed as linear functions of the state variable and its squared value there exists a unique quadratic relation between the quantities. When using two different factor it becomes impossible to find a relation between the quantities in this way and therefore also impossible to find any alternative form of the expectation hypotheses.

[4] See Cheridito et al. (2007) for an alternative approach and a thorough discussion of the modelling of the change of measure as well as empirical results.

Bibliography

Backus, David, Silverio Foresi, Abon Mozumdar, and Liuren Wu, 1998, Predictable changes in yields and forward rates, NBER Working paper series 6379, National Bureau of Economic Research.

Bekaert, Geert, Robert J. Hodrick, and David A. Marshall, 1997, On biases in tests of the expectations hypothesis of the term structure of interest rates, *Journal of Financial Economics* 44, 309–348.

Björk, Tomas, 2004, *Arbitrage Theory in Continuous Time* (Oxford University Press), 2nd edition.

Brigo, Damiano, and Fabio Mercurio, 2006, *Interest Rate Models – Theory and Practice* (Springer).

Campbell, John Y., 1986, A defence of traditional hypotheses about the term structure of interest rates, *Journal of Finance* 41, 183–193.

Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Cheridito, Patrick, Damir Filipović, and Robert L. Kimmell, 2007, Market price of risk specifications for affine models: Theory and evidence, *Journal of Financial Economics* 83, 123–170.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1981, A re-examination of traditional hypotheses about the term structure of interest rates, *The Journal of Finance* 36, 769–799.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Fama, Eugene F., 1984, The information in the term structure, *Journal of Financial Economics* 13, 509–528.

Fama, Eugene F., 1990, Term structure forecasts of interest rates, inflation, and real returns, *Journal of Monetary Economics* 25, 59–76.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *The American Economic Review* 77, 680–692.

Fisher, Irving, 1896, *Appreciation and Interest* (Macmillan).

Frachot, Antoine, and J.P. Lesne, 1993, Expectations hypothesis and stochastic volatilities, notes d'études et de recherche 24, Banque de France.

Geman, Helyette, Nicole El Karoui, and J.C. Rochet, 1995, Change of numéraire, changes of probability measure and option pricing, *Journal of Applied Probability* 32, 443–458.

Hicks, John, 1939, *Value and Capital* (Oxford University Press).

Hull, John, and Alan White, 1994, Numerical procedures for implementing term structure models ii: Two-factor models, *The Journal of Derivatives* 2, 37–48.

Jadoui, Mohammed, 1998, *Taux D'Intérêt et Taux de Change. Valorisation d'Options*, Ph.D. thesis, Université d'Evry Val d'Essonne.

Longstaff, Francis A., 2000, Arbitrage and the expectation hypothesis, *Journal of Finance* 55, 989–994.

Lutz, F. A., 1940, The structure of interest rates, *Quaterly Journal of Economics* 55, 36–63.

Platen, Eckhard, and David Heath, 2006, *A Benchmark Approach to Quantitative Finance* (Springer).

Vasiček, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 177–188.

EFI, The Economic Research Institute

Published in the language indicated by the title. A complete publication list can be found at *www.hhs.se/efi* Books and dissertations can be ordered from EFI via e-mail: EFI.Publications@hhs.se

Reports since 2004

2008

Books

Kraus, Kalle. *Sven eller pengarna? Styrningsdilemman i äldrevården.*
Portnoff, Linda. *Musikbranschens styrningsproblematik.*

Dissertations

Benson, Ilinca. *Organisering av övergångar på arbetsmarknaden : en studie av omställningsprogram.*
Farooqi Lind, Raana. *On capital structure and debt placement in Swedish companies*.
Granström, Ola. *Aid, drugs, and informality : essays in empirical economics.*
Hvenmark, Johan. *Reconsidering membership : a study of individual members' formal affiliation with democratically governed federations.*
Johansson, Marjana. *Engaging resources for cultural events : a performative view.*
Kviselius, Niklas Z. *Trust-building and communication in SME internationalization : a study of Swedish-Japanese business relations.*
Schilling, Annika. *Kan konsulter fusionera? : en studie av betydelsen av identitet vid en fusion mellan konsultföretag.*
Schriber, Svante. *Ledning av synergirealisering i fusioner och förvärv*.
Sjödin, Henrik. *Tensions of extensions : adverse effects of brand extension within consumer relationship.*
Strömqvist, Maria. *Hedge funds and international capital flow.*
Söderström, Johan. *Empirical studies in market efficiency.*
Thorsell, Håkan. *The pricing of corporate bonds and determinants of financial structure.*
Ulbrich, Frank. *The adoption of IT-enabled management ideas : insights from shared services in government agencies*.

2007

Books

Andersson, Per, Ulf Essler and Bertil Thorngren (eds). *Beyond mobility.* EFI Yearbook 2007. EFI/Studentlitteratur.
Ericsson, Daniel. *Musikmysteriet : organiserade stämningar och motstämningar.*
Samuelson, Lennart (red). *Bönder och bolsjeviker : den ryska landsbygdens historia 1902–1939.*
Wijkström, Filip och Torbjörn Einarsson. *Analysmodell för sektorsöverskridande statistik : fallet vård och omsorg.*

Dissertations

Ahlersten, Krister. *Empirical asset pricing and investment strategies.*
Alexius, Susanna. *Regelmotståndarna : om konsten att undkomma regler.*
Andersson, Magnus. *Essays in empirical finance.*
Berg, Bengt Åke. *Volatility, integration and grain bank : studies in harvests, rye prices and institutional development of the parish magasins in Sweden in the 18th and 19th centuries.*
Bianchi, Milo. *Of speculators, migrants and entrepreneurs : essays on the economics of trying your fortune.*

Brodin, Karolina. *Consuming the commercial break : an ethnographic study of the potential audiences for television advertising.*

Elger, Max. *Three essays on investment-specific technical change.*

Hagberg, Axel. *Bankkrishantering : aktörer, marknad och stat.*

Hinnerich, Mia. *Derivatives pricing and term Structure modeling.*

Hjalmarson, Hanna. *En växande marknad : studie av nöjdheten med konsumtionsrelaterade livsområden bland unga konsumenter*.

Hjelström, Tomas. *The closed-end investment company premium puzzle : model development and empirical tests on Swedish and British data*.

Kraus, Kalle. *Sven, inter-organisational relationships and control : a case study of domestic care of the elderly.*

Lindqvist, Erik. *Essays on privatization, identity, and political polarization.*

Macquet, Monica. *Partnerskap för hållbar utveckling : systrar av Oikos och guvernanten som blev diplomat.*

Melian, Catharina. *Progressive open source*

Nilsson, Daniel. *Transactions in cyberspace : the continued use of Internet banking.*

Petrelius Karlberg, Pernilla. *Den medialiserade direktören.*

Portnoff, Linda. *Control, cultural production and consumption : theoretical perspectives, empirical dilemmas, and Swedish music industry practices.*

Sköld, Martin. *Synergirealisering : realisering av produktsynergier efter företags-sammanslagningar.*

Sonnerby, Per. *Contract-theoretic analyses of consultants and trade unions.*

Tyrefors, Björn. *Institutions, policy and quasi-experimental evidence.*

Valiente, Pablo. *Re-innovating the existing : a study of wireless IS capabilities to support mobile workforces.*

2006

Books

Lundeberg, Mats, Pär Mårtensson and Magnus Mähring (eds) *IT & business performance : a dynamic relationship.* EFI Yearbook 2006. EFI / Studentlitteratur.

Thodenius, Björn. *Organisering av kunskap : en studie av Wallenberg Consortium North.*

Wijkström, Filip och Torbjörn Einarsson. *Från nationalstat till näringsliv?: det civila samhällets organisationsliv i förändring.*

Wijkström, Filip, Einarsson, Stefan och Larsson, Ola. *Staten och det civila samhället : idétraditioner och tankemodeller i den statliga bidragsgivningen till ideella organisationer.*

Östman, Lars. *Lysande ögonblick och finansiella kriser : Dramaten under ett sekel.*

Dissertations

Argenton, Cedric. *Quality provision in duopoly.*

Beckerman, Carina. *The clinical eye : constructiong and computerizing an anesthesia patient record.*

Borglund, Tommy. *Aktievärden i fokus : internationell påverkan på intressentrelationer vid förvärv och fusion.*

Breman, Anna. *The Economics of altruism, paternalism and self-control.*

Edquist, Harald. *Technological breakthroughs and productivity growth.*

Eklund, Jana. *Essays on forecasting and bayesian model averaging.*

Frostenson, Magnus. *Legitimitetskontrollen : en studie av etiska värderingars roll i gränsöverskridande förvärv och fusioner.*

Gaspar, Raquel M. *Credit risk and forward price models.*

Gustafsson, Peter. *Essays on trade and technological change.*

Hopkins, Elisabeth. *Is a higher degree of local currency pricing associated with lower exchange rate pass-through?: a study of import pricing in 51 Swedish industries*.

Kling, Ragnar. *Developing product development in times of brutal change.*

Langenskiöld, Sophie. *Peer influence on smoking : causation or correlation?*

Lychnell, Lars-Olof. *"Och fungerar det inte, gör vi på något annat sätt" : en klinisk fallstudie av IT-relaterat förändringsarbete i småföretag*

Meitz, Mika. *Five contributions to econometric theory and the econometrics of ultra-high-frequency data.*

Mendicino, Caterina. *Financial market imperfections, business cycle fluctuations and economic growth.*

Ovanfors, Anna. *Essays on nonlinear time series analysis and health economics.*

Paltseva, Elena. *Essays on commitment and inefficiency in political economy.*

Rogberg, Martin. *Den modeföljande organisationen : om acceptansen av TQM och andra populära managementmodeller.*

Silvennoinen, Annastiina. *Essays on autoregressive conditional heteroskedasticity.*

Sjögren, Ebba. *Reasonable drugs : making decisions with ambiguous knowledge.*

Slinko, Irina. *Essays in option pricing and interest rate models.*

Wilander, Fredrik. *Essays on exchange rates and prices.*

2005

Books

Andersson, Per, Susanne Hertz and Susanne Sweet (red). *Perspectives on market networks : boundaries and new connections.*

Charpentier, Claes. *IT inom omsorgen. Förväntade effekter av införande av IT-system för utförarna inom äldre- och handikappomsorgen.*

Dembrower, Maria. *Entreprenörskap i industriella nätverk.*

Lind, Johnny och Göran Nilsson (red). *Ekonomistyrningens metoder, sammanhang och utveckling : en vänbok till Lars A Samuelson.*

Samuelson, Lars A. *Organizational governance and control : a summary of research in the Swedish society.*

Svedberg Nilsson, Karin, Roger Henning och Karin Fernler (red). *En illusion av frihet? : företag och organisationer i det nya regelsamhället.* EFIs Årsbok 2005. EFI/Studentlitteratur.

Dissertations

Andersson, Martin. *Making a difference : project result improvement in organizations.*

Arvidsson, Per. *Styrning med belöningssystem : två fallstudier om effekter av belöningssystem som styrmedel.*

Bems, Rudolfs. *Essays in international macroeconomics.*

Berg-Suurwee, Ulrika. *Nya trender, nya nämnder : effekter av en stadsdelsnämndsreform inom kultur och fritid.*

Björkman, Hans. *Learning from members : tools for strategic positioning and service innovation in trade unions.*

Bodnaruk, Andriy. *Essays in empirical corporate finance and portfolio choice.*

Clapham, Eric. *Essays in real estate finance*.

Dareblom, Jeanette. *Prat, politik och praktik : om individers möten med strukturer i en kommunal satsning på kvinnors företagande.*

Fromm, Jana. *Risk denial and neglect : studies in risk perception.*

Hjelström, Anja. *Understanding international accounting standard setting : a case study of the process of revising IAS 12 (1996), income tax*.

Hortlund, Per. *Studies on Swedish banking 1870–2001.*

Lindahl, Therese. *Strategic and environmental uncertainty in social dilemmas*.

Linnarsson, Håkan. *Alliances for innovation : a structural perspective on new business development in cooperative ventures.*

Madestam, Andreas. *Developing credit markets.*

Nilsson, Roland. *The market impact of short-sale constraints.*

Nordfält, Jens. *Is consumer decision-making out of control ? : non-conscious influences on consumer decision-making for fast moving consumer goods.*

Nordin, Fredrik. *Externalising services : walking a tightrope between industrial and service logics.*

Parmler, Johan. *Essays in empirical asset pricing.*

Simbanegavi, Witness. *Price discrimination, advertising and competition.*

Thodenius, Björn. *Användning av ledningsinformationssystem : en longitudinell studie av svenska storföretag.*

Tolis, Christofer. *Framing the business : business modelling for business development*.

Östberg, Per. *Corporate disclosure and investor recognition.*

2004

Books

Ahrne, Göran och Nils Brunsson (red). *Regelexplosionen*.

Lind, Johnny. *Strategi och ekonomistyrning : en studie av sambanden mellan koncernstrategi, affärsstrategi och ekonomistyrning.*

Lind, Johnny och Walter Schuster (red). *Redovisningens teori, praktik och pedagogik : en vänbok till Lars Östman*.

Sevón, Guje och Lennart Sjöberg (red). *Emotioner och värderingar i näringslivet.* EFIs Årsbok 2004.

Wijkström, Filip and Stefan Einarsson. *Foundations in Sweden : their scope, roles and visions.*

Dissertations

Anderson, Anders. *Essays in behavioral finance.*

Balsvik, Gudrun. *Information technology users : studies of self-efficacy and creativity among Swedish newspaper journalists.*

Blix, Magnus. *Essays in mathematical finance : modeling the futures price.*

González Gómez, Andrés. *Nonlinear dynamics and smooth transition models.*

Grönqvist, Erik. *Selection and moral hazard in health insurance : taking contract theory to the data.*

Ivarsson Westerberg, Anders. *Papperspolisen : den ökande administrationen i moderna organisationer.*

Jutterström, Mats. *Att påverka beslut : företag i EUs regelsättande*.

Jönsson, Kristian. *Macroeconomic aspects of capital flows to small open economies in transition.*

Larsson, Pär. *Förändringens villkor : en studie av organisatoriskt lärande och förändring inom skolan.*

Lagerwall, Björn. *Empirical studies of portfolio choice and asset prices.*

Malmsten, Hans. *Properties and evaluation of volatility models.*

Marshall, Cassandra. *Dating for innovation : recognizing and creating opportunities in fluid environments through collaborative interorganizational relationship.*

Mattsson, Susanna. *På gränsen mellan ordning och oordning – tingens betydelse vid marknadsombildningar : en studie av svenska postväsendets ombildning under 1990-talet.*

Nilsson, Charlotte. *Studies in environmental economics : numerical analyis of greenhouse gas policies.*

Nilsson, Hans. *Medborgaren i styrsystemet : beskrivning av VAD och HUR i styrning av kommunal verksamhet*.

Nystedt, Jens. *Competition, regulation and integration in international financial markets.*

Pajuste, Anete. *Corporate governance and controlling shareholders.*

Richtnér, Anders. *Balancing knowledge creation : organizational slack and knowledge creation in product development.*

Salabasis, Mickael. *Bayesian time series and panel models : unit roots, dynamics and random effects.*

Sandberg, Rickard. *Testing the unit root hypothesis in nonlinear time series and panel models.*

Skallsjö, Sven. *Essays on term structure and monetary policy.*

Strikholm, Birgit. *Essays on nonlinear time series modelling and hypothesis testing.*

Söderström, John. *Från produkt till tjänst : utveckling av affärs- och miljöstrategier i produktorienterade företag.*

Talia, Krim. *The Scandinavian Currency Union, 1873–1924 : studies in monetary integration and disintegration*.

ISBN 978-91-7258-761-8