Multidimensional Markov-Functional and
Stochastic Volatility Interest Rate Modelling

Multidimensional Markov-Functional and Stochastic Volatility Interest Rate Modelling

Linus Kaisajuntti

Dissertation for the Degree of Doctor of Philosophy, Ph.D
Stockholm School of Economics 2011.

ISBN 978-91-7258-859-2

Keywords:
interest rate derivatives, Markov-functional model, LIBOR market model, multidimensional, stochastic volatility.

Printed by:
Ineko, Göteborg 2011

Distributed by:
The Research Secretariat
Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden
www.hhs.se

To My Family

Acknowlegdements

First and foremost I would like to thank my supervisors Tomas Björk and Joanne Kennedy for your inspiration, guidance and encouragement. Tomas is a master at explaining complicated things in a straightforward and clear manner and I have benefited greatly from his (many of them private) lectures about mathematics and mathematical finance as well as advices on how to tackle research. When I started the research part of my studies I got in contact with Jo who generously shared ideas and taught me about her unique approach to derivatives research. This has had a marked impact on this thesis with papers 1 and 3 co-authored with Jo and paper 2 inspired by discussions in connection with paper 1. Besides excellent supervision I have felt a friendship from both Tomas and Jo and it has been a pleasure spending time with both of you.

I would like to thank the Department of Finance for providing a stimulating environment. I have highly appreciated the time spent with colleagues and fellow PhD students discussing both serious research as well as less formal matters. In particular I am grateful to the 'The Math Finance Gang' for providing a fun and inspirational environment to study and learn mathematical finance. Extra thanks to Daniel Sunesson, Reimo Juks, Agatha Murgoci, Linus Siming and Sam Lee for all daily fun and various debates over the years.

I am grateful to the Statistics department at the University of Warwick for the several opportunities to visit. I would like to thank everyone at the department for taking so well care of me during my stays and in particular Martin Klimmek for the many shared beers and evenings.

Financial support from the Jan Wallander & Tom Hedelius foundation is gratefully acknowledged.

Special thanks to the design guru Jonas 'q:' Halai.

Most of all I would like to thank my parents for their inspiration, love and support, Emma for her love and patience and the shining star Smilla for all joy. To you I dedicate this thesis.

Stockholm, September, 2011
Linus Kaisajuntti

Contents

Introduction and Summary

This thesis consists of three papers in the area of interest rate derivatives modelling. The pricing and hedging of (exotic) interest rate derivatives is one of the most demanding and complex problems in option pricing theory and is of great practical importance in the market. Models used in production at various banks can broadly be divided in three groups: 1- or 2-factor instantaneous short/forward rate models (such as Hull & White (1990) or Cheyette (1996)), LIBOR/swap market models (introduced by Brace, Gatarek & Musiela (1997), Miltersen, Sandmann & Sondermannn (1997) and Jamshidian (1997)) and the one or two-dimensional Markov-functional models of Hunt, Kennedy & Pelsser (2000)).

In brief and general terms the main characters of the above mentioned three modelling frameworks can be summarised as follows. Short/forward rate models are by nature computationally efficient (implementations may be done using PDE or lattice methods) but less flexible in terms of fitting of implied volatility smiles and correlations between various rates. Calibration is hence typically performed in a 'local' (product by product based) sense. LIBOR market models on the other hand may be calibrated in a 'global' sense (i.e. fitting close to everything implying that one calibration may in principle be used for all products) but are of high dimension and an accurate implementation has to be done using the Monte Carlo method. Finally, Markov-functional models can be viewed as designed to combine the computational efficiency of short/forward rate models with flexible calibration properties.

The defining property of a Markov-functional model is that each rate and discount factor at all times can be written as functionals of some (preferably computationally simple) Markovian driving process. While this is a property of most commonly used interest rate models Hunt et al. (2000) introduced a technique to numerically determine a set of functional forms consistent with market prices of vanilla options across strikes and expiries. The term a 'Markov-functional model' is typically referring to this type of model as opposed to the more general meaning, a terminology that is adopted also in this thesis.

Although Markov-functional models are indeed a popular choice in practise

there are a few outstanding points on the practitioners' wish list. From a conceptual point of view there is still work to be done in order to fully understand the implications of various modelling choices and how to efficiently calibrate and use the model. Part of the reason for this is that while the properties of the short/forward rate and the LIBOR market models may be understood from their defining SDEs this is less clear for a Markov-functional model. To aid the understanding of the Markov-functional model Bennett & Kennedy (2005) compares one-dimensional LIBOR and swap Markov-functional models with the one-factor separable LIBOR and swap market models and concludes that the models are similar distributionally across a wide range of viable market conditions. Although this provides good intuition there is still more work to be done in order to fully understand the implications of various modelling choices, in particular in a two or higher dimensional setting.

The first two papers in this thesis treat extensions of the standard Markov-functional model to be able to use a higher dimensional driving process. This allows a more general understanding of the Markov-functional modelling framework and enables comparisons with multi-factor LIBOR market models. From a practical point of view it provides more powerful modelling of correlations among rates and hence a better examination and control of some types of exotic products.

Another desire among practitioners is to develop an efficient way of using a process of stochastic volatility type as a driver in a Markov-functional model. A stochastic volatility Markov-functional model has the virtue of both being able to fit current market prices across strikes and to provide better control over the future evolution of rates and volatilities, something which is important both for pricing of certain products and for risk management. Although there are some technical challenges to be solved in order to develop an efficient stochastic volatility Markov-functional model there are also many (more practical) considerations to take into account when choosing which type of driver to use. To shed light on this the third paper in the thesis performs a data driven study in order to motivate and develop a suitable two-dimensional stochastic volatility process for the level of interest rates. While the main part of the paper is general and not directly linked to any complete interest rate model for exotic derivatives, particular care is taken to examine and equip the process with properties that will aid use as a driver for a stochastic volatility Markov-functional model.

Below follows brief summaries of each paper.

Paper 1: An n-Dimensional Markov-Functional Interest Rate Model
This paper develops an n-dimensional Markov-functional interest rate model in the spot measure, i.e. a model driven by an n-dimensional state process and constructed using Markov-functional techniques. It is shown numerically that

this model is very similar to an n-factor LIBOR market model hence allowing intuition from the LIBOR market model to be transferred to the Markov-functional model. This generalises the results of Bennett & Kennedy (2005) from one-dimensional to n-dimensional driving state processes.

The model is suitable for pricing certain type of exotic interest rate derivative products such as TARNs on LIBORs or CMS spreads. For these products, the n-dimensional Markov-functional model may be used as a benchmark model allowing for powerful and flexible control of both correlations between different rates as well as skews/smiles in implied volatilities.

Paper 2: A Parametric n-Dimensional Markov-Functional Model in the Terminal Measure
As in paper 1 this paper develops an n-dimensional Markov-functional interest rate model. While parts of the construction are similar to the model in paper 1 this model is formulated in the terminal measure and is based on parametric functional forms of exponential type. The parametric functional forms enable analytical expressions for forward discount bonds and forward LIBORs at all times and allows for calibration of the model to caplet prices given by a displaced diffusion Black model. These analytical expressions provide a theoretical tool for understanding the structure of Markov-functional models and comparisons with the LIBOR market model.

In particular it is shown that for 'typical' market data the model is close enough to the LIBOR market model to be able to calibrate using the LIBOR market model calibration setup and machinery. This provides further information about the similarities (as well as some of the differences) between Markov-functional and LIBOR market models.

The parametric n-dimensional Markov-functional model may be used for products that require high-dimensional models for appropriate pricing and risk management. Compared with an n-factor LIBOR market model it has the virtue of being (much) faster for certain types of products.

Paper 3: Stochastic Volatility for Interest Rate Derivatives
This paper uses an extensive set of market data of forward swap rates and swaptions covering 3 July 2002 to 21 May 2009 to identify a two-dimensional stochastic volatility process for the level of rates. The process is identified step by step by increasing the requirement of the model and introduce appropriate adjustments.

The first part of the paper investigates the smile dynamics of forward swap rates at their setting dates. Comparing the SABR model (with different βs) of Hagan, Kumar, Lesniewski & Woodward (2002) and the Heston (1993) stochastic volatility models informs about what different specifications of the driving SDEs has to offer in terms of reflecting the dynamics of the smile across dates.

The outcome of the analysis is that a normal SABR model ($\beta = 0$) satisfactorily passes all tests and seems to provide a good match to the market. In contrast we find the Heston model does not.

The next step is to seek a model of the forward swap rates (in their own swaption measure) based on only two factors that enables a specification with common parameters. It turns out that this can be done by extending the SABR model with a time-dependent volatility function and a mean reverting volatility process. The performance of the extended (SABR with mean-reversion) model is analysed over several historical dates and is shown to be a stable and flexible choice that allows for good calibration across expiries and strikes.

Finally a time-homogeneous candidate stochastic volatility process that can be used as a driver for all swap rates is identified and used to construct a simple terminal Markov-functional type model under a single measure. This candidate process may in future work be used as a building block for a separable stochastic volatility LIBOR market model or a stochastic volatility Markov-functional model.

Paper 1

Paper 1

An n-Dimensional Markov-functional Interest Rate Model

Joint work with Joanne Kennedy. A shorter version of this paper is accepted for publication in Journal of Computational Finance. Comments from Vladimir Piterbarg and two anonymous referees are gratefully acknowledged.

1 Introduction

The pricing of exotic interest rate derivatives is one of the most demanding and complex problems in option pricing theory, being of great practical importance in the market. Over the past decade the LIBOR market model has become increasingly popular and serves as the benchmark model for pricing and risk management of these products. The market models were the first models that modelled observable rates directly and could calibrate perfectly to Black's formula for caplets. This approach also enabled direct control over the modelling of the correlations between LIBORs, which is important for the pricing of exotic derivatives. While the theory underlying the market models is straightforward, implementing them is not, and the literature devoted to efficient implementation and calibration is vast. This is due to their high dimensionality, which is a feature of these models even when only one stochastic driver is used.

An alternative class of models that can also calibrate to Black's formula (or indeed any arbitrage-free formula) for caplets and also gives control over the modelling of the observable market rates is the class of Markov-functional models. The spirit of the Markov-functional modelling approach is to base the model on a low-dimensional (usually one or two) Markov process which allows for efficient implementation of the model on a lattice, providing fast computation of prices and, importantly, risk sensitivities.

As shown in Bennett & Kennedy (2005), one-dimensional Markov-functional models and one-factor (i.e. one Brownian motion driver) separable LIBOR market models are very similar distributionally and virtually indistinguishable numerically for short maturities and typical market conditions. Although not in the spirit of the original Markov-functional approach, from a theoretical point of view it is a natural question to ask if this is true more generally, i.e. whether we can formulate a model using Markov-functional ideas that is similar to a multi-factor LIBOR market model, and if we can, how similar will it be to the corresponding LIBOR market model?

This paper develops a full-rank Markov-functional model, that is, a model driven by a Markov process of the same dimension as the number of LIBORs being modelled. Inspired by the techniques developed in Hunt et al. (2000), we model LIBORs at their setting dates as functions of an n-dimensional driving Markov process. As a consequence of the high dimension we may no longer implement the model on a lattice representing a low-dimensional process and instead we use the Monte Carlo method to be able to cope with high-dimensional integrals. Moreover we will be working under the rolling spot measure and use forward induction instead of the standard terminal measure and backward induction.

For a LIBOR market model an accurate implementation also has to be done

using the Monte Carlo method. The challenge in this model is to cope with the stochastic integral appearing in the drift term in each of the SDEs which specify the model. At first sight an accurate path-wise computation of these stochastic integrals would imply a very slow model. However, by clever implementations that approximate the drift over fairly large time-steps one may achieve a significant speed up; see Joshi & Stacey (2006), for an overview. Drift approximations of this type are also used for tractable calibration of the correlation structure of the model to the market.

The primary motivation for the development of Markov-functional models was to produce a class of models that could be implemented efficiently on a lattice. This is not the case for the models introduced here which have been developed, as we have stated above, for purely theoretical reasons. However, if one is willing to accept the limitations of working with a model that can only be implemented using Monte Carlo techniques (as is the case for the LIBOR market model), then the models that we introduce have some advantages. First, it allows for flexible alterations of the marginal distributions of the LIBORs at their setting dates. Secondly, approximations are usually used for the model's correlation structure when calibrating. It turns out that the correlation structure within the Markov-functional model is closer to that of the standard approximation than is that of the LIBOR market model, particularly for long maturities. This will become clearer later in the paper, see the motivation Section 3.1 for intuition and Figures 1.2 and 1.13 for numerical examples. Moreover, for products only depending on (the interaction of) LIBORs at their setting dates we do not need to keep track of a large number of processes through time and thus, for the same accuracy, the model is faster than the LIBOR market model.

As for the low-dimensional Markov-functional model, the properties of the higher-dimensional version we develop are not transparent. We perform several tests comparing the n-dimensional Markov-functional and the n-factor LIBOR market models and find that as, in the low-dimensional case, they are very closely connected and that the intuition from the LIBOR market model SDEs may be transferred over to the Markov-functional model. We first study the connection from a distributional point of view and then look at the pricing of Targeted Accrual Redemption Notes (TARNs). As shown in Section 6 the n-dimensional Markov-functional model seems like a suitable benchmark model for this type of products allowing to explore the impact of skews/smiles and different correlations structures in an efficient, flexible and accurate manner.

For products depending on (the interaction of) LIBORs before their setting dates the model is less efficient as it requires repeated computations of a high-dimensional conditional expectation. However, as shown in Section 5, the conditional expectation may be approximated with decent accuracy implying a reasonably efficient and accurate model. As the approximation is not depen-

dent on the choice of marginal distribution the model is hence a flexible and powerful alternative also for this type of product. As an example CMS spread TARN swaps are priced in Section 7.

The paper is organised as follows. Section 2 introduces notation and reviews the LIBOR market model. Section 3 develops the n-dimensional Markov-functional model. In sections 4 and 6 we test some properties of the models numerically, with Section 4 focusing on the distributional similarities and differences for LIBORs at their setting dates and Section 6 treating terminal correlations and the pricing of TARNs. Section 5 develops an approximation of LIBORs before setting dates and 7 applies it to the pricing of CMS spread TARN swap. Finally, Section 8 presents our conclusions.

2 Preliminaries

Consider a fixed set of increasing maturities

$$\text{today} = T_0 < T_1 < \cdots < T_n < T_{n+1}$$

and let, for simplicity, the accrual factors be given by $\alpha_i = T_{i+1} - T_i$. By letting D_{tT} denote the price, at time t, of a zero-coupon bond paying a unit amount at time T, LIBOR forward rates contracted at time t for the period $[T_i, T_{i+1}]$ are defined as

$$L_t^i = \frac{1}{\alpha_i} \cdot \frac{D_{tT_i} - D_{tT_{i+1}}}{D_{tT_{i+1}}}, \quad i = 1, ..., n. \tag{1}$$

We let the uncertainty in the economy be resolved over a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ and we assume the existence of forward measures, $\mathbb{Q}^i$, $i = 1, \ldots, n$ i.e. the equivalent martingale measures using $D_{\cdot T_{i+1}}$ as numeraire. Note that this implies that under Q^i, L^i is a martingale with respect to the filtration $\{\mathcal{F}_t\}$.

The models used in this paper will be constructed under what is commonly referred to as the rolling spot measure (henceforth denoted by $\mathbb{N}$). This is the equivalent martingale measure associated with the discrete savings account as numeraire. Normalising to unit value at time T_1 the numeraire process is given by

$$N_t = D_{tT_1}, \qquad 0 \leq t \leq T_1, \tag{2}$$

$$N_t = D_{tT_{i+1}} \cdot \prod_{j=1}^{i} (1 + \alpha_j L_{T_j}^j), \quad T_i \leq t \leq T_{i+1}. \tag{3}$$

We will further on assume the existence of an n-dimensional Wiener process under the measure $\mathbb{N}$

$$W = \left(W^1, \ldots, W^n\right)$$

with a given correlation structure

$$\mathrm{d}W_t^i \mathrm{d}W_t^j = \rho^{ij}\mathrm{d}t, \quad i, j = 1, \ldots, n. \tag{4}$$

We assume that the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is the augmented natural filtration generated by W.

2.1 The n-factor LIBOR market model

This section reviews the definition of the lognormal n-factor LIBOR market model of n continuous LIBORs under the rolling spot measure and using the 'scalar' formulation.

Suppose we are given n deterministic scalar volatility functions

$$\sigma_t^i, \quad t \leq T_i. \tag{5}$$

The LIBOR market model under the rolling spot measure is specified by a system of correlated SDEs. In order to have a straightforward calibration to market volatilities the setup is done such that under a measure change $\mathbb{N} \to \mathbb{Q}^i$, L^i is a lognormal martingale. Then the price today (i.e. at $T_0 = 0$) for a caplet that expiries at T_i will be given by Black's formula using the volatility

$$\bar{\sigma}^i := \sqrt{\frac{\int_0^{T_i} (\sigma_s^i)^2 \mathrm{d}s}{T_i}}, \tag{6}$$

and hence calibration of the LIBOR market model to the caplet market is trivial.

The LIBOR market model can be formally defined as follows.

Definition 2.1 *Suppose the LIBORs have the dynamics*

$$\mathrm{d}L_t^i = L_t^i \sigma_t^i \mathrm{d}\tilde{W}_t^i, \qquad i = 1, \ldots, n, \tag{7}$$

where $\tilde{W}^i$ is the Q^i-Wiener process generated by W^i under the Girsanov transformation $\mathbb{N} \to Q^i$. Then we have a discrete tenor LIBOR market model with volatilities $\sigma^1, \ldots, \sigma^n$.

From the above definition it is not obvious that, given a specification of $\sigma^1, \ldots, \sigma^n$, there exist a corresponding LIBOR market model. The following proposition states that, under the weak condition of bounded volatility, this is the case. For a proof see Bjork (2004).

Proposition 2.1 *Consider a given volatility structure $\sigma^1, \ldots, \sigma^n$ where each σ^i is assumed to be bounded. Denote by $\beta(t)$ the smallest integer such that $t \leq T_{\beta(t)}$. Now, define, under the rolling spot measure, $\mathbb{N}$, the processes $L^1, \ldots, L^n$ by*

$$\mathrm{d}L_t^i = L_t^i \left(\sum_{j=\beta(t)}^{i} \frac{\alpha_j L_t^j}{1+\alpha_j L_t^j} \sigma_t^i \sigma_t^j \rho_{ij} \right) \mathrm{d}t + L_t^i \sigma_t^i \mathrm{d}W_t^i, \quad t \leq T_i. \tag{8}$$

Then the Q^i-dynamics of L^i are given by (7) and thus there exists a LIBOR market model with the given volatility structure.

3 An n-dimensional Markov-functional model

Markov-functional interest rate models (Hunt et al. (2000)) were introduced as an alternative to short rate and LIBOR market models. The main advantages with low-dimensional Markov-functional models are that while they allow for a flexible and powerful calibration they may be implemented on a lattice and derivatives may be priced using backward induction implying fast and stable computations. The basic idea underlying Markov-functional modelling is to choose some Markov process x defined on $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$, and then develop a model where discount factors, numeraires and/or rates are written as functions of this process. The Markov process could be viewed as a process containing information about the state of the economy but have in general no economic meaning.

This section develops and discusses a Markov-functional model driven by an n-dimensional state process. When constructing higher-dimensional Markov-functional models the key is to ensure that

- the univariate and monotonicity properties which are required to make the functional fitting efficient in the one-dimensional case are retained, and
- the desired correlation/covariance structure is captured.

The standard trick when constructing multi-dimensional Markov-functional models is to introduce a pre-model that expresses each LIBOR as a function of the driving Markov process (which we now assume to be of dimension k). An example could be to take a k-factor LIBOR market model and replace the time-dependent drift with the time zero values. This would result in a model with something very close to the desired correlation structure and in which all LIBORs are lognormally distributed, but for which there is significant arbitrage. This arbitrage is finally removed by a Markov-functional sweep that works as a small perturbation of the pre-model.

In particular one assumes that the functional dependence of $L^i_{T_i}$ on the multi-dimensional vector x_{T_i} is only via the pre-model $\tilde{L}^i_{T_i}$. Thus

$$L^i_{T_i} = f^i\left(\tilde{L}^i_{T_i}(x_{T_i})\right), \tag{9}$$

for some monotonic function f^i. This specialisation is what enables the univariate and monotonicity properties in this higher dimensional setting to be retained. The derivation of the functional forms can then be done in an almost identical manner to the one-factor case.

However, computing non-trivial expectations by numerical integration and constructing a grid for evaluation of exotic derivatives is practically impossible in dimensions higher than two. But, if we are willing to use the Monte Carlo method we can use Markov-functional ideas to formulate a high dimensional model suitable for certain products. In the next section we turn first to the LIBOR market model introduced earlier for motivation on how to formulate the model and then we discuss how the model may be constructed.

3.1 Motivation

Consider an n-dimensional state vector process under the measure $\mathbb{N}$,

$$x = \left(x^1, x^2, \ldots, x^n\right), \tag{10}$$

This paper will analyse the case where x is n-dimensional Gaussian, but the construction carries over in principle also to other types of n-dimensional processes. For $i = 1, \ldots, n$ let,

$$x^i_t = \int_0^t \sigma^i_s \mathrm{d}W^i_s, \tag{11}$$

where the W^is are, as earlier, Brownian motions under $\mathbb{N}$, with instantaneous correlation $\mathrm{d}W^i_t \mathrm{d}W^j_t = \rho^{ij}\mathrm{d}t$.

The benchmark model for the type of products we have in mind is the n-factor LIBOR market model and we will start from this model to motivate the construction of the n-dimensional Markov-functional model. First note that the formal solution to the SDE defining the n-factor LIBOR market model in (8) can be written, for $t \leq T_i$, as

$$L^i_t = L^i_0 \cdot \exp\left(\int_0^t \sum_{j=\beta(s)}^{i} \frac{\alpha_j L^j_s}{1+\alpha_j L^j_s} \sigma^i_s \sigma^j_s \rho^{ij} \mathrm{d}s - \int_0^t \frac{(\sigma^i_s)^2}{2} \mathrm{d}s + x^i_t\right)$$

The computational challenge with this model is to take care of the (stochastic) integral in the drift term. As a motivation for the n-dimensional Markov-functional model we will consider a very crude drift approximation of the LIBOR market model. This is done by taking a single very long Euler step from times 0 to T_i, freezing all the L_t^j values at their time zero values. Hence, by performing this very crude drift approximation one approximates L_t^i using

$$\hat{L}_t^i = L_0^i \cdot \exp\left(\sum_{j=1}^{i} \frac{\alpha_j L_0^j}{1+\alpha_j L_0^j} \int_0^{\min(t,T_j)} \sigma_s^i \sigma_s^j \rho^{ij} \mathrm{d}s - \int_0^t \frac{(\sigma_s^i)^2}{2} \mathrm{d}s + x_t^i \right). \quad (12)$$

This gives a computationally very fast model but unfortunately there is a risk for significant errors in bond and caplet prices and the joint distribution properties of the approximation might be too far from the intended model to be of practical use. However, this approximation is very useful when calibrating the LIBOR market models to 'terminal correlations' because it allows us to compute an analytical approximation formula. Since our main focus is modelling LIBORs at their setting dates we will below give the formula for this case. Similar versions holds for other times as well, see for example Brigo & Mercurio (2006). Hence, when referring to 'terminal correlations' we will mean the correlation between $\log(L_{T_i}^i)$ and $\log(L_{T_j}^j)$. Using (12)

$$\begin{aligned} \mathrm{Corr}(\log(L_{T_i}^i), \log(L_{T_j}^j)) &\approx \mathrm{Corr}(\log \hat{L}_{T_i}^i), \log(\hat{L}_{T_j}^j)) \\ &= \frac{\int_0^{\min(T_i,T_j)} \sigma_t^i \sigma_t^j \rho^{ij} \mathrm{d}t}{\sqrt{\int_0^{T_i} (\sigma_t^i)^2 \mathrm{d}t} \sqrt{\int_0^{T_j} (\sigma_t^j)^2 \mathrm{d}t}}, \end{aligned} \quad (13)$$

which provides a connection between the instantaneous correlation parameters (ρ^{ij}) with something that may be estimated in the market. The danger of using this formula is obvious; if the approximation is not very good, the resultant LIBOR market model will not accurately reflect the intended terminal correlations, something which could have a large impact on correlation dependent products. Figure 1.2 in Section 4.1 leads us to guess that this formula will work well for maturities up to around 10 years, but could break down for longer maturities. In connection with pricing TARNs (Section 6.2) we will study this formula more closely and confirm that this is the case. Note however that even though terminal correlations are measure dependent this approximation formula is not and would look the same irrespective of measure for the LMM.

The drift approximation displayed above would be very crude and would typically not be used in a practical implementation of the LIBOR market model. Beveridge, Denson & Joshi (2009) compares several proposed approximations with respect to induced errors in bond prices as well as caplet prices. Their findings are that predictor-corrector type discretisations have a better overall

performance than the arbitrage-free discretisation method of Zhao & Glasserman (2000). The Predictor-corrector method is a two-step method based on the Euler technique. At each step forward in time one first take a standard Euler step forward holding the L_t^js constant at their previous values. This gives the predicted $L_{t+\delta}^i$. These are then used in the corrector step to provide a more accurate estimate of the L^j in the time interval $[t, t+\delta]$. The predictor-corrector technique allows for quite long steps to be made with reasonable errors in bonds and caplets prices and could mean a significant computational speed-up compared with the standard Euler technique. When implementing the LIBOR market model this is the technique we are using.

This paper will, however, not focus on drift approximations of the LIBOR market model. Instead we develop an alternative approach which takes as its focus the modelling of the joint distributions of the LIBORs at their setting dates. When performing the crude drift approximation above what one is effectively doing is writing each LIBOR at its setting date as a monotonic function of $x_{T_i}^i$ and that leads us to ask the question whether it's possible to construct a model which is similar to the crude drift approximation, in the sense that $L_{T_i}^i$ is a function of $x_{T_i}^i$ only, but unlike the drift approximation has, in an accurate implementation, negligible errors in bonds and caplets prices.

We will show below that by using the techniques introduced by Hunt et al. (2000) it is possible to construct an n-dimensional model with the above properties. This model can be seen as a small perturbation of the drift approximation model using a Markov-functional sweep i.e. for some monotonic function $\tilde{f}$ letting

$$L_{T_i}^i = \tilde{f}^i(\hat{L}_{T_i}^i). \tag{14}$$

From the construction of the model we will see that, once the process x is chosen, in order for the model to be arbitrage free and calibrated to Black's model for caplets the monotonic function of $x_{T_i}^i$ must be unique (see Theorem 3.1) and hence we may postulate

$$L_{T_i}^i = f^i(x_{T_i}^i), \tag{15}$$

directly.

Remark 3.1 *Intuitively, the Markov-functional sweep on the drift approximation model can be seen as a small perturbation that removes the induced discretisation errors and re-fits the marginal distributions. Since it's only a small perturbation the idea is that the terminal correlations produced by the Markov-functional model should be close to the terminal correlations produced by the analytical formula (13). The n-dimensional Markov-functional model is hence an arbitrage free model that is modelling terminal correlations as a first-order approximation of the LIBOR market model and has the right marginal distributions.*

3.2 Construction

The program is to construct a LIBOR Markov-functional model, modelling n LIBORs as functions of the above state vector process. As motivated above it would make sense to model each LIBOR as a function of a single component of the state vector. This is supported also from a construction point of view since it assures that the important univariate and monotonicity properties are retained. Thus let, under $\mathbb{N}$,

$$L^i_{T_i} = f^i(x^i_{T_i}), \qquad i = 1, \ldots, n, \tag{16}$$

for some monotonically increasing function f^i. The functional forms will be found in a similar fashion to standard Markov-functional modelling, but using forward instead of backward induction. For a one-dimensional Markov-functional model implemented using the rolling spot measure setup, see Fries & Rott (2004). In this setting one must use digital caplets in arrears instead of the standard digital caplets as calibration instruments (see Fries & Rott (2004) for a proof of why calibrating an arbitrage free model to digital caplet in arrears is equivalent to calibrating it to caplets).

Denote by $V^i(K)$ the value today (at time $t = 0$) of a digital caplet in arrears with strike K that expires at time T_i. By risk-neutral valuation under the measure $\mathbb{N}$,

$$V^i(K) = N_0 E^{\mathbb{N}} \left[\frac{\mathbf{1}\{L^i_{T_i} \geq K\}}{N_{T_i}} \right], \tag{17}$$

Now consider the ith step in the forward fitting procedure, i.e. assume the functional forms of $L^k_{T_k}$ are already known for $k = 1, \ldots, i-1$ and that the functional form of $L^i_{T_i}$ is sought. Start by choosing a grid of values (scalars) x^* and for each x^* calculate (note that this is typically a high-dimensional integral)

$$J^i(x^*) := N_0 E^{\mathbb{N}} \left[\frac{\mathbf{1}\{x^i_{T_i} \geq x^*\}}{N_{T_i}} \right]. \tag{18}$$

Then search for the strike K^* such that

$$V^i(K^*) = J^i(x^*) \tag{19}$$

and finally conclude that

$$f^i(x^*) = K^*. \tag{20}$$

The last conclusion follows from

$$\begin{aligned} N_0 E^{\mathbb{N}}\left[\frac{\mathbf{1}\{L^i_{T_i} \geq K^*\}}{N_{T_i}}\right] &= \\ V^i(K^*) &\equiv J^i(x^*) = \\ N_0 E^{\mathbb{N}}\left[\frac{\mathbf{1}\{x^i_{T_i} \geq x^*\}}{N_{T_i}}\right] &= N_0 E^{\mathbb{N}}\left[\frac{\mathbf{1}\{f^i(x^i_{T_i}) \geq f^i(x^*)\}}{N_{T_i}}\right], \\ &= N_0 E^{\mathbb{N}}\left[\frac{\mathbf{1}\{L^i_{T_i} \geq f^i(x^*)\}}{N_{T_i}}\right], \end{aligned} \tag{21}$$

where the monotonicity assumption is crucial for the second to last equality.

While the above construction is fine theoretically there is however one crucial practical problem to solve, namely computing $J^i(x^*)$ in equation (18). Recall that the numeraire, N_{T_i}, is a function of the first $i-1$ LIBORs and hence, by construction, a function of the first $i-1$ components of the state vector, x. This implies that one needs to compute a multidimensional integral under a non-trivial joint distribution function. The standard numerical technique for doing this is Monte Carlo integration. Subsection 3.5 outlines the setup for implementing the model.

Remark 3.2 *Note how using digital caplets in arrears and not standard digital caplets allows efficient calibration under* $\mathbb{N}$. *Since digital caplets in arrears both sets and pays at time* T_i, *the functional form of the numeraire* N_{T_i} *is all that is required to determine* $J^i(x^*)$ *in equation (18) under* $\mathbb{N}$. *Since* N_{T_i} *is a function of* $L^1_{T_1}, ..., L^{i-1}_{T_{i-1}}$ *it is determined using the functional forms deduced in the previous steps.*

3.3 Marginal distribution

Note that the construction algorithm presented in the previous subsection is not dependent on the particular choice of prices of digital caplets in arrears which dictates the marginal distributions. In principle one may choose any set of arbitrage free prices of digital caplets in arrears and hence enforce any marginal distribution on LIBORs at their setting dates. In this paper the main assumption will be that prices are given by the Black model implying lognormal marginal distributions of each forward LIBOR at their setting dates under their respective forward measures. This case is extensively numerically investigated in Section 4.

However, easy control over the marginal distributions is one of the major advantages with the Markov-functional model compared with LIBOR market models

and this section will display two possibilities to enforce implied volatility skews of caplets into the model.

Displaced diffusions

LIBOR market models with displaced diffusion marginal distributions have been very popular over the last decades, see for example Rebonato (2002) and Brigo & Mercurio (2006). Displaced diffusion marginal distributions are easily incorporated into the Markov-functional model. To get displaced diffusion marginals one models $L^i_{T_i}$ under Q^i as

$$L^i_{T_i} = Y_i - a_i \tag{22}$$

where Y_i is a lognormal random variable and $a_i \in \Re$. More precisely, write $Y_i := \exp(N_i)$ and set $E^{Q^i}[Y_i] = y_i$. Then N_i has a normal distribution with $E^{Q^i}[N_i] = \log(y_i) - \frac{\eta_i^2}{2}$ where $\eta_i^2 = \mathrm{Var}^{Q^i}[N_i]$. Note that under the condition $E^{Q^i}\left[L^i_{T_i}\right] = L^i_0$, y_i must be chosen as $L^i_0 + a_i$. For $a_i = 0$ we are back at the lognormal case and as a_i tends to ∞, $L^i_{T_i}$ tends to a normally distributed random variable.

One benefit of the displaced diffusion marginal distributions is that caplet prices (and hence also digital caplets in arrears prices) are given by a simple extension of the Black formula (simply add a_i to L^i_0 and to the strike). Moreover, a_i strictly positive leads to a strictly decreasing implied volatility skew, something which is roughly in line with typical observations in the market. The direct disadvantage is that the support of $L^i_{T_i}$ is $(-a_i, \infty)$, i.e. for positive a_i there is a possibility of negative rates.

In the displaced diffusion LIBOR market model the SDE for L^i under $\mathbb{N}$ is given by

$$\mathrm{d}L^i_t = (L^i_t + a_i)\left(\sum_{j=\beta(t)}^{i} \frac{\alpha_j (L^j_t + a_j)}{1 + \alpha_j L^j_t} \sigma^i_t \sigma^j_t \rho_{ij}\right)\mathrm{d}t + (L^i_t + a_i)\sigma^i_t \mathrm{d}W^i_t, \quad t \le T_i. \tag{23}$$

In the motivation section above the case $a_i = 0$ was studied. Note that by the structure of the above SDE everything in the motivation section carries through as it is if one puts focus on $L^i_t + a_i$ instead of L^i_t. Section 4.4 provides a numerical investigation/comparison of the displaced diffusion LIBOR market model and the n-dimensional Markov-functional model calibrated to displaced diffusion marginal distributions.

Modified displaced diffusions

Although displaced diffusion marginal distributions are easy to handle and may give market-like implied volatility skews the fact that it allows for negative LIBORs is a clear objection against it. In a Markov-functional setting there are robust ways around this since one does not have to define a continuous process for L^i, only the distribution of $L^i_{T_i}$. In general one may choose to model LIBOR at it's setting date as

$$L^i_{T_i} = g(Y_i) \tag{24}$$

for some function g. The function g may be then be chosen to enforce desired properties of the marginal distribution, such as implied volatility skews, positive rates, desirable look of the probability density function etc.

We will now display one possible choice of g that is similar to the displaced diffusion case but does not allow negative rates. Let,

$$L^i_{T_i} = E^{Q^i}\left[(Z_i Y_i - a_i)^+ \middle| Y_i\right] \tag{25}$$

where Z_i is a lognormal random variable with $E^{Q^i}[Z_i] = 1$ and $\mathrm{Var}[\ln(Z_i)] = \nu_i^2$ and with Y_i and a_i as above. The expectation may be computed in a Black-Scholes type fashion and leads to the formula

$$L^i_{T_i} = Y_i N(d_1) - a_i N(d_2) \tag{26}$$

$$d_1 = \frac{\log\left(\frac{Y_i}{a_i}\right) + \frac{\nu_i^2}{2}}{\nu_i} \tag{27}$$

$$d_2 = \frac{\log\left(\frac{Y_i}{a_i}\right) - \frac{\nu_i^2}{2}}{\nu_i} \tag{28}$$

Moreover, given the requirement $E^{Q^i}\left[L^i_{T_i}\right] = L^i_0$, y_i is endogenously given by the choice of a_i and ν_i. Roughly $y_i = L^i_0 + a_i$ as in the displaced diffusion case. Since the functional form of $L^i_{T_i}$ under Q^i is then specified, prices of caplets may be computed using standard one-dimensional numerical quadrature. The roles of the parameters are:

- a_i: The main driver of the slope of the implied volatility skews (as in the displaced diffusion case).
- η_i: Determines the level of the ATM implied volatility skew (as in the displaced diffusion case).
- ν_i: Shifts the probability mass away from zero. Note that for $\nu_i = 0$ one has a displaced diffusion model with floor at zero. Here $\nu_i > 0$ shifts the

probability mass away from zero, the larger the ν_i the further the mass moves.

Note that moving the mass from zero will lead to a somewhat less steep implied volatility smile for the modified displaced diffusion model compared with the displaced diffusion case. To match up the smiles one typically chooses a_i larger for the modified displaced diffusion model. Figure 1.1 displays an example of implied volatility skews and probability density functions given by the modified and standard displaced diffusion models.

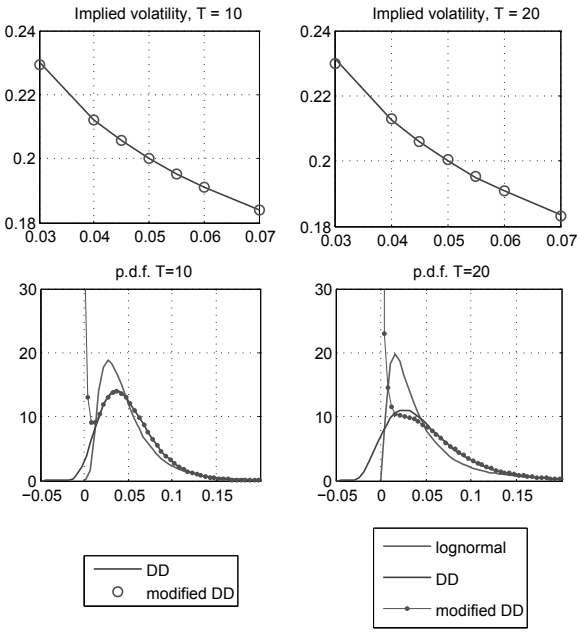

Figure 1.1: Implied volatility skews and probability density functions for the displaced diffusion and modified displaced diffusion models. $L_0^{10} = L_0^{20} = 0.05$. For the DD model $a_{10} = a_{20} = 0.05, \eta_{10} = 0.0988, \eta_{20} = 0.0975$ and for the modifiifed DD model $a_{10} = 0.055, a_{20} = 0.08, \eta_{10} = 0.0949, \eta_{20} = 0.0793$, $\nu_{10} = \nu_{20} = 0.1, x_{10} = 0.105, x_{20} = 0.128$.

3.4 Uniqueness

The following result is helpful when providing intuition about the structure of the n-dimensional Markov-functional model. Basically it says that once the joint distribution of $(x^1_{T_1}, \ldots, x^n_{T_n})$ has been specified then the functional forms f^i, $i = 1, \ldots, n$ in (16) are uniquely determined.

Theorem 3.1 *Consider an n-dimensional Markov-functional model based on the tenor structure $T_1, \ldots, T_n$ which satisfies the following conditions:*

1. *The driving Markov process is of the form specified in (11).*
2. *The ith LIBOR at time T_i, $L^i_{T_i}$, is a monotonically increasing function of the variable $x^i_{T_i}$, i.e.*
$$L^i_{T_i} = f^i(x^i_{T_i}).$$
3. *The model is calibrated to caplet prices (given by any arbitrage free model) corresponding to the rates $L^1, L^2, \ldots, L^n$ setting at dates $T_1, T_2, \ldots, T_n$.*

If such a model exists then it's unique, that is, the functional forms $f^1, f^2, \ldots, f^n$ are uniquely determined.

Proof. This follows from the construction of the n-dimensional Markov-functional model described in the previous section.

Start with the case $i = 1$. Pick any $x^* \in \Re$ and compute $J^1(x^*)$. Since for an arbitrage free model the value of a digital caplet in-arrears must be strictly monotonically decreasing in strike, the strike s.t. $V^1(K^*) = J^1(x^*)$ is unique and hence also $f^1(x^*)$ is uniquely determined.

For the case $i = 2$, uniqueness of f^1 implies uniqueness of f^2 using the same arguments as above. By induction the functional forms at all times in the above tenor structure must be unique. ∎

Note that this result explains why we can write (15) directly instead of (14).

Remark 3.3 *Note the implications of this result in practice. Any approximation to an n-factor LIBOR market model that is designed to be approximately arbitrage free and models each LIBOR at it's setting date, $L^i_{T_i}$, as a monotonic function of $x^i_{T_i}$ is, in effect, also an approximation to the unique arbitrage free n-dimensional Markov-functional model that calibrates to Black's formula for pricing the corresponding vanilla caplets.*

3.5 Implementation and efficiency

This section will describe in some detail how we chose to implement the n-dimensional Markov-functional model for pricing products depending on LIBORs at their setting dates. For products depending on LIBORs before setting dates, see Section 5. The main difference compared with a one-dimensional Markov-functional models is that, due to the high dimension, one can not use fast and accurate numerical integration methods such as Gaussian quadrature to compute expectations such as equation (18). Instead we choose to use Monte Carlo integration. While being straightforward to implement, Monte Carlo integration is inherently slow and since it will be used to find the functional forms this is a potential hazard to efficient implementation. Two methods we employ to mitigate this are as follows.

First, due to the fairly low dimension (the number of rates modelled) it is possible to speed up the convergence rate significantly by using low-discrepancy sequences, e.g. the Sobol sequence. The convergence rate for standard Monte Carlo integration is of the order $O(\frac{1}{\sqrt{M}})$, where M is the number of realisations. Jaeckel (2002), states a lower bound for the convergence rate using low-discrepancy sequences of the order $O(\frac{\log^d M}{M})$, where d is the dimension, but also argues that this lower bound is quite crude and that a careful implementation using Sobol sequences and Brownian bridge path construction methods could achieve far faster convergence, with a lot weaker dependence on dimension. Note at this point that for an accurate implementation of the n-factor LIBOR market model the dimension will become rather large (unless single steps out to each setting date is used) and hence the benefits of using Sobol sequences tends to decrease.

The second way to speed up the n-dimensional Markov-functional model is to use the same simulated state vector for calibration (i.e. pricing of digital caplet in arrears) and for pricing (of the intended exotic derivative). This will make up for major parts of the statistical error in a similar way as using caplets as control variates. We will give some examples of this at the end of this section.

Below follows a step by step outline of how we have implemented the model including some hints of how to speed up the implementation. All numerical routines are described (and implemented) in Press, Teukolsky, Vetterling & Flannery (2002).

1. Simulate a set of M realisations of the (n-dimensional Gaussian) state vector process

$$\tilde{x}(\omega_k) = \left(x^1_{T_1}(\omega_k), x^2_{T_2}(\omega_k), \ldots, x^n_{T_n}(\omega_k)\right), \qquad k = 1, \ldots, M. \tag{29}$$

   To do this we simulate a set of independent normally distributed variables

(using either pseudo-random (Mersenne Twister) or Sobol sequences) and then use spectral decomposition.

2. For each $i = 1, \ldots, n$ choose a set of reference grid points at which the functional form values will be computed. For this we have tested two different approaches;

   - The first method forces the probability mass for the gap between the grid points to be roughly equally large. To do this, sample every, say, kth number out of the (sorted) M simulated values of $x^i_{T_i}$. For example, if $M = 10000$, sort the simulated values and sample every 100th of them yielding a total of 100 reference grid points.
   - With the second method we divide the difference between the simulated largest and smallest values with the number of intended grid points and spread the grid evenly.

   In our experience the second method has been more efficient, implying a reduction in the number of grid points and hence speeding up the model.

3. Fix $i \in \{1, \ldots, n\}$. Then to find the functional form value f^i at the grid point x^* proceed as follows.

   (a) Compute $J^i(x^*)$ by Monte Carlo integration

   $$J^i(x^*) \approx N_0 \frac{1}{M} \sum_{k=1}^{M} \frac{\mathbf{1}\{x^i_{T_i}(\omega_k) \geq x^*\}}{N_{T_i}(\omega_k)}. \tag{30}$$

   The value of the numeraire at time T_i in state of the world ω_k, $N_{T_i}(\omega_k)$, is computed and stored in step $i-1$. Note that by sorting the state vectors a significant speed-up can be gained by not having to check whether or not $x^i_{T_i}(\omega) \geq x^*$ for all states.

   (b) Find the strike that matches the Black model value of a digital caplet in arrears with each of the $J^i(x^*)$. This is done in a fast and robust way using Brent's one-dimensional root finding algorithm. Remember that this strike is exactly the functional form value at point x^*

   (c) Having computed the functional form values for all grid points, use cubic splines to interpolate over the whole range of $x^i_{T_i}$. To our knowledge around 100 grid points interpolates with good precision.

   (d) Finally update the values of the numeraire at time T_{i+1} by

   $$N_{T_{i+1}}(\omega_k) = \left(1 + \alpha_i f^i(x^i_{T_i}(\omega_k))\right) N_{T_i}(\omega_k), \quad k = 1, \ldots, M. \tag{31}$$

   To get the values $f^i(x^i_{T_i}(\omega_k))$ the cubic splines fitted in the previous step are used. Even if calls to get the cubic spline interpolated values

are fairly quick it might at this point be useful to store (if memory allows) the realisations of the values $f^i(x^i_{T_i}(\omega_k))$ and reuse them for pricing later on.

4. Use the simulated state vector processes with corresponding LIBORs to price exotic derivatives using the Monte Carlo method.

As for efficiency we believe that the n-dimensional model is potentially, for the same degree of accuracy, quite a bit faster than the n-factor LIBOR market model. Computational times will of course highly depend on the skills of the person who implements the model and should always be taken somewhat lighthearted. Nevertheless, for products depending on (the interaction of) LIBORs at their setting dates we have found computational gains of at least a factor 10 (and potentially up to a factor 50) over an n-factor LIBOR market model implemented using one year long steps with the predictor-corrector scheme.

4 Numerical comparison of the models

The n-dimensional Markov-functional model defined in the previous section is formulated very differently from the n-factor LIBOR market model. While the behaviour of the LMM is fairly transparent due to its defining SDE this is not the case for the MFM, whose properties might be quite hard to grasp. Yet, given the motivation for the construction of the Markov-functional model (see Section 3.1) we wouldn't expect the models to be very different distributionally.

Note that what matters for the main type products we will consider is the joint distribution of $L^i_{T_i}$, $i = 1, \ldots, n$ under $\mathbb{N}$; if this distribution were the same for the two models we would get identical prices. To see this recall that for pricing a payoff depending on $L^i_{T_i}$ under $\mathbb{N}$ we will need to model also $L^j_{T_j}$, $j < i$, due to their presence in the numeraire N_{T_i}. This section will investigate the structure of the differences between the joint distributions (of the LIBOR at their setting dates) produced by the two models, through a number of comparisons. The main focus will be on providing intuition about in which sense the models are different and whether it is possible to transfer the intuition given from the LMM SDE over to the less transparent Markov-functional model. Most results deals with comparing the standard LMM with the MFM calibrated to the marginal distributions given by the Black model for caplet prices, however there is also a section displaying what happens in the displaced diffusion case.

Let's start with the following observation about marginal distributions in the standard case where caplet prices are given by the Black model. Recall that by construction the models' marginal distributions under the respective forward measures are fixed to be identical ($L^i_{T_i}$ has a lognormal distribution under Q^i).

However, this doesn't hold, in general, under $\mathbb{N}$. For time T_1 we have

$$\begin{aligned}\mathbb{N}(L^1_{T_1} \leq y) &= E^{\mathbb{N}}\left[\mathbf{1}\{L^1_{T_1} \leq y\}\right] \\ &= \frac{D_{0T_2}}{D_{0T_1}} E^{Q^1}\left[(1+\alpha_1 L^1_{T_1})\mathbf{1}\{L^1_{T_1} \leq y\}\right], \end{aligned} \tag{32}$$

where we have used the fact that the Radon-Nikodym derivative for the measure change Q^1 to $\mathbb{N}$ is

$$\left.\frac{\mathrm{d}\mathbb{N}}{\mathrm{d}Q^1}\right|_{T_1} = \frac{N_{T_1}}{N_0} \cdot \frac{D_{0T_2}}{D_{T_1T_2}} = \frac{D_{0T_2}}{D_{0T_1}}(1+\alpha_1 L^1_{T_1}). \tag{33}$$

Since we have made sure that both models have the same lognormal distribution under Q^1 for $L^1_{T_1}$ it's clear that the models will imply the same marginal distributions for $L^1_{T_1}$ under $\mathbb{N}$ as well. However, for the general case we get

$$\mathbb{N}(L^i_{T_i} \leq y) = \frac{D_{0T_{i+1}}}{D_{0T_1}} E^{Q^i}\left[\prod_{j=1}^{i}(1+\alpha_j L^j_{T_j})\mathbf{1}\{L^i_{T_i} \leq y\}\right]. \tag{34}$$

For this case the situation is different since, depending on the joint distribution of the $L^j_{T_j}$'s, the marginals under $\mathbb{N}$ might not be identical.

The most apparent difference between the models is that the MFM models each LIBOR at its setting date as a function of $x^i_{T_i}$ only, whereas for the LMM, $L^i_{T_i}$ depends not only on the value of $x^i_{T_i}$ but also, through the drift term, on the exact trajectories of x^j_t, $j \leq i$, $t \in [0, T_j]$ (see equation (3.1) in Section 3.1).

The process x which links the models under $\mathbb{N}$ will be key to making sensible comparisons between the models. In Section 4.1 we will provide scatter plots for the LMM overlayed on plots of the functional forms of the MFM. In Section 4.2 we study the distributions of $L^i_{T_i}$ conditional on some outcome for $x^i_{T_i}$ and in Section 4.3 we condition on the whole vector $\left(x^1_{T_1}, \ldots, x^n_{T_n}\right)$.

All results presented in this section are based on a scenario with flat initial LIBORs at 5%, flat term structure of volatility at 20% and instantaneous correlations given by $\rho^{ij} = \exp\{-0.1|T_i - T_j|\}$. We have performed tests also with other scenarios, in particular scenarios corresponding to the ones in Bennett & Kennedy (2005), Table 1. For these other tests we take initial forward LIBORs in the range 0.02 to 0.1 and volatilities in the range 0.1 to 0.3 (the term structures could be both flat, upward or downward sloping). As understood from the LMM SDE the dispersion of the LMM realisations will increase with increasing initial forward LIBORs or volatilities and the magnitude in the level of the rates will change according to the initial forward LIBOR levels. However, for these other scenarios, qualitatively nothing changes compared with the above base scenario. Since it is the qualitative features we are after in this

section we have chosen to use only this scenario, being fairly representative of typical market conditions.

4.1 Scatter plots

As a first investigation about how the models are coupled via the process x we plot in Figure 1.2 the functional form of $\log(L_{T_i}^{i,\mathrm{mfm}})$ as well as scatter plots for $\log(L_{T_i}^{i,\mathrm{lmm}})(\omega)$ vs $x_{T_i}^i(\omega)$. The scatter plots are constructed by simulating a number of realisations of $L_{T_i}^{i,\mathrm{lmm}}$ (with high accuracy) and for each realisation plotting the natural logarithm of the value vs the realised $x_{T_i}^i$. Once this is done the functional form for $L_{T_i}^{i,\mathrm{mfm}}$, f^i is plotted over the same range as for the simulated scatter plots.

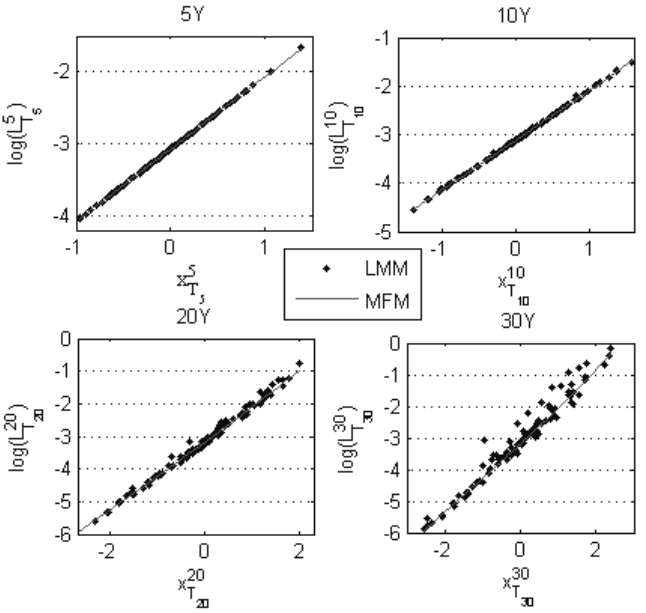

Figure 1.2: Scatter plots of $\log(L_{T_i}^{i,\mathrm{lmm}})$ vs $x_{T_i}^i$ as well as the functional form $\log f^i(x_{T_i}^i)$ for the cases 5, 10, 20 and 30 years.

From Figure 1.2 we note the following:

- The functional forms, f^i, of the Markov-functional model seem to be very close to log-linear in $x_{T_i}^i$.
- For maturities up to at least 10 years the LIBOR market model also seem

very close to log-linear in $x^i_{T_i}$. For longer maturities, however, the impact of the stochastic drift term starts to become noticeable.

Let's first note what this implies in terms of terminal correlations. Since the MFM functional forms are very close to log-linear we would expect the correlation structure of $(\log(L^1_{T_1}), \ldots, \log(L^n_{T_n}))$ and $(x^1_{T_1}, \ldots, x^n_{T_n})$ to be very similar. However, for the LIBOR market model the fairly large scatter for longer maturities implies that we wouldn't necessarily expect the approximation formula (13) to be very accurate for these cases. This view is supported by a closer investigation in connection with the pricing of TARNs which we discuss in the next section.

With regards to joint distributions note that Figure 1.2 motivates writing

$$\log(L^{i,\text{lmm}}_{T_i})(\omega) = \log(L^{i,\text{mfm}}_{T_i})(\omega) + \epsilon^i(\omega) = \log(f^i(x^i_{T_i}(\omega))) + \epsilon^i(\omega), \qquad (35)$$

where for each i, ϵ^i is a random variable that, at least for shorter maturities, has mean close to zero and comparably small variance. Hence,

$$L^{i,\text{lmm}}_{T_i}(\omega) \approx f^i(x^i_{T_i}(\omega)) \qquad (36)$$

which would lead us to expect that the joint distribution of $L^i_{T_i}$, $i = 1, \ldots, n$, will be close for the two models. For longer maturities however the scatter from the LIBOR market model is quite significant implying that the variance of ϵ^i starts getting sizeable. Moreover, it seems like we can no longer expect the mean to be zero. We will investigate this further in the next section, trying to display how the mean and variance of ϵ^i depend on maturity.

4.2 Conditioning on $x^i_{T_i}$

The scatter plots above give us realizations from the joint distributions of $(x^i_{T_i}, L^{i,\text{lmm}})$ and $(x^i_{T_i}, L^{i,\text{mfm}})$. Since $x^i_{T_i}$ is common for both models and seems to describe a large part of the variation also for $L^{i,\text{lmm}}_{T_i}$ it would be interesting to study the distribution of the $L^{i,\text{lmm}}_{T_i}$ conditional on $x^i_{T_i}$. To do this we fix an $x^i_{T_i}$ value and then conditionally fill in the whole path needed for a full simulation of the LIBOR market model (since the x vector is Gaussian this is straightforwardly done, albeit a bit messy, using some standard results from Statistics). In Figure 1.3 we plot the mean together with the 5th and 95th percentile of 100 realisations of $L^i_{T_i}$, $i = 5$, 10, 20, and 30, conditional on $x^i_{T_i} = y_j$, $j = 1, \ldots, 25$, where the values y_j are defined by $\mathbb{N}(x^i_{T_i} \leq y_j) = \frac{j}{26}$, $j = 1, \ldots, 25$. We also plot the MFM functional form values $f^i(y_j)$.

The plots in Figure 1.3 confirm the findings of the scatter plots. The conditional variance of ϵ^i is extremely small for short maturities and even for longer maturities we still consider it small.

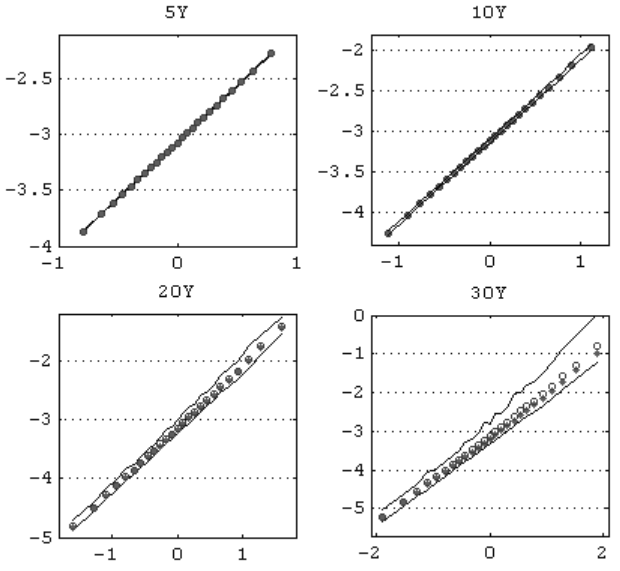

Figure 1.3: Plots of the mean (circles) together with the 5th and 95th percentiles of realisations from the LMM conditional on values for $x^i_{T_i}$ as described in the text. We also plot the MFM functional form values (dots) at the conditioned values.

To investigate this further we subtract the MFM values from the LMM ones and thus, from equation (35), we can estimate the conditional mean of ϵ^i:

$$E\left[\epsilon^i \middle| x^i_{T_i} = y_j\right] \approx \frac{1}{100}\sum_{k=1}^{100} \log L^{i,\mathrm{lmm}}_{T_i}(\omega_k)|_{x^i_{T_i}(\omega_k)=y_j} - \log f^i(y_j). \tag{37}$$

Results from this test are given in Figure 1.4. Note that the conditional mean seems to be positively biased for longer maturities. Also note that the conditional variance increases with magnitude of $x^i_{T_i}$.

To summarize the findings of this section we argue that the conditional mean and variance of ϵ^i are very close to zero and very small, respectively, for shorter maturities. Hence we would expect the joint distributions of the models to be numerically very similar for short maturities. For longer maturities, however, the conditional mean is positive (though small) and the variance is certainly sizeable. Given this we can no longer be certain that the models will imply very similar joint distributions. However, the models are definitely still very closely connected, implying that the intuition from the LMM SDE would be transferable to the MFM.

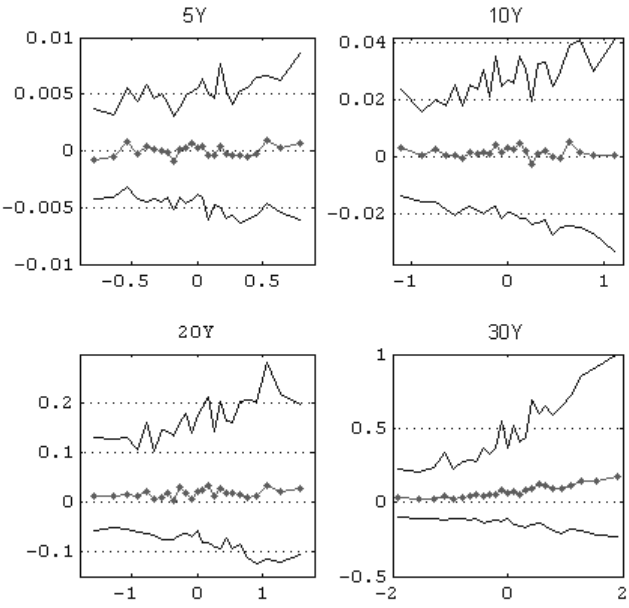

Figure 1.4: Estimation of the conditional mean of ϵ^i introduced in equation (35) for maturities $i = 5, 10, 20$ and 30. The plots also contain the 5th and 95th percentile of the simulated values.

4.3 Conditioning on $(x^1_{T_1}, \ldots, x^n_{T_n})$

In the discussion above, our starting point was the fact that for the MFM, $L^i_{T_i}$ is governed by $x^i_{T_i}$ and hence we studied the LMM conditional on $x^i_{T_i}$. To highlight differences between the models we now consider what happens when conditioning on $(x^1_{T_1}, \ldots, x^n_{T_n})$. More precisely we will choose $i = 10$ and study the behaviour of the LIBOR market model when the vector $(x^1_{T_1}, \ldots, x^n_{T_n})$ is fixed and L^{10} is simulated over the interval $[0, T_{10}]$.

Figure 1.5 displays four different state vector scenarios that we condition on when simulating L^{10} from time zero up to it's setting time $T_{10} = 10$ years. The first three scenarios are chosen to represent 'typical' realisations, and the fourth displays an usual scenario having a largely negative state vector. We fill in intermediate values needed to simulate a LIBOR market model using half-year long steps. We get realisations of L^{10}_t, where t goes from zero to T_{10}=10 years, as in Figure 1.6.

These plots are at first sight rather remarkable. Despite the fact that the

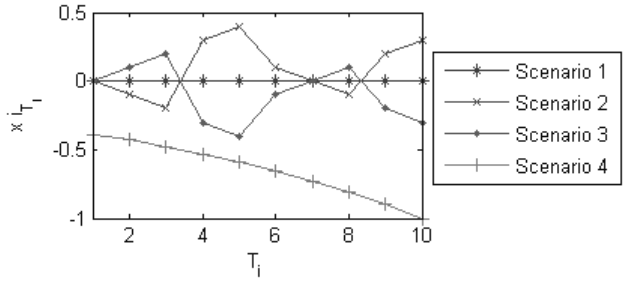

Figure 1.5: Four different scenarios for the fixed state vectors

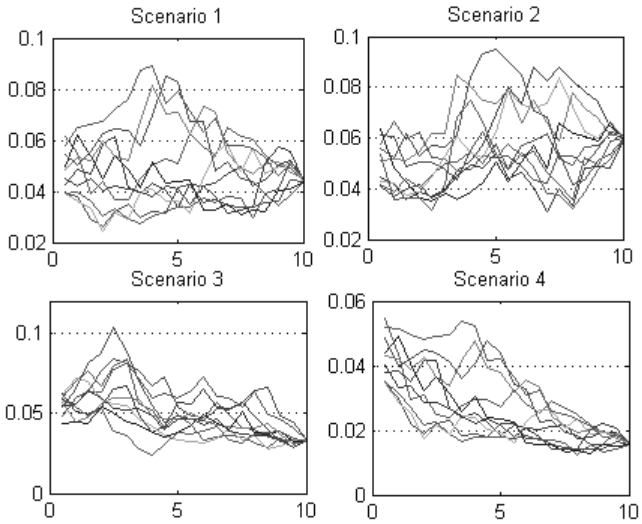

Figure 1.6: Ten different realisations of L_t^{10}, where t goes from 0 to T_{10}=10 years, conditional on the four different state vector scenarios in Figure 1.5

realised LIBORs can be very different before the setting date they all come together to what seems like almost the same spot at the end. This is consistent with the results we have seen in Figure 1.4 where we only conditioned on the end points.

An interesting next step from here would be to condition on some not so probable paths and see if they are all still coming together nicely. Three scenarios are plotted in the upper left graph in Figure 1.7. First note that all three scenarios end up at the points $x^{10}_{T_{10}} = 0.76$ but the other 9 components of the state vector are hugely different. The first scenario is a highly improbable sce-

nario with the $x^i_{T_i}$s waving from largely positive to largely negative values, the second scenario fixes all other $x^i_{T_i}$s to be zero and the third scenario fixes all $x^i_{T_i}$s to y_is s.t. $\mathbb{N}(x^i_{T_i} \leq y_i) = \mathbb{N}(x^{10}_{T_{10}} \leq 0.76)$.

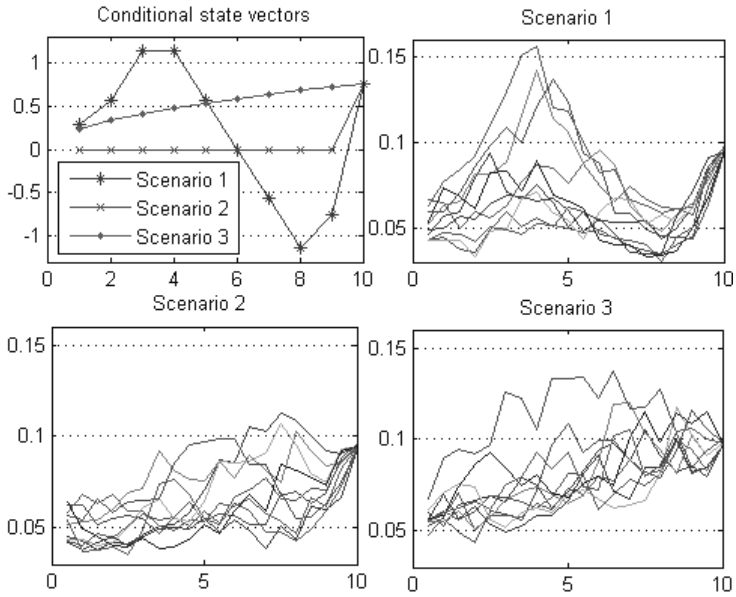

Figure 1.7: 10 different realisations of the L^{10}_t, where t goes from 0 to T_{10}=10 years, conditional on the three state vector scenarios displayed in the graph in the upper left corner.

These plots look in effect very similar to the ones in Figure 1.6, with the difference that the scatter is a bit larger and that the pattern in the conditional x vectors may readily be spotted in the realisations. We will now compare the outcomes of the $L^{10}_{T_{10}}$ values for these scenarios with the Markov-functional ones. The idea is that for these quite unusual scenarios there is a possibility that the models might not be very close anymore. To do this we generalise the scenarios (one waving, one zero and one fixed percentiles, exact values depends on the value of x^{10}_{10}) such that we can insert them into Figure 1.4 where we only conditioned on the final outcome, x^{10}_{10}. This gives rise to the graph in Figure 1.8.

From Figure 1.8 we note that even if we force the paths to take on very strange trajectories up to time T_{10} we end up with values not very different from the Markov-functional ones. Some intuition for the look of the respective paths could be as follows. Assume that $x^{10}_{T_{10}}$ is fixed to a positive value and consider the second scenario. Forcing all $x^i_{T_i}$s, $i = 1, \ldots, 9$ to be zero implies that the

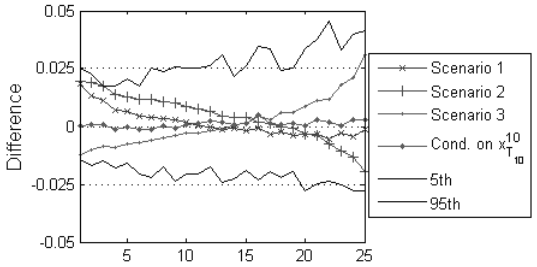

Figure 1.8: The graph in Figure 1.4 with added realisations from conditioning on generalisations of the scenarios used in Figure1.7.

L^i_t, $i = 1, \ldots, 9$ on average will be lower than they 'should' given the positive correlation between the components of x. Too low values implies a too low drift term and hence in the end a too low $L^{10}_{T_{10}}$. Fixing $x^{10}_{T_{10}}$ to a negative value reverses the arguments. These too low and too high values give rise to the biased pattern seen in the plots. For Scenario 3 similar arguments holds however since for scenario 1 the $x^i_{T_i}$ are both negative and positive it's not clear what will happen.

4.4 The displaced diffusion case

This section will explain how the above gained intuition is transferred to the case of the displaced diffusion LIBOR market model and the displaced diffusion Markov-functional models. The data used in this section is the same as in the displaced diffusion case in the TARNs pricing part, Section 6.

Figure 1.9 displays the displaced diffusion version of figure 1.2. In this plot both the functional forms in the standard Black case and the displaced diffusion case are given. The main differences between the two cases are that the range of the LIBOR realisations is smaller in the displaced diffusion case and, perhaps most interestingly, the LMM LIBOR scatter is significantly smaller. To understand why the scatter is smaller recall the LMM DD SDE, equation (23). Note that if one chooses $a_i = \frac{1}{\alpha_i}$ the drift no longer depends on earlier LIBORs and the n-dimensional DD Markov-functional model and the n-factor DD LIBOR market models would be identical! The typical DD models will have a_i somewhere in between zero and $\frac{1}{\alpha_i}$ and hence have scatter somewhere in between these two cases.

Redoing the other plots in the previous subsections for the displaced diffusion case provides no extra interesting information. For $a_i > 0$ the scatter is smaller

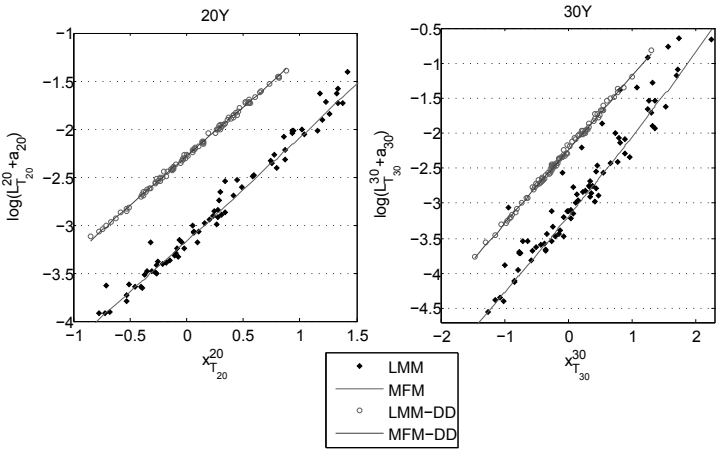

Figure 1.9: This plot is built up in the same manner as in figure 1.2. In addition to the Black case scatter plots of the displaced diffusion cases $\log(L^{i,\text{lmm}}_{T_i} + a_i)$ vs $x^i_{T_i}$ as well as the functional form of $\log(f^i(x^i_{T_i}) + a_i)$ is plotted. Only the cases 20 and 30 years are displayed.

and the figures look similar although with different scales.

4.5 Summary

From the tests in this section we have seen that $L^i_{T_i}$ is largely determined by $x^i_{T_i}$ for the LIBOR market model. However, as seen in Figure 1.2, for longer maturities the values of x^j_t, $t < T_j$, $j < i$ have a more noticeable effect on the LIBOR market model realisation of the ith LIBOR at its setting date T_i.

Remember that ultimately it is the intuition about the model behaviour we are after. We know a priori that by construction the differences displayed in the above tests will be there and our main concern was that there would be some bias somewhere which would affect the model behaviour too much. Despite the slight bias for longer maturities we still argue that the models are close enough to be able to transfer the intuition from the LMM SDE over to the Markov-functional model.

The differences displayed in the test should actually be of greater concern for someone implementing a speedy version of the LIBOR market model using drift approximations over large steps. The arbitrage and distorted calibration prop-

erties introduced by the drift approximation are hard to control and, for longer maturities, the exact trajectories in the LIBOR market model is important and has a non-negligible effect. With the Markov-functional model we are only concerned about whether we can still transfer the intuition, since the model by construction is arbitrage free and perfectly calibrated to market marginal distributions and terminal correlations. This should imply that at least for longer maturities the Markov-functional model should have a clear advantage over drift approximation models.

5 Computing the forward LIBOR curve

We have so far only considered properties of the model at the setting dates of the modelled LIBORs. A natural question to ask is how the model is specified at a general time t, not only the setting dates. Via the martingale property of L^i under its associated forward measure we have (using a change of measure from Q^i to $\mathbb{N}$ and the abstract Bayes' formula)

$$L_t^i = E^{Q^i}\left[L_{T_i}^i \middle| \mathcal{F}_t\right] = \frac{E^{\mathbb{N}}\left[\prod_{j=1}^i (1+\alpha_j L_{T_j}^j)^{-1} L_{T_i}^i \middle| \mathcal{F}_t\right]}{E^{\mathbb{N}}\left[\prod_{j=1}^i (1+\alpha_j L_{T_j}^j)^{-1} \middle| \mathcal{F}_t\right]}. \tag{38}$$

Observe that since x is a Markov process under $\mathbb{N}$, once we have chosen the process x (and therefore the conditional distribution of $(x_{T_1}^1, ..., x_{T_i}^i)$ given x_t) and the calibrating distributions, the Markov-functional model is fully specified at all times. Noting the product form in the equations and using the 'take out what is known technique' we can take out the LIBORs setting before time t and thus L_t^i will be a function of $(x_t^{\beta(t)}, ..., x_t^i)$. Hence, it is not possible to formulate an arbitrage free model where L_t^i depends only on x_t^i at all times (not only the setting time T_i) something which holds for the extreme drift approximation in equation (12).

While being informative, the above expression (38) might not be the best choice to work from in order to compute L_t^i. By using the martingale property of numeraire rebased assets one may derive the following two expressions which are slightly nicer to work from. Let $T_k \le t < s \le T_{k+1}$. Then for $i = k+1, \ldots, n$

$$E^{\mathbb{N}}\left[\prod_{j=k+1}^{i} \left(1+\alpha_j L_s^j\right)^{-1} \middle| \mathcal{F}_t\right] = \prod_{j=k+1}^{i} \left(1+\alpha_j L_t^j\right)^{-1}. \tag{39}$$

Alternatively, one may also start from LIBORs at their setting dates to get

$$E^{\mathbb{N}}\left[\prod_{j=k+1}^{i} \left(1+\alpha_j L_{T_j}^j\right)^{-1} \middle| \mathcal{F}_t\right] = \prod_{j=k+1}^{i} \left(1+\alpha_j L_t^j\right)^{-1} \tag{40}$$

That is, one could work from either the (already known) functional forms for LIBORs at their setting dates or iteratively using the time T_{k+1} functional forms. In either case one will at each time, say, T_k need to solve $n-k$ expectations to get the full forward LIBOR curve.

Note that while for the case $i = k+1$, $L^{k+1}_{T_k}$ is a function of $x^{k+1}_{T_k}$ and may hence be computed using numerical quadrature, the general case requires, due to the high dimension, the use of Monte Carlo integration. Computing the conditional expectations with Monte Carlo integration will be very computationally intense. To understand this suppose that we are constructing the model using m state matrices $x^j_{T_k}, 1 \leq k \leq j \leq n$. Then, to get all rates in each realisation one will need to compute $n-2+n-3+\ldots+1 = \frac{(n-2)(n-1)}{2}$ expectations by Monte Carlo integration.

To avoid using Monte Carlo integration of the conditional expectation the next subsection develops an approximation. In what follows we will restrict ourselves to specifying the model forward LIBORs at the canonical dates. Although the model is not in any way restricted to these dates it simplifies the presentation and still allows for pricing of most common derivative types.

5.1 Approximating $L^i_{T_k}$

Note that to get the forward LIBOR $L^i_{T_k}$ one would need to compute the conditional expectations in (38), (39) or (40). To speed up the computations this subsection outlines one possible approximation based on the expression (40).

Let, for a general set of times $[t_1, \ldots, t_n]$,

$$G^i_k(x^{k+1}_{t_{k+1}}, \ldots, x^i_{t_i}) := \prod_{j=k+1}^{i} \left(1 + \alpha_j L^j_{T_j}(x^j_{t_j})\right)^{-1} \tag{41}$$

To enable analytical computation of the expectation of $G^i_k(x^{k+1}_{T_{k+1}}, \ldots, x^i_{T_i})$ conditional on x_{T_k} one may expand G^i_k to second order in x around x_{T_k} and then take the conditional expectation. A few rather straightforward computations give

$$E^{\mathbb{N}}\left[G^i_k(x^{k+1}_{T_{k+1}}, \ldots, x^i_{T_i}) \middle| x_{T_k}\right] \approx G^i_k(x^{k+1}_{T_k}, \ldots, x^i_{T_k})\left(1 + A^i_k\right). \tag{42}$$

where

$$\begin{aligned} A_k^i &= \sum_{j=k+1}^{i} \left(\left(\frac{\alpha_j L'^j_{T_j}(x^j_{T_k})}{1+\alpha_j L^j_{T_j}(x^j_{T_k})} \right)^2 - \frac{1}{2} \frac{\alpha_j L''^j_{T_j}(x^j_{T_k})}{1+\alpha_j L^j_{T_j}(x^j_{T_k})} \right) C^j_j(T_k, T_j) \\ &+ \sum_{j=k+1}^{i} \sum_{u=j+1}^{i} \frac{\alpha_j \alpha_u L'^j_{T_j}(x^j_{T_k}) L'^u_{T_u}(x^u_{T_k})}{(1+\alpha_j L^j_{T_j}(x^j_{T_k}))(1+\alpha_u L^u_{T_u}(x^u_{T_k}))} C^j_u(T_k, (T_j, T_u)^-), \end{aligned}$$

and $C^i_j(t,T) := \mathrm{covar}[x^i_T - x^i_t, x^j_T - x^j_t]$, $L'^i_{T_i}(x) = \frac{\partial L'^i_{T_i}(x)}{\partial x}$ and $L''^i_{T_i}(x) = \frac{\partial^2 L'^i_{T_i}(x)}{\partial x^2}$. Solving for an explicit expression for $L^i_{T_k}$ gives

$$1+\alpha_i L^i_{T_k} = \frac{E^{\mathbb{N}}\left[G^{i-1}_k(x^{k+1}_{T_{k+1}}, \ldots, x^{i-1}_{T_{i-1}}) \middle| x_{T_k} \right]}{E^{\mathbb{N}}\left[G^i_k(x^{k+1}_{T_{k+1}}, \ldots, x^i_{T_i}) \middle| x_{T_k} \right]} \approx (1+\alpha_i L^i_{T_i}(x^i_{T_k})) \frac{1+A^{i-1}_k}{1+A^i_k}. \tag{43}$$

This expression provides quite good intuition about how $L^i_{T_k}$ depends on x. The main driver of $L^i_{T_k}$ is $x^i_{T_k}$ and the other parts of x are of secondary importance. What they do is basically providing some values of $L^i_{T_k}$ that are scattered around $L^i_{T_i}(x^i_{T_k})$, see Section 5.3 for more intuition.

Note that the approximation requires the partial derivatives $L'^i_{T_i}$ and $L''^i_{T_i}$. Since the construction of the Markov-functional model is based on finding the functional forms of $L^i_{T_i}$ these are in principle easy to find. However, for high numerical precision we recommend fitting a parametric function of the type $L^i_{T_i}(x) = \hat{a}_i \exp(\hat{b}_i x + \hat{c}_i x^2)$ to the functional forms and then get L' and L'' from this.

This is of course not the only possible way of approximating $L^i_{T_k}$. Among the ones we have tested it performs quite well and is reasonably simple, efficient and flexible and is therefore our preferred choice. In particular, note that the approximation is not contingent on the type of functional form of $L^i_{T_i}$ and it hence works for all of the marginal distributions introduced in the previous sections. Test results for the case with lognormal marginal distributions of each forward LIBORs at their setting dates under their respective forward measures are provided in sections 5.3 and 7.

5.2 Improving the approximation

In general the approximation in the previous subsection might not be accurate enough for practical purposes. As a way to improve the approximation one

may parameterise the approximation for $L^i_{T_k}$ in (43) as

$$\begin{aligned} 1+\alpha_i L^i_{T_k}(x_{T_k}) &\approx 1+\alpha_i \tilde{L}^i_{T_k}(x_{T_k}) \\ &:= \left(1+\alpha_i a_k \exp\left(b_k x^i_{T_k} + c_k \left(x^i_{T_k}\right)^2\right) L^i_{T_i}(x^i_{T_k})\right) \frac{1+d_k A^{i-1}_k}{1+e_k A^i_k}. \end{aligned} \tag{44}$$

This gives five parameters $(a_k, b_k, c_k, d_k, e_k)$ to play with at each T_k. The idea is hence to choose the parameters in order to reduce the approximation errors. Note that for higher accuracy it is of course possible to parameterise each $L^i_{T_k}$ independently, for example by using $(a^i_k, b^i_k, c^i_k, d^i_k, e^i_k)$ instead of $(a_k, b_k, c_k, d_k, e_k)$. We have however found that the computational cost and added complexity of doing this is not motivated by the (marginally) improved performance.

Suppose the model is set up using M realisations of the state matrix x. In order to control the accuracy of the approximation $\tilde{L}^i_{T_k}$ it may be compared with both accurately Monte Carlo computed $L^i_{T_k}$ values as well as the errors in the 'bond pricing relations'

$$E^{\mathbb{N}}\left[\frac{D_{T_k T_i}}{N_{T_k}}\right] = \frac{D_{0T_i}}{N_0}, \quad i = k+2, \ldots, n. \tag{45}$$

As the above is a consequence of the martingale property of numeraire rebased assets it is of high importance to assure small errors in this relation. This suggests using the following algorithm to find the parameters $(a_k, b_k, c_k, d_k, e_k)$.

1. Pick a 'base' set of $\hat{M}$ (random or user-defined) vectors $\hat{x}_{T_k}(\hat{m})$, $\hat{m} = 1 \ldots, \hat{M}$, with $\hat{M} << M$. Typically a $\hat{M}$ of about 100 seems to work well in our test applications.
2. Compute the forward LIBORs values $L^i_{T_k}$, $i = k+2, \ldots, n$ at each of the $\hat{M}$ chosen vectors using an accurate Monte Carlo integration (this could be using the same number of state matrices M as in setting up the model or fewer depending on the desired accuracy). Denote these values by $\hat{L}^i_{T_k}(\hat{x}_{T_k}(\hat{m}))$, $\hat{m} = 1, \ldots, \hat{M}$.
3. Pick some initial parameter values $(a_k, b_k, c_k, d_k, e_k)$. The default alternative would be to take $a_k = d_k = e_k = 1$ and $b_k = c_k = 0$ but an experienced user of the model may of course also pick parameters that are believed to be close to the calibrated ones.
4. Define the squared errors

$$E^1_k := \sum_{\hat{m}=1}^{\hat{M}} \sum_{i=k+2}^{n} \left(\hat{L}^i_{T_k}(\hat{x}_{T_k}(\hat{m})) - \tilde{L}^i_{T_k}(\hat{x}_{T_k}(\hat{m}))\right)^2. \tag{46}$$

and

$$E_k^2 := \sum_{i=k+2}^{n} \left(E^{\mathbb{N}} \left[\frac{\tilde{D}_{T_k T_i}}{N_{T_k}} \right] - \frac{D_{0T_i}}{N_0} \right), \tag{47}$$

where $\tilde{D}_{T_k T_i}$ is the discount bond prices obtained using the approximation $\tilde{L}^i_{T_k}$. Compute this expectation using M or potentially fewer realisations depending on the desired accuracy.

5. Define the weighted total error as $E^{tot} := w_1 E_k^1 + w_2 E_k^2$.
6. Minimise E_k^{tot} over the set of parameters $(a_k, b_k, c_k, d_k, e_k)$.

An example of the outcome of this exercise is given in Section 5.3. As the speed of the calibration is determined largely by the time it takes to Monte Carlo compute the accurate estimates $\hat{L}^i_{T_k}$ before entering the optimisation routine as well as the expectation in E_k^2 within the optimisation routine the next subsection outlines how these computations may be sped up.

Speed up using control variates

To speed up the Monte Carlo integrations of $\hat{L}^i_{T_k}$ and the expectation in E_k^2 we propose to use quantities based on the approximation as control variates. For a good explanation of the control variate technique, see Glasserman (2004).

To find $\hat{L}^i_{T_k}$, i.e. the accurate Monte Carlo computed estimates of $L^i_{T_k}$, we will work from equation (40). Note that computing these conditional expectations gives the (forward) discount factors and hence to find all $\hat{L}^i_{T_k}, i = k+1, \ldots, n$ we need to compute all the $n-k$ expectations (using Monte Carlo integration) and then divide through as in equation (43). Recall that in equation (42) we approximated these expectations by expanding the product inside the expectation to second order in x around x_{T_k}. Having done this we could then compute the conditional expectation analytically.

The idea here is hence to use the second order approximation of the product (used to get equation (42)) as control variate when computing the conditional expectations in (40). Since this approximation is correct to second order we expect it to be highly correlated with the true product. In our tests the correlation is, depending on i and k, in the set $[0.95, 1]$ implying a reduction of the required number of paths in the Monte Carlo integration of at least a factor 10 and in many cases up to a factor above 100.

The control variates technique may also be used to compute the expectation in E_k^2 using Monte Carlo integration and also in this case the approximation may be used to define a control variate. Unfortunately the value of the expectation in the expression for E_k^2 is not given analytically for the approximation and

instead a proxy for this will be computed before entering the optimisation routine. The procedure is a follows.

1. Pick some initial parameter values $(a_k, b_k, c_k, d_k, e_k)$ as in point 3. in the above algorithm.

2. Let $\frac{\tilde{D}^{cv}_{T_k T_i}}{N_{T_k}}$ denote the ith numeraire rebased discount bond at time T_k using the initial parameter values. Before entering the optimisation routine, compute the values of

$$E^{\mathbb{N}}\left[\frac{\tilde{D}^{cv}_{T_k T_i}}{N_{T_k}}\right], \quad i = k+2, \ldots, n,$$

by Monte Carlo integration with the M paths used for setting up the model or potentially fewer depending on the desired accuracy.

3. Within the optimisation routine, compute the values of $E^{\mathbb{N}}\left[\frac{\tilde{D}_{T_k T_i}}{N_{T_k}}\right]$, $i = k+2, \ldots, n$ using $\frac{\tilde{D}^{cv}_{T_k T_i}}{N_{T_k}}$ as control variates.

As within the optimisation routine each new set of 'proposed' (reasonable) parameters are typically close to the initial guesses set in point 1, the random variable $\frac{\tilde{D}^{cv}_{T_k T_i}}{N_{T_k}}$ will be highly correlated with $\frac{\tilde{D}_{T_k T_i}}{N_{T_k}}$ and hence a significant reduction of paths may be achieved. In our tests these correlations are in $[0.99, 1]$ implying a reduction of paths of at least a factor 50 and in many cases several hundred.

5.3 Testing the approximation

This subsection will test the quality of the basic as well as calibrated approximations. To test the accuracy of the approximations the model forward LIBORs $L^{k+j}_{T_k}$ are computed for $k = 1, \ldots, 20$ and $j = 0, \ldots, 10$. The test is based on the same times, rates, volatilities and correlations as in the pricing tests in Section 6 and 7 and under the assumption that prices of caplets are given by the Black formula. In general the approximation behaves well with very small errors for low k and j values and reasonable errors for large k and j. To get an idea about the accuracy we will display the cases $L^{20}_{T_{10}}$ and $L^{15}_{T_{10}}$. When calibrating the approximations we have used $\hat{M} = 100$ (randomly chosen) base realisations out of the $M = 100000$ realisations used to set up the model.

Figure 1.10 displays 100 randomly chosen (other than the 100 base paths used for calibrating the approximations) values of $L^{20}_{T_{10}}$ computed using the Monte

Carlo method, the base approximation (ie $a_k = d_k = e_k = 1$ and $b_k = c_k = 0$) and two different calibrated approximations. The first calibrated approximation ('Cal Approx 1') is calibrated using $w_1 = 1$ and $w_2 = 10$, ie a high weight (10 times more important than matching the Monte Carlo computed $\hat{L}^i_{T_k}$) on the 'bond pricing relation' (45) whereas the second calibrated approximation ('Cal Approx 2') puts $w_1 = 1$ and $w_2 = 0$, ie it focuses solely on matching the Monte Carlo computed $\hat{L}^i_{T_k}$. Figure 1.11 displays the differences between the Monte Carlo computed values and the (calibrated) approximations for the cases $L^{15}_{T_{10}}$ and $L^{20}_{T_{10}}$.

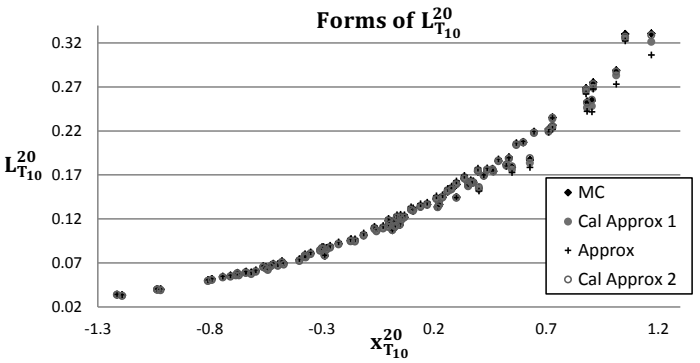

Figure 1.10: 100 randomly chosen values of $L^{20}_{T_{10}}$ computed using Monte Carlo and various approximations. 'Approx' refers to the non-calibrated approximation, 'Cal Approx 1' is calibrated using a high weight on the 'pricing relation' and 'Cal Approx 2' solely focuses on matching the Monte Carlo computed base values.

First note that the intuition about $L^{20}_{T_{10}}$ coming from the approximation (43) seems correct; $L^{20}_{T_{10}}$ is determined to a large extent by $x^{20}_{T_{10}}$. The dependence on other bits of the x-vector give rise to some scatter and this scatter seem to be reasonably covered by the approximation. Moreover, as may be seen by studying the larger rate values, note that the calibrated approximations seem to be able to adjust the basic approximation quite well.

From the difference plots note that the case $L^{15}_{T_{10}}$ has errors in the size of just a few basis points. Moreover, in both displayed examples, note that the calibrated approximations are centered around zero and one could hence hope for that the errors in a pricing situation will cancel out reasonably well.

As a further display of the accuracy of the approximation it is constructive to check the 'bond pricing relation' (45). To get something to compare with we have tested this relation also for the LIBOR market model implemented

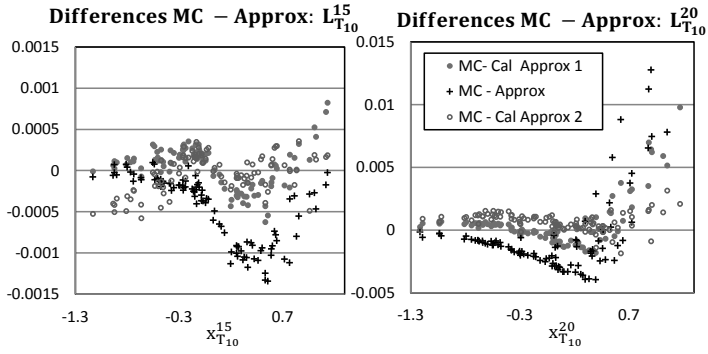

Figure 1.11: The differences between 100 randomly chosen values of $L^{20}_{T_{10}}$ computed using Monte Carlo and various approximations. 'Approx' refers to the non-calibrated approximation, 'Cal Approx 1' is calibrated using a high weight on the 'pricing relation' and 'Cal Approx 2' solely focuses on matching the Monte Carlo computed values.

using one predictor-corrector step in between every setting date. One year long steps with the predictor-corrector method seems as far as we know to be generally accepted in practice. Figure 1.12 displays the errors at time T_{20} for $i = 20, \ldots, 30$. As expected the calibrated approximation with a high weight on reducing this particular error does very well. It is however interesting to note that the errors of the non-calibrated approximation is similar in size (but with reversed sign) to the LMM errors. As a general comment note that the size of the approximation errors are increasing in time and at $T_i < T_{20}$ the errors are in general much smaller. To the benefit of the LMM it should however also be noted that at earlier times it seems to do better than the non-calibrated approximation.

6 TARNs

Targeted Accrual Redemption Notes (TARNs) provide some tough challenges for appropriate pricing due to that they are highly sensitive to correlation between the LIBORs at their setting dates. For this reason we believe it is an informative product to use in order to display properties of the n-dimensional Markov-functional model. Piterbarg (2004) provides a thorough introduction to TARNs that covers both pricing, risk sensitivities and Monte Carlo speed-up techniques in a LIBOR market model setting. Having done this he studies the deal-specific volatility structure components of a LIBOR market model and

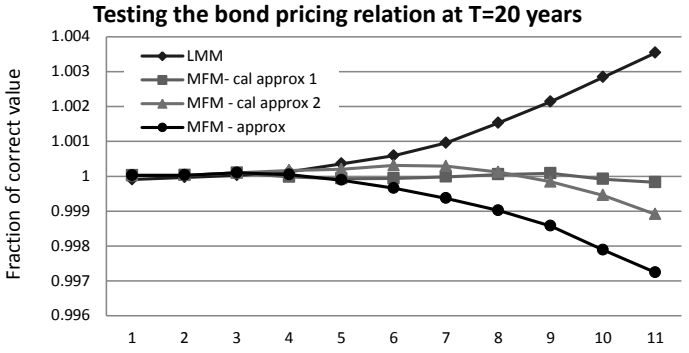

Figure 1.12: Testing the 'bond pricing relation' (45) at time T_{20} for $i = 21, \ldots, 31$. On the x-axis label i corresponds to the quantity related to $D_{\cdot i}$.

projects the main characteristics down on a low-dimensional Markov-functional model, the stochastic volatility Cheyette short rate model. Papers dealing with fast and stable computations of risk sensitivities of TARNs in a LIBOR market model setting are Fries & Joshi (2008), and Pietersz (2005).

TARNs can be structured both as notes and swaps. We have chosen to focus on a swap based product defined on the tenor structure $T_1 < T_2 < ... < T_n < T_{n+1}$, with accrual factors $\alpha_i = T_{i+1} - T_i$. For simplicity we will assume a unit notional in the description of the swap. For $i = 1, \ldots, n$, at time T_{i+1}, if the swap is not yet knocked-out, the investor receives

$$C_i = \alpha_i \max(a - bL^i_{T_i}, 0) \tag{48}$$

and pays

$$\alpha_i L^i_{T_i},$$

where a and b are some constants, which will typically depend on the level and slope of the yield curve. The swap is knocked-out when the total coupon payments reaches the target return level R. For $i = 2, \ldots, n+1$, define the aggregate coupon payments, at time T_i, as

$$Q_i = \sum_{j=1}^{i-1} C_j, \tag{49}$$

(with $Q_1 = 0$). Then, for $i \geq 1$, the time T_{i+1} swap payment is

$$X_i = \alpha_i \left(\min(R - Q_{i-1}, C_i) - L^i_{T_i} \right), \tag{50}$$

where the term $\min(R - Q_{i-1}, C_i)$ represents that the total return to the investor is capped by R. Thus, from the investors viewpoint, the value of the TARN swap at time 0 working under some risk-neutral measure $\mathbb{Q}$ with numeraire N is

$$V = N_0 E^{\mathbb{Q}} \left[\sum_{i=2}^{n+1} \frac{X_{i-1} \mathbf{1}\{Q_{i-1} < R\}}{N_{T_i}} \right]. \tag{51}$$

6.1 Pricing TARN swaps

The TARN swaps priced in this section are of the type presented above with parameter values as in Table 1.1. As initial data we use the forward LIBOR structure $L^i_{T_0} = \min(2\% + 0.5\% T_i, 10\%)$, i.e. linearly increasing from 2.5% at year 1 up to 10% at year 16 and then constant at 10% until year 30, constant implied Black volatilities at 20% and instantaneous correlations given by

$$\rho^{ij} = \exp\{\beta |T_i - T_j|\}, \tag{52}$$

where β is set at 0.05. For the displaced diffusion model we take $a_i = L^i_0$ for all i and choose η_i in order to match up an ATM implied volatility of 20%.

Tenors	1Y
Maximum length	5Y, 10Y, 15Y, 20Y, 25Y, 30Y
Pay	$L^i_{T_i}$
Receive	$\max(10\% - 2L^i_{T_i}, 0)$
Target level	10%
Notional	10 000

Table 1.1: Specifications of the TARN swaps

TARN swaps such as the one above will be very sensitive to correlations between the LIBORs. To see this consider the extreme cases full correlation and zero correlation and market data as above for a 30 year deal. Given the inverse floater coupon and the increasing LIBOR structure we get that with full correlation we will terminate early (during the first 5 years) on all 'downward' realisations (i.e. LIBORs setting lower than today's forward values) whereas in the 'upward' scenarios there is a high risk that we will continue the swap until year 30. The latter case implies a huge cost for the investor, having to pay LIBOR for 30 years, but, due to the inverse floater, basically not receiving anything until year 30. With zero correlation the chance of early termination is much higher, since it will basically be enough if 3 out of the first 6 LIBORs set below their time-zero values. Hence, the value of the TARN swap to the investor will be significantly higher for the case with zero correlation than with

full correlation. As we will see later, prices can change dramatically already for comparably small changes in instantaneous correlation.

Which measure to choose when implementing the LIBOR market model is a moot question, but in our view the rolling measure is most suitable as the bias from discretising the drift term is more evenly distributed among the different LIBORs. When pricing TARNs with a model designed for up to 30Y deals this problem is quite acute and at least for our implementation of the LMM, the rolling measure works much better.

Values to an investor of entering TARN swaps using an n-factor LIBOR market model and an n-dimensional Markov-functional model as well as the corresponding displaced diffusion versions, all implemented under the rolling measure $\mathbb{N}$, are reported in Table 1.2. For both models 100 000 paths were used. For the LIBOR market model the SDEs were evolved using 10 in-between steps between every setting date using the Predictor-Corrector method. For the Markov-functional model 1000 reference points were used to deduce the functional forms. Note that these choices are made to make sure that both the discretisation and statistical errors are very small and one could typically get away with a lot fewer steps, grid points and realisations in practice.

	5	10	15	20	25	30
LMM	-187.6	-729.2	-1095.3	-1270.5	-1338.0	-1362.5
MFM	-186.5	-719.2	-1067.4	-1230.9	-1294.0	-1316.7
LMM-DD	-213.3	-798.1	-1206.7	-1410.0	-1492.8	-1525.3
MFM-DD	-212.2	-792.4	-1191.8	-1388.6	-1467.9	-1498.8
s	5.8	12.6	17.4	19.7	20.6	21.0
vega	10.3	21.6	20.6	14.9	10.3	7.5
corr	-1.7	-9.6	-18.2	-22.4	-24.4	-25.2

Table 1.2: Values to an investor of entering into the TARN swaps specified above. Here s is defined as the number such that a 95% confidence interval for the Monte Carlo estimated value, say y, is given by $[y - s, y + s]$. The vegas are defined as the change in value when bumping all initial implied Black volatilities upwards by 1% (i.e. using 21 instead of 20%). Finally corr is defined as the change in value when using instantaneous correlations given by equation 52 using $\beta = 0.045$ instead of $\beta = 0.05$. The s, vega, and corr are all computed using the standard LMM.

The table displays some interesting features. Consider first the standard LMM and MFM models. In this case values of TARN swaps with maximum lengths up to around 10-15 years are within one vega and hence similar enough for practical purposes. However, for longer lengths the models produce values that are similar, but a bit too different to be able to say that model choice wouldn't matter with respect to uncertainty in input volatilities. Moreover, bumping the

instantaneous correlations affects prices a lot, confirming the intuition provided earlier. The prices given by the displaced diffusion models are quite different compared with the standard case and it seems like the ability to incorporate skews in the pricing model is important for TARNs pricing. The differences between the displaced diffusion LMM and MFM is smaller than for the standard cases, but qualitatively similar.

Note that when computing the above results we have chosen the same instantaneous correlation matrix for all models. However, this does not necessarily imply that the terminal correlations are the same and it is really the terminal correlations that will affect prices of TARNs. Recall the approximation formula for terminal correlations given by equation (13). If this formula works very well for both the LMM and MFM models, then we would have similar terminal correlations. If terminal correlations are in fact not the same, then this might be the main reason for the above price differences and should be investigated further. This is done in the following two subsections.

6.2 Terminal Correlations

This section studies terminal correlations, i.e. the correlation between $L^i_{T_i}$ and $L^j_{T_j}$ produced by the two models. With respect to terminal correlations, calculations and intuition will be clearer by transformation of the LIBORs through the natural log

$$\mathrm{Corr}(\log(L^i_{T_i}), \log(L^j_{T_j})).$$

Recall the approximation formula (13) in Section 3.1. This approximation is standard in the literature and commonly used in practice in order to calibrate the LIBOR market models to terminal correlations quickly and efficiently. However, in order for the calibrated LMM to actually perform as intended when calibrating it, the accuracy of this approximation is crucial. Recall the plots in Figure 1.2. For the 5 and 10 years cases the drift term in the LIBOR market model seem to have a very small impact and the natural log of $L^{i,\mathrm{lmm}}_{T_i}$ seem to be fairly well approximated as a linear function of $x^i_{T_i}$. We therefore expect the approximation formula to work well in these cases. However for later LIBORs the scatter is quite significant and the variation of $L^{i,\mathrm{lmm}}_{T_i}$ is to some extent not described only by $x^i_{T_i}$, implying that the approximation formula might break down in these cases. On the other hand, for the Markov-functional model the functional forms seem to be almost exactly linear in $x^i_{T_i}$ for all times and we will hence expect the approximation formula to work very well for this model.

Figure 1.13 confirms the intuition described above. For the Markov-functional model the approximation formula is almost perfect and essentially provides a sharp lower bound. However, for the LIBOR market model the approximation

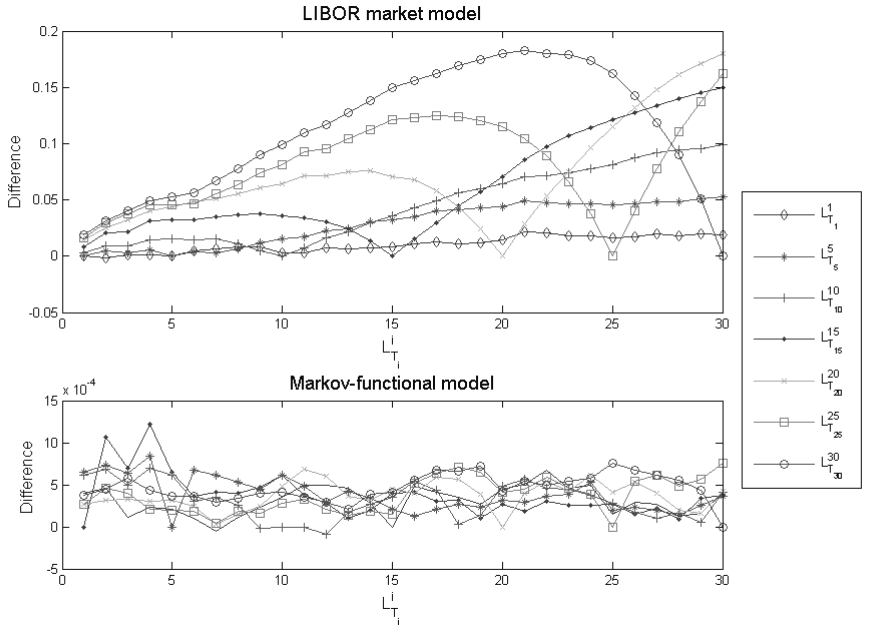

Figure 1.13: terminal correlations for the LMM and the MFM, respectively. Each line corresponds to the difference between the model terminal correlations and the approximation formula, i.e the line $L^{10}_{T_{10}}$ displays $\mathrm{Corr}(\log(L^{10}_{T_{10}}), \log(L^{i}_{T_i})) - \mathrm{Corr}(\log(\hat{L}^{10}_{T_{10}}), \log(\hat{L}^{i}_{T_i}))$ for $i = 1, \ldots, n$, for the two models respectively. The data used to produce these plots are the same as in the TARNs pricing part.

formula is systematically overestimating correlation in the calibrated LMM compared with a target correlation matrix. For longer maturities the problem seems to be quite acute and worthy of note, given that it seems to be the formula of choice for market participants.

In the displaced diffusion case the story is similar to the one above. Recall from Section 4.4 that the DD LMM scatter is decreasing in the displacement parameter. This leads to the terminal correlation formula working better for the DD LMM model. In the specific numerical case studied here, the errors from using the approximation formula are roughly half as large in the displaced diffusion LMM as in the standard LMM. For the DD MFM the terminal correlation formula continues to work very well.

6.3 Matching the models

This section will deal with providing a better match between the two model frameworks in order to potentially produce similar TARN prices. We will focus on two different but connected questions:

1. given a calibrated LIBOR market model, how can we come up with a Markov-functional model that matches it?
2. given market volatilities and terminal correlations of the LIBOR at their setting dates, how can we construct models that reflect these?

For question 1 we take the stand point that someone hands us a calibrated LIBOR market model and asks us to produce a Markov-functional model that mirrors it in terms of TARN prices. To do this we start with highlighting the following three empirical facts:

- increasing instantaneous correlations implies decreasing prices (see Table 1.2).
- the LMM produces consistently higher terminal correlations than the MFM using the same driving process x (see Figure 1.13).
- the LMM TARN swap prices are consistently lower than the MFM ones when using the same driving process.

These three facts motivates using a different state vector process than the obvious one (x) to drive the MFM, in particular a process with higher instantaneous correlation. Since we know that the approximation formula works very well for the MFM we could back out the ρ^{ij}s needed for matching the terminal correlations of the MFM with an estimated terminal correlation matrix created by Monte Carlo simulating the target LMM. Hence, define the state vector process $\hat{x}$ exactly as the original x process but with the instantaneous correlations between the Brownian motions as

$$\hat{\rho}^{ij} = \text{LMMTC}(i,j)\frac{\sqrt{\int_0^{T_i}(\sigma_t^i)^2\mathrm{d}t}\sqrt{\int_0^{T_j}(\sigma_t^j)^2\mathrm{d}t}}{\int_0^{\min(T_i,T_j)}\sigma_t^i\sigma_t^j dt}, \tag{53}$$

where LMMTC is the terminal correlation matrix obtained by simulation of the target LMM.

To match up the correlations also of the displaced diffusion versions we first compute the target LMM terminal correlations of logged $L^i_{T_i} + a_i$. For the DD MFM we then matched up these correlations using the approximative terminal

correlations formula whereas for the DD LMM we needed to use repeated Monte Carlo runs, since the terminal correlation formula was not accurate enough in this case. Having done this we would then have a good match between the all the models terminal correlations (remember that they are all calibrated to market marginals through the caplet market) and it would be interesting to see the effect of this on TARN swap prices. Table 6.3 displays the results. Note that the standard LMM and the MFM driven by $\hat{x}$ are now very close. This holds also in the displaced diffusion case. There is however still a quite large difference in price depending on the choice of marginals distributions.

	5	10	15	20	25	30
LMM	-187.6	-729.2	-1095.3	-1270.5	-1338.0	-1362.5
MFM	-186.5	-719.2	-1067.4	-1230.9	-1294.0	-1316.7
$\widehat{\text{MFM}}$	-187.0	-726.5	-1090.3	-1266.6	-1337.9	-1365.9
LMM-DD	-213.3	-798.1	-1206.7	-1410.0	-1492.8	-1525.3
MFM-DD	-212.2	-792.4	-1191.8	-1388.6	-1467.9	-1498.8
$\widehat{\text{LMM}}$-DD	-220.2	-833.1	-1276.4	-1492.7	-1588.0	-1625.7
$\widehat{\text{MFM}}$-DD	-219.3	-832.2	-1271.7	-1493.7	-1586.2	-1623.5
s.e	5.8	12.6	17.4	19.7	20.6	21.0
vega	10.3	21.6	20.6	14.9	10.3	7.5
corr	-1.7	-9.6	-18.2	-22.4	-24.4	-25.2

Table 1.3: Values to an investor of entering into the TARN swaps specified above. A hat such as, $\widehat{MFM}$, means the MFM or LMM in the case of matched up terminal correlations with the standard LMM case.

Let's now turn to question 2, i.e. given a view on market volatilities and terminal correlations, how can we construct a model that reflects these?

Since the terminal correlation approximation formula works very well for the Markov-functional model, calibrating the model to market volatilities and terminal correlations is straightforward. Recall that if we force the Markov-functional model and the LIBOR market model to agree on the terminal correlations of LIBORs at their setting dates, their prices are in almost perfect agreement. Moreover, in Section 4 we saw that the MFM and LMM are closely connected distributionally. Hence we are led to believe that with the n-dimensional Markov-functional model we have a model that is comparably fast computationally, calibrates easily and accurately and produces prices of derivatives that are very close to prices calculated with an accurately calibrated LIBOR market model.

For the LIBOR market model things are not as straightforward as for the Markov-functional model. As we have seen the approximation formula for terminal correlations breaks down for longer maturities and the prices produced

by using this model could be quite different from the intended ones (recall Table 6.3). A possible route to resolve this would be to use the Monte Carlo method in the calibration routine, however, this would be very slow and intractable. Hence, there seems to be a role for a better approximation formula to implement a LIBOR market model that is as accurate as the n-dimensional Markov-functional model.

7 CMS spread TARN swaps

This section uses the model to price CMS spread TARN swaps. A CMS spread TARN swap behaves in the same way as a standard TARN but with the difference that the coupon is based on the spread between two swap rates instead of the LIBOR. For this type of product one thus needs to model the joint behaviour over time of two swap rates as well as the LIBORs. Note that for this product we require knowledge of LIBORs at dates other than their setting dates and we will thus use the approximation developed in Section 5.

The CMS spread TARN swap considered is specified in Table 1.4. Initial forward LIBORs, volatilities and correlations as given in the TARN section, (Section 6) and caplets are assumed to be priced using the Black formula. For this setting, prices are given in Table 5. Note that the differences in prices between the LMM and calibrated MFM approximations are in general outside the 95% confidence interval for the LMM prices. Also the differences between the LMM and the calibrated MFM are about the same size as the differences between the MFM approximation and calibrated MFM approximation. The differences correspond to a 1% shift in implied volatility or a 0.005 change in the parameter β which controls the correlation via equation (52). This price difference is not unexpected for similar reasons to those dicussed in Section 6.

Tenors	1Y
Maximum length	5Y, 10Y, 15Y, 20Y
Pay	$0.3\ L^i_{T_i}$
Receive	$\max(\text{CMS}10Y - \text{CMS}2Y, 0)$
Target level	10%
Notional	10 000

Table 1.4: Specifications of the CMS spread TARN swaps

A final note about computational efficiency might be in place. For the pricing of standard TARNs we argued that the n-dimensional Markov-functional model is an efficient and flexible alternative to the n-factor LIBOR market model. As the pricing of CMS spread TARNs requires knowledge of more than the spot LIBOR at each future date some of the gain in efficiency is lost. In our

Model	5	10	15	20
LMM	389.2	258.2	235.0	229.1
MFM-Cal 1	387.1	260.1	241.1	237.3
MFM-Cal 2	385.8	259.8	241.7	238.2
MFM-Approx	389.0	267.5	250.3	246.9
s	1.1	2.0	2.3	2.4
vega	-3.9	-2.8	-2.0	-1.8
corr	-0.7	-5.8	-7.2	-7.4

Table 1.5: Values to an investor of entering into the CMS spread TARN swaps specified in Table 1.4. Here s is defined as the number such that a 95% confidence interval for the Monte Carlo estimated value, say y, is given by $[y - s, y + s]$. The vegas are defined as the change in value when bumping all initial implied Black volatilities upwards by 1% (i.e. using 21 instead of 20%). Finally corr is defined as the change in value when using instantaneous correlations given by equation (52) using $\beta = 0.045$ instead of $\beta = 0.05$. The s, vega, and corr are all computed using the standard LMM.

implementation, for a similar level of accuracy, computational times are roughly similar in both types of models. Note however that the n-dimensional Markov-functional model is a flexible option for incorporating skew/smile effects.

8 Conclusions

This paper has developed a full rank model of n LIBORs where the ith LIBOR at it's setting date is written as a function of the ith component of an n-dimensional state vector process. The functional forms are found using Markov-functional techniques originally developed for one-dimensional state vector processes by Hunt et al. (2000). This model allows us to extend the study of Bennett & Kennedy (2005), from the one-dimensional case with a separable volatility structure into this n-dimensional setting. It is shown numerically that there is a close connection between the n-dimensional Markov-functional model and the n-factor LIBOR market model even in the absence of separability. Although the Markov-functional models are models in their own right and not some type of approximation of LIBOR market models this connection is very useful since it allow us to transfer the intuition from the transparent LIBOR market model SDE formulation over to the, perhaps less transparent, Markov-functional model.

From the construction of the model we are able to prove a uniqueness result, Theorem 3.1 in Section 3.4, that has an interesting implication for n-factor LIBOR market models. In particular it means that any drift approximation of

an n-factor LIBOR market model such that each LIBOR at its setting date, $L^i_{T_i}$ is modelled as a monotonic function of $x^i_{T_i}$ is an approximation to the n-dimensional Markov-functional model that allows for arbitrage.

If one is willing to use Monte-Carlo techniques the n-dimensional Markov-functional model turns out to be potentially suitable as a benchmark model for certain types of exotic interest rate derivatives. In particular it provides a flexible and efficient way to study the effects of skews/smiles in implied volatility as well as the correlations between LIBORs for products depending on (the interaction of) LIBORs at their setting dates. For TARNs it is shown that it is of high importance to model both correlations and marginals correctly. The n-factor LIBOR market model is less flexible in modelling marginal distributions of LIBORs at their setting dates and also fails to fit the common approximation formula used for controlling correlation between the LIBORs with acceptable accuracy.

Paper 2

Paper 2

A Parametric n-Dimensional Markov-functional Model under the Terminal Measure

1 Introduction

Over the past decade the LIBOR market model (LMM) and its various extensions has become increasingly popular for exotic interest rate derivatives. For certain complex products that may require multi-factor or full rank models the LMM seems to serve as a benchmark for pricing and risk management. The market models were the first models that modeled observable rates directly and could calibrate perfectly to Black's formula for caplets and the approach enabled direct control over the modeling of the correlations between LIBORs, which is important for the pricing of exotic derivatives. While the theory underlying the market models is straightforward, implementing them is not, and the literature devoted to efficient implementation and calibration is vast. This is due to their high dimensionality, which is a feature of these models even when only one stochastic driver is used.

An alternative class of models that can also calibrate to Black's formula (or indeed any arbitrage-free formula) for caplets and also gives control over the modeling of the observable market rates is the class of Markov-functional models (MFM). The idea of the Markov-functional modeling approach is to base the model on a low-dimensional (usually one or two) Markov process which allows for efficient implementation of the model on a lattice, providing fast computation of prices and, importantly, risk sensitivities. As shown in Bennett & Kennedy (2005), one-dimensional Markov-functional models and one-factor (i.e. one Brownian motion driver) separable LIBOR market models are very similar distributionally and virtually indistinguishable numerically for short maturities and typical market conditions.

Although not in the spirit of the original Markov-functional approach Kaisajuntti & Kennedy (2011) constructed a full rank model under the spot measure of n LIBORs using Markov-functional techniques. The key in the construction was to write the ith LIBOR at its setting date as a function of the ith component of an n-dimensional Markovian state vector process. This model allowed for extending the study of Bennett & Kennedy (2005), from the one-dimensional case with a separable volatility structure into this n-dimensional setting and it was shown numerically that there is a close connection between the n-dimensional Markov-functional model and the n-factor LIBOR market model even in the absence of separability. Although the Markov-functional models are models in their own right and not some type of approximation of LIBOR market models this connection is useful since it allow transfer of the intuition from the transparent LIBOR market model SDE formulation over to the, perhaps less transparent, Markov-functional model.

Although the model by Kaisajuntti & Kennedy (2011) was primarily developed for theoretical reasons in order to be able to compare and transfer intuition from

the LMM it is a flexible and reasonably efficient alternative for products depending on LIBORs at their setting dates. However, for products depending on (the interaction of) LIBORs before their setting dates the model is less efficient as it requires repeated computations of a high-dimensional conditional expectation. Although an algorithm is outlined for how the conditional expectation may be approximated with decent accuracy it is a rather cumbersome exercise as the setup of the model is not particularly well designed for these type of products.

Inspired by Kaisajuntti & Kennedy (2011) this paper develops a full rank model by backward induction under the terminal measure. Also in this model the construction is based on writing the ith LIBOR at its setting date as a function of the ith component of an n-dimensional time-changed Brownian motion. The key here is that the choice of the terminal measure and forcing a specific parametric form for LIBORs at their setting dates implies that the required high-dimensional expectations may be computed analytically. Hence all LIBORs at all times are represented by the underlying n-dimensional Markov process through an explicit analytical function.

The chosen parametric functional form resembles the form obtained by fitting an n-dimensional MFM to prices given by a displaced diffusion Black model. Moreover, since the model provides an explicit analytical representation of all LIBORs at all times it provides a clearer understanding of the behaviour of Markov-functional models than is provided by the standard MFM construction. In particular, by the way the model is constructed it seems reasonable to believe that it should be close to the displaced diffusion LIBOR market model and the explicit functional forms provides a link to understanding the similarities and differences between the two model frameworks.

The close connection suggests using the same instantaneous volatilities and correlations in the driving process of both models. In fact, as displayed in Section 4.2, the standard correlation and swaption approximations used when calibrating the LMM seems to work also for the MFM, suggesting both that the models are close (in the sense that the differences in terms of calibration products are within acceptable calibration errors of the LMM) and that the machinery of calibration of the LMM may be used also for the MFM.

Due to the high dimension of the driving process pricing and risk computations must be performed using the Monte Carlo method. The analytical representation of LIBORs at all dates in terms of an n-dimensional time-changed Brownian motion is then a highly desirable feature as the computations may be performed by the Monte Carlo method without discretisation error. However, as we will see, while the functional forms for LIBORs before setting dates are on analytic form they can consist of a large number of terms and hence for efficient implementation, for some of the LIBORs, the forms need to be

approximated. Section 3.5 outlines a couple of straightforward and easy to implement approximations.

For the LIBOR market model an accurate implementation also has to be done using the Monte Carlo method. The challenge in this model is to cope with the stochastic integral appearing in the drift term in each of the SDEs which specify the model. At first sight an accurate pathwise computation of these stochastic integrals would imply a very slow model. However, by clever implementations that approximate the drift over fairly large time-steps one may achieve a significant speed up. The key for efficient use of the LMM is hence to chose a drift approximation enabling an as long as possible time-step without inducing too large discretisation errors. If a single step out to each date of interest is used, then the drift approximated LMM is similar in spirit to the MFM in the sense that each LIBOR is explicitly represented as a function of an n-dimensional time-changed Brownian motion.

In terms of pricing caplets and FRAs using a single step out to each expiry Joshi & Stacey (2006) studies the induced errors of several different drift approximations. By the setup of the n-dimensional parametric MFM it is particularly well designed for pricing products depending on (the interaction of) LIBORs at their setting dates. However, the use of the specific parametric functional form implies that prices of caplets are only approximately given by the displaced diffusion Black formula. Under the same setup, Section 4.1 compares some of the drift approximation methods in Joshi & Stacey (2006) with the n-dimensional parametric MFM and concludes that the parametric MFM is both much faster and provides prices of caplets that are closer to the displaced diffusion Black model than any of the drift approximated LMMs.

For products depending on (the interaction of) LIBORs before their setting dates any of the approximations developed in Section 3.5 need to be used in order to improve the computational performance. Section 4.2 argues that the performance of the approximations is quite good in terms of both computational speed and induced errors in prices of (forward) discount bonds and swaptions.

The paper is organised as follows. Section 2 sets notation, reviews the displaced diffusion LIBOR market model, the standard Markov-functional models and the basics of functional form modeling in this setting. Section 3 develops the n-dimensional Markov-functional model and discusses calibration and approximations. Then Section 3.6 discusses the relation with other models, Section 4 tests some properties of the model numerically and finally Section 5 presents the conclusions.

2 Preliminaries

2.1 Setup

Consider a fixed set of increasing maturities

$$\text{today} = T_0 < T_1 < \cdots < T_n < T_{n+1}$$

and let, for simplicity, the accrual factors be given by $\alpha_i = T_{i+1} - T_i$. By letting D_{tT} denote the price at time t of a zero-coupon bond paying a unit amount at time T, LIBOR forward rates contracted at time t for the period $[T_i, T_{i+1}]$ are defined as

$$L_t^i = \frac{1}{\alpha_i} \cdot \frac{D_{tT_i} - D_{tT_{i+1}}}{D_{tT_{i+1}}}, \quad i = 1, ..., n. \tag{1}$$

The uncertainty in the economy is resolved over a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ and we assume the existence of forward measures, $\mathbb{Q}^i$, $i = 1, \ldots, n$ i.e. the equivalent martingale measures using $D_{\cdot T_{i+1}}$ as numeraire. Note that this implies that under Q^i, L^i is a martingale with respect to the filtration $\{\mathcal{F}_t\}$.

Furthermore, assume the existence of an n-dimensional Wiener process under the measure Q^n

$$W = \left(W^1, \ldots, W^n\right)$$

with a given correlation structure

$$\mathrm{d}W_t^i \mathrm{d}W_t^j = \rho^{ij} \mathrm{d}t, \quad i, j = 1, \ldots, n. \tag{2}$$

The filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is taken as the augmented natural filtration generated by W.

Finally, for later reference, let $x_t = \left(x_t^1, \ldots, x_t^n\right)$ and define the following quantities

$$x_t^i := \int_0^t \sigma_s^i \mathrm{d}W_s^i, \tag{3}$$

$$V_t^i := \mathrm{Var}\left[x_t^i\right], \tag{4}$$

$$C_i^j(s,t) := \mathrm{Covar}\left[x_t^i - x_s^i, x_t^j - x_s^j\right]. \tag{5}$$

2.2 The displaced diffusion Black model

In the displaced diffusion Black model (see for example Brigo & Mercurio (2006) or Rebonato (2004)) $L_{T_i}^i$ under Q^i is modeled as

$$L_{T_i}^i = Y_i - a_i \tag{6}$$

where Y_i is a lognormal random variable and $a_i \in \Re$. More precisely, let $Y_i = \exp(N_i)$ where N_i is a Gaussian random variable where

$$\begin{aligned} E^{\mathbb{Q}^i}[N_i] &= \log(L_0^i + a_i) - \frac{\eta_i^2}{2} \\ \mathrm{Var}^{Q^i}[N_i] &= \eta_i^2 \end{aligned}$$

In this setting it may be shown that caplet prices are given by the Black pricing formula by adding a_i to the arguments for the strike and the initial forward LIBOR value. Moreover, positive a_i leads to a strictly decreasing implied volatility skew, a feature that is roughly in line with typical observations of caplet prices in the market. The direct disadvantage of the displaced diffusion extension of the Black model is that the support of L_i is $(-a_i, \infty)$, i.e. for positive a_i there is a possibility of negative rates.

For calibration the parameter η may be chosen to fit the market ATM (or indeed any strike) caplet price and a_i may be chosen to provide an implied volatility skew that is similar to the observed market skew. Note that, of course, for $a_i = 0$, η corresponds to the standard Black implied volatility.

2.3 The DD-LMM: definition and approximations

In the displaced diffusion LIBOR market model (henceforth DD-LMM) under the measure Q^n the forward LIBORs $L_t^i, t \le T_i, i = 1, \ldots, n$ are given by (see for example Brigo & Mercurio (2006))

$$L_t^i = -a_i + (a_i + L_0^i) \exp\left(- \sum_{j=i+1}^{n} \int_0^t \frac{\alpha_j (a_j + L_s^j) \sigma_s^j \sigma_s^n \rho_{jn}}{1 + \alpha_j L_s^j} - \frac{(\sigma_s^i)^2}{2} \mathrm{d}s + x_t^i \right). \quad (7)$$

The above dynamics is induced such that under a change of measures Q^n to Q^i, L_t^i has drift less dynamics with the explicit solution

$$L_t^i = -a_i + (L_0^i + a_i) \exp\left(-\frac{(\sigma_s^i)^2}{2} \mathrm{d}s + \int_0^t \sigma_s^i \mathrm{d}\hat{W}_s^i \right),$$

where $\hat{W}^i$ is a standard Wiener process under Q^i. This implies that the price of a caplet at time 0 is given by the displaced diffusion Black formula with

$$\eta_i = \sqrt{\frac{\int_0^{T_i} \left(\sigma_t^i\right)^2 \mathrm{d}t}{T_i}}. \quad (8)$$

and hence calibration of the LIBOR market model to the caplet market is a trivial exercise.

While the drift implies straightforward calibration to caplets it provides a challenge when numerically implementing the model as the value of L_t^i under Q^n depends on the complete history of L_s^j, $j = i+1, \ldots, n$, $0 \leq s \leq t$. This is an impediment to efficient implementation since in a practical implementation the drift integral will need to be approximated. In a first attempt to implementation one could consider using the log-Euler method implying

$$\hat{L}_{t+h}^i = -a_i + (a_i + L_t^i) \exp\left(- \sum_{j=i+1}^{n} \frac{\alpha_j (a_j + L_t^j)}{1 + \alpha_j L_t^j} \int_t^{t+h} \sigma_s^j \sigma_s^n \rho_{ij} - \frac{(\sigma_s^i)^2}{2} \mathrm{d}s + \left(x_{t+h}^i - x_t^i \right) \right). \tag{9}$$

In general the step size h required for an accurate implementation is fairly small and leads to a computationally inefficient model. To improve on the naive straightforward log-Euler approximation there is a wide variety of proposed approximations for the drift integral in the literature including for example the Predictor-Corrector (henceforth, PC) method of Hunter, Jackel & Joshi (2001), the Brownian bridge (BB) method of Pietersz, Pelsser & Van Regenmortel (2004) and the arbitrage free discretisation method (GZ) by Zhao & Glasserman (2000).

Ideally, for best computational performance, one would like to take as long steps as possible using an efficient approximation method. In particular, if one is interested in L_t^i the ideal solution would be to take a single step from time zero to time t. This leads to an approximation of L_t^i of the form

$$L_t^i \approx -a_i + (a_i + L_0^i) \exp\left(\mathrm{DA}(x_t) + x_t^i \right),$$

where $\mathrm{DA}(x_t)$ is the applied drift approximation, being a function of time t, the set of initial forward LIBORs, the collection of instantaneous volatilities in $[0, t]$ and the process x_t. Note that this implies that L_t^i is now a functional of x_t as it is represented by the Markov process x_t and may hence be written on the form

$$L_t^i = f_{DA}^i(t, x_t).$$

This implies an attractive model from an implementation point of view, but, depending on the accuracy of the drift approximation, there is risk for errors in bonds and options prices. Joshi & Stacey (2006) compares most introduced drift approximations (and introduces a few more) in terms of pricing forward rate agreements and caplets using a single step. Their conclusion is that an improved version of the PC method has the overall best performance, at the cost of being a few times slower than the original PC method (which is the fastest among the tested drift approximation methods and about half as fast as the naive one-step log-Euler approximation).

In the above mentioned drift approximations the full system of forward LIBORs $L^i, i = 1, \ldots, n$ is in general of dimension n. This would be the case even if the dimension of the driving Brownian motion would be of dimension one. To reduce the dimensions further one may specify a separable volatility structure. Under a separable volatility specification each component x^i may be written as

$$x_t^i = \nu^i \hat{x}_t, \tag{10}$$

where ν^i is an m-dimensional row vector and $\hat{x}_t$ is an m-dimensional column vector where the jth entity is given as $\hat{x}_t^j = \int_0^t \hat{\sigma}_s^j \mathrm{d}W_s^j$. Under this specification it is enough to keep track of the process $\hat{x}_t$ through time and the system of drift approximated forward LIBORs may hence be represented by a Markov process of dimension m. Using $m = 1$ or 2, Pietersz et al. (2004) combines separable volatility and a Brownian bridge drift approximation to develop a model that may be implemented on a lattice, implying fast and stable computations compared with the Monte Carlo based implementation required in the general case.

2.4 Markov-functional models

An alternative to LIBOR market models and in particular to the separable single-step drift approximation models is the one- or two-dimensional Markov-functional model introduced in Hunt et al. (2000). As in the LIBOR market model a LIBOR Markov-functional model also models discretely compounded forward LIBORs but, instead of specifying the instantaneous dynamics of the forward LIBORs, it directly specifies all rates and discount bonds as functionals of some Markov process z_t. While this is a feature of most available models in the literature (put f.e. $z = (L^1, L^2, \ldots, L^n)$ in the LIBOR market model) Hunt et al. (2000) introduced an algorithm to deduce the time-t functional forms of some 'computationally simple' Markov process (for example x above) that is consistent with a set of arbitrage free market prices of European type options.

In a LIBOR Markov-functional model, prices of caplets (across strikes) with expiry at each of the setting dates of the forward LIBORs, T_i, deduce the functional forms $L_{T_i}^i(x_{T_i})$. The functional forms at earlier dates are then given by the martingale property of numeraire rebased assets,

$$\frac{D_{tT_i}}{D_{tT_{n+1}}}(x_t) = E^{Q^n}\left[\left.\frac{1}{D_{T_iT_{n+1}}(x_{T_i})}\right| x_t\right]. \tag{11}$$

In the particular case of x being a one-dimensional (or a function of a two-dimensional) process the model allows an efficient implementation on a lattice.

While the model construction and properties of the LIBOR market model (as well as, most of, the drift approximations) is quite transparent and easily under-

stood it might not be the case for the LIBOR Markov-functional model. To shed light on this Bennett & Kennedy (2005) compares the one dimensional Markov-functional and the Pietersz et al. (2004) drift approximation model with the one-factor (accurately implemented) LIBOR market model and conclude that the models are very close through a wide range of market conditions. Since the LIBOR Markov-functional model prices bonds and caplets across all strikes (almost, depending on the accuracy of the grid implementation) perfectly and has similar computational efficiency as the single-step drift approximation of Pietersz et al. (2004) they argue it might be a preferred choice in practice for products where one-factor models are appropriate.

For a state process of dimension larger than two the efficiency of the standard Markov-functional model tend to get lost due to that it may no longer be implemented on a lattice. In the spot measure Kaisajuntti & Kennedy (2011) introduces an n-dimensional model that is constructed using the Markov-functional technique. It is argued that this model is close to the n-factor LIBOR market model and that it provides reasonably fast and flexible (and accurate due to lack of drift approximation errors) pricing of certain types of derivatives. Even though computations has to be done using the Monte Carlo method it is still quite efficient computationally due to the fact that each L_t^i is a functional of the simple Markov process x_t.

The key to the construction of the n-dimensional model in Kaisajuntti & Kennedy (2011) is the particular identification that $L_{T_i}^i$ is a function of $x_{T_i}^i$ only. Note that this identification aids in the construction of an n-dimensional Markov-functional model under the terminal measure as well. The key in the construction is to compute the quantity

$$\begin{aligned} J(x^*) &= D_{0T_{n+1}} E^{Q^n}\left[\frac{\mathbf{1}\left\{x_{T_i}^i \geq x^*\right\}}{D_{T_{i+1}T_{n+1}}}\right] \\ &= D_{0T_{n+1}} E^{Q^n}\left[\mathbf{1}\left\{x_{T_i}^i \geq x^*\right\} \prod_{j=i+1}^{n} (1+\alpha_j L_{T_j}^j)\right]. \end{aligned}$$

Having computed this, the functional form value of $L_{T_i}^i(x^*)$ is then given by the strike of the digital caplet with expiry T_i whose price is equal to $J(x^*)$ (see Kaisajuntti & Kennedy (2011) for details in the spot measure, the terminal measure follows in a similar fashion). By working backwards the functional forms of $L_{T_i}^i$ may be deduced subsequently and, since the functional forms at the setting dates are all that is needed to compute $J(x^*)$, finding the functional forms of the LIBORs at their setting dates may be done reasonably efficient using a set of pre-simulated vectors $\left(x_{T_1}^1, \ldots, x_{T_n}^n\right)$.

As is the case in any Markov-functional model, the values before setting needs to be found using the conditional expectation in (11), which due to the high-

dimensional setting, will be very time consuming to compute. In the spot measure Kaisajuntti & Kennedy (2011) developed an approximation that either could be used as it is (if the approximation errors would seem acceptable for the product at hand) or preferably optimised to fit a set of correctly computed values. While this is a possible route to take also in the terminal measure, this paper takes a different route to the model construction by forcing the functional form of (displaced) LIBORs at their setting dates to be of exponential type. The reason for assuming this particular form is that the conditional expectation in (11) may then be computed analytically. The driving relation used in the model construction is displayed in the next subsection.

2.5 Functional form modeling

Recall the martingale property of numeraire rebased bonds (11) and its use in the construction of Markov-functional models. In effect this relation is also determining the drift term in the LMM. To see this, recall that for the LIBOR market model by assumption

$$\mathrm{d}L_t^i = \mu^i(t, L_t)L_t^i\mathrm{d}t + \sigma_t^i L_t^i \mathrm{d}W_t^i, \tag{12}$$

where $L_t = (L_t^1, \ldots, L_t^n)$. Also recall that since L^n is a martingale under Q^n it has drift less dynamics. From (11) note that for $t \leq T_{n-1}$

$$(1+\alpha_{n-1}L_t^{n-1})(1+\alpha_n L_t^n) = E^{Q^n}\left[(1+\alpha_{n-1}L_{T_{n-1}}^{n-1})(1+\alpha_n L_{T_{n-1}}^n)\Big|\,\mathcal{F}_t\right]$$

implying that the product

$$Z_t := (1+\alpha_{n-1}L_t^{n-1})(1+\alpha_n L_t^n)$$

is a martingale with respect to the filtration $\mathcal{F}_t$. Hence, by the Ito formula,

$$\begin{aligned}
\mathrm{d}Z_t &= (1+\alpha_{n-1}L_t^{n-1})\alpha_n \mathrm{d}L_t^n + (1+\alpha_n L_t^n)\alpha_{n-1}\mathrm{d}L_t^{n-1} \\
&+ \alpha_n\alpha_{n-1}\mathrm{d}L_t^{n-1}\mathrm{d}L_t^n \\
&= \left((1+\alpha_n L_t^n)\alpha_{n-1}\mu^{n-1}(t,L_t) + \alpha_n\alpha_{n-1}L_t^n L_t^{n-1}\sigma_t^n\sigma_t^{n-1}\rho^{n,n-1}\right)\mathrm{d}t \\
&+ \ldots \mathrm{d}W_t^n + \ldots \mathrm{d}W_t^{n-1}
\end{aligned}$$

which is drift less if

$$\mu^{n-1}(t,L_t) = -\frac{\alpha_n L_t^n}{1+\alpha_n L_t^n}\sigma_t^n\sigma_t^{n-1}\rho^{n,n-1}.$$

This is the LIBOR market model drift condition for L^{n-1}. This exercise may of course be repeated for all $i = 1, \ldots, n$ and hence, given the assumption that the

dynamics is of the form specified in (12), the martingale property (11) implies the form of $\mu^i(t, L_t)$ for the LMM.

The model developed in this paper is constructed using a slight reformulation of (11) including only the conditional expectation of LIBORs at their setting dates.

Proposition 2.1 *For all* $0 \leq t \leq T_i$

$$\frac{D_{tT_i}}{D_{tT_{n+1}}} = \prod_{j=i}^{n} \left(1 + \alpha_j L_t^j\right) = E^{Q^n}\left[\prod_{j=i}^{n} \left(1 + \alpha_j L_{T_j}^j\right) \middle| \mathcal{F}_t\right] \tag{13}$$

Proof.

$$\begin{aligned}
E^{Q^n}\left[\prod_{j=i}^{n} \left(1 + \alpha_j L_{T_j}^j\right) \middle| \mathcal{F}_t\right] &= \\
E^{Q^n}\left[\prod_{j=i}^{n-1} \left(1 + \alpha_j L_{T_j}^j\right) E^{Q^n}\left[1 + \alpha_n L_{T_n}^n \middle| \mathcal{F}_{T_{n-1}}\right] \middle| \mathcal{F}_t\right] &= \\
E^{Q^n}\left[\prod_{j=i}^{n-1} \left(1 + \alpha_j L_{T_j}^j\right) \frac{D_{T_{n-1}T_n}}{D_{T_{n-1}T_{n+1}}} \middle| \mathcal{F}_t\right] &= \\
E^{Q^n}\left[\prod_{j=i}^{n-2} \left(1 + \alpha_j L_{T_j}^j\right) \frac{1}{D_{T_{n-1}T_{n+1}}} \middle| \mathcal{F}_t\right] &= \\
E^{Q^n}\left[\prod_{j=i}^{n-2} \left(1 + \alpha_j L_{T_j}^j\right) \frac{D_{T_{n-2}T_{n-1}}}{D_{T_{n-2}T_{n+1}}} \middle| \mathcal{F}_t\right] &= \\
E^{Q^n}\left[\frac{1}{D_{T_iT_{n+1}}} \middle| \mathcal{F}_t\right] &= \frac{D_{tT_i}}{D_{tT_{n+1}}}
\end{aligned} \tag{14}$$

∎

A few manipulations shows that L_t^i is then given explicitly as

$$L_t^i = \frac{1}{\gamma_t^{i+1}} E^{Q^n}\left[L_{T_i}^i \prod_{j=i+1}^{n} \left(1 + \alpha_j L_{T_j}^j\right) \middle| \mathcal{F}_t\right] \tag{15}$$

where $\gamma_t^i := \prod_{j=i}^{n} \left(1 + \alpha_j L_t^j\right)$ and $\gamma_t^{n+1} := 1$

Given this relation a model may be build through specifying the Q^n marginal distributions of the random variables $L^i_{T_i}$ as well as the law of the Wiener process that generates the filtration $\mathcal{F}_t$, in a similar fashion to the procedure of a standard Markov-functional model. As mentioned before, in high dimensions the conditional expectations will in general need to be computed using Monte Carlo integration implying a very inefficient model. However, as will be shown in the next section, by assuming a particular exponential form for $L^i_{T_i}$ the above conditional expectation may be computed analytically implying a much more efficient model construction.

3 A parametric MFM under Q^n

This section develops a full rank arbitrage free model under the terminal measure that is similar to the full rank LMM but with the difference that L^i_t may be directly represented as a function of x_t. The model is based on a postulated simple parametric functional forms for (spot) LIBORs at their setting dates and the no arbitrage form of the (forward) LIBORs at dates prior to setting will subsequently be derived using the relation (15).

3.1 Motivation

Recall that in the LMM only the terminal LIBOR, L^n, allows to be represented by x_t since there

$$L^n_t = L^n_0 \exp\left(x^n_t - \frac{V^n_t}{2}\right) \tag{16}$$

for all $t \in [0, T_n]$. Expressions for the other forward LIBORs follow from the assumption of a particular form of the dynamics which is then explicitly pinned down by the requirement to fulfill certain measure dependent martingale relations. As we have seen this leads to a conceptually appealing model, but for which efficient implementation in the general case is quite challenging.

As a crude approximation to the LIBOR market model one may fix the drift at the time zero values

$$L^i_t = L^i_0 \exp\left(x^i_t - \frac{V^i_t}{2} - \mu^i(0, L_0)t\right). \tag{17}$$

This 'model' has the advantage of being computationally efficient but admits arbitrage and has errors in bond prices. Moreover, caplet prices are only approximately given by the Black formula with volatility $\sqrt{\frac{V^i_{T_i}}{T_i}}$. However, since the variability and size of the drift term is fairly small one would not expect it

to be too far away from the correctly specified LIBOR market model distributionally. The model developed in this section is intuitively based on the above crude approximation but with the arbitrage removed by 'filtering' it through the martingale type relation (15).

3.2 The model

Assume that the functional forms at the setting dates are

$$L^i_{T_i} = b_i \exp\left(d_i x^i_{T_i} - \frac{d_i^2 V^i_{T_i}}{2} \right), \quad i = 1, \ldots, n \tag{18}$$

where b_i and d_i are some constants. As we will see later the constants b_i and d_i may be chosen to assure that bonds and caplets are correctly priced.

Recall that the relation (15) determines the time-t value of the forward LIBORs given the functional forms at their respective setting dates and hence, given (18), the model is completely specified at all times. Note that in principle the functional forms at the setting dates may be quite freely chosen and the reason for choosing this particular form is in order to be both close to the LMM and allow analytical computations of the conditional expectation in (15). The next subsection generalises this form to the displaced diffusion case.

To understand the procedure in the computation of (15) the cases L^n_t, L^{n-1}_t and L^{n-2}_t are computed explicitly. Using the notation

$$\hat{L}^i_t := b_i \exp\left(d_i x^i_t - \frac{d_i^2 V^i_t}{2} \right) \tag{19}$$

and following (15) gives

$$\begin{aligned} L^n_t &= E^{Q^n}\left[L^n_{T_n} \middle| \mathcal{F}_t \right] = E^{Q^n}\left[L^n_{T_n} \middle| x_t \right] = \hat{L}^n_t, && (20) \\ L^{n-1}_t &= \frac{1}{\gamma^n_t} E^{Q^n}\left[L^{n-1}_{T_{n-1}} (1 + \alpha_n L^n_{T_n}) \middle| x_t \right] \\ &= \frac{1}{\gamma^n_t}\left(\hat{L}^{n-1}_t + \alpha_n \hat{L}^{n-1}_t \hat{L}^n_t \exp\left(d_n d_{n-1} C^n_{n-1}(t, T_{n-1}) \right) \right) && (21) \end{aligned}$$

with $C^n_{n-1}(t,T_{n-1})$ as in (5) and

$$\begin{aligned} L^{n-2}_t &= \frac{1}{\gamma^{n-1}_t} E^{Q^n}\left[L^{n-2}_{T_{n-2}}(1+\alpha_{n-1}L^{n-1}_{T_{n-1}})(1+\alpha_n L^n_{T_n}) \middle| x_t \right] = \\ &= \frac{1}{\gamma^{n-1}_t}\Big(\hat{L}^{n-2}_t + \alpha_{n-1}\hat{L}^{n-2}_t \hat{L}^{n-1}_t \exp\left(d_{n-1}d_{n-2}C^{n-1}_{n-2}(t,T_{n-2})\right) \\ &+ \alpha_n \hat{L}^{n-2}_t \hat{L}^n_t \exp\left(d_n d_{n-2} C^n_{n-2}(t,T_{n-2})\right) \\ &+ \alpha_{n-1}\alpha_n \hat{L}^{n-2}_t \hat{L}^{n-1}_t \hat{L}^n_t \exp\left(d_{n-1}d_{n-2}C^{n-1}_{n-2}(t,T_{n-2}) \right. \\ &+ \left. d_n d_{n-2} C^n_{n-2}(t,T_{n-2}) + d_n d_n C^n_{n-1}(t,T_{n-1})\right)\Big) \end{aligned} \tag{22}$$

Note that the structure is such that moving one step further from n implies an addition of all possible combinations of terms including also the next intermediate LIBOR.

To display the form for the general case L^i_t it is useful to introduce the vector of integers $K^{i,n}$. $K^{i,n}$ refers to a particular choice of k integers out of the integers $(i+1,\ldots,n)$. The lth integer in the vector is referred to as $K^{i,n}[l]$. Note that in total it is possible to choose k elements out of m in $\binom{m}{k}$ unique ways. For some given i,n all possible vectors of integers may hence be referred to as $K^{i,n}_j, j=1,\ldots,\binom{n-i}{k}$. Using this notation one may express L^i_t as in the below proposition.

Proposition 3.1 *Given the choice of functional forms at the reset dates as in (18), the ith forward LIBOR at time t is given by*

$$L^i_t = \frac{\hat{L}^i_t}{\gamma^{i+1}_t}\left(1 + \sum_{k=1}^{n-i} h(L^{i+1}_t,\ldots,L^n_t;k)\right) \tag{23}$$

where

$$h(L^{i+1}_t,\ldots,L^n_t;k) = \sum_{j=1}^{\binom{n-i}{k}} \prod\left(\alpha_{K^{i,n}_j}\right)\prod\left(\hat{L}^{K^{i,n}_j}_t\right)\exp\left(\bar{C}^{K^{i,n}_j}_t\right), \tag{24}$$

with

$$\prod\left(\alpha_{K^{i,n}_j}\right) = \alpha_{K^{i,n}_j[1]}\cdot\ldots\cdot\alpha_{K^{i,n}_j[k]}, \tag{25}$$

$$\prod\left(\hat{L}^{K^{i,n}_j}_t\right) = \hat{L}^{K^{i,n}_j[1]}_t\cdot\ldots\cdot\hat{L}^{K^{i,n}_j[k]}_t, \tag{26}$$

and

$$\begin{aligned}\bar{C}_t^{K_j^{i,n}} &:= \sum_{u=1}^{k} d_i d_{K_j^{i,n}[u]} C_i^{K_j^{i,n}[u]}(t, T_i) \\ &+ \sum_{u=1}^{k} \sum_{v=u+1}^{k} d_{K_j^{i,n}[u]} d_{K_j^{i,n}[v]} C_{K_j^{i,n}[u]}^{K_j^{i,n}[v]}\left(t, T_{K_j^{i,n}[u]}\right), \end{aligned} \tag{27}$$

under the convention that empty sums are zero.

Proof. First note that the expression holds for the cases L^n, L^{n-1}, L^{n-2} given above. Straightforward, but rather tedious, calculations for L^{n-3} and L^{n-4} reveals the structure in the result from which the general case follows. ∎

While the expression in the above proposition might look fairly greasy due to all indices it is in effect just a sum of terms and is relatively straightforward to implement on a computer. Note however that the number of terms in the expression for L_t^i is given as 2^{n-i}. Hence for i far from n evaluating this expression is rather time consuming. Section 3.5 outlines a possible approximation that reduces the computation time.

3.3 The displaced diffusion case

In the displaced diffusion setting the functional form at the setting date is

$$L_{T_i}^i = -a_i + b_i \exp\left(d_i x_{T_i}^i - \frac{d_i^2 V_{T_i}^i}{2}\right). \tag{28}$$

To understand the structure in the difference from the case with $a_i = 0$ note that, using the notation

$$\delta_i := 1 - \alpha_i a_i \tag{29}$$

L_t^{n-2} is given by

$$
\begin{aligned}
L_t^{n-2} &= \frac{1}{\gamma_t^{n-1}} E^{Q^n}\left[L_{T_{n-2}}^{n-2}(1+\alpha_{n-1}L_{T_{n-1}}^{n-1})(1+\alpha_n L_{T_n}^n)\Big| x_t\right] = \\
&= -\frac{a_{n-2}\gamma_t^{n-1}}{\gamma_t^{n-1}} + \frac{1}{\gamma_t^{n-1}} E^{Q^n}\left[\hat{L}_{T_{n-2}}^{n-2}(\delta_{n-1}+\alpha_{n-1}\hat{L}_{T_{n-1}}^{n-1})(\delta_n+\alpha_n\hat{L}_{T_n}^n)\Big| x_t\right] \\
&= -a_{n-2} + \frac{1}{\gamma_t^{n-1}}\Bigg(\delta_{n-1}\delta_n \hat{L}_t^{n-2} \\
&+ \delta_n\alpha_{n-1}\hat{L}_t^{n-2}\hat{L}_t^{n-1}\exp\left(d_{n-1}d_{n-2}C_{n-2}^{n-1}(t,T_{n-2})\right) \\
&+ \delta_{n-1}\alpha_n\hat{L}_t^{n-2}\hat{L}_t^{n}\exp\left(d_n d_{n-2}C_{n-2}^{n}(t,T_{n-2})\right) \\
&+ \alpha_{n-1}\alpha_n\hat{L}_t^{n-2}\hat{L}_t^{n-1}\hat{L}_t^{n}\exp\Big(d_{n-1}d_{n-2}C_{n-2}^{n-1}(t,T_{n-2}) \\
&+ d_n d_{n-2}C_{n-2}^{n}(t,T_{n-2}) + d_n d_{n-1}C_{n-1}^{n}(t,T_{n-1})\Big)\Bigg)
\end{aligned}
\tag{30}
$$

Powered with this one may observe the structure in the expressions by studying the similarities with the $a_i = 0$ case above and perhaps also writing the expression out for the case L^{n-3}. To write the expression in a bit more condensed form let the vector $\bar{K}^{i,n}$ contain the integers between $i+1,\ldots,n$ that are not in the vector $K^{i,n}$. The general expression is given in the below proposition.

Proposition 3.2 *Given the choice of functional forms at the reset dates as in (28), the ith forward LIBOR at time t is given by*

$$
L_t^i = -a_i + \frac{\hat{L}_t^i}{\gamma_t^{i+1}}\left(\delta_{i+1}\cdot,\ldots,\cdot\delta_n + \sum_{k=1}^{n-i} g(L_t^{i+1},\ldots,L_t^n;k)\right) \tag{31}
$$

where

$$
g(L_t^{i+1},\ldots,L_t^n;k) = \sum_{j=1}^{\binom{n-i}{k}} \prod\left(\delta_{\bar{K}_j^{i,n}}\right)\prod\left(\alpha_{K_j^{i,n}}\right)\prod\left(\hat{L}_t^{K_j^{i,n}}\right)\exp\left(C_t^{K_j^{i,n}}\right), \tag{32}
$$

with

$$
\prod\left(\delta_{\bar{K}_j^{i,n}}\right) = \delta_{\bar{K}_j^{i,n}[1]}\cdot\ldots\cdot\delta_{\bar{K}_j^{i,n}[n-i-k]}, \tag{33}
$$

and the other entities as in proposition 3.1.

3.4 Calibration

When calibrating the parametric MFM it is possible to borrow most of the already developed calibration techniques for the (DD) LIBOR market model,

see for example Brigo & Mercurio (2006) for a wide variety of different approaches and possible choices of the instantaneous volatility functions σ_t^i and instantaneous correlations ρ^{ij}.

In principle, to calibrate the LMM, one first chooses the instantaneous volatilities as either constant, piecewise constant or some parametric time-dependent (preferably time to expiry dependent) function. These are then calibrated such that market prices of (ATM or some other strike) caplets are fitted by matching the relation (8). The instantaneous correlations ρ^{ij} are then fitted to match a view about terminal correlations and/or prices of swaptions and/or other option market prices through various approximations.

For terminal correlations in the LMM a standard approximation is based on freezing the drifts at time zero and then computing the correlations between the natural log of the (forward) LIBORs leading to

$$\begin{aligned} \operatorname{corr}\left[\log\left(L_{T_k}^i + a_i\right), \log\left(L_{T_k}^j + a_j\right)\right] &\approx \operatorname{corr}\left[x_{T_k}^i, x_{T_k}^j\right] \\ &= \frac{\int_0^{T_k} \sigma_t^i \sigma_t^j \rho^{ij} \mathrm{d}t}{\sqrt{\int_0^{T_k} \left(\sigma_t^i\right)^2 \mathrm{d}t \int_0^{T_k} \left(\sigma_t^j\right)^2 \mathrm{d}t}} \end{aligned} \tag{34}$$

For swaptions there exists a variety of approximations (Brigo & Mercurio (2006) covers several of them) and one of the simplest, most straightforward and popular choices is the Rebonato approximation. In this approximation the implied squared volatility multiplied with time to expiry for an option expiring at T_i to enter a β periods swap is given as

$$\left(\Sigma_{T_i,\beta}\right)^2 T_i = \sum_{k,l=i}^{i+\beta-1} \frac{w_0^k w_0^l L_0^k L_0^l}{y_0^{i,\beta}} \rho^{kl} \int_0^{T_i} \sigma_t^k \sigma_t^l \mathrm{d}t \tag{35}$$

where w_0^k is a function of the time zero discount bond prices and $y_0^{i,\beta}$ is the time zero forward swap rate for the swap under consideration. Note how this formula is linked to the above approximation of the terminal correlations by the integrals $\int_0^{T_i} \sigma_t^k \sigma_t^l \mathrm{d}t$ implying that there is a rather close link between swaption prices and terminal correlations/covariances in the LMM.

For the parametric MFM note that by construction

$$\operatorname{corr}\left[\log\left(L_{T_i}^i + \alpha_i\right), \log\left(L_{T_j}^j + \alpha_j\right)\right] = \operatorname{corr}\left[x_{T_i}^i, x_{T_j}^j\right]$$

and hence the above formula (34) is exact and not an approximation for the correlations between (logged) LIBORs at their setting dates in this model. For correlations between L^i and L^j at some time T_k the link is less direct as there are higher order terms to take into account as well. However, given the

expressions in propositions 3.1 and 3.2, it seems reasonable to believe that most of the variation of $L^i_{T_k}$ comes from $x^i_{T_k}$ and that the formula (34) may be used also for the parametric MFM. Moreover, given the similarities in the setup of the parametric MFM and the LMM it seems reasonable to use the swaption approximation formula (35) also for the parametric MFM. These choices are tested and justified in Section 4.2.

While calibration of the instantaneous volatilities and correlations (as well as the displaced diffusion coefficients a_i) may be performed in the same manner as for the LMM, the parameters b_i and d_i also need to be set for the parametric MFM. The constants d_i may be chosen in order to exactly fit caplet prices and a way to do this is outlined in section 4.1. Finally, given d_i, b_i is uniquely determined by the requirement that the model should hit the time zero forward LIBORs and may be found by putting $t = 0$ and inverting the expressions (23) or (31) such that L^i_0 equals the initial forward LIBOR values. Obviously, the b_i:s should be determined backwards since for given $b_j, j = i+1, \ldots, n$, b_i is linear in the equations (23) and (31).

3.5 Approximating $L^i_{T_k}$

Recall from propositions 3.1 and 3.2 that there are in total 2^{n-i} terms to sum up in the expression for L^i_t. In a practical implementation, when $n >> i$, computing all terms exactly would be too time-consuming and hence this expression needs to be approximated somehow. Luckily, in most situations, terms of order larger than 3-4 (i.e. terms involving the product of more than 3-4 LIBORs) are very small and the main contribution is coming from the first few terms of the sum. At this point note that the terms $\hat{L}^i_t$ is similar to the crude LMM one-step Euler drift approximation and hence the approximation $L^i_t \approx \hat{L}^i_t$ would imply a model of similar accuracy. Powered with this a first approximation could hence be to simply remove the contribution of terms of higher order than some maximum level M. While this seems to work O.K. a better method is outlined below which requires small extra computational costs.

Suppose we are interested in approximating L^i_t in the general displaced diffusion case and that L^j_t, $j > i$ are already determined. Recall that each of the terms in the sum producing $g(L^{i+1}_t, \ldots, L^n_t; k)$ consists of four terms; products of $1 - \alpha_i a_i$, accrual factors and forward LIBORs and the exponential of the sum of covariances. One way to approximate g is provided in the below four steps.

1. Compute the first M terms in (31) exactly.

2. Note that the terms

$$\begin{aligned}\prod\left(\delta_{\bar{K}_j^{i,n}}\right) &= \delta_{\bar{K}_j^{i,n}[1]}\cdot\ldots\cdot\delta_{\bar{K}_j^{i,n}[n-i-k]},\\ \prod\left(\alpha_{K_j^{i,n}}\right) &= \alpha_{K_j^{i,n}[1]}\cdot\ldots\cdot\alpha_{K_j^{i,n}[k]}\end{aligned}$$

are more or less the same for each j, the only difference being due to day counting conventions (note that typically the dates in the dates structure are roughly equally spaced). Hence, for each $k = M+1,\ldots,n-i$ take

$$\prod\left(\delta_{\bar{K}_j^{i,n}}\right) \approx \delta_{\bar{K}^{i,n}} := \left(\delta_{i+1}\cdot\ldots\cdot\delta_n\right)^{\frac{n-i-k}{n-i}}, \quad j=1,\ldots,\binom{n-i}{k} \tag{36}$$

$$\prod\left(\alpha_{K_j^{i,n}}\right) \approx \alpha_{K^{i,n}} := \left(\alpha_{i+1}\cdot\ldots\cdot\alpha_n\right)^{\frac{k}{n-i}}, \quad j=1,\ldots,\binom{n-i}{k}. \tag{37}$$

Note that this approximation has typically very good precision and all products may be pre-computed and stored in the local cache before performing the simulations.

3. Then, inspired by the above case, take for each $k = M+1,\ldots,n-i$

$$\prod\left(\hat{L}_t^{K_j^{i,n}}\right) \approx \hat{L}_t^{K^{i,n}} := \left(\hat{L}_t^{i+1}\cdot\ldots\cdot\hat{L}_t^{n}\right)^{\frac{k}{n-i}}, \quad j=1,\ldots,\binom{n-i}{k}. \tag{38}$$

which is obviously a much cruder approximation. Note, however, that the contributions from these terms are typically rather small if M is sufficiently high and since all forward LIBORs typically are strongly correlated (this depends of course on the choice of instantaneous correlations ρ^{ij}), there is hence hope that this is not a too bad of an approximation.

4. Note that in the previous step each of the products of forward LIBOR realisations are approximated to be the same for all j. As the sum of covariances are all deterministic all relevant sums may be pre-computed and stored in the local cache before performing the simulations. As there still can be quite many terms to sum up this may, depending on desired computational speed and accuracy, be done either exact or approximated. This step may hence be split in two parts, either

   (a) compute the sum of all exponentials of covariance terms exactly, or

   (b) first compute the sum of all relevant covariance terms

$$\bar{C}_t^{i,n} := \sum_{u=i+1}^{n} d_i d_u C_i^u\left(t,T_i\right) + \sum_{u=i+1}^{n}\sum_{v=u+1}^{n} d_u d_v C_u^v\left(t,T_u\right), \tag{39}$$

and then approximate

$$\bar{C}_t^{K_j^{i,n}} \approx \bar{C}_t^{K^{i,n}} := \bar{C}_t^{i,n} \frac{k(k+1)}{(n-i+2)(n-i+1)}, \quad j = 1, \ldots, \binom{n-i}{k}, \tag{40}$$

where the multiplication factor stems from that there are in total $\frac{k(k+1)}{2}$ terms in each of the $\bar{C}^{K_j^{i,n}}$ and $\frac{(n-i+2)(n-i+1)}{2}$ terms in $\bar{C}^{i,n}$.

To sum up, the above steps define two potential approximations of the k:th order terms

$$g(L_t^{i+1}, \ldots, L_t^n; k) = \delta_{\bar{K}^{i,n}} \alpha_{K^{i,n}} \hat{L}_t^{K^{i,n}} \cdot \sum_{j=1}^{\binom{n-i}{k}} \exp\left(\bar{C}^{K_j^{i,n}}\right), \tag{41}$$

$$g(L_t^{i+1}, \ldots, L_t^n; k) = \delta_{\bar{K}^{i,n}} \alpha_{K^{i,n}} \hat{L}_t^{K^{i,n}} \cdot \binom{n-i}{k} \exp\left(\bar{C}^{K_t^{i,n}}\right), \tag{42}$$

In tests of the above approximations on market like data it seems like an M between 0 to 4 is enough where the order depends on the difference $n - i$. In an actual implementation of the parametric MFM in practice we would recommend using an adaptive algorithm that starts approximating the product of the forward LIBORs when the previous order only implies a marginally (as defined by the user) change in the value of L_t^i. Note that it is also to some extent possible to use information of what was good for L_t^{i+1} when deciding at which order to start approximating. Finally, note that while the cruder of the above approximations (42) is tested on data in Section 4.2 and is found to behave reasonably well there are most probably better approximations to be found in a careful study.

Remark 3.1 *Note the similarities in intuition with an implementation of the LMM in practice. There the induced errors are due to the size of the time-step and the choice of drift approximation and there has, as mentioned previously, been a large number of papers introducing and comparing different drift approximations. For the parametric MFM there is no time-discretisation error, but instead there is an induced error by approximating the exact form for L_t^i.*

3.6 Relationships with other models

The LIBOR market model

Recall that by construction the parametric MFM is setup to be close to the LMM in the sense that it has the same structure as a crude one-step log-Euler approximation of the LMM. For further understanding of the connection

consider the Q^n dynamics of L_t^{n-1} under the parametric MFM model

$$\begin{aligned} \mathrm{d}L_t^{n-1} &= \hat{L}_t^{n-1} \frac{\alpha_n L_t^n}{1+\alpha_n L_t^n} \left(\exp\left(C_n^{n-1}(t,T_{n-1})\right) - 1\right) \sigma_t^n d_n \mathrm{d}W_t^n \\ &+ L^{n-1}\sigma_t^{n-1} d_{n-1} \mathrm{d}W_t^{n-1} + \ldots \mathrm{d}t \end{aligned} \tag{43}$$

Note that while the W^{n-1} diffusion parts of the models are the same the parametric MFM also depends on W_t^n, a feature that by assumption is restricted in the setup of the LMM. Also note that as t gets close to the setting date T_{n-1} the contribution from the W_t^n bit decreases as $C_n^{n-1}(t,T_{n-1})$ tends to 0. The finite variation terms are rather greasy and consists of a few terms that are not particularly instructive and we confine ourself to mention that it contains a term which is similar to the LMM drift term.

Intuitively one can view the parametric MFM and the LMM as models sharing the same foundation but with the LMM enforcing exact recovery of Black caplet prices (at the sacrifice of functional representation of the forward LIBORs in terms of x) and the parametric MFM enforcing functional representation of the forward LIBORs in terms of x (at the sacrifice of exact recovery of Black caplet prices). To get a feeling for the similarities and differences, Section 4 provides some numerical comparisons between the LMM and the parametric MFM model.

Markov-functional models

Consider a 'properly implemented' n-dimensional Markov-functional model, i.e. a model constructed using the construction technique briefly outlined in Section 2.4, fitted to caplet prices given by the displaced diffusion Black formula. While the parametric Markov-functional model will provide very similar functional forms of LIBORs at their setting dates (and hence prices) they will not be exactly the same as the simple exponential functional form of the parametric MFM is not flexible enough.

On the other hand, note that if a 'properly implemented' n-dimensional Markov-functional model is fitted to the caplet prices given by the parametric Markov-functional model they would be exactly the same as this will return exactly the exponential functional forms of the parametric MFM.

The numerical tests in Section 4 will provide some evidence for the closeness between the 'properly' implemented n-dimensional MFM and the parametric MFM by studying the resulting caplet prices.

The Hull-White model

Note that if one puts $a_i = \frac{1}{\alpha_i}$ and $d_i = 1, \forall i$ in the displaced diffusion version of the MFM type model then the ratios of discount factors for all i are

$$\begin{aligned}\frac{D_{tT_i}}{D_{tT_{n+1}}} &= \prod_{j=i}^{n}\left(1+\alpha_j L_t^j\right) = E^{Q^n}\left[\prod_{j=i}^{n}\left(1+\alpha_j L_{T_j}^j\right)\middle|\mathcal{F}_t\right] \\ &= E^{Q^n}\left[\prod_{j=i}^{n}\left(\hat{L}_{T_j}^j\right)\middle|\mathcal{F}_t\right] = \prod_{j=i}^{n} b_j \exp\left(x_t^j - \frac{V_t^j}{2}\right). \end{aligned} \tag{44}$$

Inverting this expression subsequently and then evaluating it at $t = 0$ gives that $b_i = \frac{1}{\alpha_i} + L_0^i$ and

$$L_t^i = -\frac{1}{\alpha_i} + \left(\frac{1}{\alpha_i} + L_0^i\right)\exp\left(x_t^i - \frac{V_t^i}{2}\right), \quad i = 1, \ldots, n. \tag{45}$$

Compared with the general expression given in proposition 3.2 this expression is of course much easier and faster to work with as the expression for L_t^i does not contain any other forward LIBORs. It is interesting to note that also for the DD-LMM, taking $a_i = \frac{1}{\alpha_i}$ for all i, gives exactly the same expression for the L_t^i:s as then the drift term disappears. Hence, in this particular case the MFM and the LMM are exactly the same.

The particular simple product form of the ratio of discount factors seems like something that should have been introduced previously in the literature. In fact this is effectively the n-dimensional Hull-White model for which it can be shown (see Hunt & Kennedy (2000) for the details in the one-dimensional case) that under the terminal forward measure the ratios of discount factors are given as

$$\frac{D_{tT_i}}{D_{tT_{n+1}}} = \frac{D_{0T_i}}{D_{0T_{n+1}}}\prod_{j=i}^{n}\exp\left(\int_0^t \hat{\sigma}_s^i \mathrm{d}W_s^i - \int_0^t \frac{\left(\hat{\sigma}_s^i \mathrm{d}s\right)^2}{2}\right) \tag{46}$$

for some function $\hat{\sigma}_t^i$ that depends on the (time-dependent) parameters of the Hull-White model.

3.7 Rank reduction

As mentioned in the introduction to the LMM the rank may be reduced by letting the vector x_t be driven by an m-dimensional instead of n-dimensional process. However, also in this case the dimension of the model would still be n

and to reduce the dimension a separable volatility structure along the lines of (10) is needed. As is the case for the LMM from a computational perspective this will only reduce computational speed marginally unless $m \leq 2$ for which the model may be implemented on a lattice or solved by PDE methods.

This paper will only focus on the full rank case and investigations of the rank reduced models are left to future research.

4 Implementation, calibration and numerics

This section studies the pricing of caplets and swaptions as well as the terminal correlations between the (forward) LIBORs. First we will study the modeling of LIBORs at their setting dates, something which the parametric MFM is well suited at. Recall that by construction of the model

$$\operatorname{corr}\left[\log L^i_{T_i}, \log L^j_{T_j}\right] = \operatorname{corr}\left[x^i_{T_i}, x^j_{T_j}\right] = \rho^{ij} \tag{47}$$

and hence there is perfect control of these correlations in the model. Recall that the motivation for the choice of the parametric form was that it should be close to a crude drift approximation of the LMM and hence should give prices that are close to the ones for the LMM. In particular we want the caplet prices to be close to the caplet prices obtained using the displaced diffusion Black formula under the appropriate forward measures. Clearly, caplet prices depend on the constants d_i but, even if these constants are chosen to match up the ATM volatility, it is not clear how caplets are priced across strikes (i.e. how close distributionally the parametric MFM is from the displaced diffusion Black formula). Calibrating the d_i:s to match the ATM volatility and testing caplet pricing are done in the first subsection in this section.

After this the parametric MFM is investigated in terms of terminal correlations between (forward) LIBORs and swaption pricing. The start point for these investigations are that the parametric MFM is so close to the LMM that all standard LMM calibration techniques and intuition can be transferred. Hence, by matching terminal correlations as well as the pricing of caplets and/or swaptions to the LMM the idea is that then also prices of exotics will be similar in the two models.

In the spot measure Kaisajuntti & Kennedy (2011) introduces a similar model but takes a different route in terms of testing the model. Instead of studying correlations and the pricing of simple products Kaisajuntti & Kennedy (2011) investigates pricing of the exotic products TARNs and CMS spread TARNs and compares with the LIBOR market model. Conclusions are that if terminal correlations and volatilities are matched, then also prices of exotics are matched in the two model frameworks. While these products would be suitable also for

the parametric MFM in the terminal measure we believe that there is not much extra insights to gain by pricing these products also in this paper and hence put focus on the 'basic' quantities instead.

Throughout this section only the case with all displacement factors $a_i = 0$ are considered. The reason for this is that positive displacement constants (negative would not be interesting as market skews are 'always' declining in strike) reduce the variation in the models coming from the drift term in the LMM and the higher order terms in the MFM and hence this case is the 'hardest' to deal with. For some numerical checks in a similar model using a positive displacement factor, see Kaisajuntti & Kennedy (2011).

4.1 Caplet pricing

The purpose of this section is to provide intuition about the parametric MFM, to display a way to fit the model to ATM caplet prices and to numerically compare the prices across strikes with the Black formula. Note that even though for the LMM caplet prices across strikes should theoretically be in perfect agreement with the Black formula with constant 'volatility' η this would not be the case in a practical implementation of the model. The reason is the induced discretisation error when implementing the model. As mentioned previously there is a variety of drift approximations introduced to reduce this error and throughout this section a single-step drift approximation of the LMM using the Glasserman-Zhao (GZ) and the predictor-corrector (PC) methods will be used as a comparison.

The reason for choosing to use a single step is that this leads to a model which is similar in spirit to the Markov-functional model in the sense that L_t^i is then a functional of x_t and does not depend on values of x_s for $s < t$. Also, from a computational perspective the LMM needs to be simulated using a single step in order to compete at all with the parametric MFM in terms of speed. Moreover, note that while in the parametric Markov-functional model L_t^i is a function of only x_t^i, in the single-step drift approximation using the PC and GZ methods L_t^i is a function of the whole vector $(x_t^i, \ldots, x_t^n)$, although x_t^i provides by far most of the variation.

Joshi & Stacey (2006) compares several of the introduced LMM drift approximation methods when using a single time step out the final expiry. The two chosen methods are due to their popularity and the fact that the GZ method has no bond pricing error (as is the case for the parametric MFM) and that the PC is one of the fastest and most flexible of the available methods and still, according to Joshi & Stacey (2006) provides among the best in terms of accuracy. The results in this section are based on the same inputs as in Joshi & Stacey (2006) i.e. flat rates at 5%, flat volatilities at 20% and instantaneous

correlation $\rho_{ij} = \exp(-\beta|T_i - T_j|)$, with $\beta = 0.05$ for a set of forward LIBOR rates with one year tenors out to 20 years.

Recall that the parameters b_i are used to perfectly calibrate the MFM to bond prices and that the parameters d_i may be chosen to fit the prices of ATM caplets. Before displaying how this may be done it will be informative to start by studying the case when all $d_i = 1$.

Putting all $d_i = 1$

Figure 2.1 displays the prices of ATM caplets converted into implied volatilities using 50 million (antithetic) pseudorandom paths implying that the size of a 95% confidence intervals of the caplet prices are all less than 0.1 bp. Note that both the GZ and the PC methods for the LMM gives errors in the caplet prices. These errors are consistent with the errors reported in Joshi & Stacey (2006) although they display caplet prices instead of corresponding implied volatilities. Also note that in contrast with Joshi & Stacey (2006) we do not get any bias with the GZ method for short caplet expiries.

In terms of computational time the pricing of caplets is very fast in the parametric MFM. In Joshi & Stacey (2006) the PC method is the fastest available method and the GZ one of the slowest (about 5 times slower). Generating a set of $L^i_{T_i}$, for $i = 1, \ldots, 20$ is in our implementation about 40 times faster in the parametric MFM compared with the PC method. The reason for this is that given a value of $x^i_{T_i}$ the parametric MFM only needs one call to the exponential function to evaluate $L^i_{T_i}$ whereas the PC LMM first needs to recall $x^i_{T_i}$, $L^j_0, j = i+1, \ldots, n$, the (previously computed) $L^j_{T_i}, j = i+1, \ldots, n$, the integral of the covariance functions up to time T_i, evaluate the sum in the drift terms twice and finally input it to the exponential function.

Note that with all $d_i = 1$ the parametric MFM caplet prices are larger than the 'desired' 20%, with the largest bias at the 13 year expiry. This difference is a bit too large to be of practical use and hence the model needs to be calibrated by changing the d_is. This is the task of the next subsection.

Calibrating d_i

As noted in the previous subsection the constants d_i needs to be modified in order to perfectly fit the model to ATM caplet prices. Ideally one would like to have an analytical (or a good approximate) expression for caplet prices, as is the case in the LMM (by switching measures). Switching measures in the MFM model would not help with this as the expression will then (still) include a drift term. While this drift term perhaps could be approximated with decent

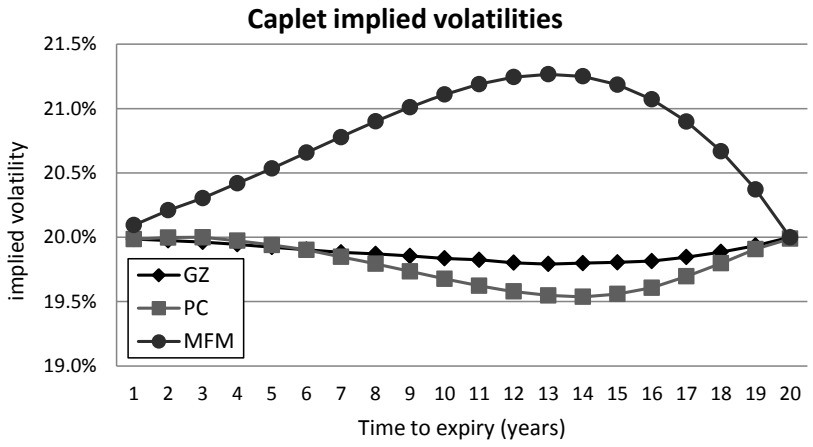

Figure 2.1: ATM caplet prices in terms of implied volatilities for the parametric MFM with all $d_i = 1$ as well the single-stepped LMM using the PC and GZ drift approximations.

precision we have not found any good generic way to do this such that it would be of practical use.

Instead we propose a numerical way of calibrating the model to caplet prices. Since caplet pricing is fast and efficient in the parametric MFM one could perform a calibration using the Monte Carlo method in reasonable computation time. To speed up the convergence of the Monte Carlo method we have chosen to use FRAs as control variates. As FRAs are highly correlated with caplets this work very efficiently as is motivated in Figure 2.2. This figure displays the differences between the model and the 'market' caplet prices (ie the Black model using a volatility of 20%). Note the huge improvement in convergence by using the FRAs as control variates. To achieve the same type of error using standard Monte Carlo one needs to use about 60 times more realisations.

By the use of FRAs as controls variates one may hence calibrate the d_i to ATM caplets in a reasonably fast way by the following simple algorithm:

1. Set $d_i = 1$ for all i.
2. Set b_i for all i.
3. For each i compute ATM caplet prices and convert into implied volatility.
4. Put $d_i = d_i \frac{\Sigma^i_{target}}{\Sigma^i_{model}}$, where Σ^i_{target} and Σ^i_{model} are the target and model

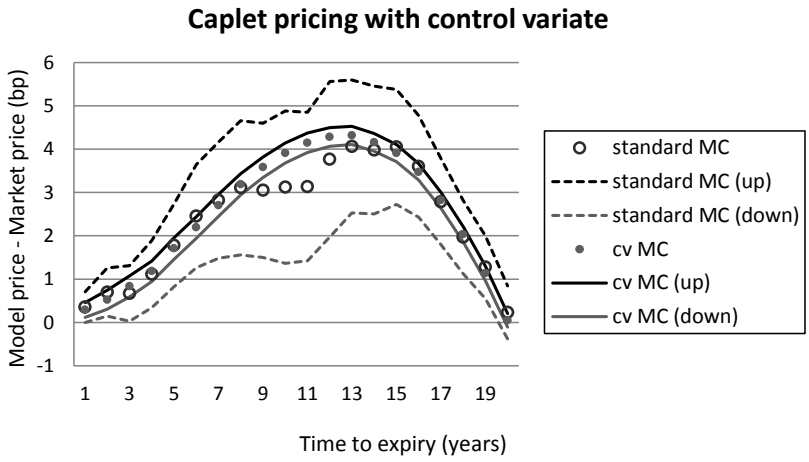

Figure 2.2: ATM caplet pricing with the FRA as control variate in the parametric MFM.

implied volatilities.

5. Reset b_i.

6. Repeat point 2 to 4 a few times. Typically two times are enough.

Figure 2.3 displays the implied volatilities after round 0, 1 and 2 using 10 million paths. Note that 10 million paths are a lot more than needed in practice but are chosen to make sure that the statistical error of the Monte Carlo method is redundant. Figure 2.4 displays the differences between market and model caplets and FRAs using 50 million paths. Note that the accuracy of the calibration algorithm is of the same order as the statistical FRA and caplet pricing errors using 50 million implying that the simple algorithm described above does a good job. Also note how correlated the FRA and caplet pricing (without control variates) errors are.

Remark 4.1 *Calibration using the Monte Carlo method would seem a bit inefficient. Note however that due to that Monte Carlo caplet pricing using control variates is very efficient this is not particularly prohibitive. Suppose for example that the model would be used for pricing a standard LIBOR TARN and that m realisations are needed for good accuracy. Then, three rounds with the above algorithm using, say, $\frac{m}{10}$ realisations (which is perfectly fine if control variates are used) takes shorter time than generating the paths for pricing the TARN and hence the total computation time would still be less than for single-stepped LMM using for example the PC drift approximation.*

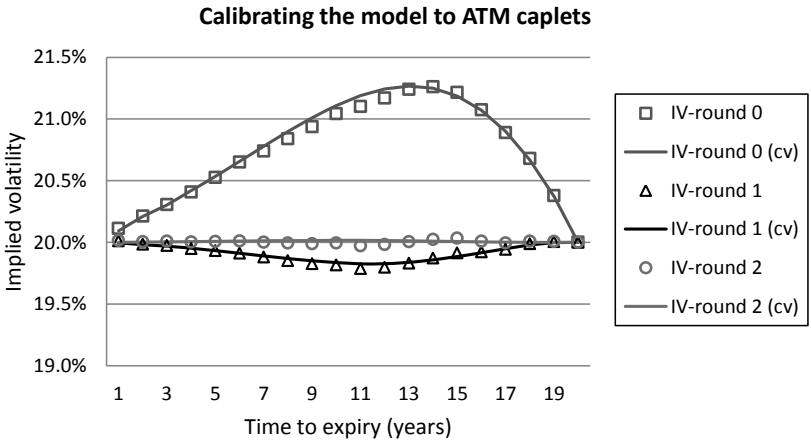

Figure 2.3: Calibration of the parametric MFM to ATM caplet prices using the simple algorithm described in the text. Round 0 refers to caplet prices using all $d_i = 1$ and round 1 and round 2 refers to caplet prices using the adjusted d_i after one and two rounds of adjustment respectively.

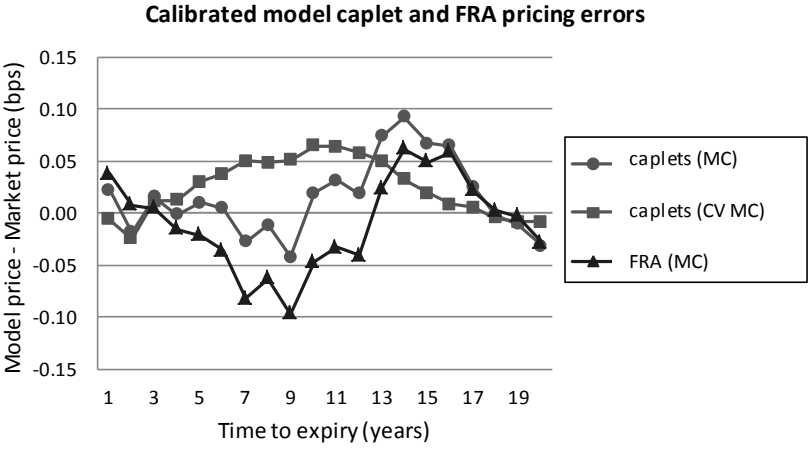

Figure 2.4: Results from the calibration exercise. Note that two rounds of adjustments are enough to get prices in line with the accuracy from using 50 million antithetic paths. Also note how correlated the caplet pricing errors without control variates and the FRA pricing errors are.

Remark 4.2 *Note that in a typical LMM setup the instantaneous volatilities are modeled by some time to expiry dependent function $g(T_i - t)$. To get a perfect fit to the ATM volatility it is then common to introduce multiplying constants k_i. The d_i used in the parametric MFM could be seen as having the same goal and effect as these constants.*

Non ATM strikes

The previous subsection displayed how to calibrate the parametric MFM to ATM (or indeed any particular strike of interest at each T_i) caplets. Since the model is not constructed to be perfectly consistent with the Black model a natural question to ask is far off is it? To get something to compare with and to put the results into context the GZ and PC 'single-step' drift approximation methods will once again be used as references.

Figure 2.5 displays caplet prices converted into implied volatility for the parametric MFM for some different strikes and figures 2.6 and 2.7 display the same thing for the PC and GZ drift approximations. First note that the GZ drift approximation has quite distorted marginal distributions with caplet prices that are scattered around the correct prices (in effect producing a skew in the implied volatility smiles). The PC has less scattered prices but has a negative bias on average. The calibrated parametric MFM on the other hand has both fairly small scatter as well close to correct ATM implied volatilities. Hence, compared with the GZ and the PC 'single-step' drift approximations it seems fair to say that the parametric MFM is the most consistent with the Black model. Needless to say, shorter steps with the GZ and the PC methods will of course improve the performance. However, since the parametric MFM is already at this point much faster, the LMM drift approximations will then be quite significantly slower.

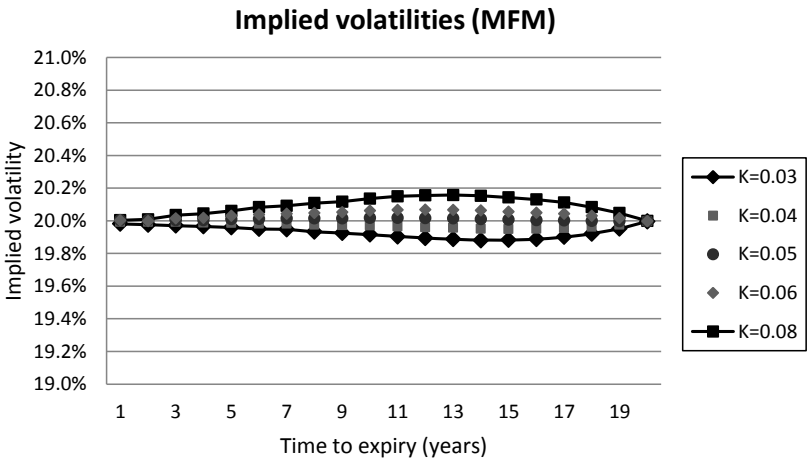

Figure 2.5: Caplet prices in terms of implied volatility across strikes for the calibrated parametric MFM. Note the small systematic difference in implied volatility between the different strikes.

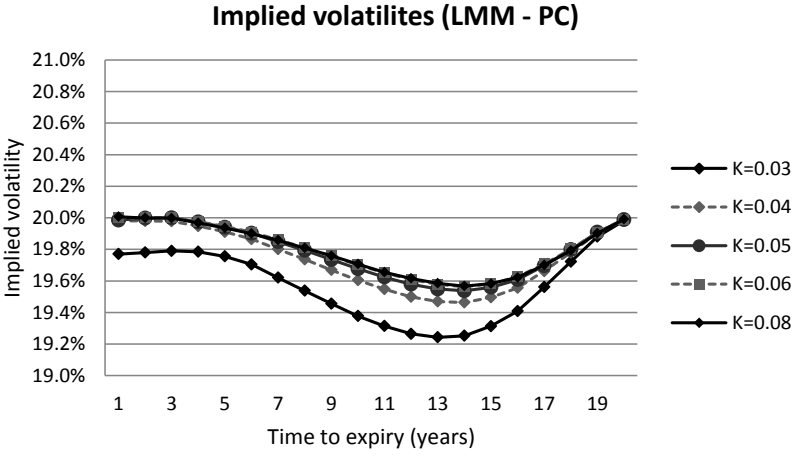

Figure 2.6: Caplet prices in terms of implied volatility across strikes for the calibrated single-stepped LMM using the PC drift approximation. Note that this method misprices ATM caplet prices by about 50 implied volatility basis points for the 13 years to expiry caplet. There is also a systematic difference in implied volatility between the different strikes of similar size as in the parametric MFM.

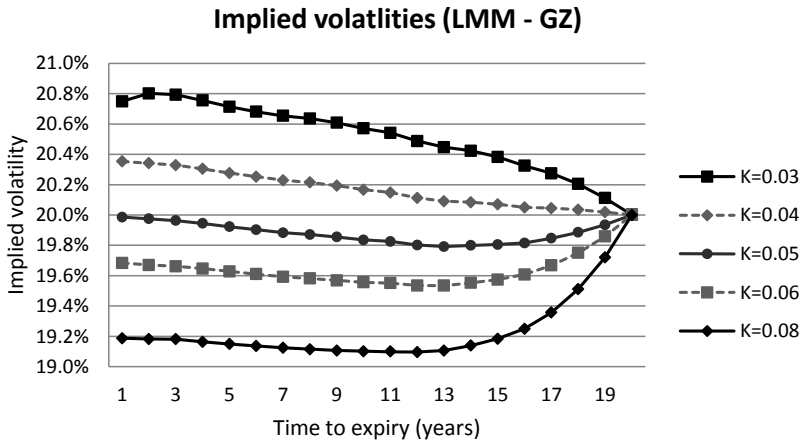

Figure 2.7: Caplet prices in terms of implied volatility across strikes for the calibrated single-stepped LMM using the GZ drift approximation. Note the big systematic difference difference in implied volatility between the different strikes.

4.2 DF, TC and swaptions

This section studies LIBORs before their setting dates. Recall that while the parametric MFM is an efficient choice for products depending on LIBORs at their setting dates, computing LIBORs before setting could be rather inefficient if the full expression (recall propositions 3.1 and 3.2) for the L_t^i:s (in particular when $i << n$) would be used. As an attempt to speed up the computations Section 3.5 argued that these expressions could be computed by using the exact values up to some order M and then approximating the higher order terms. Hence, as opposed to when pricing caplets, in this case the values for both the parametric MFM and the LMM are affected by both statistical and approximation (discretisation) errors. To keep the statistical error low 5 million antithetic paths are used for both models in all computations in this section.

As test products the prices of (forward) discount bonds and swaptions as well as terminal correlations will be computed and compared with either the true values or the approximation formulae, (34) and (35). In order for these results to be comparable with the previous tests the same data as in Section 4.1 are used.

All studied results are dependent on the forward discount curve at $T_{10} = 10$ years and for the parametric MFM the orders $M = 3, 4$ and 5 are used whereas for the LMM either 1 or 8 PC steps per year are used. Although only the results are displayed at $T = 10$ years we were computing the full discount curves at all dates in the date structure and this case is displayed due to that it is one of the hardest to approximate (in both models) and is hence displaying the largest errors. The computational times in our implementation are about the same for $M = 4$ with the parametric MFM and one step per year in the LMM.

The first test is based on the martingale property

$$E^{Q^n}\left[\frac{D_{T_kT_i}}{D_{T_kT_n}}\right] = \frac{D_{0T_i}}{D_{0T_n}}, \quad i = k, \ldots, n. \tag{48}$$

Note that this relation is important to fulfill in order to not induce errors in bond and FRA prices. For both the parametric MFM and the LMM Figure 2.8 displays the ratios of the left and right hand sides of (48) for $T_k = T_{10} = 10$ years under all of the above mentioned cases. To get something to compare with, the figure also displays the boundaries of a 95% confidence interval for the statistical error (centered around the value 1) computed using 1 million paths in the LMM with one PC step per year. The statistical errors using 5 million paths are rather small (but sizeable in the test) and a 95% confidence interval is given by about plus/minus 0.0002-0.0005. Also note that for the MFM there is no approximation error and only statistical error at the first two points and the values are hence the same irrespective of the order M.

First note that one year long PC steps in the LMM performs by far the worst and is outside the one million paths confidence interval, implying that the discretisation error is, with more than 95% probability, larger than the statistical error achieved using one million paths. In fact, the discretisation error is of about the same size as the limits of a 95% confidence interval obtained using about 100 000 paths. While this might be an O.K. accuracy in a practical application note that the MFM using $M = 4$ or 5 are both much better and that also $M = 3$ performs better. Using $M > 5$ gives very small differences compared with the $M = 5$ case. Also note that using eigth steps per year in the LMM effectively reduced the disretisation error.

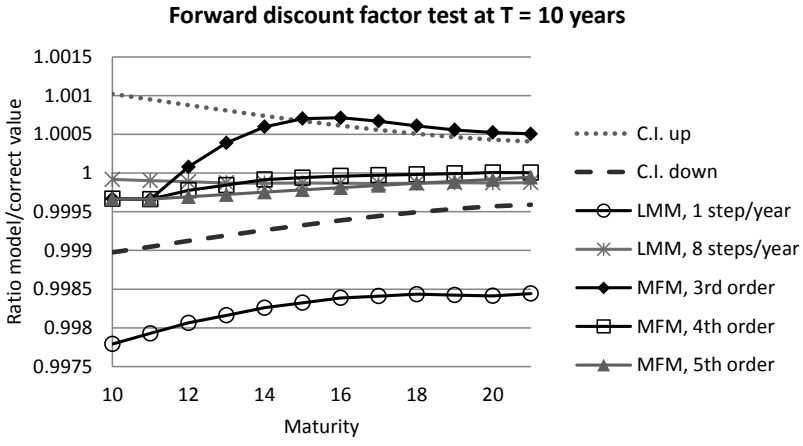

Figure 2.8: Ratios of the left and right hand sides of (48) for $T_k = T_{10} = 10$ years. The figure displays the ratios using the LMM with both one and eight PC steps per year as well as the parametric MFM using the correct orders up to $M = 3, 4$ and 5. The figure also displays the boundaries of a 95% confidence interval (centered around the value 1) computed using 1 million paths in the LMM with one PC step per year.

The next performed test treats the accuracy of the terminal correlation formula (34) proposed to be used for calibration or check-ups of both models. First recall that for the parametric MFM this formula is exact for LIBORs at their setting dates. For the LMM the errors when using the formula for LIBORs at their setting dates are given in Figure 2.9. Note that while the errors do display a somewhat worrying pattern they are quite small and likely smaller than the uncertainty in a traders estimate of desired terminal correlations. Note however that increasing the volatility or time to expiry will increase these errors.

Figure 2.10 displays the terminal correlation errors between different forward

LIBORs at $T_{10} = 10$ years for both the parametric MFM and the LMM. Note that also these are quite well approximated although the errors for the LMM is in general about twice as large as for the parametric MFM. Also note that the approximation errors have a different sign in the MFM versus the LMM.

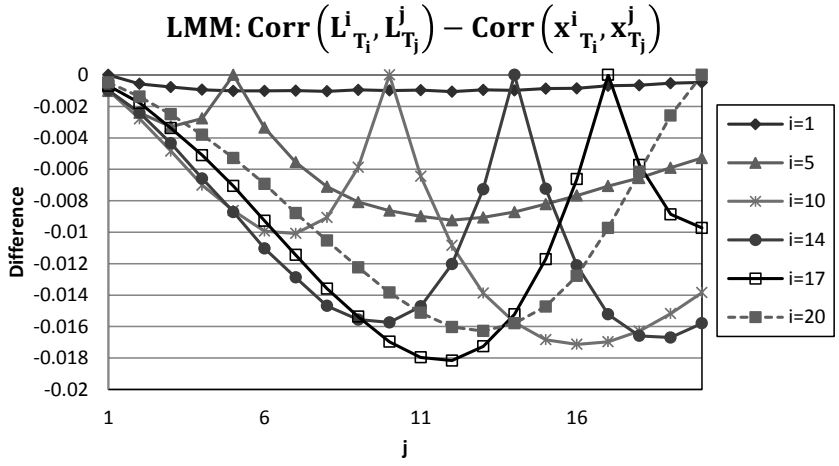

Figure 2.9: Correlations between the LIBORs at their setting dates for the LMM.

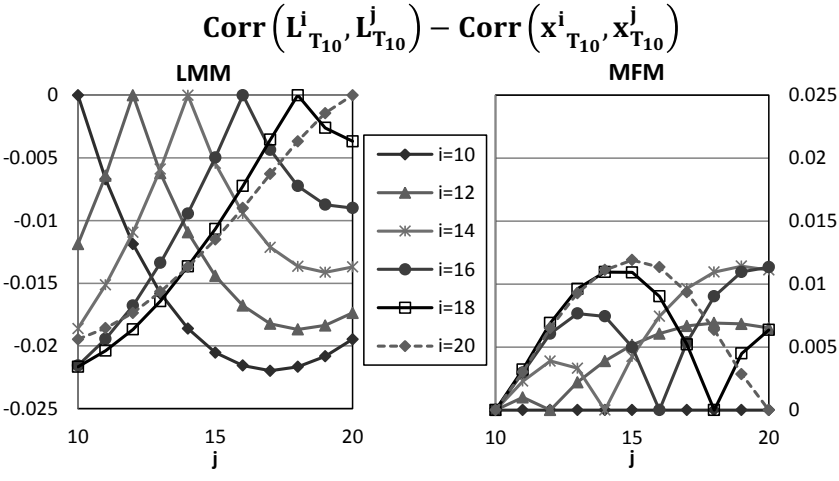

Figure 2.10: Correlations between the forward LIBORs at some dates T_j.

Prices of swaptions (in terms of implied volatilities) expiring at T_{10} with tenors

of one up to eleven years are reported in Figures 2.11 and 2.12. Figure 2.11 displays the prices for the parametric MFM with $M = 4$, the LMM with one year long PC steps as well as the swaption approximation (35) equipped with the 95% bounds for an LMM simulation with 1 million paths. Note that while the LMM prices are outside the confidence interval the parametric MFM is almost within the confidence interval in all cases. Figure 2.12 displays the differences between the model prices and the swaption approximation. From this one may conclude that while the parametric MFM seems to be rather close to the swaption approximation the model does seem to give a slightly different behaviour in terms of time to expiry. For the LMM note that whereas the correctly computed prices with eight steps per year are very close to the approximation the one year steps case gives a consistent bias at all expiries.

The intuition provided in these tests points at that calibrating the parametric MFM using the machinery developed for the LMM seems to work reasonably well. For the swaption approximation there might be a need for a slightly better approximation than provided by the Rebonato LMM swaption approximation. While this and further tests are left for future research note that a start of a better swaption approximation could be to write down the full dynamics of the generating SDE in the parametric MFM case and then follow the same arguments leading to the approximation in the LMM case.

4.3 Summing up

The results of this sections points at that compared with a drift-approximated LMM over long time-steps the parametric MFM is an appealing alternative as it provides small errors in bond and caplet prices (across strikes) and may be calibrated to terminal correlations and/or swaptions with good accuracy using the already developed calibration approximations of the LMM. This supports the view of Bennett & Kennedy (2005) and Kaisajuntti & Kennedy (2011) that MFM and LMM are closely connected.

In terms of exotic instruments this means that if both the parametric MFM and the LMM are calibrated using the outlined approximations then it seems reasonable to believe that prices should be quite close. Recall, however, that the errors of calibrating the terminal correlations seem to have different signs and the behaviour of the swaption implied volatilities across expiry are slightly different in the two models. For some types of products, and in particular for larger volatility levels or long maturities, one could expect these 'calibration errors' to have a sizeable effect on prices. For some intuition about this note that Kaisajuntti & Kennedy (2011) compares prices of LIBOR and CMS spread TARNs using the same type of models (albeit under the spot measure) and argues that if terminal correlations are matched up (exactly, not approx-

imately via some approximation formula) then prices of LIBOR TARNs are very similar.

Finally, note that we have performed some tests also on other term structures of forward LIBORs and volatilities such as different levels as well as upward and downward sloping market like term structures. The results display similar patterns as the above ones, albeit scaled roughly with the level of volatility and forward LIBOR values.

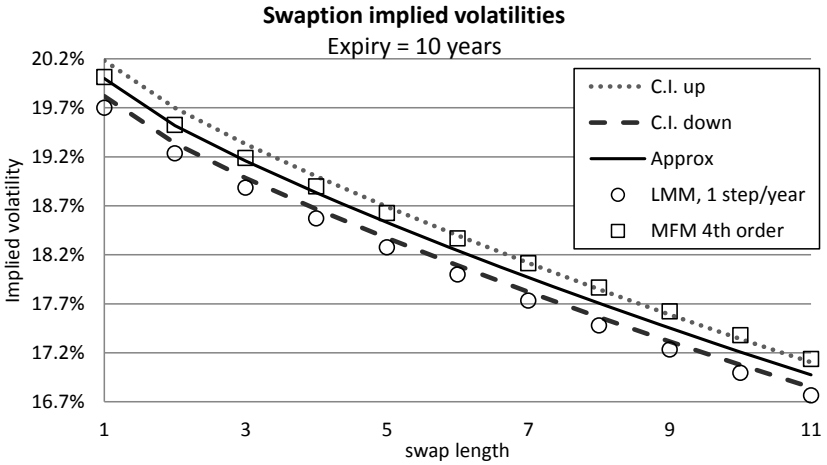

Figure 2.11: ATM swaption prices in terms of implied volatility (expiry 10 years and tenors of one up to eleven years) for the LMM, the parametric MFM and the Rebonato swaption approximation.

5 Conclusions

This paper has developed and tested an n-dimensional Markov-functional interest rate model in the terminal measure based on parametric functional forms of exponential type. The parametric functional forms enable analytical expressions for forward discount bonds and forward LIBORs at all times and allows for (approximate) calibration of the model to caplet prices given by a displaced diffusion Black model. By studying the prices of caplets across strikes it is argued that the parametric functional forms are very close to the forms obtained by a standard non-parametric Markov-functional model calibrated to caplet prices given by a displaced diffusion Black model. Hence, the analytical expressions of the model provides a theoretical tool for understanding the structure of standard Markov-functional models.

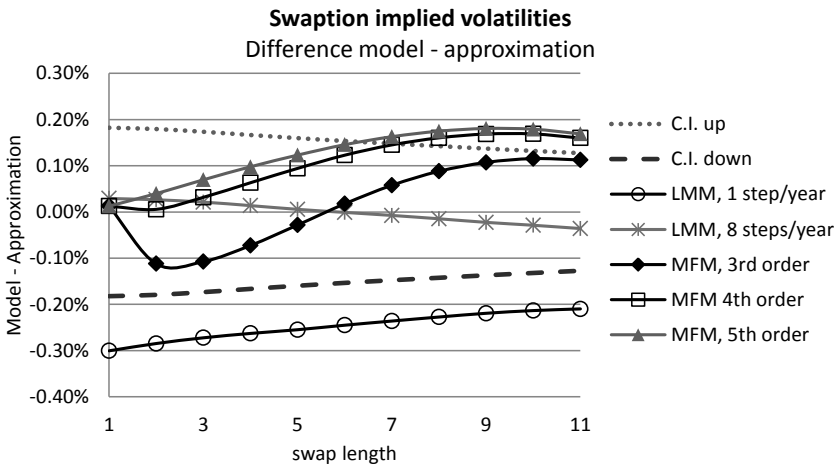

Figure 2.12: The difference between ATM Swaption prices (expiry 10 years and tenors of one up to eleven years) in terms of implied volatility for various approximations of the parametric MFM and the LMM and the Rebonato swaption approximation given in refswpnapprox.

The setup of the model suggests it to be similar to an n-factor LIBOR market model. This is supported numerically by studying the prices of caplets and swaptions and the terminal correlations between the (forward) LIBORs. In particular it is shown that for 'typical' market data the models are close enough to be able to use the machinery developed for calibration and understanding of the LMM as a drop in replacement in the parametric MFM.

The similarities between Markov-functional and LIBOR market models are also reported in Bennett & Kennedy (2005) and Kaisajuntti & Kennedy (2011). In particular the latter paper, that is similar in structure to this paper, the models are compared in the spot measure and with the main focus on path wise comparisons of rates at their setting dates and prices of exotic products such as (LIBOR and CMS spread) TARNs. In addition to that paper this paper also explicitly compares the models at times before settings by studying prices of swaptions and terminal correlations. Taken together this provides further information about the similarities (and differences) of Markov-functional and LIBOR market models.

The parametric n-dimensional Markov-functional model may be used for products that require high-dimensional models for appropriate pricing and risk management. Compared with an n-factor LIBOR market model it has the virtue of being (much) faster for certain types of products.

Future research may focus on using a more general and flexible form than the exponential function for LIBORs at their setting dates as well as extending the driving process to allow for jumps or stochastic volatility. Moreover, as the parametric MFM model has explicit forms for LIBORs at all dates and is close to both the standard MFM and LMM it might be used to derive further links between the models as well as approximation formulae for correlations and swaptions in the standard MFM setting. In particular it could be interesting to compare the rank reduced model with the model in Pietersz et al. (2004) as well as properly implemented one- and two-dimensional LMM and Markov-functional models.

Paper 3

Paper 3

Stochastic Volatility for Interest Rate Derivatives

Joint work with Joanne Kennedy.

1 Introduction

The objective of this paper is to perform a data-driven investigation in order to find a good model for pricing interest rate derivatives. This is, of course, a hugely complex problem, in general, requiring an understanding of the number of factors needed in the model and which copula is most appropriate, amongst other things.

Given the reality of the data available in practice (lots of asset/swap prices that could inform us about the infinitessimal behaviour of rates, but relatively few prices that would inform us about the distribution of rates at discrete times in the future) and given the non-stationary properties of the market in general, we do not attempt such an ambitious task. Instead we consider only what market data tells us about the overall level of rates (its joint distribution at a set of future times). Thus an output of our analysis is not a model suitable for pricing Bermudan-callable CMS-spread options. But it could be considered adequate for pricing simpler level-dependent derivatives, the canonical example of which is the standard Bermudan swaption.

To this day many banks continue to price level-dependent derivatives using a one-factor model. However it is clear that such a model does not adequately capture the joint distribution of the level of rates. Indeed Hagan et al. (2002) showed that in order to adequately capture the distribution of a single swap rate at its setting date an extra factor is required. Our objective in this paper is to formulate a model for the level of rates that will improve on this situation, without creating a model with unnecessary complexity (too many factors). It turns out that we can do this with just one more factor which represents stochastic volatility.

Given our objective of modelling only the level of rates, we proceed by first choosing a rate to represent the overall level of rates and investigate how to model it. We later change this choice to ensure that the results of our analysis are not dependent on this somewhat arbitrary initial selection. Our initial investigations were made for a model for 10 year swap rates of start dates 2 to 30 years. Start dates shorter than this may require a different class of models to capture the behaviour appropriately. By looking at 10 year tenors we were purposely choosing rates that had a large overlap so the models for different rates could be viewed as a block. From this we would expect parameters driving the model to be comparable between rates and one would have good hope of finding an appropriate common driver for all rates. Having fitted a model with common parameters that captured the (macro distributions of the) market well we took the ten year spot rate as our proxy when reading off a suitable model for the level of rates. Later checks using 20 and 30 year tenors show the results to be qualitatively similar.

Let y^i denote the swap rate corresponding to start date T_i in the future and having a 10 year tenor. In seeking a suitable model for the 10 year swap rates themselves we worked within the class of stochastic volatility models specified by a system of SDEs of the form

$$\begin{aligned} \mathrm{d}y_t^i &= f^i(T_i - t, y_t^i, \sigma_t^i)\mathrm{d}W_t^i, & (1)\\ \mathrm{d}\sigma_t^i &= h^i(\sigma_t^i)\mathrm{d}t + g(\sigma_t^i)\mathrm{d}V_t^i, & (2) \end{aligned}$$

under $\mathbb{S}^i$ corresponding to $P_\cdot^i$ (the PVBP corresponding to y^i) as numeraire and where W^i, V^i are Brownian motions with

$$\mathrm{d}W_t^i dW_t^j = \mathrm{d}V_t^i \mathrm{d}V_t^j = \mathrm{d}t, \quad \mathrm{d}W_t^i \mathrm{d}V_t^i = \rho \mathrm{d}t, \quad j \neq i.$$

Theoretically it is not necessary to restrict to this class in our search for a model but we found this class to be adequate for our purpose at each stage of the modelling process. In our investigations we put aside considerations of tractability and let the data dictate the precise final form of the model as we increased the modelling demands made of it. The judgement of whether the model was successful at any stage was based on how well it could represent information on the macro distributions contained in the data. Given the quality of the data available from the market we believe this was the most sensible yardstick to judge a model by.

The first part in our search for an appropriate building block is to find a good model for the $\mathbb{S}^i$ marginal distribution

$$\left(y_{T_i}^i, \sigma_{T_i}^i | y_0^i, \sigma_0^i\right) \tag{3}$$

for each $i = 1, \ldots, N$. Note that fitting the implied volatility smile for swaptions with time to expiry T_i implies fitting the $\mathbb{S}^i$ marginal distribution

$$\left(y_{T_i}^i, \sigma_{T_i}^i | y_0^i = y, \sigma_0^i = \sigma\right). \tag{4}$$

This may be done with good accuracy for a large variety of models in the class of stochastic volatility models. However, models agreeing on (4) do not necessarily agree on (3) (where we are assuming the parameters of the model remain fixed). To be able to analyse this we assume that market data is generated by a model such that the conditional distribution is roughly stable over time. This enables us to use observations of swaption prices (implied volatility smiles) with time *to* expiry T_i at several trading dates to inform about (3).

A model assigning an appropriate distribution (3) provides a '**smile dynamic**' that is similar to the market one and hence calibrated parameters would be nearly stationary across dates. Section 4 provides an investigation of the SABR model with different β as well as the Heston model in order to inform about what the different specifications of the SDE have to offer in terms of reflecting

the available information on the market one step conditional distributions. We found that the SABR model with $\beta = 0$ gave very good results so we did not need to extend our search beyond these two popular models. We note in passing that the results of this section are of independent interest in terms of offering a data driven analysis of SABR versus Heston.

What we have identified so far is a model that gives a good fit to the distribution of the swap rate at its setting date with each start date treated separately. Having a model which gives a good representation of the conditional distribution at T_i does not tell us anything about whether the model gives a good representation of the (conditional) '**forward smiles**' i.e. the distribution of $\left(y^i_{T_i}, \sigma^i_{T_i} | y^i_t, \sigma^i_t\right)$, $t \in (0, T_i)$. Note that a stochastic volatility model for y^i does specify all the conditional distributions and hence does imply both a certain smile dynamic and forward smile. In particular the stochastic volatility models studied in this first stage are time homogeneous whereas we expect the distribution of y^i_t to depend on time to maturity $T_i - t$. Never the less we use the SDE identified at this stage, the SABR model, as our starting point for the next step and address the issue of appropriate time dependence of the model when fitting to data at several expiries requires it.

Our next objective in our search for a stochastic volatility model for the level of rates is to seek a model of the swap rates (in their own swaption measure) based on only two factors and that enables a specification of parameters that is *common* to all forward swap rates having ten year tenor and start dates from 2 to 30 years. This simultaneous fit to all expiries is done in Section 5 (for a single date) and can be achieved by replacing the geometric Brownian motion for the volatility process in the SABR model by the exponential of an Ornstein-Uhlenbeck process and a term exponential in $T_i - t$ in the equation for the forward rate. Section 6 provides further study of this model by analysing its performance over several historical dates and the model is shown to be a good representation of the market smile dynamics. In particular it is investigated if using a separable volatility structure is a major limitation and to what extent the model could be simplified in order to allow efficient implementation.

Our end goal is to identify a two-dimensional time homogeneous process of stochastic volatility type for the level of rates. Up to this point the focus has been on finding a suitable volatility structure linking the rates y^i under their respective measures. Given that the swap rates we are looking at have a large overlap and that we have been able to represent the data well with common parameters the expectation is that this procedure should give us a good guage on a process which captures the level of rates. Note that as we are assuming the process we are seeking is time homogeneous in principle, we only need to fit to one expiry. However in practice as we are unable to observe the volatility process directly we need the simultaneous fit for all expiries to pick up the right signal from the data.

In Section 7 by considering the corresponding spot rate process and ignoring the effect of measure changes we identify a candidate model for the level of rates. Denoting this candidate process by X we have that

$$\begin{aligned} \mathrm{d}X_t &= -cX_t\mathrm{d}t + \exp(U_t)\mathrm{d}W_t,\ X_0 = 0, \\ \mathrm{d}U_t &= \kappa U_t\mathrm{d}t + \nu\mathrm{d}V_t, \\ \mathrm{d}W_t\mathrm{d}V_t &= \rho\mathrm{d}t, \end{aligned}$$

where c, ν, κ are positive constants, $\rho \in (-1, 1)$ and W and V are correlated Brownian motions under S^N say, with T_N being the final expiry. We carry out a simple check on our candidate by using it as a two dimensional driver for a model of the swap rates in the one measure and show we can recover the prices of swaptions with different expiries.

2 Preliminaries and market data

Consider the discrete date structure $0 = T_0 < T_1 < \ldots T_N < T_{N+1}$ with $\delta_i = T_i - T_{i-1}$ and let $D_{T_iT_j}$ denote the zero-coupon bond price at time T_i for a unit payoff at the maturity time T_j. The time t equilibrium forward swap rate for a swap that starts at time T_i and ends at time T_j is given by

$$y_t(T_i, T_j) := \frac{D_{tT_i} - D_{tT_{j+1}}}{P_t(T_i, T_j)},$$

where P is the the present value of a basis point, PVBP, of the swap

$$P_t(T_i, T_j) := \sum_{k=i}^{j} \delta_{i+1} D_{tT_{k+1}}.$$

For ease of notation, when the tenor of the swap is either obvious or irrelevant, the forward swap rate setting at date T_i will be referred to as y_t^i and the PVBP as P_t^i.

In the market prices of swaptions are tracked using the implied Black volatility as this allows for easier comparison of swaptions at different strikes, tenors and expiries as well as across dates. To refer to the implied volatility we will use Σ equipped with specific arguments when needed (such as $\Sigma(K, y^i)$). Moreover, the term implied volatility smile refers to the implied volatilities across strikes and likewise the implied volatility surface/cube for implied volatilities across strikes, tenors and/or expiries.

Finally, note that the distribution of $y_{T_i}^i$ (given the initial conditions) under $\mathbb{S}^i$ completely specifies the time-0 implied volatility smile across strikes for swaptions with expiry T_i. Vice versa, market implied volatilities with expiry T_i at some strikes provides partial information about the $\mathbb{S}^i$ distribution of $y_{T_i}^i$.

2.1 Market data

The empirical investigations in this paper are based on an extensive data set of Euro swap rates and swaption prices. Each snapshot of data contains forward swap rates with expiries 2Y, 3Y ..., 10Y, 12Y, 15Y, 20Y, 25Y, and 30Y and tenors 1Y, 2Y, ..., 10Y, 12Y, 15Y, 20Y, 25Y, and 30Y. All rates are based on annual payment frequencies. For each rate the corresponding swaption implied Black volatilities are given for strikes -2, -1, -0.5, -0.25, 0, 0.25, 0.5, 1 and 2 percent away from the at-the-money (ATM) strike. As the data covers the period 3 July 2002 to 21 May 2009 it is produced using a one-curve approach as opposed to the more recent multi-curve approach taken by many banks as a consequence of the changed market environment due to the financial crisis starting in the autumn of 2008.

The data consists of two series, the first contains 66 roughly monthly spaced snapshots of the market spanning the period 3 July 2002 to 9 March 2007 and the second 556 daily samples of data covering 9 March 2007 to 21 May 2009. To our knowledge this is one of the very first papers that uses such an extensive data set. From almost daily samples of USD market data covering June 2002 to June 2005 the performance of a variety of models for dynamic swaption hedging is performed in Henrard (2005). While the mean is different the goal and the conclusions of the article is partly similar to the analysis performed in Section 4. Using data of Euro swaptions covering 15 December 2004 to 5 October 2007 Rebonato, McKay & White (2010) makes similar observations as we do and proposes a SABR/LIBOR market model to account for the observed effects. While their aim and to some extent approach is similar to ours their end goal is different and they limit the study to the LIBOR market model framework. In a recent paper Trolle & Schwartz (2010) uses a very extensive data set of both Euro and USD swaptions across strikes, tenors and expiries. Their end goal is however quite different from ours as they focus on displaying stylized facts of the distributions and its macroeconomic drivers.

Figure 3.1 displays a typical set of ten year tenor swaption implied volatility smiles from 27 October 2007. Note that the smiles look quite similar in shape although the curvature of the implied volatility smile seems to be decreasing in expiry. Figure 3.2 displays the forward swap rates as well as the ATM level, slope and curvature of the implied volatility smiles at 101 dates spanning the available period of data (66 dates up to 9 March 2007 and then another 35 evenly spaced samples from the daily data period). As a proxy for the ATM slope and curvature finite difference approximations are computed using the ATM and the plus/minus 1% strikes. These quotes are used as they are, according to traders, the most liquid strikes in the market.

First note that, beginning September 2008, the market turmoil is clearly spotted with large movements in rates and volatility measures and in particular

the 30 year rate gets very low compared with previous dates. Also note that, except for the turmoil period the different expiries moves largely in parallel and that there seem to be a positive correlation between the forward spot rate and the ATM level and curvature of the implied volatility smile, implying that as rates go down, implied volatilities and curvature increases. Moreover, while in the calm period the ATM level and curvature of the implied volatility smiles are declining in expiry this is not the case in the turmoil period where things seems much more disconnected.

For the slope of the implied volatility smiles there is no clear relation with time to expiry. There does however seem to be negative correlation between the rates and slope implying that as rates go down, the slope of the implied volatility smile gets steeper.

Although the above analysis just provides a rough idea about the macro (large time steps) behaviour of rates and implied volatilities it provides an idea about the challenge and requirements on a sound model for the level of rates. In particular, one would expect that during the turmoil period any model would struggle to fit the market data and market moves.

Section 4 provides a further and cleaner look at slopes and curvature dynamics of the smiles at one expiry at a time using the daily data series and Sections 5 and 6 develops and tests a model for all expiries.

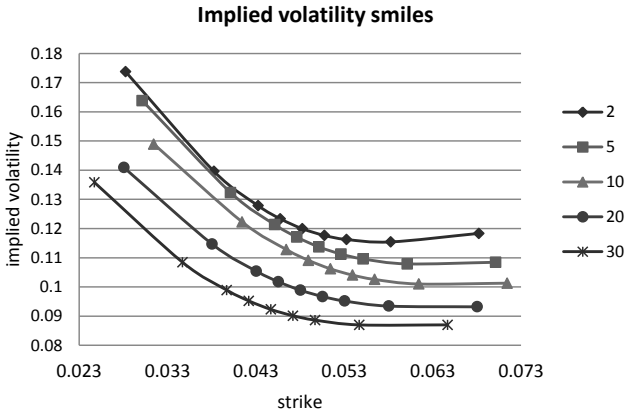

Figure 3.1: The implied volatility smiles for 10 year tenor swaptions with 2, 5, 10, 20, and 30 years to expiry on October 27, 2007.

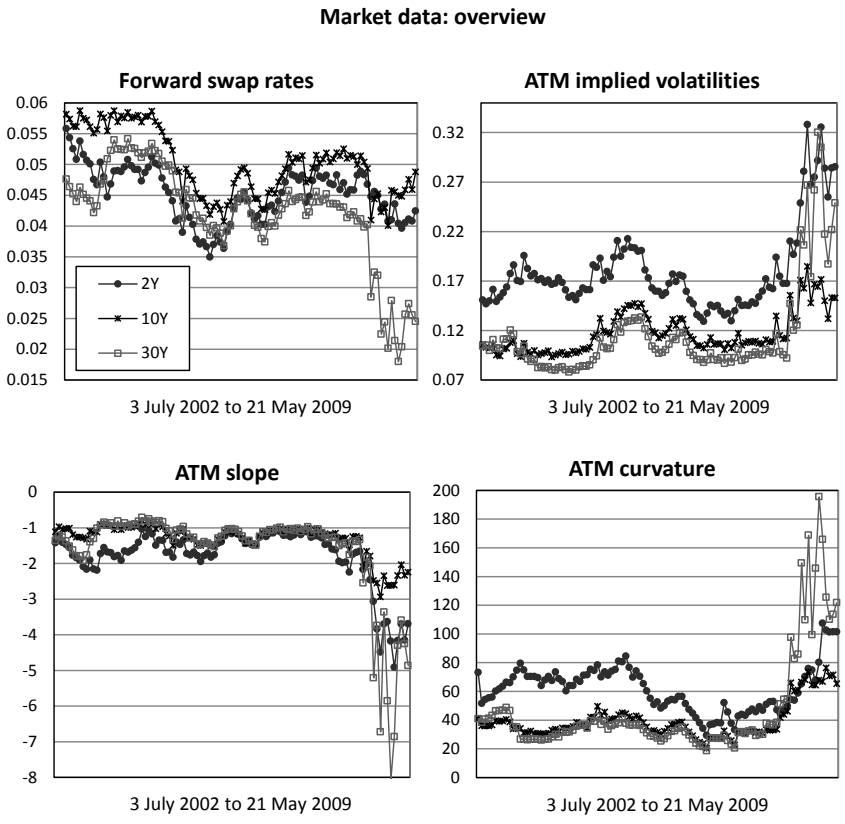

Figure 3.2: The forward swap rates and the implied volatility smile levels, slopes and curvatures at 101 evenly sampled dates in the data set. Swaptions with 10 year tenors. Dates are in chronological order, i.e. the left most data points refers to 3 July 2002.

3 Objective

As outlined in the introduction our approach to finding a suitable model for the level of rates is to build the model one step at a time, at any stage trying to understand where our model fails to reflect the data and adding complexity as necessary to capture the right qualitative behaviour.

Our first task is to specify a model that gives a good representation of (3), the one step marginal distributions of the swap rates at their setting dates. The paper Hagan et al. (2002) makes clear that even for the first step we need to

introduce at least a second factor to adequately reflect the behaviour of the market. Technically a second factor can only be introduced as a multiplicative factor if we are to specify our model as a diffusion. Given this we choose to start our analysis with an assessment of how well a stochastic volatility model can capture the one step macro distributions of the data where the volatility is specified by an autonomous equation.

The two most well known stochastic volatility models are of course the SABR model introduced by Hagan et al. (2002) and the Heston model (Heston (1993)). In Section 4 we carefully examine the ability of each of these models to reflect the features of our data at a single expiry. This was not a case of deciding which model is best but whether either does the job. We found the SABR model to pass all the tests we put it through and so took this as our building block for the next stage. One concern regarding the SABR model is that its widespread use in the market in general and its influence on practitioners in filling in the implied curve from liquid points in particular could make this choice of model a self-fulfilling one. We address this point in our study and argue why we believe it to be a sound choice for our purposes though there is room for improvement in capturing behaviour away from at the money.

We should mention here that there are several stochastic volatility models for interest rate derivatives introduced in the literature, see for example Andersen & Brotherton-Ratcliffe (2005), Piterbarg (2005) and Rebonato et al. (2010) in a LIBOR/swap market model setting or Andersen & Andreasen (2002) and Albanese & Trovato (2005) for HJM type models. Our work differ from the above in that we initially put considerations of tractability aside and investigate historical data to dictate the final form of the model and we only concern ourselves with capturing the level of rates process, a (potentially) low dimensional problem. Our approach may in future studies give some insight into what the most appropriate building block would be for a high dimensional full term-structure model such as a separable LIBOR market model.

4 Smile dynamics: SABR vs Heston

Recall our objective of formulating a model that is data driven. However we have to work within the constraints of the quality of the available data as we do not have available reliable and liquid option prices for each swaption at every strike. On any given day, apart from the ATM strike, only a few of the implied volatilities at various strikes will correspond to liquidly traded swaptions, with the other prices coming from some form of interpolation carried out by the bank. Given this we need to impose some structure to begin the modelling process and we have chosen to work within the class of stochastic volatility models specified by (1) - (2).

The first test any proposed model within this class must pass is that it should be able to fit each of the one step marginal distributions of

$$\left(y^i_{T_i}, \sigma^i_{T_i} | y^i_0, \sigma^i_0\right), \quad i = 1, \ldots, N, \tag{5}$$

under the measures $\mathbb{S}^i$, respectively. In this section we will make the assumption that for each i this conditional distribution is stable over time. Through this assumption we can link our model to the observed changes in the market smile as y^i_0 and σ^i_0 change, referred to as the smile dynamics. We will then investigate whether the one step marginals distributions given by the most popular models, the SABR model (for different βs) and the Heston model are consistent with the one step marginal distributions observed using historical data of forward swap rates and implied volatility smiles. All test in this section are based on the daily data series.

4.1 The SABR model

The SABR model dynamics for some underlying forward process f_t under some forward martingale measure is given by

$$\begin{aligned} \mathrm{d}f_t &= f_t^{\beta}\sigma_t \mathrm{d}W_t, \quad f_0 = f, & (6)\\ \mathrm{d}\sigma_t &= \nu\sigma_t \mathrm{d}V_t, \quad \sigma_0 = \sigma, & (7)\\ \mathrm{d}W_t \mathrm{d}V_t &= \rho \mathrm{d}t, & (8) \end{aligned}$$

for $\beta \in [0,1]$, $\nu \geq 0$ and W, V Brownian motions with correlation $\rho \in (-1,1)$. For this model Hagan et al. (2002) derives a reasonably accurate approximation of the implied volatility as a function of the initial values of the variables f and σ, the parameters β, ρ and ν and time to expiry T. This approximation is in widespread use to calibrate the SABR model to option prices in an efficient manner.

In addition to the higher order approximation a simpler approximation (eq. (3.1) in Hagan et al. (2002)) is also derived from which it is easier to understand the behavior of the model. For this 'crude' approximation the implied volatility for an option with strike K when the underlying forward is at f is given by

$$\begin{aligned} \Sigma(K,f) &\approx \frac{\sigma}{f^{1-\beta}} \left\{ 1 - \frac{1}{2}\left[1 - \beta - \frac{\rho\nu f^{1-\beta}}{\sigma}\right] \log\frac{K}{f} \right. \\ &+ \left. \frac{1}{12}\left[(1-\beta)^2 + (2-3\rho^2)\nu^2 \frac{f^{2-2\beta}}{\sigma}\right] \log^2\frac{K}{f} \right\}. \end{aligned} \tag{9}$$

Note that using this approximation the ATM implied volatility is given by

$$\Sigma(f,f) = \frac{\sigma}{f^{1-\beta}}. \tag{10}$$

Inserting this in (9) gives a formula for the implied volatility across strikes in terms of the ATM implied volatility

$$\begin{aligned}\Sigma(K,f) &\approx \Sigma(f,f)\left\{1-\frac{1}{2}\left[1-\beta-\frac{\rho\nu}{\sigma(f,f)}\right]\log\frac{K}{f}\right. \\ &+ \left.\frac{1}{12}\left[(1-\beta)^2+(2-3\rho^2)\frac{\nu^2}{\sigma^2(f,f)}\right]\log^2\frac{K}{f}\right\}.\end{aligned} \tag{11}$$

For a pure fitting of the implied volatility smile the most interesting effects of the parameters are that β and ρ both have a direct affect on the slope of the smile whereas ν has a direct affect on the curvature. The overlap of the parameters β and ρ implies that fitting the market smile for some expiry across a reasonably wide range of strikes may be done with similar precision for any value of β and hence the exact value of β would seem redundant. This is not the case.

Even for pure inter/extrapolating of implied volatilities β has an effect since while SABR models with different β may be fitted to more or less exactly agree on some strikes around the ATM point they do in general not agree exactly on all other strikes. In particular for large strikes this is an issue and market prices of CMS products can depend quite strongly on the choice of β.

Another aspect on the choice of β is that for a change in f or σ, the change in the implied volatility smile depends on β. This difference in 'smile dynamics' will have an effect on hedging or pricing of anything more exotic than vanilla swaptions. As our data set does neither contain market prices for high strikes nor prices of CMS products the focus of our investigations are on this latter aspect of the choice of β.

In terms of distributions, recall that fitting the implied volatility smile for some particular values of f and σ fixes the distribution of

$$(f_T, \sigma_T | f_0 = f, \sigma_0 = \sigma)\,. \tag{12}$$

While this can be done with similar precision for any $\beta \in [0,1]$ it does not imply that the distributions of

$$(f_T, \sigma_T | f_0, \sigma_0)\,. \tag{13}$$

are the same (for a fixed set of parameters). This is easily seen by matching up the implied volatility smiles for some different β and then vary the variables f and σ. A natural question to ask is hence; is the SABR model consistent with observed market 'martingale measure' marginal distributions and, if so, which β is appropriate?

In Hagan et al. (2002) the authors suggest choosing β by fitting the 'backbone' to historical data, where the backbone is defined as the curve traversed by the

ATM volatility, $\Sigma(f, f)$, as the underlying f varies. Indeed, as shown in figure 3.3, there is typically an inverse relation between the ATM volatility and the underlying. For the SABR model Hagan et al. (2002) identifies the backbone as approximately $\frac{\sigma}{f^{1-\beta}}$ and proposes to estimate β by a linear regression of the relation

$$\log \Sigma(f, f) = \log \sigma - (1 - \beta) \log f \tag{14}$$

using historical observations of $(f, \Sigma(f, f))$ pairs. This estimation procedure is unfortunately flawed as within a stochastic volatility model, as f changes, so does σ. If the market data was really from a SABR model we can't treat σ as a fixed parameter, since it is highly linked to f through the correlation ρ. That is, observations of $(f, \Sigma(f, f))$ do not provide sufficient information to estimate β via the macro approach (i.e. via considering the distribution of the forward at time T).

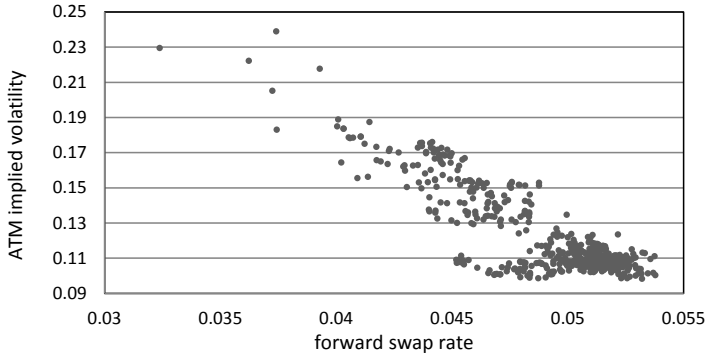

Figure 3.3: Daily quotes of ATM implied volatility plotted versus the forward swap rate for the period 9 March 2007 to 21 May 2009. Swaptions on the 10Yx10Y forward swap rate.

Let's now consider how observations of $(f, \Sigma(f, f))$ could be used to estimate β by taking an infinitesimal approach. First note that one may rewrite the SDE for the volatility as

$$\mathrm{d}\sigma_t = \rho\nu \frac{\mathrm{d}f_t}{f_t^{\beta}} + \nu\sigma\sqrt{1 - \rho^2}\mathrm{d}Z_t, \tag{15}$$

where Z_t is a Brownian motion uncorrelated to W and V. Letting $\Sigma_t := \Sigma(f_t, f_t)$ and using Ito's formula on the approximation (10) gives (after a few

manipulations)

$$\begin{aligned} \mathrm{d}\Sigma_t &= [\rho\nu + (\beta-1)\Sigma_t]\,\frac{\mathrm{d}f_t}{f_t} + \nu\sqrt{1-\rho^2}\Sigma_t \mathrm{d}Z_t \\ &+ (\beta-1)\Sigma_t\left[\frac{\beta-2}{2}\Sigma_t^2 + \rho\nu\right]\mathrm{d}t. \end{aligned} \tag{16}$$

Dividing through with Σ_t gives that $\rho\nu$ and $(\beta-1)$ may be estimated by regressing on the factors $\frac{\mathrm{d}f_t}{\Sigma_t f_t}$ and $\frac{\mathrm{d}f_t}{f_t}$ using a set of historical observations. Using a large amount of simulated data from a SABR model we have found this procedure to work decently. However, for estimation using market data one would need a lot more data than we have available and, moreover, as we can not expect market data to be generated exactly by a SABR model the estimation would probably be a rather noisy exercise.

Returning to the macro approach note that from (9) the ATM slope of the implied volatility smile is approximately

$$\frac{\partial\Sigma(K,f)}{\partial K}\,|_{K=f} \approx \frac{1}{2f}(\Sigma(f,f)(\beta-1)+\rho\nu). \tag{17}$$

Estimating β from a regression of

$$2f\frac{\partial\Sigma(K,f)}{\partial K}\,|_{K=f} = \Sigma(f,f)(\beta-1)+\rho\nu, \tag{18}$$

requires estimation of the ATM slope of the implied volatility smile and thus information about the implied volatility at more than the ATM strike is needed. While ATM swaptions are liquidly traded this is not the case for all strikes and in particular not necessarily the strikes that are very close to ATM and needed for an accurate computation of the ATM slope. The next subsection deals with informing about β (as well, if needed, ρ and ν) from slope and curvature of the implied volatility smile like relations.

4.2 Estimating β for the SABR model

As argued above we would not expect to be able to estimate β from historical observations of ATM implied volatilities and forward swap rates with good precision. To extract information about β we instead propose to investigate the slope of the implied volatility smile. Note that from (11) the first order difference between the implied volatility at strikes $f+h$ and $f-h$ is approximately

$$\Sigma(f+h,f)-\Sigma(f-h,f) \approx \frac{1}{2}\left[(\beta-1)\Sigma(f,f)+\rho\nu\right]\left(\log\frac{f+h}{f}-\log\frac{f-h}{f}\right). \tag{19}$$

After some manipulations, define

$$S(h,f) := 2\frac{[\Sigma(f+h,f)-\Sigma(f-h,f)]}{\log(\frac{f+h}{f-h})} \approx (\beta-1)\Sigma(f,f)+\rho\nu. \tag{20}$$

Since, for small h, $\log(\frac{f+h}{f-h}) \approx 2\frac{h}{f}$ the l.h.s. in the above relation is tending to $2f$ times the ATM implied volatility slope as h tends to 0. This means that the above relation is relating a, suitably scaled, finite difference approximation of the ATM implied volatility slope to the parameters of the SABR model. Hence, if the SABR model is a good representation of the market then β (and the product $\rho\nu$) could be estimated by a linear regression of the l.h.s. on $\Sigma(f,f)$ from a set of market snapshots.

To further investigate (or indeed estimate the parameters ρ and ν) whether the SABR model fits market data one may study a relation linked to the curvature of the implied volatility smile. Using the implied volatility at the strikes $f+2h, f$ and $f-2h$ and (11) up to second order gives

$$\begin{aligned} C(h,f) &:= \frac{2\Sigma(f,f)}{F_2}\left[\Sigma(f+2h,f)+\Sigma(f-2h,f)-2\Sigma(f,f)\right] \\ &\approx \left[\frac{(1-\beta)^2}{6}-(1-\beta)\frac{F_1}{F_2}\right]\Sigma^2(f,f)+\rho\nu\frac{F_1}{F_2}\Sigma(f,f)+\frac{2-3\rho^2}{6}\nu^2, \end{aligned} \tag{21}$$

where $F_1 = \log(1+\frac{2h}{f})+\log(1-\frac{2h}{f})$ and $F_2 = \log^2(1+\frac{2h}{f})+\log^2(1-\frac{2h}{f})$.

Note that for small h, $F_2 \approx 2\left(\frac{2h}{f}\right)^2$ and hence the l.h.s. above is closely related to the finite difference approximation of the ATM curvature of the implied volatility smile scaled with the factor $f^2\Sigma(f,f)$. Moreover, for small h, $\frac{F_1}{F_2} \approx -\frac{1}{2}$ and hence as h tends to 0 the above relation tends to the, suitably scaled, ATM curvature of the implied volatility smile calculated from the approximation (11).

To sum up; if the SABR model is a good representation of the market, then the parameters of the quadratic polynomial in $\Sigma(f,f)$ in the r.h.s. of (21) may be estimated using a set of market snapshots. Alternatively, β may be fixed from (20) and a linear regression may be performed to work out ρ and ν.

4.3 Investigating the SABR model using market data

The previous subsection outlines a method to estimate the parameters of a SABR model using a set of implied volatilities for the ATM plus two other strikes. However, typically, for a fixed β, ρ and ν are calibrated using current market data and there is hence no need to estimate them using historical data. What we are mainly interested in is investigating whether there is a β such

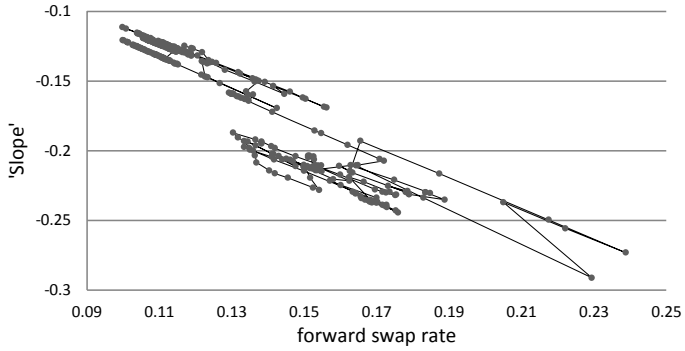

Figure 3.4: Market values of the l.h.s. of the 'slope' relation (20) versus the ATM implied volatility. $h = 0.01$. Daily values from the period 9 March 2007 to 21 May 2009 of 10Yx10Y swaptions.

that the marginal distributions of (5) from the SABR model is consistent with the market. As we can't extract much information from using only the ATM implied volatilities we will use the dynamics of the rescaled slope and curvature relations to inform about appropriate marginal distributions.

This subsection tests the SABR model using the set of daily market data of swaptions covering the period 9 March 2007 to 21 May 2009. To evaluate the SABR model we have chosen to use the 10 years tenors and the strikes ATM, ATM - 1% and ATM + 1%. According to traders these are the most liquid strikes and tenors and hence the most reliable quotes. Main focus is on the 10 years expiry but we also provides results for the 2, 5 20 and 30 years expiries.

Slope

Recall that if the market is well represented by a SABR model then daily values of the lhs of (20) plotted versus the ATM implied volatility would give, approximately, an affine function. Figure 3.4 plots $S(0.01, y_0^{10})$ vs $\sigma(y_0^{10}, y_0^{10})$ for the 10Yx10Y swaptions at all available dates. Note that the points seems to form 'lines' between which there at some dates is a jump. In terms of equation (20), $(\beta - 1)$ would then be the slope of the 'lines' and the jumps would correspond to a change in the product $\rho\nu$. By inspection the slope of the lines seems to correspond to a β of around 0.

To be able to make a more precise investigation one may look at the difference between the approximate slope relations between dates. Define the change in

'slope' between date t_i and t_j as

$$\Delta S(h,i,j) := S(h,f_{t_j}) - S(h,f_{t_i}) \tag{22}$$

and the change in ATM implied volatility as

$$\Delta\Sigma(i,j) := \Sigma(f_{t_j},f_{t_j}) - \Sigma(f_{t_i},f_{t_i}). \tag{23}$$

For a fixed β, but allowing a change in $\rho\nu$, the difference in 'slope' between dates i and j is then (using (20)) approximately

$$\Delta S(h,i,j) \approx (\beta-1)\Delta\Sigma(i,j) + \rho(j)\nu(j) - \rho(i)\nu(i). \tag{24}$$

Hence, if the SABR model is a good representation of the market marginal distributions, by using market data to plot the above lhs versus the change in ATM implied volatility we would expect to see points lining up along a line with slope $\beta - 1$ as well as potentially some scattered points corresponding to changes in $\rho\nu$. Figure 3.5 confirms that this is the case for 10Yx10Y swaptions. In this case a linear regression using (24) would estimate β to be -0.04 with an R^2 of 0.84.

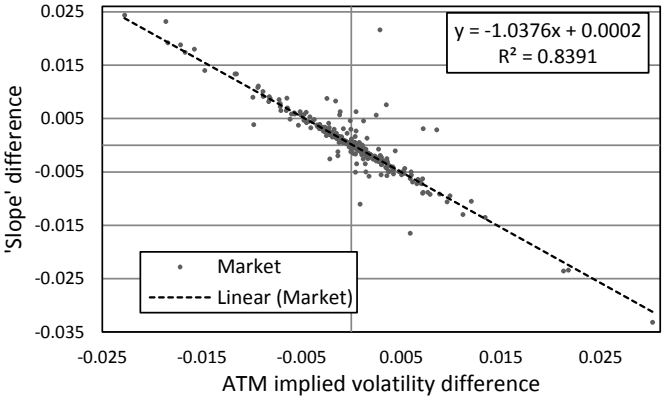

Figure 3.5: Market values of 'slope' difference for 10Yx10Y swaptions plotted versus the ATM implied volatility difference estimated using daily steps from 9 March 2007 to 21 May 2009. The linear regression line of the market values are also displayed implying a β estimate of about -0.04.

The accuracy of informing about an appropriate β from the above procedure depends mainly on two things. First, one needs a reasonably large sample of reliable market data. The historical data set of swaption data used in this paper consists of 556 dates. Hence, using daily steps to get data for (24) gives 555 points. By stepping more than one date forward one would get more points

(as well as larger differences) but in our experiments one date forward gives enough points to provide estimates with decent accuracy.

Second, the approximation leading to the slope expression (24) needs to be reasonably accurate. As the expression is based on a rather crude approximation this might of course be far from the case but there might also be hope as we are only concerned with the slope and not the full implied volatility smile across strikes. To address this we started by fixing a set of parameters (β, ρ, ν), forward swap rates, ATM implied volatilities and expiries and then compared the rhs of equation (24) with the values of the lhs produced using the full approximation of Hagan et al. (2002). Our findings points at that the rhs of expression (24) is reasonably accurate but slightly overestimates the slope. More precisely, to match up the true value of the lhs we needed to adjust β downwards by a factor 0-0.12, depending on the specific values of the parameters, rates, volatilities and expiries.

To approximately correct for this bias we have used a two-step procedure. In the first step we find an estimate β^* using linear regression and equation (24). Note that as the rhs of (24) overestimates the slope we expect β^* to be underestimated. In the second step we set $\beta = \beta^*$ and calibrate the parameters $\rho^*(i)$ and $\nu^*(i)$ at each of the dates in the sample (using the full Hagan approximation). We then search for a correction constant c that minimises

$$\sum_{i=1}^{556} \left(S(0.01, f_{t_i}) - (\beta^* - c - 1)\Sigma(f_{t_i}, f_{t_i}) - \rho^*(i)\nu^*(i)\right)^2 \tag{25}$$

Finally, we take c as a proxy for the underestimation of β and set $\beta_c^* = \beta^* + c$ as our corrected estimate of β.

It should at this point be noted that even though we do produce point estimates for β we are mainly interested in informing about a reasonably good choice of β and for our purposes we are hence fine with this level of accuracy.

Table 3.1 displays the estimates β^*, β_c^* the standard errors, $\rho\nu$ and the R^2 obtained by regressing on (24) using data of 2Yx10Y, 5Yx10Y, 10Yx10Y, 20Yx10Y and 30Yx10Y swaptions. The regression is performed for both all historical observations as well as the first one and a half year of the data that covers a period of quite calm markets.

Note that the 2, 5 and 10 years expiries suggests a β of about zero irrespective of estimation period. For the 20 and 30 years expiries β zero seems appropriate during the calm period whereas during the turmoil it seems more appropriate with a larger β. Actually, as the turmoil period provides most of the variation in the data, estimation during the turmoil period gives an estimate very close to estimation over the complete set of data. Other tenors are checked as well and give similar estimates and accuracy.

expiry	period	β^*	β_c^*	s.e.	$\rho\nu$	R^2
2Y	full	-0.07	0.00	0.040	9.42E-05	0.56
2Y	calm	0.01	0.07	0.043	8.15E-05	0.58
5Y	full	-0.08	-0.02	0.018	1.23E-04	0.87
5Y	calm	-0.07	-0.01	0.026	2.73E-05	0.81
10Y	full	-0.04	0.03	0.019	1.58E-04	0.84
10Y	calm	-0.07	-0.01	0.025	4.53E-05	0.82
20Y	full	0.19	0.29	0.013	1.86E-04	0.88
20Y	calm	-0.07	0.00	0.026	4.68E-05	0.82
30Y	full	0.30	0.41	0.014	1.27E-04	0.81
30Y	calm	-0.05	0.03	0.026	4.76E-05	0.80

Table 3.1: Regressing (24) on $\Sigma(f, f)$ to estimate β for different expiries and estimation periods.

The change in β for the longer expiries during the period of turmoil does not appear to have its basis in a short term change in perception of the market (where we might expect β to alter for the shorter expiries). Rather it is indicative of a period of market instability and we would need a richer model than SABR to capture anything useful. Even with a richer model it would be difficult to extract a coherent signal. Indeed this period corresponds to the subprime credit crunch. From the plots in Section 2.1 it can be seen that the period of turmoil covers the last six months of the seven year period of the monthly series. Our focus in this paper will be on building a model for the period when the markets are stable.

Remark 4.1 *The plot of the slope relation in Figure 3.5 looks remarkably like what one would expect if the data came from a SABR model where the parameters ρ and ν were altered from time to time. In searching for an appropriate model it is important to be careful not to carry out an investigation whose conclusions are a consequence of a SABR like implied volatility curve being used by traders to interpolate between liquid strikes. That is why we worked with the most liquid strikes in studying the data.*

Remark 4.2 *It is interesting to note however that from talking to traders they seem to conclude that while $\beta = 0$ is appropriate they tend to prefer using a larger β for day to day use. The reason for this is that though any β provides a good fit near at the money, the prices of CMS products suggest that away from the money a β closer to one tends to be more appropriate. The chosen value of β is hence a compromise between these cases. This would suggest that if it were possible to identify the liquid points away from ATM one might choose a more general functional form than that of a CEV model in the equation for the forward or choosing a different scaling of volatility of volatility.*

Curvature

As displayed in the previous subsection a SABR model with an appropriately chosen β provides a good reflection of the market dynamics of slope of the implied volatility smile. As a further check of whether the SABR model with a suitably chosen β provides appropriate marginal distributions one may study the dynamics of the curvature. As showed above by looking at scaled 'curvature' type relations one could estimate also ρ and ν using historical data, provided sufficient accuracy of the approximation leading to (21). In general, however, this is not needed since ρ and ν will be decided by fitting the model to option values across strikes and this relation is only used to provide further intuition for the behavior of the model.

As a first check of whether the SABR model provides a reasonable reflection of the curvature it is instructive to relate the market changes in curvature to the model changes. Note from (21) that for a SABR model the change in the rescaled finite difference approximation of the curvature will depend on β. Hence, plotting the market changes in the lhs of (21) versus the model changes for different βs will display how well the model reflects the curvature dynamics.

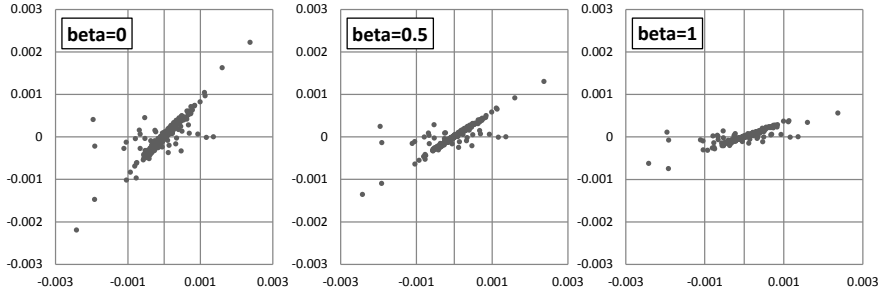

Figure 3.6: Scatter plots of changes in 'curvature' as defined in (21) for the market (x-axis) and the SABR model (y-axis) with different β. Changes are for 10Yx10Y swaption and taken over one date during the calm period, implying a total of 388 points. See the text for a more thorough explanation.

To get the SABR model values the following procedure is used. First the higher order approximation of the SABR model is calibrated at date i to the strikes $K = f, f+0.01$ and $f-0.01$. Note that this can be done perfectly for all β and will hence exactly agree with the market values. Then, at date j the model is recalibrated to the ATM implied volatility by updating (the variable) σ. All parameters (β, ν and ρ) are kept fixed at the date i values. The model values of the change in the lhs of (21) may then be computed. Figure 3.6 displays this for the 10Yx10Y swaptions. Changes are computed during the calm period and by stepping one date forward.

Note that for $\beta = 0$ most points line up along the 45 degree line implying a close relation between the market and the model. For $\beta = 1$ the points line up along a line with significantly less steep slope and $\beta = 0.5$ is in between. The reason for choosing the calm period is that the scatter is quite small and hence good for displaying the effect without too much noise. However, other periods and expiries displays the same effect albeit in some cases with a bit more scatter. In general the β estimated from the 'slope' relations above seems to work best also for the curvature.

As a further test it is instructive to plot the curvature changes versus the ATM implied volatility changes as made in the 'slope' case above. Recall from (21) that a finite difference approximation of the rescaled curvature is approximately given by a quadratic polynomial in the ATM implied volatility. Hence, one can not expect a linear relation between curvature and ATM implied volatility changes. However, as for δ small the difference $(x+\delta)^2 - x^2$ is approximately linear it could still be instructive to plot the change in curvature versus the change in the ATM implied volatility for a period of small changes. Figure 3.7 plots the market change and the model changes with $\beta = 0, 0.5$ and 1 during the calm period for the 10Yx10Y swaption.

Note that $\beta = 0$ gives a pattern similar to the market whereas for $\beta = 0.5$ and 1 the points line up with too flat slope.

4.4 The Heston model

This section investigates whether the Heston model (Heston (1993)) is consistent with the market marginal distributions. The Heston model is defined by the system of SDE:s

$$\begin{aligned} \mathrm{d}f_t &= f_t \sigma_t \mathrm{d}W_t, \quad f_0 = f, && (26)\\ \mathrm{d}\sigma_t^2 &= \kappa(\bar{\sigma}^2 - \sigma_t^2)\mathrm{d}t + \nu\sigma_t \mathrm{d}V_t, \quad \sigma_0^2 = \sigma^2, && (27)\\ \mathrm{d}W_t \mathrm{d}V_t &= \rho \mathrm{d}t. && (28) \end{aligned}$$

For the above system option prices may be found by numerical inversion of an associated Fourier transform implying reasonably efficient pricing and calibration. In terms of fitting option prices across strikes the effects of ρ and ν are similar as in the SABR model. In addition to the SABR model there is also a mean-reversion term where the mean-reversion level $\bar{\sigma}$ determines the long run ATM implied volatility. The mean-reversion speed κ also has an effect on the ATM implied volatility but the more interesting effect is a decline in curvature of the smile which is roughly inversely proportional to the time to expiry. Hence, ν and κ have similar, but opposite, effects. Through the combination of ν and κ one may hence control the curvature of the smile across

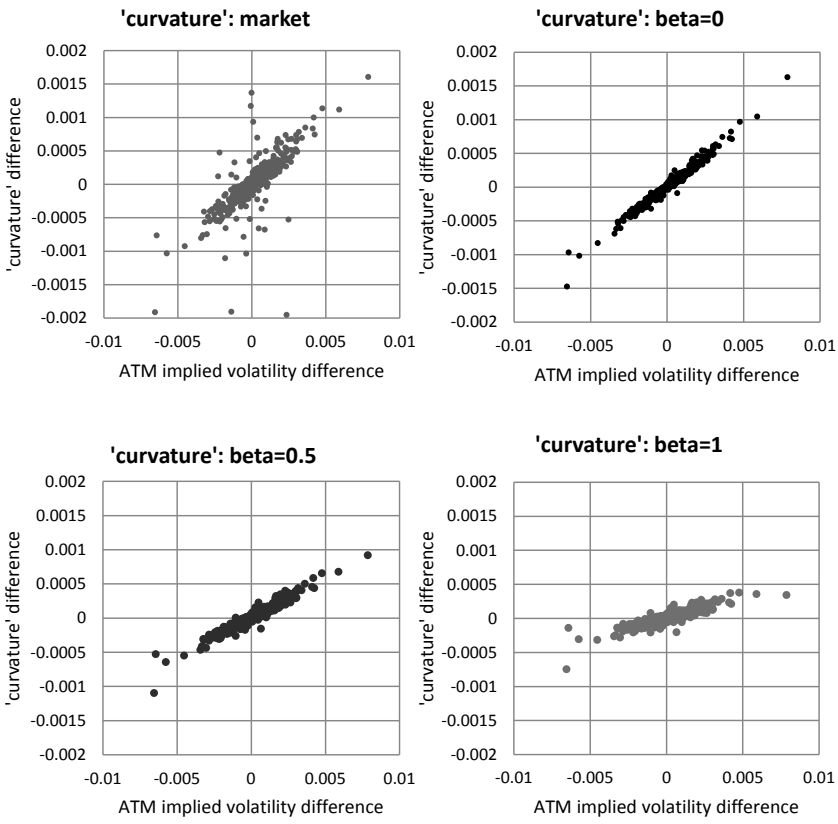

Figure 3.7: Changes in 'curvature' as defined in (21) for the market and the SABR model with different β plotted against the change in the ATM implied volatility. Changes are for the 10Yx10Y swaption and taken over one date during the calm period, implying a total of 388 points. See the text for a more thorough explanation.

several expiries. However, in terms of fitting one expiry only this may be done at virtually similar accuracy for any $\bar{\sigma}$ and κ.

In the standard Heston model there is no β term. Extending the model with a CEV type local volatility function is of course in theory an easy task however this implies that the Fourier transform technique can no longer be used and hence efficient pricing of European type option is lost. To get around this issue one may use a local volatility function of displaced diffusion type for which the Fourier transform technique is still applicable, see Andersen & Piterbarg

(2010) for details. To limit the extent of this study we have however chosen to only investigate the standard Heston model and the fairest comparison with the SABR model is hence the $\beta = 1$ case.

For the Heston model there does not, to our knowledge, exist a simple and reasonably accurate approximation of implied volatilities as is the case for the SABR model. For short expiries and strikes close to ATM Durrleman (2004) provides an expression which is of similar complexity as (9) above. Expanding this expression in $\log \frac{f}{K}$ and a few manipulations give

$$\begin{aligned}\Sigma(K,f) &\approx \Sigma(f,f) + \frac{\rho\nu}{4\Sigma(f,f)} \log\frac{K}{f} + \\ &+ \left(\frac{\nu^2}{24\Sigma(f,f)\sigma^2}(1 - \frac{7}{4}\rho^2) - \frac{\nu^2\rho^2}{32\Sigma^3(f,f)} \right) \log^2 \frac{K}{f}. \end{aligned} \tag{29}$$

Note that this expression does not contain $\bar{\sigma}$ or κ and is hence, if at all, only expected to work for very short expiries. While this approximation is lot a worse than the SABR approximation in terms of pricing it does provide a reasonably accurate intuition for the behavior of the Heston model. For the 'slope' expression $S(h,f)$ defined in (20) one gets that the 'slope' is inversely proportional to $\Sigma(f,f)$. Note that since interest rates implied volatility smiles are typically (always) downward sloping, ρ is typically negative. Hence, for an increase in implied volatility, the implied volatility slope will also increase (less negative). Note that this is the opposite behavior to what is observed in the market.

For the curvature expression (21) the story is similar to the one for the slope. To analyse this case it is instructive to make a further approximation. Using that for short expiries $\sigma^2 \approx \Sigma^2(f,f)$ (see Durrleman (2004)) the 'curvature' expression $C(h,f)$ defined in (21) above is for the Heston model

$$C(h,f) \approx \nu\rho\frac{F_1}{F_2} + \frac{\nu^2}{96\Sigma^2(f,f)}\left(4 - 10\rho^2\right). \tag{30}$$

Hence, the effect of an increase in $\Sigma(f,f)$ depends on the sign of $(4 - 10\rho^2)$. It turns out that, across all dates and expiries in our sample, the calibrated $(4 - 10\rho^2)$ is negative and hence an increase in $\Sigma(f,f)$ leads to an increase in the 'curvature' $C(h,f)$ of the implied volatility smile. Also this is opposite to what is observed in the market.

Since the above approximation is quite crude it is not certain that this effect is valid for the true Heston model. As a check consider Figure 3.8 in which 10Yx10Y swaptions are investigated during the calm period of the daily data set. The upper plots in the figure display the Heston model 'slope' and 'curvature' changes plotted versus the market 'slope' and 'curvature' changes. As for the SABR model, the Heston changes are produced by fitting the parameters

at one date and then at the next date updating σ^2 such that the ATM implied volatility is matched. The lower plots display the changes in 'slope' and 'curvature' versus the changes in the ATM implied volatility. Note that the 'curvature' changes are almost exactly opposite to the market ones and that the 'slope' changes have the wrong sign.

To isolate the effect of $\bar{\sigma}$ and κ these parameters are fixed at the initial value (σ) as well as 0.01, respectively, for all dates. A κ of 0.01 is a typical value resulting from a calibration of the Heston model to expiries ranging from 2 to 30 years simultaneously. Different values of $\bar{\sigma}$ and κ were tested and provided very similar results.

In general, for shorter expiries, results are a bit less noisy and more pronounced than the ones displayed in the figure whereas for longer expiries investigated during the full period of available historical observation there is more noise. However, all performed test reveals the same issue; the Heston model provides wrong changes in slope and curvature. Compared with SABR using $\beta = 1$ the curvature case is the most different.

Remark 4.3 *Following Durrleman (2004) it may be seen that extending the Heston model with a local volatility function of CEV type (or a matched up displaced diffusion transformation) adds the same β-terms that are present in the 'crude' approximation of implied volatilities in the SABR model. Hence, approximately, one may expect to modify the base case ($\beta = 1$) behavior in a similar magnitude as in the SABR model. However, as the base case ($\beta = 1$) is a lot closer to the market behavior in the SABR model compared with the Heston model we conjecture that matching the market smile dynamics would be challenging also when using an extended Heston model. In particular, as the β terms have smaller effect on the curvature, the market curvature changes would seem hard to match.*

4.5 Summing up

The tests performed in this section provide evidence that the SABR model with an appropriately chosen β is a good reflection of the marginal distributions of swaptions observed in the market. The Heston model on the other hand seems to provide a worse representation of the market.

Note that these two models are of course not the only ones in the class (1 - (2) and there could be other stochastic volatility models that perform as well as or better than SABR. In particular one may consider using different local volatility functions for the forward or, for example, a power function in σ in the SDE for the stochastic volatility process. For the latter choice note that

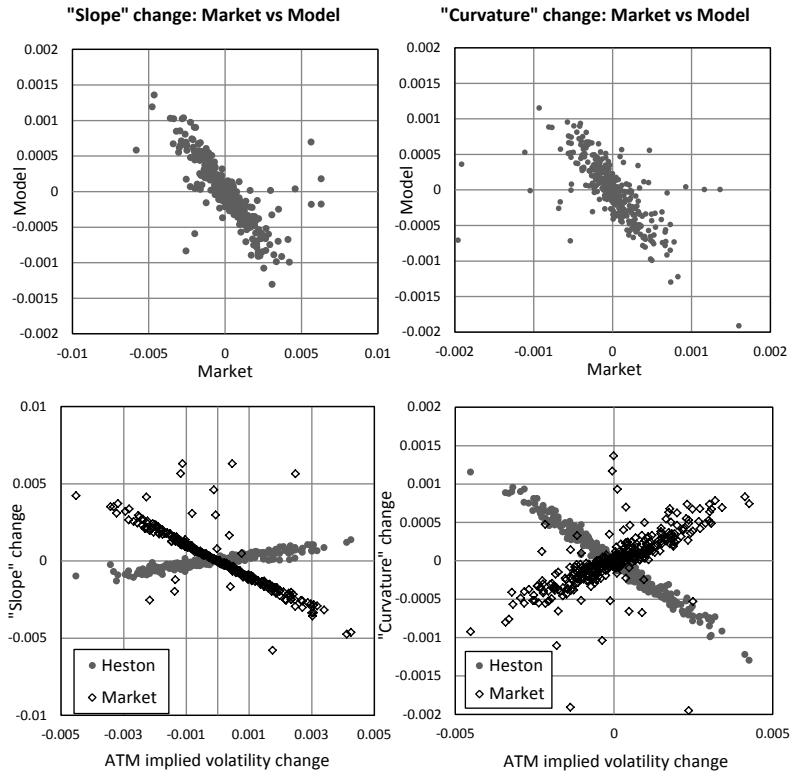

Figure 3.8: Changes in 'slope' and 'curvature' in the Heston model for 10Yx10Y swaptions investigated during the calm period. The tests were performed in the same manner as for the SABR model.

in the Heston model the power constant would be 0.5 whereas in the SABR model it would be 1. Again, Durrleman (2004) could be used to get intuition about how the dynamics of the 'slope' and 'curvature' would look like in this more general case and it might be possible to find a structure that matches up both dynamics of the smile and market prices of CMS products.

However, given the constraints of our data and how well the SABR model with an appropriate β performs in the outlined tests we have chosen the SABR model as a building block for (5).

5 One date, many expiries: the SABR-MR model

In the previous section we found that the SABR model with appropriate β provides a good fit to the distribution of a forward swap rate at its start date. For our objective of finding a stochastic volatility model for the level of rates the contribution of this section is to engineer a model linking the forward swap rates (although still under their own swaption measure) based on only two factors that enables a specification of parameters that is *common* to all forward swap rates having ten year tenor and start dates from 2 to 30 years.

Note that a model with a good representation of the conditional distribution at T_i does not necessarily provide a good representation of the (conditional) forward smiles i.e. the distribution of $\left(y^i_{T_i}, \sigma^i_{T_i} | y^i_t, \sigma^i_t\right)$, $t \in (0, T_i)$. Using the SABR model as our starting point we will in this section address the issue of appropriate time dependence as fitting to data at several expiries requires it.

For our general model setup

$$\begin{aligned} \mathrm{d}y^i_t &= f^i(T_i - t, y^i_t, \sigma^i_t)\mathrm{d}W^i_t, \\ \mathrm{d}\sigma^i_t &= h^i(\sigma^i_t)\mathrm{d}t + g(\sigma^i_t)\mathrm{d}V^i_t, \end{aligned}$$

we start by assuming the function f^i is of the form

$$f^i(T_i - t, y^i_t, \sigma^i_t) = l^i(T_i - t)\phi^i(y^i_t)\sigma^i_t \tag{31}$$

Note that this setup is not particularly restrictive and still covers most of the previously introduced stochastic volatility models in the literature. The coming sections discusses the implications of various choices of the functions f^i, g, h^i and introduces the SABR with mean-reversion model.

5.1 Many expiries with the SABR model

Table 3.2 displays the parameters of disconnected SABR models calibrated to expiries ranging from 2 up to 30 years. Note that for both $\beta = 0$ and $\beta = 0.5$ the calibrated parameter values of σ_0 and ν are steadily decreasing in expiry whereas the ρ values, although scattered, seems less dependent on expiry. The patterns in σ_0 and ν are a general feature in our data sample (albeit the decline are at some dates less smooth and nice) whereas for ρ there are also dates with more monotone patterns. In general, though, the dependence between expiry and ρ is, if at all there, less clear and less pronounced than the one in σ_0 and ν. As we are after a model for the level of rates and hence need common parameters for all swap rates we need to modify our model to avoid such systematic changes in parameters.

Expiry	β	σ_0	ν	ρ	β	σ_0	ν	ρ
2Y	0.5	0.026	0.326	-0.160	0	0.0057	0.309	0.060
3Y	0.5	0.025	0.309	-0.183	0	0.0056	0.290	0.042
4Y	0.5	0.025	0.299	-0.204	0	0.0056	0.278	0.023
5Y	0.5	0.025	0.297	-0.220	0	0.0055	0.275	0.003
6Y	0.5	0.024	0.289	-0.215	0	0.0054	0.268	0.007
7Y	0.5	0.024	0.283	-0.211	0	0.0054	0.263	0.012
8Y	0.5	0.024	0.276	-0.209	0	0.0053	0.255	0.017
9Y	0.5	0.023	0.269	-0.207	0	0.0053	0.249	0.022
10Y	0.5	0.023	0.263	-0.205	0	0.0052	0.243	0.027
12Y	0.5	0.022	0.253	-0.207	0	0.0050	0.233	0.028
15Y	0.5	0.022	0.241	-0.211	0	0.0048	0.221	0.031
20Y	0.5	0.020	0.223	-0.219	0	0.0044	0.203	0.035
25Y	0.5	0.019	0.212	-0.221	0	0.0041	0.192	0.038
30Y	0.5	0.018	0.203	-0.223	0	0.0038	0.183	0.041

Table 3.2: Calibration of the SABR model on 27 October 2007. Expiry in years. Options to enter 10 year swaps.

5.2 Extending the SABR model

This section puts the issues with calibrating the SABR model to several expiries simultaneously into context and discusses some potential extensions. To get some insight into how the SABR model can be suitably modified the displaced diffusion SABR model (DD-SABR) will be helpful. In this model the dynamics of y_t^i under $\mathbb{S}^i$ are given as

$$\begin{aligned} \mathrm{d}y_t^i &= (y_t^i + \delta^i)\sigma_t^i \mathrm{d}W_t^i, && (32)\\ \mathrm{d}\sigma_t^i &= \nu^i \sigma_t^i \mathrm{d}V_t^i, && (33)\\ \mathrm{d}W_t^i \mathrm{d}V_t^i &= \rho^i \mathrm{d}t, && (34) \end{aligned}$$

for some displacement constant $\delta^i \in \Re$. For models with deterministic volatility Marris (1999) showed that by choosing (for $\beta \in (0,1]$)

$$\begin{aligned} \delta^i &= y_0^i \frac{1-\beta}{\beta}, && (35)\\ \sigma_t^i &= \sigma_t^{cev,i} \beta (y_0^i)^{\beta-1} && (36) \end{aligned}$$

the CEV and displaced diffusions models produce very similar prices of European calls across strikes. In a stochastic volatility setting Kennedy, Mitra & Pham (2011) find that the standard SABR (CEV-SABR) and DD-SABR models continue to be very close qualitatively (matched up using (35) and (36) and taking the same ν^i and ρ^i parameters). To keep the link between the CEV-SABR and DD-SABR we will continue to refer to the parameter β also for the

DD-SABR model and will hence implicitly assume that δ^i from equation (35) is used.

For the DD-SABR model with $\beta \in (0,1]$ the distribution of $y^i_{T_i}$ can be written (see Kennedy et al. (2011) or Appendix 9.1) as an exponential of a function of $V^i_{T_i} = \int_0^{T_i} (\sigma^i_s)^2 \mathrm{d}s$ and an independent standard normal random variable G

$$y^i_{T_i} \stackrel{d}{=} (y^i_0 + \delta^i) \exp\left(\frac{\rho^i}{\nu^i}(\sigma^i_{T_i} - \sigma^i_0) - \frac{1}{2}V^i_{T_i} + \sqrt{V^i_{T_i}}G\right) - \delta^i. \tag{37}$$

For $\beta = 0$ the equivalent relation is

$$y^i_{T_i} \stackrel{d}{=} y^i_0 + \frac{\rho}{\nu}(\sigma^i_{T_i} - \sigma^i_0) + \sqrt{V^i_{T_i}}G. \tag{38}$$

To analyse the (DD-) SABR model we will consider the case $\rho^i = 0$ for simplicity. The case $\rho^i \neq 0$ is similar in essence but with greasier expressions. Note that for $\beta = 0$ taking $\rho^i = 0$ is not a too bad choice compared with the outcome of a market calibration (see Table 3.2).

Note for $\beta \in (0,1]$, conditional on $V^i_{T_i}$, the distribution of $y^i_{T_i}$ is displaced log-normal and the hence the price of a swaption with strike K can be written as an expectation of the standard Black's pricing formula where the implied volatility is a function of $V^i_{T_i}$. For $\beta = 0$ the conditional distribution of $y^i_{T_i}$ is instead Gaussian and so the Bachelier formula applies in this case. Without stochastic volatility the implied volatility smiles are monotone, downward sloping (steepness depending on β) and with no or very little curvature. Adding stochastic volatility provides an increase in the overall level of the smile as well as an increase in curvature (and if ρ is not zero a tilt in the slope). Hence the particular distribution of $V^i_{T_i}$ is linked to these features.

Although we do not know the distribution of $V^i_{T_i}$ it is informative to study its first two moments

$$E^{\mathbb{S}^i}\left[V^i_{T_i}\right] = \int_0^{T_i} (\sigma^i_0)^2 \exp((\nu^i)^2 s)\mathrm{d}s = \frac{(\sigma^i_0)^2}{(\nu^i)^2}\left(\exp\left((\nu^i)^2 T_i\right) - 1\right) \tag{39}$$

$$\begin{aligned} E^{\mathbb{S}^i}\left[\left(V^i_{T_i}\right)^2\right] &= 2\int_0^{T_i}\int_0^t E^{\mathbb{S}^i}\left[\left(\sigma^i_s\right)^2\left(\sigma^i_t\right)^2\right]\mathrm{d}s\mathrm{d}t \\ &= \frac{(\sigma^i_0)^4}{15(\nu^i)^4}\left[\exp\left((\nu^i)^2 T_i\right)\left(\exp\left(5(\nu^i)^2 T_i\right) - 6\right) + 5\right]. \end{aligned} \tag{40}$$

Note that in the limit $\nu \to 0$ the first moment collapses to $(\sigma^i_0)^2 T_i$ and the second moment is zero and we are back at the Black/Bachelier models. In terms of the implied volatility smile a rough rule of thumb is that the first moment controls the overall level of the smile and the second moment is linked to the curvature of the smile.

Now recall that the calibration of the CEV-SABR model to market data returned decreasing σ_0^i and ν^i with time to expiry. If we calibrate a DD-SABR model instead we will get similar parameter values and, in particular, qualitatively the results will be exactly the same. For the purpose of this discussion we can hence assume at this point that the calibrated parameters in Table 3.2 are from a calibrated DD-SABR model.

For our objective of finding a model with common parameters for all forward swap rates the first and second moments are helpful in understanding how the SABR model should be extended. Both moments in (39) and (40) are growing at exponential pace in expiry (or really in $(\nu^i)^2 T_i$). Compared with market data exponential growth seems to be too fast (as calibration returns decreasing ν^i) and in particular one would not expect this to be reasonable for long expiries.

To match the (market induced) moments of $V_{T_i}^i$ we start with some intuition from the deterministic case. Put $\nu^i = 0$ and let $\hat{\sigma}_0^i$ denote the outcome of a calibration to the market price of an expiry-T_i ATM swaption. If we take $\sigma_t^i = l(T_i - t)$ then equations (37) and (38) still hold and we need to choose the parameters of $l(T_i - t)$ such that $V_{T_i}^i = \int_0^{T_i} l^2(T_i - t)\mathrm{d}t = (\hat{\sigma}_0^i)^2 T_i$ for all i to calibrate the model to the market prices of ATM swaptions across expiries.

If we extend the model in the stochastic volatility setting and set

$$\mathrm{d}y_t^i = (y_t^i + \delta^i) l(T_i - t)\sigma_t^i \mathrm{d}W_t^i, \tag{41}$$

with σ_t^i as in the usual SABR model but with $\sigma_0^i = 1$ and $V_{T_i}^i := \int_0^{T_i} l^2(T_i - t)\sigma_t^i \mathrm{d}t$, then the distribution of $y_{T_i}^i$ can still be written as in equations (37) and (38) with the first and second moments of $V_{T_i}^i$ given by

$$E^{\mathbb{S}^i}\left[V_{T_i}^i\right] = \int_0^{T_i} l^2(T_i - t)\exp((\nu^i)^2 s)\mathrm{d}s \tag{42}$$

$$E^{\mathbb{S}^i}\left[\left(V_{T_i}^i\right)^2\right] = 2\int_0^{T_i} l^2(T_i - t)\exp((\nu^i)^2 t)\int_0^t l^2(T_i - s)\exp(5(\nu^i)^2 s)\mathrm{d}s\mathrm{d}t \tag{43}$$

Note that the function $l(T_i - t)$ affects both moments in a similar fashion. However, as ν^i does not have the same effect on each of the moments the extended model is still not powerful enough to match up both moments and in particular it would struggle with matching up the second moment. At an intuitive level the function $l(T_i - t)$ can be thought of as targeted to fitting the first moment (or the ATM level of the smiles) under the respective measures $\mathbb{S}^i$.

To match both moments we hence need to extend the model further. To account for the time to expiry effect of ν, Rebonato et al. (2010) uses a similar approach as for σ_0^i also for the volatility of volatility and extends the SABR model constant ν with a function $n(T_i - t)$. By using rather general and flexible

functions $l(T_i - t)$ and $n(T_i - t)$ they manage to fit their model quite well across strikes and expiries and then constructs a LIBOR/SABR market model. Although this approach is adequate for the (high-dimensional) LIBOR market model it does not enable using a common stochastic volatility driver for the level of rates and is hence not suitable for our purposes. Instead we choose to focus on an alternative route that alters the stochastic volatility process and introduces mean-reversion.

While mean-reversion may be added to the SABR model in several ways we aim for a specification that is similar to the standard SABR model but at the same time flexible enough to be able to fully appreciate the extra degree of freedom. Our approach builds on a model by Fouque, Papanicolaou & Sircar (2000) and is based on taking

$$\begin{aligned} \sigma_t^i &= \sigma_0^i \exp(U_t^i), && (44) \\ \mathrm{d}U_t^i &= -\kappa U_t^i \mathrm{d}t + \nu \mathrm{d}V_t^i, \quad U_0 = 0. && (45) \end{aligned}$$

This modeling choice is also done in Jackel & Kahl (2007) although they choose to focus on a minor modification of the exponential function of hyperbolic type. With this choice of equation for the volatility process replacing the usual log-normal equation in the SABR model again we have the distribution of $y_{T_i}^i$ as in (37) and (38) (see Appendix 9.1). If we allow for i-dependent σ_0^i but choose common ν and κ parameters we get

$$\begin{aligned} E^{\mathbb{S}^i}\left[V_{T_i}^i\right] &= (\sigma_0^i)^2 \int_0^{T_i} \exp\left(\frac{\nu^2}{\kappa}\left(1 - \mathrm{e}^{-2\kappa t}\right)\right) \mathrm{d}t && (46) \\ E^{\mathbb{S}^i}\left[\left(V_{T_i}^i\right)^2\right] &= 2(\sigma_0^i)^4 \int_0^{T_i} \int_0^t \exp\left(\frac{\nu^2}{\kappa}\left(5 - 4\mathrm{e}^{-2\kappa s} - \mathrm{e}^{-2\kappa(t-s)}\right)\right) \mathrm{d}s\mathrm{d}t. && (47) \end{aligned}$$

As both moments are controlled by the combination of ν and κ there is an extra degree of freedom that will aid in fitting both moments across expiries with a single set of parameters. Note that ν directly effects the level of the moments whereas κ controls the dampening over time. Finally note that the growth of the first moment in T_i is now linear and the second quadratic compared to the exponential growth for these moments in the SABR model.

5.3 The DD-SABR-MR model

The previous two subsections argued that the standard SABR model is not adequate as a choice of model for the level of rates and displayed two types of extensions of the SABR model; deterministic time to expiry dependent instantaneous volatility and mean-reversion. Combining both approaches gives the displaced diffusion SABR with mean-reversion (DD-SABR-MR) model where

each forward swap rate is modelled as

$$\begin{aligned} \mathrm{d}y_t^i &= l^i(T_i - t)\exp(U_t^i)\left(y_t^i + \delta_i\right)\mathrm{d}W_t^i, \quad y_0^i = y_0^i, && (48)\\ \mathrm{d}U_t^i &= -\kappa U_t^i \mathrm{d}t + \nu \mathrm{d}V_t^i, \quad U_0^i = 0, && (49)\\ \mathrm{d}W_t^i \mathrm{d}W_t^j &= \mathrm{d}V_t^i \mathrm{d}V_t^j = \mathrm{d}t, \quad \mathrm{d}W_t^i \mathrm{d}V_t^i = \rho \mathrm{d}t. \end{aligned}$$

under $\mathbb{S}^i$ corresponding to $P_\cdot^i$ as numeraire and with $\delta \in \Re$. Recall at this point the link (35) between the CEV and displaced diffusion formulations. For the remainder of this paper we will continue to refer to the parameter β which means that the corresponding δ_i from equation (35) are being used. Moreover, note from (35) that for large δ β is close to zero ((or indeed zero as $\delta \to \infty$) and hence in this case we use the stochastic volatility Gaussian model (the $\beta = 0$ case in the CEV formulation) with y^i dynamics

$$\mathrm{d}y_t^i = l^i(T_i - t)\exp(U_t^i)\mathrm{d}W_t^i, \quad y_0^i = y_0^i. \tag{50}$$

To get an intuitive feeling for the model note that U mean-reverts around zero implying that the volatility of y_t^i mean-reverts around the function $l^i(T_i - t)$. Indeed, Ito's formula gives

$$\mathrm{d}\sigma_t = \mathrm{d}\exp(U_t) = -\kappa\sigma_t \ln\left(\frac{\sigma_t}{\sigma_0}\right)\mathrm{d}t + \nu\sigma_t \mathrm{d}Z_t + \sigma_t \frac{\nu^2}{2}\mathrm{d}t, \tag{51}$$

which is why we have chosen to denote the model as the SABR with mean-reversion model.

Note that we have chosen to use the displaced diffusion formulation instead of the CEV one in the model. The main reason is that the CEV type local volatility function $\phi(x) = x^\beta$ is quite cumbersome from an implementation point of view. In particular, for $\beta \in (0, 1)$ the behaviour around $x = 0$ is tricky to deal with due to the property of absorption at 0. For example, for the Monte Carlo method the behaviour close to zero needs special (ad-hoc) treatment and, in particular for longer expiries, different (but all reasonable) choices can imply rather different values of certain moments and products.

The displaced diffusion setting leads to a much more stable and effective implementation as in this case one may write out the distribution of each of the y_t^i on explicit form, see Appendix 9.1 for details. The direct disadvantage with this choice is that the domain of the underlying is $(-\delta, \infty)$ implying a possibility of negative rates. Note however that while we will focus on the DD case most results are generic in the sense that they would look very similar if the CEV type formulation would be used instead.

Modeling $l^i(T_i - t)$

In models with deterministic volatility, for example the LIBOR market model, a popular choice of $l^i(T_i - t)$ is to model this by an instantaneous volatility function of the type (see for example Brigo & Mercurio (2006))

$$\hat{l}^i(T_i - t) := (a + b(T_i - t)) \exp(-c(T_i - t)) + d. \tag{52}$$

The function $\hat{l}^i$ links the different rates and induces appropriate time-dependence by calibrating the constants a, b, c and d and is a rather powerful and flexible choice. It is also the choice of Rebonato et al. (2010) in their LMM/SABR model.

However, for our end goal of finding a common stochastic volatility driver for all rates we need the volatility structure to be *separable*. Separability appears in the literature when requiring some high-dimensional system to be represented by a low-dimensional Markov process, see for example Pietersz et al. (2004) for the use of separability to represent an approximate LIBOR market model by a one- or two-dimensional driving Markov process. Separability is in this context achieved if we may represent each of the time-dependent functions $l^i(T_i - t)$ as

$$l^i(T_i - t) = \Lambda(T_i)\lambda_t \tag{53}$$

Putting $b = d = 0$ in (52) gives

$$l^i(T_i - t) = \Lambda(T_i)\lambda_t := a \exp(-cT_i) \exp(ct) \tag{54}$$

which is of separable form. Note that if we are willing to add another factor we could achieve the flexibility of (52) while retaining separability. However, by using market data it is shown in Sections 5.4 and 6 that this is not necessary as (54) does a good job provided the the 'right' choice of β.

To be able to perfectly calibrate to ATM implied volatilities we have furthermore chosen to equip the model with either i-dependent constants k_i or using a piecewise constant function

$$a_t = a_i, \quad T_{i-1} < t \leq T_i. \tag{55}$$

Ideally, for a good model performance all k_i should be close to 1 or all a_i of similar level. While we do believe that the piecewise constant formulation is a better choice in practise the constants k_i setting is good for displaying intuition and behaviour of the model.

Simulation and calibration

The displaced diffusion formulation of the model implies a comparably fast and stable implementation as in this case the distribution of y^j_T may be written out

on explicit form, see Appendix 9.1 for details. This allows for straightforward and accurate simulation as in order to simulate from 0 to T all that is required is simulating the U-process (which has Gaussian increments) and computing a few integrals involving U using numerical quadrature. As is displayed in figure 3.13 in Appendix 9.1 this is much faster and more accurate than an Euler based implementation.

Calibration of the model is outlined in Appendix 9.2. As there is not a good enough approximation of the (CEV/DD)-SABR-MR model available and deriving one is out of the scope of this paper we are using the Monte Carlo method in the calibration routine. Although there is a fast and accurate way to simulate the model, Monte Carlo calibration is of course a slow process and would not be applicable in a live trading environment. Note however that by following the approach of Kennedy et al. (2011) for the DD-SABR model it seems probable that one could derive a good approximation also for the DD-SABR-MR model.

Comment on smile dynamics

Recall that in section 4 we deduced that the smile dynamics of the SABR model with an appropriate β matched the market dynamics quite well. Even though the DD-SABR-MR model is supposedly close to the SABR model in terms of dynamics it is not clear that it would inherit these properties. To address this first note that applying the results of Durrleman (2004) gives that the first order 'small time' approximation (11) is in fact the same in this and the standard SABR model. As argued in Section 4 in connection with the Heston model note that while this means that while the mean-reversion term is not even present in this expression, and hence the expression is not appropriate for pricing, it is a reasonably good approximation for the dynamical relations studied in Section 4. Numerical investigations of the same type as performed when testing the appropriateness of using the approximate 'slope' expression (24) reveals that the 'slope' and 'curvature' dynamics are almost exactly the same in the SABR and the DD-SABR-MR model. Hence we believe that the results and intuition from Section 4 may be transferred to the DD-SABR-MR model.

5.4 Calibration results

This section displays the results from calibrating the DD-SABR-MR model to market data of options to enter 10 years swaps on 27 October 2007 and 9 March 2003 across the same expiries as in Section 5.1.

The calibrated parameters on 27 October 2007 using $\beta = 0, 0.5$ and 1 are given in Tables 3.3 and 3.4 and the calibration errors (market minus model implied

volatility) are given in Figure 3.9. First note that $\beta = 0$ seems to produce a very good fit to the market implied volatilities. For $\beta = 0.5$ the calibration is also very satisfactory whereas for $\beta = 1$ the errors are (although still quite acceptable) about three times larger than for $\beta = 0$. In fact, at this date, for $\beta = 0$ the calibration errors are about the same size as the outcome from standard SABR models calibrated at each expiry individually.

β	a	c	κ	ν	ρ
1	0.122	0.078	0.015	0.283	-0.476
0.5	0.027	0.068	0.050	0.278	-0.233
0	0.006	0.065	0.098	0.317	0.033

Table 3.3: Calibrated parameters of the SABR-MR model on 27 October 2007. 2 to 30 years options to enter 10 year swaps.

	k_i			a_i		
Expiry	$\beta = 0$	$\beta = 0.5$	$\beta = 1$	$\beta = 0$	$\beta = 0.5$	$\beta = 1$
2Y	0.999	1.008	1.016	0.00579	0.02684	0.1241
3Y	0.996	1.004	1.011	0.00574	0.02655	0.1224
4Y	0.996	1.003	1.007	0.00578	0.02659	0.1221
5Y	0.998	1.002	1.004	0.00582	0.02661	0.1216
6Y	0.995	0.995	0.995	0.00571	0.02589	0.1177
7Y	0.997	0.994	0.991	0.00583	0.02636	0.1195
8Y	1.003	0.997	0.992	0.00595	0.02685	0.1216
9Y	1.005	0.997	0.991	0.00588	0.02653	0.1205
10Y	1.005	0.996	0.989	0.00583	0.02636	0.1200
12Y	1.004	0.994	0.988	0.00580	0.02635	0.1205
15Y	1.004	0.997	0.996	0.00582	0.02665	0.1229
20Y	1.000	1.000	1.005	0.00578	0.02668	0.1235
25Y	0.994	0.999	1.003	0.00573	0.02654	0.1224
30Y	0.996	1.008	1.011	0.00579	0.02706	0.1243

Table 3.4: Calibrated k_i and levels of the piecewise constant function a_t for the SABR-MR model on October 27, 2007. Expiry in years. Options to enter 10 year swaps.

Note that with a single set of parameters one would expect the slopes of the implied volatility smiles to be too steep/flat and have too much/little curvature at some expiries. This effect is spotted in Figure 3.9 by studying the plus/minus 2% offset strikes. Moreover, note that at this date the constants k_i are all close to 1 and the piecewise constant function, a_t, is close to constant implying a very good fit to the full implied volatility surface.

As the results look remarkably good it should be noted that this particular date

is one of the better dates in our sample. As a comparison, Figure 3.10 displays the calibration results on 9 March 2003. Also at this date $\beta = 0$ performs best in terms of fitting errors with 2-3 times smaller errors than the other two cases. Moreover, note that for $\beta = 0$ it seems like a single ρ works very well whereas the shorter (longer) expiries would need slightly smaller (larger) ν with the medium term expiries displaying close to zero fitting errors. For $\beta = 1$ one would on the other hand need to change both ρ and ν individually at almost all expiries to effectively reduce the errors.

For this date studying the constants k_i provides more information about appropriate choice of β and why choosing to use (54) instead of (52) as parametrization of l_t is not necessarily particularly restrictive. Figure 3.11 displays the forward swap rates, the ATM implied volatilities and the calibrated k_i for the different β. First note that at this date the ATM implied volatilities are steadily declining until the 15 years expiry where it starts increasing. As, approximately, for $\beta = 1$ the function l_t needs to fit this dependence, the restricted function in (54) is not flexible enough, something which is clearly displayed by the widely varying constants k_i.

However, as seen in Figure 3.11 for $\beta = 0$ or 0.5 the constants k_i are all close to 1. To understand why this is the case recall that in the SABR model (as well as the deterministic CEV model or a matched up displaced diffusion transformation) the ATM volatility is approximately given by $\Sigma = \frac{\sigma_0}{y_0^\beta}$. Hence, there is a direct relation between the parameters σ_0^i, the forward swap rates and the market implied ATM volatilities and it turns out that using an appropriate β stabilizes the σ_0^i and makes the function (54) flexible enough to fit the model.

To provide more information about the suitability of the model the next section displays calibration results across all dates in our sample.

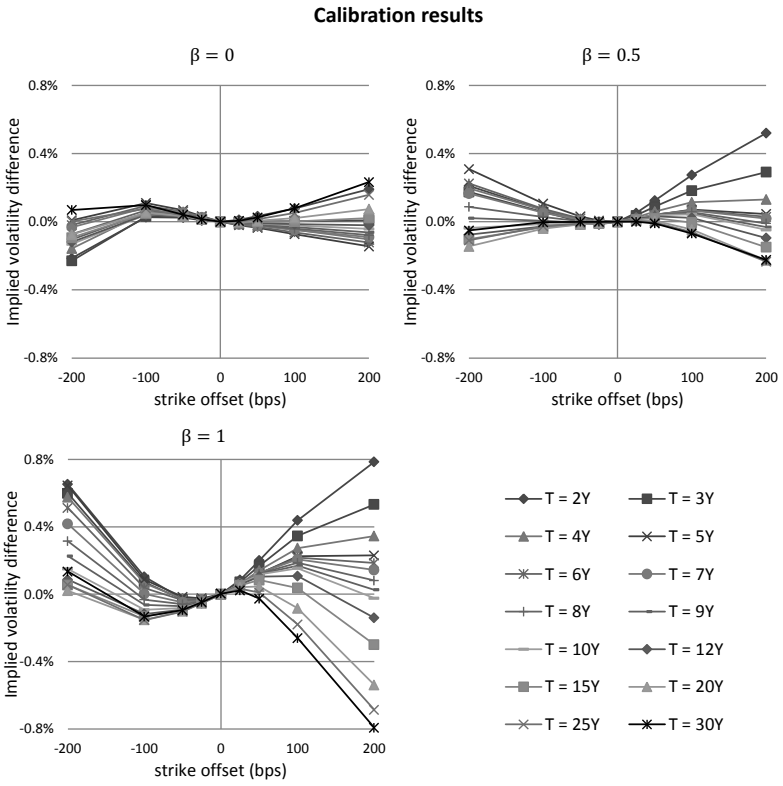

Figure 3.9: Calibration errors (market minus model implied volatilities) for the SABR-MR model on 27 October 2007. 2 to 30 years options to enter 10 year swaps.

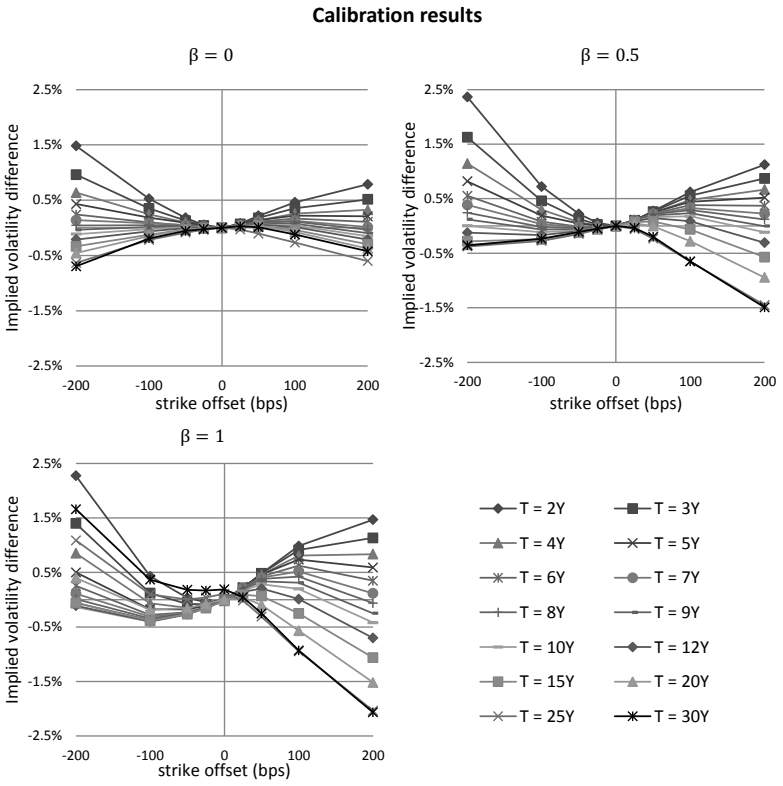

Figure 3.10: Calibration errors (market minus model implied volatilities) for the SABR-MR model on 9 March 2003. 2 to 30 years options to enter 10 year swaps.

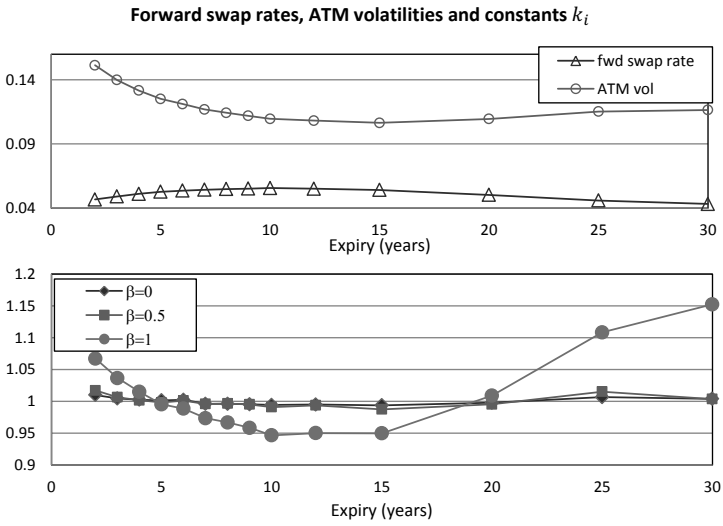

Figure 3.11: The forward swap rates, the ATM implied volatilities and the resulting constants k_i on 9 March 2003. 2 to 30 years options to enter 10 year swaps.

6 Testing the model on all dates

Recall that a well calibrated model has small calibration errors across expiries and strikes as well as close to constant function a_t or all constants k_i close to 1. As seen in the previous section a SABR-MR model with appropriate β seems to be able to fulfill both of of these properties with good accuracy. To test if this is a general feature and whether $\beta = 0$ performs well in general this section tests the calibration performance of the SABR-MR model at the 101 dates plotted in Figure 3.2 covering the period 3 July 2002 to 21 May 2009. As the calibration errors during the turmoil will inevitably be larger and provides a tough challenge for any model the data set is divided into two parts with 89 dates up to turmoil and 12 dates during the turmoil.

As the case $\beta = 1$ is shown in previous sections to be inferior both in terms of smile dynamics and calibration this section only deals with $\beta = 0$ and 0.5. Though $\beta = 0.5$ is not consistent with the value of β we found in Section 4 for the period when the markets were stable it is a value often used by traders which is another reason why we include it here.

6.1 Calibration errors, 'normal' markets

Table 3.5 provides summary statistics of the calibrated k_i values at the first 89 dates up to the start of the turmoil. Note that the means of all k_i (in total 89 dates with 14 expiries implying 1246 values) are close to 1 (as they should given appropriate calibration). Moreover, for both models the standard deviation and the maximum and minimum values are quite close to zero and one respectively implying a sound model from this perspective across all dates. This is an appealing feature as is tells us that the sacrifice of using the simple form for $l^i(T_i - t)$ given in (54) is minor and hence separability may be imposed without sacrificing the calibration performance of the model.

β	mean	stdev	max	min
0	1.0002	0.0108	1.0455	0.9593
0.5	1.0002	0.0113	1.0389	0.9567

Table 3.5: Summary statistics of calibrated constants k_i on 89 dates from 2002 up to the turmoil period. 2 to 30 years options to enter 10 year swaps.

Summary statistics for the calibration errors in implied volatility are given in Table 3.6. As expected the mean is quite close to zero however the standard deviation, mean of the absolute errors and the maximum and minimum errors are a bit smaller for $\beta = 0$ compared with $\beta = 0.5$. Even though plus/minus a couple of percent in implied volatility is not too a too bad fitting note that if only studying the plus/minus 1% strike offset or expiries ranging from 5 to 20 years then all fitting errors are for $\beta = 0$ within 1%.

At first sight the maximum and minimum fitting errors might seem a bit prohibitive. However, as the overall level of the implied volatility surface changes over time it is in relative terms not too bad. The date with the maximum error has ATM implied vols ranging from 15 to 21 percent and while 2.76% is a bit off it is not much worse in relative terms compared with other dates (for example compared with the 9 March 2003 date studied in the previous section). Moreover, the particular point with the error 2.76% is for the 2 year expiry with -2% strike and a market implied volatility of 26%.

Recall that at each date the model is trying to fit $14 \cdot 9 = 126$ data points with rather few parameters. As the parameters a, c and the constants k_i or the function a_t mainly focuses on the ATM part there is in effect only three parameters to play with for the away from the money strikes. In principle, the differences between the market and model implied volatilities are due to either slope or curvature mismatch of the smiles.

Recall that for a fixed β, ρ fits the slope of the smile. If the slopes are 'different' for different expiries one parameter is not enough to fit all expiries perfectly

and the calibration will return an 'average' ρ. At some dates this is not much of an issue whereas at other it is worse. Note that in the second calibration date displayed in Section 5 this is an issue for $\beta = 1$ whereas $\beta = 0$ has basically only curvature errors.

Also recall that the combination ν and κ controls the curvature of the smiles. The parametrisation of the SABR-MR model gives a specific type of control of the curvature that seems to be a decent one at most dates. The $\beta = 0$ case in the calibration on 9 March 2003 displayed in 3.9 is a good example of what is going on with a calibration returning too little curvature for the 2 year smile and too much curvature for the 30 year smile. At some dates this effect is larger, at some dates smaller.

With only three parameters to fit all away from the money strikes we conclude that the SABR-MR does a good job in terms of fitting and seems to be well specified and parametrised. In particular we find no major reason to go into complications about extending the model with time-dependent parameters ρ, κ and ν, a route that has been suggested by some authors.

strike offset	-2	-1	-0.5	-0.25	0	0.25	0.5	1	2
β=0									
mean	-0.05	0.16	0.10	0.05	0.00	0.01	0.03	0.04	0.07
stdev	0.47	0.19	0.09	0.05	0.00	0.04	0.07	0.16	0.34
mean(abs)	0.37	0.16	0.08	0.04	0.00	0.03	0.05	0.12	0.26
max	2.76	1.08	0.43	0.21	0.00	0.19	0.39	0.74	1.59
min	-1.85	-0.51	-0.24	-0.12	0.00	-0.12	-0.15	-0.36	-0.82
β=0.5									
mean	0.15	0.10	0.05	0.01	0.00	0.04	0.08	0.10	0.04
stdev	0.62	0.24	0.10	0.05	0.00	0.05	0.10	0.25	0.55
mean(abs)	0.48	0.18	0.07	0.03	0.00	0.04	0.08	0.18	0.38
max	3.48	1.24	0.47	0.19	0.00	0.23	0.49	0.91	1.69
min	-1.90	-0.56	-0.28	-0.14	0.00	-0.11	-0.27	-0.78	-1.72

Table 3.6: Summary statistics of the errors (market minus model) in implied volatility for calibrated SABR-MR models on 89 dates covering 3 July 2002 up to the start of the turmoil in September 2008. 2 to 30 years options to enter 10 year swaps. Strike offset away from ATM and the implied volatilities are reported in percent. The row mean(abs) refers to the mean of the absolute values of the fitting errors.

6.2 Calibration errors: turmoil period

As discussed earlier the period of market turmoil provides a tough challenge for any interest rate model. While of course models with high number of parameters will perform well in terms of fitting there is a risk for overfitting when market prices are not necessarily internally sound and consistent. On the other hand, while a model with few parameters may struggle in terms of fitting it could be a more solid and stable companion when times are rough.

As observed in Figure 3.2 the yield curve as well as the ATM level, slope and curvature of the implied volatility smiles are almost inverted and some values are very different compared to 'normal' markets. As rates are very low we will disregard from the -2% strike offset since this strike is close to zero implying very high implied volatilities that will affect the calibration and choice of β a bit too much and hence blur the results.

Summary statistics of calibrated constants k_i are given in Table 3.7. Note that in this case the errors are significantly larger (about twice as large) than before and that $\beta = 0$ performs far better than $\beta = 0.5$.

β	mean	stdev	max	min
0	1.004	0.026	1.085	0.962
0.5	1.008	0.063	1.211	0.906

Table 3.7: Summary statistics of calibrated constants k_i on 12 evenly spaced dates during the turmoil period. 2 to 30 years options to enter 10 year swaps.

Summary statistics for the calibration errors in implied volatility are given in Table 3.8. Note that in this period the errors are quite a lot larger than before and that both choices of β has similar performance. Part of the increase in errors are of course due to that the level of implied volatility is larger (note from Figure 3.2 that the 2 and 30 years ATM implied volatilities are about twice as large as on previous dates and moreover for some strikes the implied volatility is above 40%) and actually the relative errors are not that much larger compared with previous results. Given the large magnitude and the vastly different slopes and curvatures of the smiles across expiries compared with previous dates results are perhaps better than one ould expect them to be. However, though we can fit the data using the model the results of Section 4 suggest that it is not consistent with how the market is behaving.

6.3 Calibrated parameters

To further investigate the performance of the SABR-MR Figure 3.12 plots time series of calibrated parameters at the 101 reference dates. Note that although

strike offset	-1	-0.5	-0.25	0	0.25	0.5	1	2
β=0								
mean	-0.08	0.10	0.06	0.00	0.06	0.09	0.05	-0.13
stdev	1.24	0.34	0.13	0.00	0.15	0.25	0.42	0.66
mean(abs)	0.98	0.22	0.08	0.00	0.13	0.22	0.32	0.44
max	1.60	0.67	0.31	0.00	0.68	1.21	1.98	2.87
min	-5.21	-1.33	-0.48	0.00	-0.18	-0.33	-0.62	-1.21
β=0.5								
mean	0.04	0.07	0.04	0.00	0.09	0.09	-0.04	-0.50
stdev	0.97	0.26	0.09	0.00	0.07	0.12	0.36	0.75
mean(abs)	0.74	0.19	0.07	0.00	0.06	0.07	0.24	0.65
max	1.93	0.76	0.33	0.00	0.35	0.29	0.63	1.41
min	-4.66	-0.98	-0.30	0.00	-0.16	-0.42	-1.19	-2.45

Table 3.8: Summary statistics of the errors (market minus model) in implied volatility for calibrated SABR-MR models on 12 evenly spaced dates during the turmoil period. 2 to 30 years options to enter 10 year swaps. Strike offset away from ATM and the implied volatilities are reported in percent. The row mean(abs) refers to the mean of the absolute values of the fitting errors.

parameter stability is one measure of a sound model it is the combination of the parameters that provides the distribution. This means that rather different combinations of some parameters (for example ν and κ combinations) may produce similar distributions and hence one should not put too much attention in simply studying parameter stability. Nevertheless it does provide some further intuition and evidence of model performance.

First note that the parameter variations are very similar for both choices of β. Although parameters do vary over time the variation is not particularly large. For example, for $\beta = 0$ minor tilts of a ν of about 0.33, a κ of 0.08 and a ρ of 0 seems to do the job. Moreover, note that during the turmoil period the parameter values are not much different from earlier which is a bit surprising keeping in mind the vastly different ATM levels, slopes and curvatures of the smiles observed in this period. This suggests that the model does a good job in explaining the implied volatility smiles conditional on the values of the forward swap rates and the instantaneous stochastic volatility.

Finally, note that for about a third of the dates ρ is negative also for $\beta = 0$ and recall that the combination of β and ρ determines the slope of the implied volatility smile. This means that in general one can't put $\rho = 0$ (which may be desired from an implementation point of view) and change β to fit the smile as then the slope of the smile will be too flat.

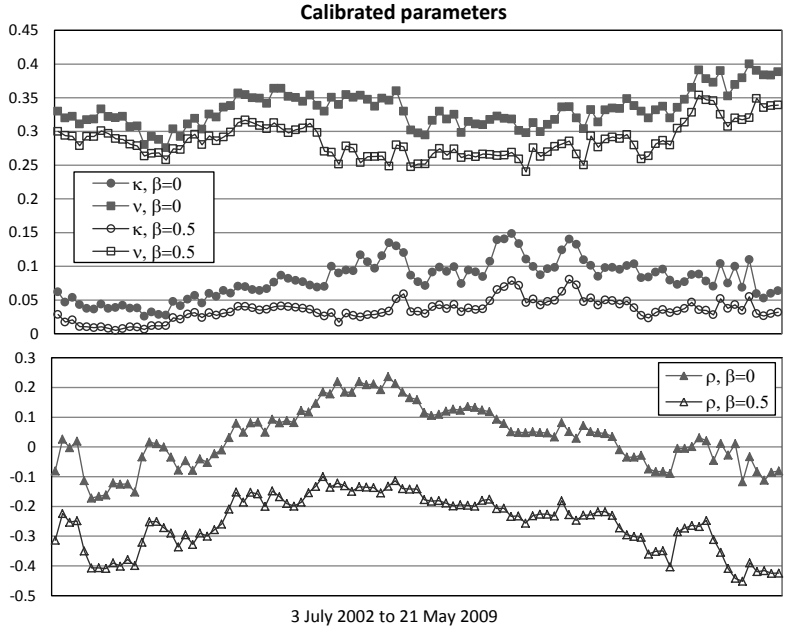

Figure 3.12: Calibrated parameters at 103 evenly spaced dates covering 3 July 2002 to 21 May 2009.

7 A stochastic volatility model for the level of rates and pricing under a single measure

The output of our investigations so far is a model for each ten year swap rate in its own swaption measure that is consistent with market implied distributions for the corresponding swaptions and which is based on two factors and a set of parameters which are common for all swap start dates. In this section our task is to identify a candidate for modelling the level of rates process which is of low dimension. We will specify our candidate process in the measure $\mathbb{S}^n$ correponding to the final expiry T_n. The discussion here is suggestive of a possible candidate process. We test that our simplifying assumptions have led to a choice that is reasonable by using it to recover the prices of vanilla swaptions under the one measure in the following subsection.

In order to understand how to link the swap rates under a single measure we begin by studying the case corresponding to $\beta = 0$ where for each i under $\mathbb{S}^i$ we have

$$\begin{aligned} y_t^i &= y_0^i + a\int_0^t \exp(-c(T_i-u))\exp(U_u^i)\mathrm{d}W_u^i, \\ &= y_0^i + \exp(-c(T_i-t))X_t^i \end{aligned}$$

where W^i, V^i are correlated Brownian motions and U^i is an Ornstein-Uhlenbeck process as specified in equation (49) and we have defined

$$X_t^i := a\int_0^t \exp(-c(t-u))\exp(U_u^i)\mathrm{d}W_u^i. \tag{56}$$

Note that the process X^i in the above equation provides a representation for the spot process at time T_i under $\mathbb{S}^i$, since $y_{T_i}^i = y_0^i + X_{T_i}^i$.

Now further suppose that $\rho = 0$. If we were to specify an expression for the process X^i in some equivalent martingale measure (here the drift of V^i and hence of U^i is left unaltered as we are working with bond based numeraires) we would have

$$\begin{aligned} X_t^i &= a\int_0^t \exp(-c(t-u))\exp(U_u)\mathrm{d}W_u + F_t^i. \\ U_t &= \nu\exp(-\kappa t)\int_0^t \exp(\kappa s)\mathrm{d}V_t, \end{aligned}$$

where (W, V) is a standard Brownian motion under our chosen EMM and F^i is a finite variation process dependent on the particular measure we have chosen. Assuming the common diffusion part is what is significant for modelling suggests a candidate for our level of rates process. Note that if ρ is not equal to

zero moving to the one measure would introduce an i-dependent drift change into the equation for U. Ignoring this complication (that to do this is a reasonable assumption is verified numerically in the next subsection) we take our candidate process for representing the level of rates when $\beta = 0$ as

$$\begin{aligned} \mathrm{d}X_t &= -cX_t\mathrm{d}t + \exp(U_t)\mathrm{d}W_t, \quad X_0 = 0, \\ \mathrm{d}U_t &= \kappa U_t\mathrm{d}t + \nu\mathrm{d}V_t, \\ \mathrm{d}W_t\mathrm{d}V_t &= \rho\mathrm{d}t, \end{aligned}$$

where W and V are correlated Brownian motions under S^n say, with T_n being the final expiry.

Remark 7.1 *Observe that our candidate model is mean reverting. Recall that the short-rate in a one factor Hull-White model is mean reverting and if the instantaneous volatility of the forward rate f_{tT} is of the form $exp(-c(T-t))\sigma_t$ then the instantaneous volatility of the short-rate r is σ_t. It is interesting to note that if we view y^i_t as a discrete analogue of f_{tT_i} and X_t as playing the role r then we have the same relationship between the instantaneous volatility of the two processes as in the Hull-White model despite the presence of stochastic volatility.*

For the case when $\beta \neq 0$ observe that taking a monotonic increasing function of the swap rates will not alter the information about the level of rates. For the CEV form of the model define

$$\begin{aligned} F(x) &= (1-\beta)x^{1-\beta}, \quad 0 \leq \beta < 1, \\ &= \log x, \quad \beta = 1. \end{aligned}$$

and for the displaced diffusion version take

$$F(x) = \log(x + \delta).$$

In the CEV version for each i under $\mathbb{S}^i$ by Ito's formula we have

$$\begin{aligned} F(y^i_t) &= F(y^i_0) + \exp(-c(T_i - t))X^i_t \\ &- \frac{a\beta}{2}\int_0^t (y^i_u)^{\beta-1}\exp(-2(c(T_i-u)\exp(2U^i_u)\mathrm{d}u. \end{aligned}$$

The displaced diffusion version for a given δ gives the same formula for $F(y^i_t)$ as above but with $\beta = 1$. Hence, for the transformed rate we are back at the $\beta = 0$ case albeit with the addition of a finite variation term. Again assuming that the common diffusion part is what is significant for modelling suggests the same candidate for our level of rates process.

Note that we have made several simplifying assumptions in arriving at our candidate process. The next section checks our choice is sensible by using it to recover swaption prices under a single measure.

7.1 Pricing under a single measure

Up to now any model based recovery of swaption prices has been done by using the model for each swap rate in its own swaption measure. Our earlier analysis showed we were able to capture the marginal distribution of each rate with good accuracy using a common set of parameters for all expiries. The swap rates we considered were chosen to have a large overlap so that our end goal of identifying a suitable model for the level of rates could be achieved. Having identified a candidate we now need to test if this process does in fact adequately summarize the state of our economy. In particular the simplifying assumptions we made about the role of the drift need to be verified numerically.

Building a full term structure model based our candidate process is beyond the scope of the paper. Instead we construct a terminal swap rate Markov-functional type model and show that we may recover the vanilla swaption prices under a single joint measure. The purpose here is to motivate that, also in the stochastic volatility case, suitable functional forms may take care of the induced drifts when switching measures and that a model may be constructed in a Markov-functional manner using our identified stochastic volatility driver. Note that though we investigate the suitability of our candidate process using the Markov-functional approach the results of our analysis are general and the proposed driver could be incorporated into a separable LIBOR market model for example.

Suppose we work in the terminal measure $\mathbb{S}^n$ corresponding to P_{T_n} as numeraire. We consider the same date structure as previously with yearly spaced dates up to 30 years, i.e. we take $n = 30$. To simplify things slightly compared with the suggestion made in the previous section we will take as our two-dimensional driver the swap rate process y^{30}, as it, of course, is a drift less process under $\mathbb{S}^{30}$.

Note that under $\mathbb{S}^n$ swaption prices are given by

$$C_0^i(K) = P_0^n E^{\mathbb{S}^n}\left[\frac{P_{T_i}^i\left(y_{T_i}^i - K\right)^+}{P_{T_i}^n}\right] \tag{57}$$

In order to calculate these prices for each i we need to model $y_{T_i}^i$ and $\frac{P_{T_i}^i}{P_{T_i}^n}$ under $\mathbb{S}^n$. For the stochastic volatility Gaussian case (i.e. $\beta = 0$) we take

$$y_{T_i}^i = A_i y_{T_i}^n + B_i, \tag{58}$$

and

$$\frac{P_{T_i}^i}{P_{T_i}^n} = (1 + c_i y_{T_i}^n + d_i)^{T_n - T_i}, \tag{59}$$

Note that the equation for $y^i_{T_i}$ follows from a first order Taylor expansion for $y^i_{T_i}$ which is viewed as a function of our driver at T_i and the second equation is what we would get if we assumed that $D_{0T_j} = (1 + cy + d)^{-T_j}$. For the displaced diffusion case we take

$$\log\left(y^i_{T_i} + \delta\right) = A_i \log(y^n_{T_i} + \delta) + B_i \tag{60}$$

To find the constants A_i, B_i, c_i and d_i we solve the system of equations

$$E^{\mathbb{S}^n}\left[\frac{P^i_{T_i}}{P^n_{T_i}}\right] = \frac{P^i_0}{P^n_0} \tag{61}$$

$$E^{\mathbb{S}^n}\left[\frac{y^i_{T_i} P^i_{T_i}}{P^n_{T_i}}\right] = \frac{y^i_0 P^i_0}{P^n_0}, \tag{62}$$

$$C^i_0(K_1) = P^n_0 E^{\mathbb{S}^n}\left[\frac{P^i_{T_i}}{P^n_{T_i}}\left(y^i_{T_i} - K_1\right)^+\right], \tag{63}$$

$$C^i_0(K_2) = P^n_0 E^{\mathbb{S}^n}\left[\frac{P^i_{T_i}}{P^n_{T_i}}\left(y^i_{T_i} - K_2\right)^+\right]. \tag{64}$$

The first two equations follow from the martingale property for rebased assets and are hence crucial in order to assure that discount factors and swap rates are recovered. The last two equations fits the model to swaption prices at two different strikes where we have chosen to use the ATM and the ATM - 1% strikes. Intuitively the last two equations can be seen as required in order to make up for the two-dimensional drift changes, although we have made no attemp to formalize this.

In the test the calibrated parameters of the DD-SABR-MR model on 27 October 2007 reported in Table 3.3 are used and we take all constants a_i and k_i as 1. To evaluate the model we calibrate the parameters A_i, B_i, c_i, d_i to the expiries $T_i = 2, 5, 10, 15, 20, 25$ and 30 years and compare the prices of swaptions across strikes with the prices coming from the DD-SABR-MR model. The expectations are computed using the Monte Carlo method as outlined in Section 9.1. In order to make sure that the statistical error is negligible we have used 200 000 (Sobol) paths and 300 steps out to each expiry for the stochastic volatility process.

The differences between prices converted into implied volatility basis points between the DD-SABR-MR model under each of the $\mathbb{S}^i$ and the terminal swap rate model under $\mathbb{S}^n$ are given in Table 3.9. Note that the size of the differences are extremely small and that the largest difference at the 2 year swaption at strike ATM - 2% is only about 0.05% in implied volatility. Also note that the case $\beta = 0$ works a bit better than $\beta = 0.5$.

Note that as the terminal swap rate model is set up as a rather crude approximation of the forward swap rates these results are quite remarkable. Even though there are four parameters that are used to remove arbitrage and match up two strikes it is interesting to note that, as far as swaption pricing across the ATM plus/minus 2% strikes goes, the model is providing correct marginal distributions for each of the forward swap rates under the joint measure $\mathbb{S}^n$. This gives good hope for the construction of a full term structure model based on our identified driver.

Expiry	-2%	-1%	-0.5%	-0.25%	ATM	0.25%	0.5%	1%	2%
$\beta = 0$									
2Y	-0.41	0.00	0.01	0.01	0.00	-0.01	-0.02	-0.04	-0.07
5Y	-0.08	0.00	0.01	0.00	0.00	0.00	-0.01	-0.02	-0.04
10Y	-0.03	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.02
15Y	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01
20Y	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01
25Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\beta = 0.5$									
2Y	-4.94	0.00	0.26	0.14	0.00	-0.22	-0.44	-0.89	-1.76
5Y	-0.71	0.00	0.07	0.07	0.00	-0.05	-0.18	-0.53	-1.25
10Y	-0.32	0.00	0.09	0.09	0.00	-0.11	-0.24	-0.64	-1.64
15Y	-0.25	0.00	0.13	0.04	0.00	-0.08	-0.14	-0.43	-0.87
20Y	0.01	0.00	-0.02	-0.08	0.00	0.07	0.18	0.34	0.99
25Y	0.09	0.00	-0.10	-0.11	0.00	0.15	0.38	0.85	2.17
30Y	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 3.9: Differences in implied volatility basis points between the DD-SABR-MR swaption prices across strikes and the terminal Markov-functional model using y^{30} as driver.

8 Conclusion

This paper has used an extensive set of historical market data of swap rates and swaptions to identify a two-dimensional stochastic volatility process for the level of rates. The procedure of the paper has been two identify this process step by step by increasing the requirement of the model and discuss how to adjust the process to take this into account.

We started off in Section 4 by identifying a suitable time-homogeneous model for swap rates at their setting dates and it turned out that the SABR model with $\beta = 0$ satisfactorily passed all our tests and that also $\beta = 0.5$ seemed to

perform reasonably well. Section 5 then identified a model of ten year swap rates under their own measures based on common parameters for all swap start dates and Section 6 tested it on market data covering 3 July 2002 to 21 May 2009. The performed tests displayed that the model is a stable and flexible choice that allows for good calibration across expiries and strikes. Section 7 identified and put into context a time-homogeneous candidate stochastic volatility process that can be used as a driver for all swap rates. Finally, Section 7.1 used this to construct a simple terminal Markov-functional type model under a single measure and displayed that prices of swaptions across strikes and expiries may be recovered with good accuracy.

Recall that the ten year swap rates were chosen in order to have some overlap between the rates and hence providing hope for that finding a model of the above type would be possible. We have also performed the above tests on the 5, 20 and 30 year swap rates and have observed similar results as the ten year case. Moreover, we have calibrated the model to co-terminal swaptions, as is required when pricing some Bermudan swaption type products, and our findings suggests that also this may be done with good accuracy.

For future work we expect that the DD-SABR-MR model would make a good building block for a stochastic volatility LIBOR market model. Moreover, as we have put some considerations at keeping dimensions low and allowing for separability we believe it would be a suitable driver of a stochastic volatility Markov-functional model.

9 Appendix

9.1 Monte Carlo implementation

This section outlines a comparably fast and efficient way of implementing the DD-SABR-MR model. Recall the dynamics (without the constants k_i or piecewise constant function a_t for clarity) under each of the measures $\mathbb{S}^i$

$$\begin{aligned} \mathrm{d}\log\left(y_t^i+\delta\right) &= F(l_t^i,U_t)(y_t^i+\delta)\mathrm{d}W_t^i - \frac{F^2(l_t^i,U_t^i)}{2}\mathrm{d}t, && (65)\\ \mathrm{d}U_t^i &= -\kappa U_t^i \mathrm{d}t + \nu \mathrm{d}Z_t^i, && (66)\\ \mathrm{d}W_t^i \mathrm{d}Z_t^i &= \rho \mathrm{d}t, && (67) \end{aligned}$$

with

$$\begin{aligned} l_t^i &= a\exp\left(-c(T_i-t)\right), && (68)\\ F(l,u) &= l\exp(u). && (69) \end{aligned}$$

First note that W_t^i may be rewritten as $W_t^i = \rho Z_t^i + \sqrt{1-\rho^2}B_t^i$, where B^i is a Brownian motion independent of W^i and Z^i. This gives

$$\begin{aligned}
\mathrm{d}\log\left(y_t^i+\delta\right) &= F(l_t^i,U_t^i)\left(\rho \mathrm{d}Z_t + \sqrt{1-\rho^2}\mathrm{d}B_t^i\right) - \frac{F^2(l_t^i,U_t^i)}{2}\mathrm{d}t \\
&= \frac{\rho}{\nu}F(l_t^i,U_t^i)\left(\mathrm{d}U_t^i + \kappa U_t^i \mathrm{d}t\right) + F(l_t^i,U_t^i)\sqrt{1-\rho^2}\mathrm{d}B_t^i \\
&- \frac{F^2(l_t^i,U_t^i)}{2}\mathrm{d}t \\
&= \frac{\rho}{\nu}F(l_t^i,U_t^i)\mathrm{d}F(l_t^i,U_t^i) - \frac{\rho}{\nu}cF(l_t^i,U_t^i)\mathrm{d}t - \frac{\rho\nu}{2}F(l_t^i,U_t^i)\mathrm{d}t \\
&+ \frac{\rho\kappa}{\nu}F(l_t^i,U_t^i)U_t^i\mathrm{d}t - \frac{F^2(l_t^i,U_t^i)}{2}\mathrm{d}t \\
&+ F(l_t^i,U_t^i)\sqrt{1-\rho^2}\mathrm{d}B_t^i. \qquad (70)
\end{aligned}$$

Now note that conditional on the σ-algebra generated by Z^i over $[0,T]$ the $\mathbb{S}^i$ conditional moment generating function of

$$\int_0^T F(l_t^i,U_t^i)\mathrm{d}B_t^i$$

is equivalent to that of

$$\sqrt{V^i(0,T)}G,$$

with $V^i(t,T) := \int_t^T F^2(l_s^i,U_s^i)\mathrm{d}s$ and G a standard normal random variable that is independent of B^i. By the tower property this holds also for the unconditional $\mathbb{S}^i$ moment generating functions and hence

$$\int_t^T F(l_s^i,U_s^i)\sqrt{1-\rho^2}\mathrm{d}B_s^i \overset{d}{=} \sqrt{1-\rho^2}\sqrt{V^i}G. \qquad (71)$$

Finally, integrating (70) from $t \geq 0$ to $T \leq T_i$ and again comparing the moment generating functions under $\mathbb{S}^i$ and using the tower law it is possible to show that

$$\begin{aligned}
y_T^i &\overset{d}{=} (y_t^i+\delta)\exp\left(\frac{\rho}{\nu}\left(F(l_T,U_T^i) - F(l_t,U_t^i)\right) + \rho I^i(t,T)\right. \\
&- \left.\frac{1}{2}V^i(t,T) + \sqrt{1-\rho^2}\sqrt{V^i(t,T)}G\right) - \delta, \qquad (72)
\end{aligned}$$

with

$$I^i(t,T) = \int_t^T \left(\frac{\kappa}{\nu}F(l_s^i,U_s^i)U_s^i - \frac{c}{\nu}F(l_s^i,U_s^i) - \frac{\nu}{2}F(l_s^i,U_s^i)\right)\mathrm{d}s. \qquad (73)$$

Hence, to simulate a realisation of y^i_T one only needs one standard normal random variable plus the integrals I^i and V^i along a particular path of U^i. For a simulated path of U^i the integrals $I^i(t,T)$ and $V^i(t,T)$ may be computed by numerical quadrature using, for example, the Simpson rule.

Moreover, note that conditional on the full path of U from 0 to T_i, $y^i_{T_i}$ is a displaced lognormal random variable and hence prices of swaptions are given in closed form. Hence, in a Monte Carlo implementation to price swaptions it is only needed to simulate realisations of U^i, compute the values of F^i, I^i and V^i using numerical quadrature and finally insert in the displaced diffusion Black formula. Finally, note that for $\beta = 0$ the above calculations goes through with just a minor modification.

To get an idea of the accuracy of using this method compared with the (log) Euler method Figure 3.13 displays prices of a 10 year ATM swaption computed using different number of steps per year. Note that while the Euler method requires lots of steps per year to get rid of the discretisation error, the new method allows for very long steps.

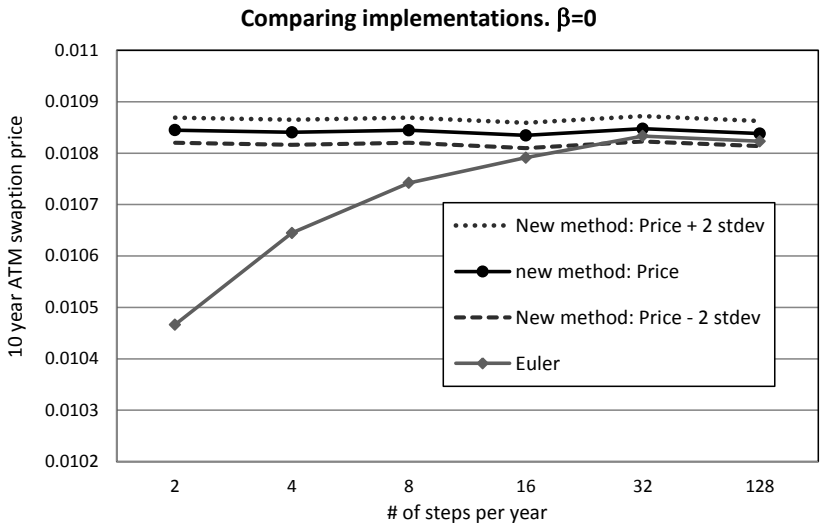

Figure 3.13: Comparing the new implementation method of y^i with the standard Euler method for pricing a 10 year ATM swaption. For the one step method the prices plus/minus 2 standard deviations are also reported. 4 million pseudo-random paths are used implying a very small statistical error. $\beta = 0$, $a = 0.009$, $c = 0.1$, $\kappa = 0.05$, $\nu = 0.36$ and $\rho = 0.2$.

9.2 Calibration

To calibrate the SABR-MR model we are minimising the squared relative error between the market and model implied volatilities at all strikes and expiries. Note that while in principle there are only five parameters to calibrate $(a, c, \kappa, \nu, \rho)$ we also need to find the parameters of the function a_t or the constants k_i. As a full global optimisation of all parameters are too time-consuming to perform in practise we perform the calibration in two steps

1. First calibrate $(a, c, \kappa, \nu, \rho)$. Since in the end we will make sure that the ATM levels are perfectly matched the goal function is structured such that first the ATM errors are found, then the model is adjusted to fit the ATM perfectly by using the constants k_i and finally the errors at the other strikes are found.

2. In the second step we found the piecewise constants needed for a_t and the constants k_i subsequently starting from time T_1.

As there is not a good enough approximation of the SABR-MR model available and deriving one is out of the scope of this paper we are using the Monte Carlo method in the calibration routine. To simulate $y^i_{T_i}$ we are using the method outlined in the previous subsection with 10 000 Sobol paths and 30 steps out to each expiry. Note that in principle it is the number of points between time 0 and expiry that is important, not the number of steps per year and we have found the 30 steps gives good accuracy for all expiries. Compared with more accurate choices that reduces the numerical errors we have found that the above choices are correct within a few basis points. On a standard computer this means that a full implied volatility surface of 14 expiries and 9 strikes per expiry is priced in about 1 second and a calibration is made in at most a few minutes.

While this is of course not fast enough to be used in a market environment note that the above structure gives hope of finding a good approximation for calibration, see Kennedy et al. (2011) for a similar approximation for the standard SABR model.

Bibliography

Albanese, C. & Trovato, M. (2005), 'A Stochastic Volatility Model for Swaptions and Callable Swaps', Working paper, Imperial College.

Andersen, L. & Andreasen, J. (2002), 'Volatile Volatilities', *Risk* **15**(12), 163–168.

Andersen, L. & Brotherton-Ratcliffe, R. (2005), 'Extended LIBOR Market Models with Stochastic Volatility', *Journal of Computational Finance* **9**(1), 1.

Andersen, L. & Piterbarg, V. (2010), *Interest Rate Modeling*, 1 edn, Atlantic Financial Press.

Bennett, M. & Kennedy, J. (2005), 'A Comparison of Markov-Functional and Market Models: The One-Dimensional Case', *The Journal of Derivatives* .

Beveridge, C., Denson, N. & Joshi, M. (2009), 'Comparing Discretisations of the Libor Market Model in the Spot Measure', *Australian Actuarial Journal* **15**(2), 231.

Bjork, T. (2004), *Arbitrage Theory in Continuous Time*, 2 edn, Oxford University Press.

Brace, A., Gatarek, D. & Musiela, M. (1997), 'The Market Model of Interest Rate Dynamics', *Mathematical Finance* **7**(2), 127–147.

Brigo, D. & Mercurio, F. (2006), *Interest Rate Models - Theory and Practice*, 2 edn, Springer.

Cheyette, O. (1996), 'Markov Representation of the Heath-Jarrow-Morton Model', *SSRN eLibrary* .
URL: `http://ssrn.com/paper=6073`

Durrleman, V. (2004), *From Implied to Spot Volatilities*, PhD Thesis, Princeton University.

Fouque, J., Papanicolaou, G. & Sircar, K. (2000), *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press.

Fries, C. & Joshi, M. (2008), 'Partial Proxy Simulation Schemes for Generic and Robust Monte-Carlo Greeks', *Journal of Computational Finance* **11**(3), 79.

Fries, C. & Rott, M. (2004), 'Cross-Currency and Hybrid Markov-Functional Models', *SSRN eLibrary* .
URL: `http://ssrn.com/paper=d=532122`

Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer.

Hagan, P., Kumar, D., Lesniewski, A. & Woodward, D. (2002), 'Managing Smile Risk', *Wilmott Magazine* (Sep/Oct), 84–108.

Henrard, M. (2005), 'Swaptions: 1 price, 10 Deltas, and...6 1/2 Gammas', *Wilmott Magazine* (November), 48–57.

Heston, S. (1993), 'A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options', *Review of Financial Studies* **6**(2), 327–343.

Hull, J. & White, A. (1990), 'Pricing Interest-Rate-Derivative Securities', *Review of Financial Studies* **3**, 573–592.

Hunt, P. & Kennedy, J. (2000), *Financial Derivatives in Theory and Practice*, Wiley.

Hunt, P., Kennedy, J. & Pelsser, A. (2000), 'Markov-Functional Interest Rate Models', *Finance & Stochastics* .

Hunter, C. J., Jackel, P. & Joshi, M. S. (2001), 'Drift Approximations in a Forward-Rate-Based LIBOR Market Model', *Risk magazine* **14**, 81–84.

Jackel, P. & Kahl, C. (2007), 'Hyp Hyp Hooray', Working paper.
URL: `http://www.jaeckel.org`

Jaeckel, P. (2002), *Monte Carlo Methods in Finance*, Wiley.

Jamshidian, F. (1997), 'LIBOR and Swap Market Models and Measures', *Finance and Stochastics* **1**, 293–330.

Joshi, M. & Stacey, A. M. (2006), 'New and Robust Drift Approximations for the Libor Market Model', *SSRN eLibrary* .
URL: `http://ssrn.com/paper=907385`

Kaisajuntti, L. & Kennedy, J. E. (2011), 'An N-Dimensional Markov-Functional Interest Rate Model', *SSRN eLibrary* .
URL: `http://ssrn.com/paper=1081337`

Kennedy, J., Mitra, S. & Pham, D. (2011), 'On the Approximation of the SABR Model: a Probabilistic Approach', Draft paper, University of Warwick.

Marris, D. (1999), 'Financial Option Pricing and Skewed Volatility', M. Phil. Thesis, University of Oxford.

Miltersen, K., Sandmann, K. & Sondermannn, D. (1997), 'Closed-Form Solutions for Term Structure Derivatives with Lognormal Interest Rates', *Journal of Finance* pp. 409–430.

Pietersz, R. (2005), 'Importance Sampling in Market Models for Targeted Accrual Redemption Notes', Working paper, ABN AMRO.

Pietersz, R., Pelsser, A. & Van Regenmortel, M. (2004), 'Fast Drift-Approximated Pricing in the BGM Model', *Journal of Computational Finance* **8**, 93–124.

Piterbarg, V. (2004), 'TARNs: Models, Valuation, Risk Sensitivities', *Wilmott Magazine* .

Piterbarg, V. (2005), 'Stochastic Volatility Model with Time-Dependent Skew', *Applied Mathematical Finance* **12**(2), 147–185.

Press, W., Teukolsky, S., Vetterling, W. & Flannery, B. (2002), *Numerical Recipes in C++*, 2 edn, Cambridge University Press.

Rebonato, R. (2002), *Modern Pricing of Interest-Rate Derivatives: the LIBOR Market Model and Beyond*, Princeton University Press.

Rebonato, R. (2004), *Volatility and Correlation*, 2 edn, Wiley.

Rebonato, R., McKay, K. & White, R. (2010), *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, Wiley.

Trolle, A. & Schwartz, E. (2010), An Empirical Analysis of the Swaption Cube, Technical report, National Bureau of Economic Research.

Zhao, X. & Glasserman, P. (2000), 'Arbitrage-Free Discretization of Lognormal Forward Libor and Swap Rate Models', *Finance and Stochastics* **4**(1), 35–68.

The Stockholm School of Economics

A complete publication list can be found at *www.hhs.se/research/publications*. Books and dissertations are published in the language indicated by the title and can be ordered via e-mail: *publications@hhs.se*.

A selection of our recent publications

Books

Alexandersson, Gunnar. (2011). *Den svenska buss- och tågtrafiken: 20 år av avregleringar*. Forskning i fickformat.

Barinaga, Ester (2010). *Powerful dichotomies.*

Engwall, Lars (2009). *Mercury meets Minerva: business studies and higher education: the Swedish case.*

Ericsson, Daniel (2010). *Den odöda musiken.*

Ericsson, Daniel (2010). *Scripting Creativity.*

Holmquist, Carin (2011). *Kvinnors företagande – kan och bör det öka?*

Holmquist, Carin (2009). *Entreprenörskap på riktigt: teoretiska och praktiska perspektiv.*

Lundeberg, Mats (2011). *Improving business performance: a first introduction.*

Melén, Sara (2010). *Globala från start: småföretag med världen som marknad.* Forskning i Fickformat.

Modig, Niklas & Åhlström, Pär. (2011). *Vad är lean?*

Mårtensson, Pär & Mähring, Magnus (2010). *Mönster som ger avtryck: perspektiv på verksamhetsutveckling.*

Ottosson, Mikael (2011). *Skogsindustrin och energiomställningen*. Forskning i fickformat.

Sjöström, Emma (2010). *Ansiktslösa men ansvarsfulla*. Forskning i fickformat.

Wijkström, Filip (2010). *Civilsamhällets många ansikten.*

Dissertations

Bjerhammar, Lena (2011). *Produktutvecklingssamarbete mellan detaljhandelsföretag och deras varuleverantörer*.

Bohman, Claes (2010). *Attraction: a new driver of learning and innovation*.

Bottero, Margherita (2011). *Slave trades, credit records and strategic reasoning: four essays in microeconomics.*

Bång, Joakim (2011). *T Essays in empirical corporate finance.*

Carlsson-Wall, Martin (2011). *Targeting target costing: cost management and inter-organizational product development of multi-technology products*.

Freier, Ronny (2011). *Incumbency, divided government, partisan politics and counciu size: essays in local political economics*.

Fries, Liv (2011). *Att organisera tjänstesektorns röst: en teori om hur organisationer formar aktörer*.

Glassér, Charlotte (2010). *The fountainhead of innovation health: a conceptualization & investigation.*

Hemrit, Maetinee (2011). *Beyond the Bamboo network: the internationalization process of Thai family business groups.*

Lundmark, Martin (2011). *Transatlantic defence industry integration: discourse and action in the organizational field of the defence market*.

Lundvall, Henrik (2010). *Poverty and the dynamics of equilibrium unemployment: [essays on the economics of job search, skills, and savings]*.

Migueles Chazarreta, Damián (2011). *The implications of trade and offshoring for growth and wages.*

Monsenego, Jérôme (2011). *Taxation of foreign business income within the European internal market: an analysis of the conflict between the objective of achievement of the European internal market and the principles of territoriality and worldwide taxation.*

Ranehill, Eva (2011). *Essays on gender, competition and status*.

Runsten, Philip (2011). *Kollektiv förmåga: en avhandling om grupper och kunskapsintegration.*

Stepanok, Ignat (2011). *Essays on international trade and foreign direct investment*.

Strid, Ingvar (2010). *Computational methods for Bayesian inference in macroeconomic models.*

Söderblom, Anna (2011). *Private equity fund investing: investment strategies, entry order and performance.*