CALIBRATION OF CHAOTIC MODELS FOR INTEREST RATES

M. R. GRASSELLI* AND T. TSUJIMOTO †

Abstract. In this paper we calibrate chaotic models for interest rates to market data using a polynomial–exponential parametrization for the chaos coefficients. We identify a subclass of one–variable models that allow us to introduce complexity from higher order chaos in a controlled way while retaining considerable analytic tractability. In particular we derive explicit expressions for bond and option prices in a one–variable third chaos model in terms of elementary combinations of normal density and cumulative distribution functions. We then compare the calibration performance of chaos models with that of well–known benchmark models. For term structure calibration we find that chaos models are comparable to the Svensson model, with the advantage of guaranteed positivity and consistency with a dynamic stochastic evolution of interest rates. For calibration to option data, chaos models outperform the Hull and White and rational lognormal models and are comparable to LIBOR market models.

Key words. Positive interest rate models, Wiener chaos, model calibration

JEL classification. E43

AMS subject classification. 91G30, 91G20, 91G70

1. Introduction. The purpose of this paper is to investigate the calibration performance of interest rate models based on the Wiener chaos expansion. Chaotic models were introduced in [16] as an axiomatic framework for interest rates satisfying both no arbitrage and positivity conditions, following the line of research initiated by [11], where zero–coupon bond prices are modeled as

$$P_{tT} = \frac{\int_T^\infty M_{ts} ds}{\int_t^\infty M_{ts} ds}, \tag{1.1}$$

for a parametrized family of positive martingales M_{ts}. It was then shown in [24] and [26] that one can focus instead on modeling a supermartingale V_t which is related to the martingales M_{ts} through

$$V_t = \int_t^\infty M_{ts} ds, \tag{1.2}$$

The chaotic approach is derived from the observation that V_t can itself be written as the conditional variance of a terminal random variable X_∞. This square integrable random variable then has a unique Wiener chaos decomposition, and studying the way its expansion coefficients affect the corresponding interest rates model becomes the subject of the theory.

Notice the successive simplifications in the class of objects that one needs to model: from an entire family of martingales M_{ts} in the Flesaker–Hughston framework, to the process V_t in the potential approach, and finally the single random variable X_∞ in chaotic models. Arguably, each step of the way reduces the arbitrariness in the modeling exercise, with presumed advantages for calibration to real data. In particular, the chaos expansion allows one to successively introduce randomness in the model

*Department of Mathematics and Statistics, McMaster University, Hamilton L8S 4K1, ON, Canada. Supported by the Natural Sciences and Engineering Research Council of Canada and MITACS.

†School of Economics and Finance, University of St Andrews, Fife KY16 9AL, Scotland, UK. Supported by the Scottish Institute for Research in Economics.

in a way that can, in principle, be made to correspond to the increased complexity of financial instruments under consideration. Such are the types of statements that we propose to put to test in this paper.

Models in the Flesaker–Hughston framework have been implemented and calibrated, for example, in [5] and [14], whereas implementations of the potential approach can be found in [19] and [25]. To our knowledge, we present here the first practical implementation and calibration to market data of chaotic interest rate models. We adopt the day-by-day calibration methodology used in [12], [18], and [23]. The motivation for this is to capture the prices of liquid interest rate derivatives such as caps and swaptions by a model as parsimonious as possible, which can then be used for pricing and hedging of exotic options, such as the Chooser flexible cap and Bermudan Swaption.

After reviewing the chaotic approach in Section 2, we implement to two separate calibration exercises. In Section 3 we consider the calibration of chaotic models to the observed term structure of interest rates. For comparison, we use the Nelson–Siegel [21] and Svensson [28] models for forward rates as benchmarks. These so called descriptive models [4] are examples of a general exponential–polynomial class of models analyzed in [1], which we use as motivation for the parametric form we adopt for the chaos coefficient functions. We find in this section that chaotic models perform comparably to descriptive models with the same number of parameters, with the advantage of avoiding problems with positivity and consistency.

We then move to a full calibration to yields and option prices in Section 4. We recall known expressions for option prices in a second chaos model, derive the corresponding formulas in a third chaos model, and calibrate them to two separate data sets in three different ways. For comparison, we also calibrate the Hull and White model [17], the rational lognormal proposed in [11] and a lognormal LIBOR market model [12]. We find that chaos models generally perform much better than the Hull and White and rational lognormal models and have fitting errors comparable to those obtained with a LIBOR model. When we apply an information criterion that takes into account the different number of parameters in the models, we find that one of our third chaos models consistently outperforms the LIBOR market model. Section 5 concludes the paper by summarizing the results and pointing to future research in the area.

2. The chaotic approach. We review the framework proposed in [16], whereby a general positive interest rate model with no arbitrage opportunities is associated with a square integrable terminal random variable X_∞, which can then be modeled using a Wiener chaos expansion.

2.1. Basic definitions. Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with the standard augmented filtration $(\mathcal{F}_t)_{0\le t<\infty}$ generated by a k–dimensional Brownian motion W_t and suppose that $X_\infty \in L^2(\Omega, \mathcal{F}, P)$ is a random variable with the property that X_∞ is not $\mathcal{F}_t$–measurable for any finite value of t and $E[X_\infty] = 0$. Using the martingale representation theorem, we can write

$$X_\infty = \int_0^\infty \sigma_s dW_s \tag{2.1}$$

for an adapted process σ_t. Denoting the conditional variance of X_∞ with respect to the σ–algebra $\mathcal{F}_t$ by

$$V_t = E\left[(X_\infty - E_t[X_\infty])^2 \middle| \mathcal{F}_t \right], \tag{2.2}$$

it follows from the Itô isometry that

$$V_t = E\left[\left(\int_t^\infty \sigma_s dW_s\right)^2 \middle| \mathcal{F}_t\right] = E\left[\int_t^\infty \sigma_s^2 ds \middle| \mathcal{F}_t\right]. \tag{2.3}$$

Defining the increasing adapted process

$$A_t = \int_0^t \sigma_s^2 ds. \tag{2.4}$$

we have that

$$V_t = E[A_\infty - A_t|\mathcal{F}_t], \tag{2.5}$$

from which it follows that V_t is a potential, that is, a positive supermartingale satisfying $E[V_t] \to 0$ as $t \to \infty$. It is then well–known (see, for example, [24]) that V_t can be used as a state–price density for an arbitrage–free interest rate model in which the price of a zero–coupon bond with maturity T is given by

$$P_{tT} = \frac{E\left[V_T|\mathcal{F}_t\right]}{V_t} = \frac{Z_{tT}}{Z_{tt}}, \qquad 0 \le t \le T < \infty, \tag{2.6}$$

where

$$Z_{tT} = E[V_T|\mathcal{F}_t], \qquad 0 \le t \le T < \infty. \tag{2.7}$$

It follows from (2.5) that for each fixed t the processes Z_{tT}, and consequently the bond prices P_{tT}, are decreasing functions of the maturity T, which in turn implies that all instantaneous forward rates

$$f_{tT} = -\frac{\partial}{\partial T} \log P_{tT} \tag{2.8}$$

are automatically positive in this framework. Furthermore, the short–rate r_t and the market price of risk vector λ_t are determined by the dynamics of V_t as follows:

$$dV_t = -r_t V_t dt - \lambda_t V_t dW_t, \qquad V_0 > 0. \tag{2.9}$$

For later use, we introduced the family of positive martingales

$$M_{ts} = E[\sigma_s^2|\mathcal{F}_t], \qquad 0 \le t \le s < \infty, \tag{2.10}$$

so that V_t can be written as in (1.2) and bond prices as in (1.1).

Returning to the squared integrable random variable X_∞, the Wiener chaos decomposition (see [22] and [31]) says that it can be written as the L^2–convergent sum

$$\begin{aligned} X_\infty = &\int_0^\infty \phi_1(s) dW_s + \int_0^\infty \int_0^s \phi_2(s, s_1) dW_{s_1} dW_s \\ &+ \int_0^\infty \int_0^s \int_0^{s_1} \phi_3(s, s_1, s_2) dW_{s_2} dW_{s_1} dW_s + \cdots, \end{aligned} \tag{2.11}$$

for deterministic square–integrable functions $\phi_1(s), \phi_2(s, s_1), \phi_3(s, s_1, s_2), \ldots$ called the chaos coefficients of X_∞. We say that an interest rate model is an n-th chaos

model if the decomposition for the random variable X_∞ can be completely determined by its first n coefficient functions.

The basic insight of the chaotic approach is that the decomposition above provides a way to add complexity to an interest rate model in a controlled manner. For example, as we will see in the next section, the first order chaos correspond to deterministic interest rate models, whereas the second order chaos give rise to stochastic interest rate models with randomness governed by a parametric family of Gaussian processes.

More importantly for calibration purposes, the increased complexity in the interest models should be related to the instruments that are available in the market. In what follows, we propose a systematic way to calibrate chaotic interest models starting with bond prices alone and then gradually increasing the complexity of market instruments included in the calibration. Our general strategy will consist of choosing general parametric forms for the deterministic functions ϕ_i, $i = 1, 2, \ldots$ and then fit the parameters to market data according to the bond and option pricing formulas emerging from the models.

2.2. Option Pricing. We now describe how the prices of the most common interest rate options can be written in terms of the quantities defined in Section 2.1. Since V_t is a state–price density, it follows that the price at time t of a derivative with payoff H_T is given by

$$H_t = \frac{E[V_T H_T | \mathcal{F}_t]}{V_t}. \tag{2.12}$$

We see that expression (2.6) for bond prices is simply a special case of (2.12) with $H_T = P_{TT} = 1$. We apply this general expression to European put options, caplets and swaptions, since these will be the only ones needed in our calibration section, although similarly results also hold for other interest rate derivatives, notably call options and floors (see [30]). We follow the standard definitions and notations in [2].

A put option with maturity t and strike price K written on a bond with maturity $T \geq t$ correspond to the following payoff

$$(K - P_{tT})^+. \tag{2.13}$$

Using (2.12), we see that the price of the a bond option at time $s \leq t \leq T$ is given by

$$p(s,t,T,K) = \frac{E[V_t (K - P_{tT})^+ | \mathcal{F}_s]}{V_s} = \frac{E[(K Z_{tt} - Z_{tT})^+ | \mathcal{F}_s]}{V_s}. \tag{2.14}$$

A caplet is a call option with strike K and maturity T written on a spot LIBOR rate $L(t,T)$ for the time interval $[t,T]$, that is, it is defined by the payoff

$$\mathcal{N}(T-t)(L(t,T) - K)^+, \tag{2.15}$$

where $\mathcal{N}$ is a fixed notional amount. Recalling that the spot LIBOR rate is defined as

$$L(t,T) := \frac{1}{T-t}\left(\frac{1}{P_{tT}} - 1\right) = \frac{1}{T-t}\left(\frac{V_t}{Z_{tT}} - 1\right) \tag{2.16}$$

we see from (2.12) that the value of a caplet at time $s \leq t$ is

$$\begin{aligned} \mathrm{Cpl}(s,t,T,\mathcal{N},K) =& \frac{E\left[V_T \mathcal{N}(T-t)(L(t,T) - K)^+ | \mathcal{F}_s\right]}{V_s} \\ =& \frac{\mathcal{N}}{V_s} E\left[\left(Z_{tt} - Z_{tT}(1 + K(T-t))\right)^+ \Big| \mathcal{F}_s\right] \end{aligned} \tag{2.17}$$

Comparing (2.17) and (2.14) it is easy to deduce the following well–known relation [2, page 41]:

$$\mathrm{Cpl}(s,t,T,\mathcal{N},K) = \mathcal{N}(1+K(T-t))p\left(s,t,T,\frac{1}{1+K(T-t)}\right), \tag{2.18}$$

which allows us to obtain prices for caplets provided we know how to calculate the prices of bond put options.

Observe that the payoff (2.15) can be rewritten as

$$\mathcal{N}(T-t)(F(t,t,T)-K)^+, \tag{2.19}$$

where the forward LIBOR rate $F(s,t,T)$ is defined as

$$F(s,t,T) := \frac{1}{T-t}\left(\frac{P_{st}}{P_{sT}}-1\right), \qquad 0 \le s \le t \le T < \infty. \tag{2.20}$$

Using this, we can define the caplet implied volatility as the value $v_{t,T}$ satisfying

$$\mathrm{Cpl}(0,t,T,\mathcal{N},K) = \mathcal{N}P_{0T}(T-t)\mathrm{Black}\left(K,F(0,t,T),v_{t,T}\sqrt{t}\right), \tag{2.21}$$

where

$$\mathrm{Black}(K,F,v) := F\Phi\left(\frac{\log(F/K)+\frac{v^2}{2}}{v}\right) - K\Phi\left(\frac{\log(F/K)-\frac{v^2}{2}}{v}\right), \tag{2.22}$$

is the Black formula and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. We say that a caplet is at-the-money (ATM) at time $t=0$ if $K = F(0,t,T)$. Given a set of maturities T_i, the so–called term structure of caplet volatility is the implied volatility curve obtained from ATM caplets of consecutive maturities, that is, a function of the form $T_i \to v_{T_{i-1},T_i}$.

A payer Swaption with maturity t is defined by the payoff

$$\left(\mathcal{N}\sum_{i=1}^{n}(T_i-T_{i-1})P_{tT_i}(F(t,T_{i-1},T_i)-K)\right)^+, \tag{2.23}$$

where $\mathcal{T} = (T_0,T_1,\ldots,T_n)$ is a set of future dates with $T_0 = t$ and $\mathcal{N}$ is the notional. The total length of the time interval $(T_n - T_0)$ is called the tenor of the swaption.

Recalling the definition of the forward LIBOR rate in (2.20), we obtain from (2.12) that the value of the payer swaption at time $s \le t$ is given by

$$\begin{aligned}\mathrm{Swp}(s,\mathcal{T},\mathcal{N},K) =& \frac{1}{V_s}E\left[V_t\left(\mathcal{N}\sum_{i=1}^{n}(T_i-T_{i-1})P_{tT_i}(F(t,T_{i-1},T_i)-K)\right)^+\middle|\mathcal{F}_t\right]\\ =& \frac{\mathcal{N}}{V_s}E\left[\left(Z_{tt}-Z_{tT_n}-K\sum_{i=1}^{n}(T_i-T_{i-1})Z_{tT_i}\right)^+\middle|\mathcal{F}_t\right].\end{aligned} \tag{2.24}$$

Observe that the payoff in (2.23) can be rewritten as

$$\left(\mathcal{N}\sum_{i=1}^{n}(T_i-T_{i-1})P_{tT_i}(S(t,t,T_n)-K)\right)^+, \tag{2.25}$$

where the forward swap rate $S(s,t,T_n)$ is defined as

$$S(s,t,T_n) := \frac{P_{st} - P_{sT_n}}{\sum_{i=1}^{n}(T_i - T_{i-1})P_{sT_i}}, \quad 0 \le s \le t \le T_n < \infty. \tag{2.26}$$

Similarly to caplets, we can use this to define the swaption implied volatility as the value v_{t,T_n} satisfying

$$\mathrm{Swp}(0,\mathcal{T},\mathcal{N},K) = \mathcal{N}\sum_{i=1}^{n}(T_i - T_{i-1})P_{0T_i}\mathrm{Black}\left(K, S(0,t,T_n), v_{t,T_n}\sqrt{t}\right). \tag{2.27}$$

The swaption is said to be at-the-money (ATM) at time $t = 0$ if $K = S(0,t,T_n)$.

This concludes our general overview of chaotic models. We will have a chance to explore these formulas in more detail after we introduce specific models of different chaos orders in the calibration sections that come next. For simplicity, we will assume from now on that the underlying Brownian motion W_t is one dimensional, although extensions to the multidimensional case should be relatively straightforward.

3. Term structure calibration. We first consider the calibration of the chaos models to the bond prices P_{0T} observed at time $t = 0$, or equivalently, to the corresponding forward rates

$$f_{0T} = -\frac{\partial}{\partial T}\log P_{0T}.$$

Historically, the problem of characterizing the term structure of interest rates implied by bond prices at time $t = 0$ has been extensively tackle by so–called descriptive models, which consists of choosing a parametric form for the forward rates f_{0T} and then fit the relatively small number of parameters to the observed yield curve. Popular examples of this approach are:

$$f_{0T} = \begin{cases} b_0 + [b_1 + b_2 T]e^{-c_1 T} & \text{(Nelson and Siegel)} \\ b_0 + [b_1 + b_2 T]e^{-c_1 T} + b_3 T e^{-c_2 T} & \text{(Svensson)} \\ b_0 + \sum_{i=1}^{4} b_i e^{-c_i T} & \text{(Cairns)}, \end{cases} \tag{3.1}$$

introduced respectively in [21], [28] and [4]. All of these examples fall within the exponential-polynomial family of forward rates analyzed in [1]:

$$f_{0T} = L_0(T) + \sum_{i=1}^{n} L_i(T)e^{-c_i T} \quad \text{where} \quad L_i(T) = \sum_{j=0}^{k_i} b_{ij}T^j. \tag{3.2}$$

As explained in [4], descriptive models provide a snapshot, but not a movie. That is to say, they do not provide a model for the stochastic evolution of interest rates. Nevertheless, they can be used to price simple instruments such as forward rate agreements and swaps, which do not depend on the dynamics of interest rates. They have also been widely used by central banks for the purposes of monetary policy (see the table presented in [10]).

For more complicated interest rate derivatives, it becomes necessary to embed such static descriptions into a fully dynamic model. For example, one can use the fitted forward rate curve f_{0T} as the initial term structure for a model within the Heath–Jarrow–Morton framework [15], supplementing it with the specification of the

volatility structure. In other words, the specification of the initial term structure used as an input is done separately from its stochastic evolution.

By contrast, instead of focusing on the forward rates f_{0T}, we propose to use the observed term structure at time $t = 0$ to calibrate the deterministic functions $\phi_i(T)$, which then completely specify the stochastic evolution of interest rates. Since these are markedly different procedures, their performance need to be compared using real observed data, as we do later in this section. Before that we need to describe the term structure of interest rates for chaotic models of different orders in mode detail.

3.1. Term structure in first chaos models. In a first chaos model we have

$$X_\infty = \int_0^\infty \phi_1(s) dW_s \tag{3.3}$$

where $\phi_1(s)$ is a deterministic function of one variable. Comparing this with (2.1) we conclude that $\sigma_s \equiv \phi_1(s)$, so that

$$V_t = E\left[\int_t^\infty \sigma_s^2 ds \,\middle|\, \mathcal{F}_t\right] = \int_t^\infty \phi_1(s)^2 ds, \tag{3.4}$$

which implies that

$$dV_t = -\phi_1(t)^2 dt.$$

Comparing this with (2.9) then gives

$$r_t = \frac{\phi_1^2(t)}{V_t} = \frac{\phi_1^2(t)}{\int_t^\infty \phi_1^2(s) ds} \tag{3.5}$$

$$\lambda_t = 0. \tag{3.6}$$

In other words, in first chaos models the state–price density is a deterministic function, implying that the interest rates associated with it are themselves deterministic. In particular, $M_{ts} = \phi_1^2(s)$ constitutes a family of constant martingales for each value of s and bond prices are given by

$$P_{tT} = \frac{E\left[V_T|\mathcal{F}_t\right]}{V_t} = \frac{\int_T^\infty M_{ts} ds}{\int_t^\infty M_{ts} ds} = \frac{\int_T^\infty \phi_1^2(s) ds}{\int_t^\infty \phi_1^2(s) ds} = e^{-\int_t^T r_s ds}. \tag{3.7}$$

It follows that instantaneous forward rates in first chaos models have the form

$$f_{tT} = -\frac{\partial}{\partial T} \log P_{tT} = \frac{\phi_1^2(T)}{\int_T^\infty \phi_1(s)^2 ds} = \frac{\phi_1^2(T)}{V_T} = r_T, \tag{3.8}$$

which are manifestly positive, as they should be in any chaotic model. Moreover, because they represent market expectations at time t for what the (deterministic) short–rate will be at time T, instantaneous forward rates in first chaos models are independent of t and always equal to r_T.

As already noticed in [16], these observations show that first chaos models characterize general arbitrage–free, positive, deterministic interest rate term structures, which should not be immediately disregarded as trivial, since they constitute the starting point for the majority of the applications of interest rate theory. Chaotic models then provide a well–defined way to embed such deterministic models into their stochastic generalizations, as we will demonstrate in the remainder of this paper.

From the point of view of calibration, we can calibrate a first chaos model to the term structure at time $t = 0$ by focusing on the equation

$$P_{0T} = \frac{\int_T^\infty \phi_1(s)ds}{\int_0^\infty \phi_1(s)ds}. \tag{3.9}$$

Because of the deterministic nature of first chaos models, this is enough to characterize bond prices at all times, since it follows from (3.7) that $P_{tT} = P_{0T}/P_{0t}$.

3.2. Term structure in second chaos models. In a second chaos model we have

$$X_\infty = \int_0^\infty \phi_1(s)dW_s + \int_0^\infty \int_0^s \phi_2(s,u)dW_u dW_s,$$

for deterministic functions $\phi_1(s)$ and $\phi_2(s,u)$. Comparing this with (2.1) gives

$$\sigma_t = \phi_1(t) + \int_0^t \phi_2(s,u)dW_u. \tag{3.10}$$

It then follows from the conditional Ito isometry that

$$\begin{aligned} M_{ts} &= E[\sigma_s^2|\mathcal{F}_t] \\ &= \left(\phi_1(s) + \int_0^t \phi_2(s,u)dW_u\right)^2 + \int_t^s \phi_2^2(s,u)du \\ &= R_{ts}^2 - Q_{ts} + Q_{ss} \end{aligned} \tag{3.11}$$

where, for each fixed value $s \in [t, \infty)$, the process

$$R_{ts} := \phi_1(s) + \int_0^t \phi_2(s,u)dW_u \tag{3.12}$$

is a martingale in the variable t with quadratic variation

$$Q_{ts} := \int_0^t \phi_2^2(s,u)du. \tag{3.13}$$

Setting $t = 0$ in (3.11) and using (1.1) we have that the initial term structure of bond prices in second chaos models is generically given by

$$P_{0T} = \frac{\int_T^\infty \left(\phi_1^2(s) + \int_t^s \phi_2^2(s,u)du\right) ds}{\int_0^\infty \left(\phi_1^2(s) + \int_0^s \phi_2^2(s,u)du\right) ds}. \tag{3.14}$$

Expression (3.14) looks uncomfortably complicated for the purpose of calibrating a relatively simple object such as P_{0T}. To achieve greater tractability, we now recall the class of *factorizable* second chaos models introduced in [16], whereby

$$\phi_1(s) = \alpha(s), \qquad \phi_2(s,u) = \beta(s)\gamma(u). \tag{3.15}$$

for deterministic functions α, β and γ. In other words, factorizable second chaos models correspond to random variables X_∞ of the form

$$X_\infty = \int_0^\infty \alpha(s)dW_s + \int_0^\infty \int_0^s \beta(s)\gamma(u)dW_u dW_s. \tag{3.16}$$

for $0 \leq u \leq s < \infty$. In this case we have that

$$R_{ts} = \alpha(s) + \beta(s)R_t, \qquad Q_{ts} = \beta^2(s)Q_t, \tag{3.17}$$

where

$$R_t := \int_0^t \gamma(u)dW_u \tag{3.18}$$

is a Gaussian martingale with quadratic variation $Q_t := \int_0^t \gamma^2(u)du$. Inserting (3.17) into (3.11) we obtain

$$\begin{aligned} Z_{tT} &= \int_T^\infty M_{ts}ds \\ &= \int_T^\infty \left[(\alpha(s) + \beta(s)R_t)^2 - \beta_s^2 Q_t + \beta^2(s)Q_s\right] ds \\ &= A(T) + B(T)R_t + C(T)(R_t^2 - Q_t) \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} A(T) &= \int_T^\infty [\alpha^2(s) + \beta^2(s)Q_s]ds, \\ B(T) &= \int_T^\infty 2\alpha(s)\beta(s)ds, \\ C(T) &= \int_T^\infty \beta^2(s)ds. \end{aligned} \tag{3.20}$$

Using the fact that $V_t = Z_{tt}$, it is easy to see that the state–price density in this case is given by

$$V_t = A(t) + B(t)R_t + C(t)(R_t^2 - Q_t), \tag{3.21}$$

which is a simple quadratic polynomial in the Gaussian process R_t. Therefore bond prices in a factorizable second chaos model are given by the following ration of quadratic polynomials in R_t:

$$P_{tT} = \frac{A(T) + B(T)R_t + C(T)(R_t^2 - Q_t)}{A(t) + B(t)R_t + C(t)(R_t^2 - Q_t)}. \tag{3.22}$$

In particular, the initial term structure is given by

$$P_{0T} = \frac{A(T)}{A(0)} = \frac{\int_T^\infty [\alpha^2(s) + \beta^2(s)Q_s]ds}{\int_0^\infty [\alpha^2(s) + \beta^2(s)Q_s]ds}. \tag{3.23}$$

Despite being much simpler than (3.14), we observe that (3.23) still offers too much freedom for the choice of parametric forms for the deterministic functions involved. To make this point more explicitly, observe that from the strict point of view of calibrating the initial term structure, we see that a factorizable second chaos model behaves exactly like a first chaos model if we identify

$$\phi_1(s) = \alpha^2(s) + \beta^2(s) \int_0^s \gamma^2(u)du, \tag{3.24}$$

for arbitrary choices of α, β and γ. This is to be expected in all chaos models, since the initial bond prices available for calibration correspond to the one–variable function P_{0T}, regardless of how many deterministic functions ϕ_i we wish to calibrate. Of course, once calibrated, chaos models of different orders give rise to very different *future* term structures: a deterministic one with $P_{tT} = P_{0T}/P_{0t}$ for first chaos models and a fully stochastic one where P_{tT} is given by (3.22) for factorizable second chaos models, for example.

These remarks point to a general yet delicate feature of chaotic models: one can achieve increased stochastic complexity in the form of higher order chaos at the expense of having to calibrate complicated expressions that are not necessarily supported by the data available. Our approach to deal with this problem consists of increasing the complexity of the models as gradually as possible, while carefully evaluating their statistical performance when calibrated to real data.

Accordingly, we propose to investigate the subclass of factorizable second chaos models with $\gamma(u) \equiv 1$ and $\beta(s)$ satisfying

$$\int_0^\infty \beta(s)^2 s ds < \infty, \tag{3.25}$$

that is, models corresponding to the terminal random variable

$$X_\infty = \int_0^\infty \alpha(s) dW_s + \int_0^\infty \beta(s) \int_0^s dW_u dW_s = \int_0^\infty \alpha(s) dW_s + \int_0^\infty \beta(s) W_s dW_s. \tag{3.26}$$

As we can see, these *one–variable* second chaos models provide the simplest stochastic extensions of first order chaos, for which we have that $R_t = W_t$ and the initial term structure of bond prices is given by

$$P_{0T} = \frac{\int_T^\infty [\alpha^2(s) + \beta^2(s) s] ds}{\int_0^\infty [\alpha^2(s) + \beta^2(s) s] ds}. \tag{3.27}$$

3.3. Term structure in higher order chaos models. In a general chaos model we that

$$\begin{aligned} Z_{0T} &= \int_T^\infty \mathbb{E}\left[\left(\phi_1(s_1) + \int_0^{s_1} \phi_2(s_1, s_2) dW_{s_2} + \cdots\right)^2\right] ds_1 \\ &= \int_T^\infty \left(\phi_1^2(s_1) + \int_0^{s_1} \phi_2^2(s_1, s_2) ds_2 + \cdots\right) ds_1. \end{aligned}$$

Therefore

$$P_{0T} = \frac{Z_{0T}}{Z_{0t}} = \frac{\int_T^\infty \psi(s) ds}{\int_t^\infty \psi(s) ds}, \tag{3.28}$$

where

$$\psi(s) = \begin{cases} \phi_1^2(s) & \text{(first chaos)} \\ \phi_1^2(s) + \int_0^s \phi_2^2(s, u) du & \text{(second chaos)} \\ \phi_1^2(s) + \int_0^s \phi_2^2(s, u) du + \int_0^s \int_0^u \phi_3^2(s, u, v) dv du & \text{(third chaos)} \\ \quad \vdots \end{cases} \tag{3.29}$$

In other words, as mentioned in the previous section, equation (3.28) shows that from the strict point of view of calibrating the initial term structure P_{0T}, all chaos models behave like first order chaos models, albeit with a complicated choice of function $\psi(s)$. The differences between chaos models will become apparent, however, when we tackle the calibration to option prices in Section 4. For now, we find unnecessary to consider fully general higher order chaos and restrict ourselves to the simplest possible third chaos models. That is to say, we start with factorizable third chaos models of the form

$$X_\infty = \int_0^\infty \alpha(s)dW_s + \int_0^\infty\int_0^s \beta(s)\gamma(u)dW_u dW_s + \int_0^\infty\int_0^s\int_0^u \delta(s)\epsilon(u)\zeta(v)dW_v dW_u dW_s,$$

for $0 \le v \le u \le s < \infty$. We then take $\gamma \equiv \epsilon \equiv \zeta \equiv 1$, β satisfying (3.25) and δ satisfying

$$\int_0^\infty \delta(s)^2 s^2 ds < \infty \tag{3.30}$$

to obtain the class of *one–variable* third chaos models of the form

$$\begin{aligned} X_\infty &= \int_0^\infty \alpha(s)dW_s + \int_0^\infty\int_0^s \beta(s)dW_u dW_s + \int_0^\infty\int_0^s\int_0^u \delta(s)dW_v dW_u dW_s \\ &= \int_0^\infty \left[\alpha(s) + \beta(s)W_s + \frac{1}{2}\delta(s)(W_s^2 - s)\right] dW_s. \end{aligned} \tag{3.31}$$

We see from (3.31) that in this case

$$\sigma_s = \alpha(s) + \beta(s)W_s + \frac{1}{2}\delta(s)(W_s^2 - s) \tag{3.32}$$

Recalling that $M_{ts} = E[\sigma_s^2|\mathcal{F}_t]$, a direct calculation shows that

$$\begin{aligned} Z_{tT} = \int_T^\infty M_{ts}ds = \widetilde{A}(T) + \widetilde{B}(T)W_t + \widetilde{C}(T)\left(W_t^2 - t\right) \\ + \widetilde{D}(T)\left(W_t^3 - 3tW_t\right) + \widetilde{E}(T)\left(W_t^4 - 6tW_t^2 + 3t^2\right), \end{aligned} \tag{3.33}$$

where the coefficients are

$$\begin{aligned} \widetilde{A}(T) &= \int_T^\infty \left(\alpha^2(s) + s\beta^2(s) + \frac{s^2\delta(s)^2}{2}\right) ds \\ \widetilde{B}(T) &= \int_T^\infty 2\beta(s)(\alpha(s) + s\delta(s))ds \\ \widetilde{C}(T) &= \int_T^\infty \left(\beta^2(s) + \alpha(s)\delta(s) + \delta^2(s)s\right) ds \\ \widetilde{D}(T) &= \int_T^\infty \delta(s)\beta(s)ds \\ \widetilde{E}(T) &= \int_T^\infty \frac{\delta^2(s)}{4}ds \end{aligned} \tag{3.34}$$

Using again the fact that $V_t = Z_{tt}$, it follows that the state–price density is

$$\begin{aligned} V_t = \widetilde{A}(t) + \widetilde{B}(t)W_t + \widetilde{C}(t)\left(W_t^2 - t\right) \\ \widetilde{D}(t)\left(W_t^3 - 3tW_t\right) + \widetilde{E}(t)\left(W_t^4 - 6tW_t^2 + 3t^2\right). \end{aligned} \tag{3.35}$$

Therefore bond prices in a one–variable third chaos model are given by the following ratio of forth degree polynomials in W_t:

$$P_{tT} = \frac{\widetilde{A}(T)+\widetilde{B}(T)W_t+\widetilde{C}(T)\big(W_t^2-t\big)+\widetilde{D}(T)\big(W_t^3-3tW_t\big)+\widetilde{E}(T)\big(W_t^4-6tW_t^2+3t^2\big)}{\widetilde{A}(t)+\widetilde{B}(t)W_t+\widetilde{C}(t)\big(W_t^2-t\big)+\widetilde{D}(t)\big(W_t^3-3tW_t\big)+\widetilde{E}(t)\big(W_t^4-6tW_t^2+3t^2\big)} \tag{3.36}$$

In particular, the initial term structure is given by

$$P_{0T} = \frac{\widetilde{A}(T)}{\widetilde{A}(0)} = \frac{\int_T^\infty \left(\alpha^2(s) + s\beta^2(s) + \frac{s^2\delta(s)^2}{2}\right) ds}{\int_0^\infty \left(\alpha^2(s) + s\beta^2(s) + \frac{s^2\delta(s)^2}{2}\right) ds} \tag{3.37}$$

3.4. Choices of the chaos coefficients. Inspired by the parametric forms used in descriptive models, we propose to model the function ϕ_1 itself in exponential–polynomial form. Observe that setting

$$\phi_1(s) = \sum_{i=1}^{n} L_i(s)e^{-c_i s}, \qquad L_i(s) = \sum_{j=0}^{k_i} b_{ij}s^j, \tag{3.38}$$

where $L_0(s) = 0$ and $c_i > 0$ to guaranteed integrability, leads to a first chaos model in which forward rates are ratios of functions in exponential–polynomial form. We regard this *rational exponential–polynomial* family as a natural extension of the exponential–polynomial class considered in [1], with added flexibility and potentially better calibration performance. In particular, this class allows us to reproduce all empirical shapes commonly observed for forward rate curves, such as increasing, decreasing and hump-shaped functions.

Having made this choice for the function ϕ_1, equation (3.24) and its generalization (3.29) naturally lead us to consider functions α, β, γ and δ all belonging to the exponential–polynomial class as well.

3.5. Calibration results. We describe the details of the term structure calibration in Appendix A. We calibrate 14 different chaos models using two distinct data set from the UK bond market: observations of bonds of different maturities at every other business day from January 1998 to January 1999 (a volatile market often exhibiting an inverted yield curve) and weekly observations from December 2002 to December 2005 (a more moderate market). For comparison, we also calibrate two of the descriptive models specified in (3.1).

We summarize our calibration results for the two data sets in Tables 3.1 and 3.2. The first column in each table labels the models, starting with the descriptive Nelson–Siegel and Svensson models defined by the first two expression in (3.1), followed by the models (A.1)–(A.14) in Appendix A. The second column characterizes the type of model, whereas the third one gives the number of calibrated parameters. The remaining columns show the average values for the negative log–likelihood function, the Root-Mean-Squared Percentage Error and the Diebold-Mariano statistics with respect to the Svensson model, as described in Appendix A.

Starting with the results in Table 3.1 for the volatile market, we first notice that, as expected, the Svensson model, which has 6 parameters, outperforms all three models that use a smaller number of parameters. On the other hand, out of five chaos models with 6 parameters, three are outperformed by the Svensson model, while the other two perform comparably. Finally, all chaos models with 7 parameters have a performance that is comparable to the Svensson model. We are led to conclude

TABLE 3.1
Term structure calibration for 1998-1999 (Volatile Market)

	Model	N	-L	RMSPE (%)	DM
Sv	Svensson	6	160	0.70	-
NS	Nelson–Siegel	4	2101	2.67	-4.45
1	1st chaos	3	4420	4.44	-11.46
2	1st chaos	5	250	0.86	-3.54
3	one-var 2nd chaos	6	162	0.82	-2.26
4	one-var 2nd chaos	7	160	0.69	0.22
5	one-var 2nd chaos	7	145	0.75	-1.05
6	factorizable 2nd chaos	6	335	0.88	-2.54
7	factorizable 2nd chaos	6	245	0.68	0.27
8	factorizable 2nd chaos	6	1245	1.26	-3.81
9	factorizable 2nd chaos	7	179	0.63	1.38
10	factorizable 2nd chaos	7	153	0.72	-1.07
11	one-var 3rd chaos	6	168	0.72	-1.24
12	one-var 3rd chaos	7	141	0.76	-1.16
13	one-var 3rd chaos	7	152	0.72	-1.19
14	one-var 3rd chaos	7	149	0.76	-1.43

TABLE 3.2
Term structure calibration for 2002-2005 (Moderate Market)

	Model	N	-L	RMSPE (%)	DM
Sv	Svensson	6	442	0.76	-
NS	Nelson–Siegel	4	541	0.97	-1.76
1	1st chaos	3	8716	3.96	-3.50
2	1st chaos	5	438	0.99	-1.99
3	one-var 2nd chaos	6	388	0.89	-1.23
4	one-var 2nd chaos	7	388	0.80	-0.38
5	one-var 2nd chaos	7	329	0.66	1.26
6	factorizable 2nd chaos	6	437	1.04	-3.33
7	factorizable 2nd chaos	6	495	0.84	-0.68
8	factorizable 2nd chaos	6	421	1.19	-2.84
9	factorizable 2nd chaos	7	365	0.82	-0.78
10	factorizable 2nd chaos	7	323	0.72	0.36
11	one-var 3rd chaos	6	388	0.87	-1.06
12	one-var 3rd chaos	7	350	0.78	-0.11
13	one-var 3rd chaos	7	367	0.68	1.24
14	one-var 3rd chaos	7	325	0.69	0.60

that chaos models do not offer a significant advantage with respect to the Svensson model under volatile market conditions, when yield curves can change from normal to inverted to humped shapes quiet rapidly. Moving on to Table 3.2, we observe smaller RMSPEs and higher DM statistics uniformly across all chaotic models, but nevertheless not high enough to reject the null hypotheses that they have the same

calibration performance as the Svensson model.

More interestingly, both tables confirm our previous remarks in Section 3.3 that all chaos models behave similarly to a first chaos model from the point of view of calibrating the initial term structure. Indeed, we do not find significant difference in the calibration performances between chaos models of different orders. If anything, we notice that factorizable second chaos models perform slightly worse than their one–variable analogues, therefore justifying our assertion that one should avoid trying to calibrate the initial term structure using a complicated expression such as (3.23).

We illustrate the results by plotting the RMSPE as a function of time in Figures 3.1 and 3.2. For comparison, we restrict ourselves to the chaotic models with 6 parameters in Tables 3.1 and 3.2, as well as the Svensson and Nelson-Siegel models.

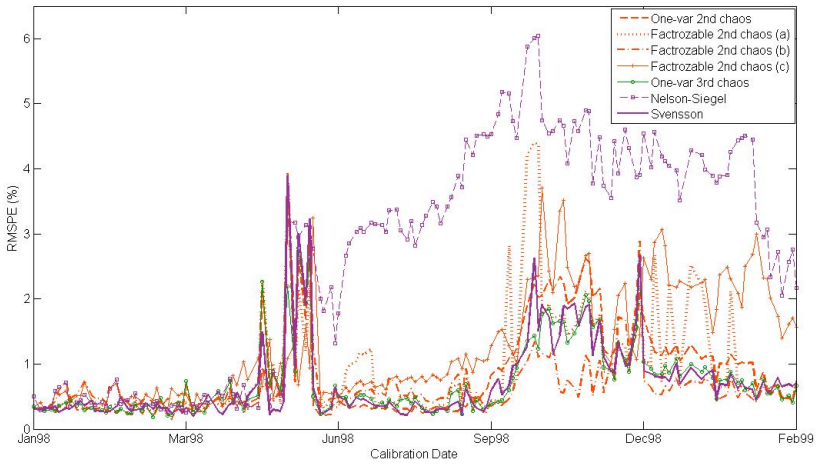

FIG. 3.1. *Error for term structure calibration in 1998-1999.*

The basic overall conclusion of this section is that, as far as the daily calibration of initial term structure is concerned, we cannot state that there exists a significant difference in the calibration performances between the Svensson and the chaotic models. However, as we mentioned before, chaotic models specify the stochastic evolution of interest rates, in addition to ensuring their positivity. In other words, the parameters calibrated from yield data may be applied directly to the volatility term structure. As we will see in the next section, this turns out to be greatly advantageous in the joint calibration of yields and option prices.

4. Option calibration. In the previous section we showed that the chaotic and descriptive models have comparable performances when calibrated to yield curves. In this section, we take into account ATM European options, in particular caplets and swaptions. Since descriptive models for yield curves are silent regarding the future evolution of interest rates, we compare chaos models of different orders with other popular interest rate models in the literature, including short–rate models, models based on the potential approach and market models. But let us first analyze in greater detail the expressions for option prices in chaos models of different orders.

4.1. Option prices in second chaos models. Since in a first chaos model the price at time 0 of an interest rate derivative with payoff H_T reduces to the determin-

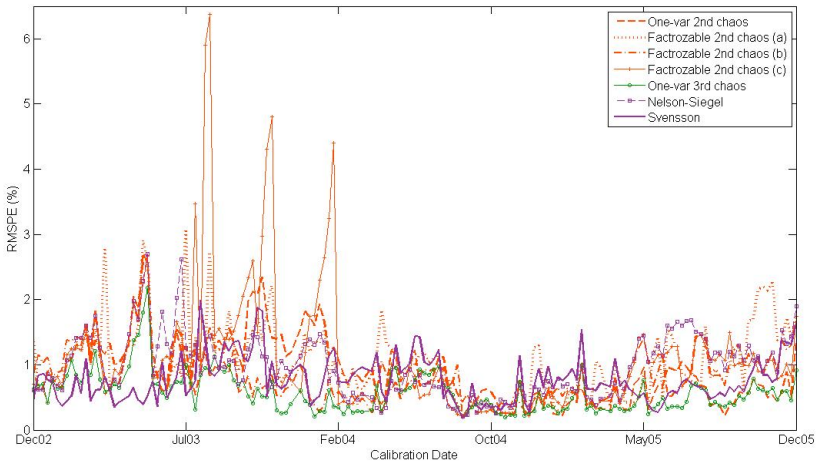

FIG. 3.2. *Error for term structure calibration in 2002-2005.*

istic expression

$$H_0 = \frac{E[V_T H_T]}{V_0} = \frac{V_T}{V_0} H_T = e^{-\int_0^T r_s ds} H_T, \tag{4.1}$$

any nontrivial option pricing should start with at least a second chaos model. We begin with bond put options in a factorizable second chaos, for which it has already been observed in [16] that (3.19) implies that

$$KZ_{tt} - KZ_{tT} = \mathcal{P}_p^{(2)}(z), \tag{4.2}$$

where $z = R_t/\sqrt{Q_t}$ is a standard normal random variable and

$$\mathcal{P}_p^{(2)}(z) := a_0 + a_1 z + a_2 z^2 \tag{4.3}$$

is a second degree polynomial with coefficients

$$\begin{aligned} a_0 &= -\big(A(T) - KA(t)\big) + \big(C(T) - KC(t)\big)Q_t \\ a_1 &= -\big(B(T) - KB(t)\big)Q^{1/2}(t) \\ a_2 &= -\big(C(T) - KC(t)\big)Q_t. \end{aligned} \tag{4.4}$$

Using (2.14), it then follows that

$$p(0,t,T,K) = \frac{1}{A(0)\sqrt{2\pi}} \int_{\mathcal{P}_p^{(2)}(z)\geq 0} \mathcal{P}_p^{(2)}(z) e^{-\frac{z^2}{2}} dz. \tag{4.5}$$

As shown in [16], once we calculate the roots of $\mathcal{P}_p^{(2)}(z)$, we have that (4.5) reduces to a simple expression given in terms of the standard normal cumulative distribution function. Having calculated the price of a put option on a bond, we can easily find the price of a caplet using expression (2.18).

Regarding the price of a payer swaption, it was also observed in [16] that

$$Z_{tt} - Z_{tT_n} - K\sum_{i=1}^{n}(T_i - T_{i-1})Z_{tT_i} = \mathcal{P}^{(2)}_{Swp}(z) \tag{4.6}$$

where z is the standard normal random variable defined above and

$$\mathcal{P}_{Swp}(z) := b_0 + b_1 z + b_2 z^2 \tag{4.7}$$

is a second degree polynomial with coefficients

$$\begin{aligned} b_0 &= \left(A(t) - A(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})A(T_i)\right) \\ &\quad - \left(C(t) - C(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})C(T_i)\right)Q_t \\ b_1 &= \left(B(t) - B(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})B(T_i)\right)Q^{1/2}t \\ b_2 &= \left(C(t) - C(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})C(T_i)\right)Q_t. \end{aligned} \tag{4.8}$$

It then follows from (2.24) that the price of a payer swaption in a factorizable second chaos model is

$$\mathrm{Swp}(0,\mathcal{T},\mathcal{N},K) = \frac{1}{A(0)\sqrt{2\pi}}\int_{\mathcal{P}^{(2)}_{Swp}(z)\geq 0}\mathcal{P}^{(2)}_{Swp}(z)e^{-\frac{z^2}{2}}dz. \tag{4.9}$$

Once more, as shown in [16], having calculated the roots of $\mathcal{P}^{(2)}_{Swp}(z)$, expression (4.9) reduces to an explicit formula in terms of the standard normal cumulative distribution and density functions.

4.2. Option prices in third chaos models. We focus on one–variable third chaos models of the form (3.31). Recalling expression (3.33) for the process Z_{tT} we have that

$$KZ_{tt} - Z_{tT} = \mathcal{P}^{(4)}_{p}(z) \tag{4.10}$$

where $z = W_t/\sqrt{t}$ is a standard normal random variable and

$$\mathcal{P}^{(4)}_{c}(z) := c_0 + c_1 z + c_2 z^2 + c_3 z^3 + c_4 z^4 \tag{4.11}$$

is a forth degree polynomial with coefficients

$$\begin{aligned} c_0 &= -\left(\widetilde{A}(T) - K\widetilde{A}(t)\right) + \left(\widetilde{C}(T) - K\widetilde{C}(t)\right)t - 3\left(\widetilde{E}(T) - \widetilde{E}(t)\right)t^2 \\ c_1 &= -\left(\widetilde{B}(T) - K\widetilde{B}(t)\right)t^{1/2} + 3\left(\widetilde{D}(T) - \widetilde{D}(t)\right)t^{3/2} \\ c_2 &= -\left(\widetilde{C}(T) - K\widetilde{C}(t)\right)t + 6\left(\widetilde{E}(T) - K\widetilde{E}(t)\right)t^2 \\ c_3 &= -\left(\widetilde{D}(T) - K\widetilde{D}(t)\right)t^{3/2} \\ c_4 &= -\left(\widetilde{E}(T) - K\widetilde{E}(t)\right)t^2 \end{aligned} \tag{4.12}$$

We then have the the price of a bond put option is a one–variable third chaos model is

$$p(0,t,T,K) = \frac{1}{\widetilde{A}(0)\sqrt{2\pi}} \int_{\mathcal{P}_p^{(4)}(z)\geq 0} \mathcal{P}_p^{(4)}(z) e^{-\frac{z^2}{2}} dz, \tag{4.13}$$

which can then be used to obtain the price of a caplet according to (2.17). As before, once we calculate the roots of the polynomial $\mathcal{P}_p^{(4)}(z)$, both of these expressions reduce to explicit formulas in terms of the standard normal cumulative distribution and density functions (see [30]).

Regarding the price of a payer swaption, we have that

$$Z_{tt} - Z_{tT_n} - K\sum_{i=1}^{n}(T_i - T_{i-1})Z_{tT_i} = \mathcal{P}_{Swp}^{(4)}(z) \tag{4.14}$$

where $z = W_t/\sqrt{t}$ a standard normal random variable and

$$\mathcal{P}_{Swp}^{(4)}(z) := d_0 + d_1 z + d_2 z^2 + d_3 z^3 + d_4 z^4 \tag{4.15}$$

is a forth degree polynomial with coefficients

$$\begin{aligned}
d_0 &= \left(\widetilde{A}(t) - \widetilde{A}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{A}(T_i)\right) \\
&\quad - \left(\widetilde{C}(t) - \widetilde{C}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{C}(T_i)\right) t \\
&\quad + 3\left(\widetilde{E}(t) - \widetilde{E}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{E}(T_i)\right) t^2 \\
d_1 &= \left(\widetilde{B}(t) - \widetilde{B}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{B}(T_i)\right) t^{1/2} \\
&\quad - 3\left(\widetilde{D}(t) - \widetilde{D}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{D}(T_i)\right) t^{3/2} \\
d_2 &= \left(\widetilde{C}(t) - \widetilde{C}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{C}(T_i)\right) t \\
&\quad - 6\left(\widetilde{E}(t) - \widetilde{E}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{E}(T_i)\right) t^2 \\
d_3 &= \left(\widetilde{D}(t) - \widetilde{D}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{D}(T_i)\right) t^{3/2} \\
d_4 &= \left(\widetilde{E}(t) - \widetilde{E}(T_n) - K\sum_{i=1}^{n}(T_i - T_{i-1})\widetilde{E}(T_i)\right) t^2
\end{aligned} \tag{4.16}$$

It then follows from (2.23) that the price of a payer swaption in a one–variable third chaos model is

$$\mathrm{Swp}(0,\mathcal{T},\mathcal{N},K) = \frac{1}{\widetilde{A}(0)\sqrt{2\pi}} \int_{\mathcal{P}_{Swp}^{(4)}(z)\geq 0} \mathcal{P}_{Swp}^{(4)}(z) e^{-\frac{z^2}{2}} dz. \tag{4.17}$$

Once more, after we find the roots of $\mathcal{P}^{(4)}_{Swp}(z)$, expression (4.17) reduces to an explicit formula in terms of the standard normal cumulative distribution and density functions (see [30]).

4.3. Benchmark models. In this section we describe the models we use for comparing the calibration performance of chaotic models to option prices.

We start with the simplest short–rate model for which the initial term structure can be freely specified in order to match observed data, namely the Hull–White [17] dynamics

$$dr_t = (\theta(t) - \kappa r_t)dt + \eta dW_t, \tag{4.18}$$

where we choose the function $\theta(t)$ in such a way that the initial term structure for forward rates has the Svensson form in (3.1), so that there are 8 parameters in total to be calibrated for this model. The prices for bond options, caplets and swaptions in a Hull–White model are given by well–known analytic expressions (see for example [2, pages 76–77]).

Next we consider the rational lognormal model introduced in [11] (see related literature in [14], [20] and [23]). This is one of the earliest examples of a model in the potential approach and consists of setting

$$\sigma_s^2 = g_1(s)M_s + g_2(s), \quad 0 \le s \le t, \tag{4.19}$$

where g_1, g_2 are nonnegative deterministic functions of time and M_t is a strictly positive continuous martingale such that $M_0 = 1$. We then see that

$$Z_{tT} = \int_T^\infty M_{ts}ds = \int_T^\infty E\left[\sigma_s^2 \middle| \mathcal{F}_t\right] ds = G_1(T)M_t + G_2(T), \tag{4.20}$$

where

$$G_1(t) := \int_t^\infty g_1(s)ds, \quad \text{and} \quad G_2(t) := \int_t^\infty g_2(s)ds. \tag{4.21}$$

This implies that bond prices are represented in the following way:

$$P_{tT} = \frac{G_1(T)M_t + G_2(T)}{G_1(t)M_t + G_2(t)}, \quad \text{for} \quad 0 \le t \le T < \infty.$$

As shown in [11], caplets and swaptions can be priced analytically in a rational lognormal model. Note however that, as also discussed in [11] and [6], for each fixed t and T, bond prices in this model are bounded above and below by the deterministic ratios $G_1(T)/G_1(t)$ and $G_2(T)/G_2(t)$. As shown in [6], it follows that the short rate r_t is also bounded above and below by the deterministic ratios $G_1'(t)/G_1(t)$ and $G_2'(t)/G_2(t)$. As a consequence, rational lognormal models are unable to accurately price deep in the money and out of the money interest rate options.

To implement the rational lognormal model, we follow [20] and parametrize the functions g_1 and g_2 as

$$g_1(t) = -k_1\frac{\partial P_{0t}}{\partial t}(P_{0t})^{k_2} \quad \text{and} \quad g_2(t) = -\frac{\partial P_{0t}}{\partial t}[1 - k_1(P_{0t})^{k_2}], \quad \text{for} \quad t \ge 0, \tag{4.22}$$

for constants $k_1, k_2 \in \mathbb{R}$, where we choose the function P_{0t} is then chosen so that the initial term structure of forward rates has the Svensson form in (3.1). Regarding the process M_t, we follow [14] and represent is as an exponential martingale of the form

$$M_t = \exp\left[\eta W_t - \frac{1}{2}\eta^2 t\right], \quad \text{for} \quad t \ge 0, \tag{4.23}$$

for some constant $\eta \in \mathbb{R}$. With this choices, there are 9 parameters in total to be calibrated in this model.

The last benchmark that we consider are LIBOR market models. We first introduce the following notation:

$$F_j(t) := F_{tT_{j-1}T_j}, \quad j = 1, 2, \ldots,$$

where $\{T_0, T_1, \ldots, T_{i-1}, T_i, \ldots, T_n\}$ is an increasing set of dates and $F(t, S, T)$ is the forward LIBOR rate defined in (2.20).

The lognormal forward market model (LFM), consists of specifying the dynamics:

$$dF_j(t) = \sigma_j(t)F_j(t)dZ^j(t), \tag{4.24}$$

where $\sigma_j(t), t \geq 0$ is a deterministic process, and Z^j is a Brownian Motion under the forward measure $\mathbb{Q}^j$, that is, the pricing measure associated with the numeraire $P(\cdot, T_j)$ (see [13]), with a correlation given by

$$dZ^i(t)dZ^j(t) = \rho_{ij}dt. \tag{4.25}$$

The motivation for introducing such models is that caplets can then be exactly priced using the Black formula introduced in (2.22). Following [2, page 224], we parametrize the volatilities as

$$\sigma_i(t) = b_1 + (b_2 + b_3(T_{i-1} - t))e^{-c_1(T_{i-1}-t)}, \tag{4.26}$$

for $b_1, b_2, b_3, c_1 \in \mathbb{R}$. Using the Svensson form in (3.1) for the initial term structure of forward rates, we see that there are 10 parameters to be calibrated to yield curves and caplet prices alone.

To price a swaption maturing at t with tenor $(T_n - t)$ we use the Rebonato approximation (see [2, page 283]) for the implied volatility

$$v_{t,T_n} = \sqrt{\frac{1}{T_n}\sum_{i,j=1}^{n} \frac{w_i(0)w_j(0)F_i(0)F_j(0)}{S(0,t,T_n)^2}\rho_{ij}\int_0^{T_n} \sigma_i(t)\sigma_j(t)dt}, \tag{4.27}$$

where the forward swap rates are assumed to be expressed as a linear combination of forward LIBOR rates as follows:

$$S(t, T_0, T_n) = \sum_{i=1}^{n} w_i(t)F_i(t) \quad \text{where} \quad w_i(t) = \frac{(T_i - T_{i-1})P_{tT_i}}{\sum_{k=1}^{n}(T_k - T_{k-1})P_{tT_k}}. \tag{4.28}$$

For the correlations appearing in (4.27), we use the Schoenmakers and Coffey [27] parametrization:

$$\begin{aligned}\rho_{ij} = \exp\Big[- \frac{|j-i|}{n-1}\Big(- \log\rho_\infty + \eta_1 \frac{i^2 + j^2 + ij - 3ni - 3nj + 3i + 3j + 2n^2 - n - 4}{(n-2)(n-3)} \\ - \eta_2 \frac{i^2 + j^2 + ij - ni - nj - 3i - 3j + 3n + 2}{(n-2)(n-3)}\Big)\Big],\end{aligned} \tag{4.29}$$

for $i, j = 1, 2, \ldots, n$ and constants $\eta_1, \eta_2, \rho_\infty \in \mathbb{R}$ satisfying the conditions $\eta_2 \geq 0$ and $0 \leq \eta_1 + \eta_2 \leq -\log\rho_\infty$. As we can see, if we use either only yields and caplets, we need to calibrate 10 parameters, whereas there are in total 13 parameters to be calibrated to yields and swaptions or yields, caplets, and swaptions together.

4.4. Calibration Results. We describe the details of the calibration to yields and option prices in Appendix B. We calibrate 6 different chaos models using two distinct data set from the UK interest rate market: weekly observations of yields and implied volatilities for ATM caplets and ATM swaptions from September 2000 to August 2001 and from May 2005 to May 2006. For comparison, we also calibrate the three benchmark models described in the Section 4.3: the Hull and White short rate model, the rational lognormal model and the lognormal forward LIBOR market model.

We can gain a first impression in Figures 4.1 and 4.2, which show the calibrated implied volatility surfaces for ATM swaptions of different maturities and tenors at two specific calibration dates, one from each of the data sets above. We can see that neither the Hull and White nor the rational lognormal models have enough flexibility to reproduce the market quotes, with the LIBOR and chaos models providing much better fits. Figures 4.3 and 4.4 show similar results for the term structure of ATM caplets. In particular, we observe that both the Hull and White and the rational lognormal fail to capture the hump shaped term structure of caplet implied volatility, whereas the LIBOR and chaos models are able to do so.

The quantitative calibration results are summarized results in Tables 4.1 to 4.6. The first column in each table labels the calibrated chaos models as listed in Appendix B and the benchmark models as listed in Section 4.3. The second column characterizes the type of model, whereas the third one gives the number of calibrated parameters. The remaining columns in each table show the errors defined in Appendix B averaged across the different calibration dates. Specifically, the fourth column in each table shows the average value for the objective function minimized in the calibration. We plot this error across time for all calibration dates in Figures 4.5 to 4.10. The next columns show the average contribution to this total error coming from yields and the options used in each calibration. In the last column of Tables 4.1, 4.2, 4.4 and 4.5, we also show the correspond to the pricing error obtained when we use the calibrated parameters to forecast the prices of options that were not used in the calibration. For example, in Table 4.1 we calibrate the model parameters to yields and caplets and use them to forecast the prices of swaptions.

We can readily see that in general all chaos models that we consider fit the observed data better than the Hull and White and the rational lognormal models, as well as having smaller pricing errors for the options that were not used in each particular calibration. When compared with the LIBOR model, however, we see that only third chaos models have competitive performance.

For a more detailed comparison, we consider relative frequencies for model selection based on the Akaike Information Criterion described in Appendix B. For each of our two data sets, we compared the relative performance of third chaos models with 7 and 9 parameters and the benchmark LIBOR model. The results are presented in Tables 4.7 and 4.8. Whereas the relative performance of the third chaos model with 7 parameters is somewhat ambiguous and depends on which type of options are considered, we find that the third chaos model with 9 parameters consistently outperforms the lognormal forward LIBOR market model with 13 parameters. For example, in the first data set, it outperforms LIBOR in 36 out of the 53 calibration dates when we use yields and caplets, 53 out of 53 calibration dates when we use yields and swaptions, and 39 out of 53 calibration dates when we use yields, caplets and swaptions. For the second data set, the corresponding number of dates are 52, 44 and 39 respectively.

TABLE 4.1
Yield and ATM caplet calibration for 2000-2001

No.	Model	N	TotalE1 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	5.1	2.0	4.6	14.9
2	one-var 2nd chaos	7	3.3	1.7	2.7	16.3
3	factorizable 2nd	6	3.8	2.1	3.1	26.5
4	one-var 3rd chaos	6	4.2	2.0	3.5	15.5
5	one-var 3rd chaos	7	3.2	1.3	2.9	15.7
6	one-var 3rd chaos	9	2.6	1.1	2.3	17.0
I	Hull-White	8	8.7	0.6	8.7	25.8
II	Rational-log	9	9.2	0.6	9.2	13.9
III	LIBOR	10	3.0	0.6	3.0	-

TABLE 4.2
Yield and ATM swaption calibration for $2000-2001$

No.	Model	N	TotalE2 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	7.1	1.8	6.8	14.5
2	one-var 2nd chaos	7	7.1	2.0	6.7	14.6
3	factorizable 2nd	6	7.1	2.1	6.8	14.3
4	one-var 3rd chaos	6	5.3	2.9	4.1	10.2
5	one-var 3rd chaos	7	3.8	1.5	3.4	8.6
6	one-var 3rd chaos	9	3.5	1.5	3.1	9.1
I	Hull-White	8	10.2	0.6	10.2	17.6
II	Rational-log	9	8.4	0.6	8.4	15.3
III	LIBOR	13	5.0	0.6	5.0	8.1

TABLE 4.3
Yield, ATM caplet and ATM swaption calibration for $2000-2001$

No.	Model	N	TotalE2 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	12.5	2.2	9.3	7.9
2	one-var 2nd chaos	7	12.1	2.4	9.3	7.3
3	factorizable 2nd	6	12.1	2.6	8.4	8.2
4	one-var 3rd chaos	6	8.2	4.3	4.4	5.2
5	one-var 3rd chaos	7	7.1	1.6	4.4	5.2
6	one-var 3rd chaos	9	5.9	2.2	4.1	3.4
I	Hull-White	8	18.4	0.6	12.2	13.7
II	Rational-log	9	14.6	0.6	10.0	10.6
III	LIBOR	13	6.5	0.6	5.5	3.1

TABLE 4.4
Yield and ATM caplet calibration for $2005 - 2006$

No.	Model	N	TotalE3 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	6.3	1.6	6.1	9.4
2	one-var 2nd chaos	7	3.4	1.5	3.0	14.0
3	factorizable 2nd	6	4.3	2.4	3.4	20.0
4	one-var 3rd chaos	6	4.9	1.9	4.4	26.2
5	one-var 3rd chaos	7	3.6	1.4	3.2	14.2
6	one-var 3rd chaos	9	3.3	1.3	3.0	17.7
I	Hull-White	8	8.4	0.4	8.4	16.3
II	Rational-log	9	9.4	0.4	9.4	10.5
III	LIBOR	10	3.5	0.4	3.5	-

TABLE 4.5
Yield and ATM swaption calibration for $2005 - 2006$

No.	Model	N	TotalE1 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	6.5	3.2	5.5	32.2
2	one-var 2nd chaos	7	5.0	1.5	4.8	11.9
3	factorizable 2nd	6	6.8	2.3	6.4	13.7
4	one-var 3rd chaos	6	4.5	2.2	3.8	21.2
5	one-var 3rd chaos	7	4.2	1.6	3.8	13.4
6	one-var 3rd chaos	9	3.1	1.3	2.8	14.2
I	Hull-White	8	9.5	0.4	9.5	11.2
II	Rational-log	9	8.2	0.4	8.2	10.9
III	LIBOR	13	3.5	0.4	3.5	14.8

TABLE 4.6
Yield, ATM caplet and ATM swaption calibration for $2005 - 2006$

No.	Model	N	TotalE1 (%)	YldE (%)	CplE (%)	SwpE (%)
1	one-var 2nd chaos	6	10.4	2.5	7.3	6.9
2	one-var 2nd chaos	7	8.6	1.5	6.3	5.6
3	factorizable 2nd	6	10.3	1.9	7.9	6.2
4	one-var 3rd chaos	6	9.1	3.3	5.8	6.1
5	one-var 3rd chaos	7	7.8	1.8	5.0	5.5
6	one-var 3rd chaos	9	5.9	1.4	4.1	4.0
I	Hull-White	8	14.0	0.4	10.1	9.5
II	Rational-log	9	13.0	0.4	8.4	9.9
III	LIBOR	13	6.2	0.4	4.8	3.8

5. Concluding remarks. We proposed and implemented a systematic way to calibrate chaotic models for interest rates to available term structure and option data. The calibration performance to initial term structures is comparable to that of traditional descriptive models for forward rates with the same number of parameters, with the advantages of guaranteed positivity and consistency with a fully stochastic model for the time evolution of interest rates. When we include option data in the form of at-the-money caplets and swaptions, we see that chaos models perform significantly better than the benchmark Hull and White and rational lognormal models, and comparably to a LIBOR market model with a higher number of parameters. If

TABLE 4.7
AIC model selection relative frequency (first dataset)

	Model	Caplets	Swaptions	Joint
comparison 1	one-var 3rd chaos (7 par)	2/53	50/53	23/53
	LIBOR (13 par)	51/53	3/53	30/53
comparison 2	one-var 3rd chaos (9 par)	36/53	53/53	39/53
	LIBOR (13 par)	17/53	0/53	14/53

TABLE 4.8
AIC model selection relative frequency (second dataset)

	Model	Caplets	Swaptions	Joint
comparison 1	one-var 3rd chaos (7 par)	14/53	23/53	7/53
	LIBOR	39/53	30/53	46/53
comparison 2	one-var 3rd chaos (9 par)	52/53	44/53	39/53
	LIBOR	1/53	9/53	14/53

we take the number of parameters into account, a conservative information criterion shows that one of our chaos models consistently outperforms the benchmark LIBOR market model.

The underlying reason for the superior performance of LIBOR market models when compared, for example, to Markovian short–rate models is the rich correlation structure they provide for LIBOR forward rates at different dates. Similarly, in chaos models the resulting short rate is in general not Markovian, and our calibration results show that an equally rich correlation structure can be achieved without having to model forward rates individually under each corresponding forward measure.

The next step in our program is to consider options that are not at-the-money. The current industry standard for this is to use stochastic volatility models for either forward or swap rates. In a paper in preparation we show that chaos models naturally give rise to rates with stochastic volatility, and explore this fact to calibrate them to the smile observe in interest rate data.

Expectedly, much work remains to be done, in particular on the interpretation of the financial interpretation of the chaos coefficients, where techniques such as principal components analysis might shed some additional light on the meaning of the most relevant calibrated parameters. Extensions to other classes of financial products, notably exchange rate derivatives, are also possible. We believe the analysis presented in this paper demonstrates the viability of chaos models as serious contenders for practical use in the financial industry and will stimulate further work in the area.

Acknowledgements. We are grateful for helpful comments by M. Avellaneda, D. Brody, D. Crisan, J-P. Fouque, L. Hughston, T. Hurd, J. Teichmann, M. Yor and the participants at the CRFMS seminar, Santa Barbara, May 2009, the Mathematical Finance and Related Topics in Engineering and Economics Workshop, Kyoto, August 2009, the Research in Options Conference, Buzios, November 2009, the Sixth World Congress of the Bachelier Finance Society, Toronto, June 2010, and the Imperial College finance seminar, October 2010, where this work was presented.

REFERENCES

[1] T. Björk and B. J. Christensen, *Interest rate dynamics and consistent forward rate curves*, Mathematical Finance, 9 (1999), pp. 323–348.

[2] Damiano Brigo and Fabio Mercurio, *Interest Rate Models - Theory and Practice*, Springer, 2nd ed., 2006.

[3] K. P Burnham and D. R Anderson, *Model selection and multi model inference : a practical information-theoretic approach*, Springer, 2nd ed., 2002.

[4] Andrew J. G. Cairns, *Descriptive bond-yield and forward-rate models for the british government securities' market*, British Actuarial Journal, 4 (1998), pp. 265–321.

[5] Andrew J. G. Cairns, *A family of term-structure models for long-term risk management and derivative pricing*, Mathematical Finance, 14 (2004), pp. 415–444.

[6] Andrew J. G. Cairns, *Interest Rate Models: An Introduction*, Princeton University Press, 2004.

[7] Andrew J. G. Cairns and Delme J. Pritchard, *Stability of descriptive models for the term structure of interest rates with application to german market data*, British Actuarial Journal, 7 (2001), pp. 467–507.

[8] F. X. Diebold and R. S. Mariano, *Comparing predictive accuracy*, Journal of Business and Economic Statistics, 13 (1995), pp. 253–263.

[9] Frank J. Fabozzi, *Interest Rate, Term Structure and valuation modeling*, John Wiley and Sons, Inc., 2002.

[10] Damir Filipović, *Exponential-polynomial families and the term structure of interest rates*, Bernoulli, 6 (2000), pp. pp. 1081–1107.

[11] B Flesaker and L. P. Hughston, *Positive interest*, Risk, 9 (1996), pp. 46–49.

[12] Dariusz Gatarek, Przemyslaw Bachert, and Robert Maksymiuk, *The LIBOR market model in practice*, John Wiley and Sons, Inc., 2006.

[13] Hélyette Geman, Nicole El Karoui, and Jean-Charles Rochet, *Changes of numéraire, changes of probability measure and option pricing*, Journal of Applied Probability, 32 (1995), pp. pp. 443–458.

[14] Lisa R. Goldberg, *Volatility of the short rate in the rational lognormal model*, Finance and Stochastics, 2 (1998), pp. 199–211.

[15] David Heath, Robert Jarrow, and Andrew Morton, *Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, 60 (1992), pp. pp. 77–105.

[16] Lane Hughston and Avraam Rafailidis, *A chaotic approach to interest rate modelling*, Finance and Stochastics, 9 (2005), pp. 43–65.

[17] John Hull and Alan White, *Pricing interest-rate-derivative securities*, Review of Financial Studies, 3 (1990), pp. 573–92.

[18] R. Jarrow, H. Li, and F. Zhao, *Interest rate caps "smile" too! but can the libor market models capture the smile?*, The Journal of Finance, 62 (2007), pp. 345–382.

[19] Tino Kluge and L. C. G. Rogers, *The potential approach in practice*. Preprint, Cambridge University, 2008.

[20] Naosuke Nakamura and Fan Yu, *Interest rate, currency and equity derivatives valuation using the potential approach*, International Review of Finance, 1 (2000), pp. 269–294.

[21] Charles R Nelson and Andrew F Siegel, *Parsimonious modeling of yield curves*, Journal of Business, 60 (1987), pp. 473–89.

[22] David Nualart, *The Malliavin calculus and related topics*, Probability and its Applications, Springer-Verlag, New York, 1995.

[23] F. Rapisarda and R. Silvotti, *Implementation and performance of various stochastic models for interest rate derivatives*, Applied Stochastic Models in Business and Industry, 17 (2001), pp. 109–120.

[24] L. C. G. Rogers, *The potential approach to the term structure of interest rates and foreign exchange rates*, Math. Finance, 7 (1997), pp. 157–176.

[25] L. C. G. Rogers and O. Zane, *Fitting potential models to interest rates and foreign exchange rates.*, in Vasicek and beyond, L. P. Hughston, ed., RISK Publications, 1997, pp. 327–342.

[26] Marek Rutkowski, *A note on the flesaker-hughston model of the term structure of interest rates*, Applied Mathematical Finance, 4 (1997), pp. 151–163.

[27] John Schoenmakers and Brian Coffey, *Systematic generation of parametric correlation structures for the libor market model.*, International Journal of Theoretical & Applied Finance, 6 (2003), p. 507.

[28] Lars E.O. Svensson, *Estimating forward interest rates with the extended nelson and siegel method*, Sveriges Riksbank Quarterly Review, 3 (1995), pp. 13–26.

[29] Anders B. Trolle and Eduardo S. Schwartz, *A general stochastic volatility model for the pricing of interest rate derivatives*, Review of Financial Studies, 22 (2009), pp. 2007–2057.
[30] Tsunehiro Tsujimoto, *Calibration of the Chaotic Interest Rate Model*, PhD thesis, University of St Andrews, 2010.
[31] Norbert Wiener, *The Homogeneous Chaos*, American Journal of Mathematics, 60 (1938), pp. 897–936.

Appendix A. Term structure calibration procedure.

We begin by listing our choices of chaos coefficients of different orders according to the general parametric form in (3.38). We start with first chaos models directly inspired by the Nelson–Siegel and Svensson parametric forms in (3.1), namely:

$$\phi_1(s) = (b_1 + b_2 s)e^{-c_1 s} \tag{A.1}$$
$$\phi_1(s) = (b_1 + b_2 s)e^{-c_1 s} + b_3 s e^{-c_2 s}. \tag{A.2}$$

Moving up one chaos order, we consider next one–variable second chaos models with

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = (b_3 + b_4 s)e^{-c_2 s} \tag{A.3}$$
$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = (b_2 + b_3 s)e^{-c_2 s} + b_4 s e^{-c_3 s} \tag{A.4}$$
$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s} + b_3 s e^{-c_2 s}, \quad \beta(s) = b_4 s e^{-c_3 s} \tag{A.5}$$

as well as factorizable second chaos models with

$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = b_2 e^{-c_2 s}, \quad \gamma(s) = (1 + b_3)e^{-c_3 s} \tag{A.6}$$
$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = (b_2 + b_3 s)e^{-c_2 s}, \quad \gamma(s) = e^{-c_3 s} \tag{A.7}$$
$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = b_3 e^{-c_2 s}, \quad \gamma(s) = e^{-c_3 s} \tag{A.8}$$
$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = (b_2 + b_3 s)e^{-c_2 s}, \quad \gamma(s) = (1 + b_4 s)e^{-c_3 s} \tag{A.9}$$
$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = b_3 e^{-c_2 s}, \quad \gamma(s) = (1 + b_4 s)e^{-c_3 s}. \tag{A.10}$$

Finally, we consider the following one–variable third chaos models:

$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = b_2 e^{-c_2 s}, \quad \delta(s) = b_3 e^{-c_3 s} \tag{A.11}$$
$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = b_2 e^{-c_2 s}, \quad \delta(s) = (b_3 + b_4 s)e^{-c_3 s} \tag{A.12}$$
$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = (b_2 + b_3 s)e^{-c_2 s}, \quad \delta(s) = b_4 e^{-c_3 s} \tag{A.13}$$
$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = b_3 e^{-c_2 s}, \quad \delta(s) = b_4 e^{-c_3 s} \tag{A.14}$$

We now describe in detail the steps we take to calibrate these chaos models to observed yield curves, which we obtain from clean prices of treasury coupon strips in the UK bond market from the United Kingdom Debt Management Office (DMO) [1] according to

$$y_{0T} := -\frac{1}{T}\log P_{0T}, \tag{A.15}$$

using an Actual/Actual day-count convention [2]. We consider the following two data sets:

[1] http://www.dmo.gov.uk/index.aspx?page=Gilts/Daily

1. Bond prices at 146 dates (every other business day) from January 1998 to January 1999, with around 49 to 62 maturities for each date.
2. Bond prices at 157 dates (every Friday) from December 2002 to December 2005, with around 100 to 130 maturities for each date.

Note that the first dataset contains a volatile market including the period of the Long-Term Capital Management (LTCM) crisis, whereas the second dataset corresponds to more moderate market conditions.

We then apply the maximum likelihood estimation (MLE) method suggested by Cairns in [4] [7] to each of the models (A.1)–(A.14) and each starting date in the data sets above. This is done as follows: given a model with parameter vector $\Theta = (b_1, \ldots, b_{N_1}, c_1, \ldots, c_{N_2})$ and a starting date $t_0 = 0$ with available maturities T_i, $i = 1, \ldots, n$, we denote the theoretical prices by $P_{0T_i}(\Theta)$ and the corresponding observed bond prices by $\overline{P}_{0T_i}$. We then assume that

$$\log \overline{P}_{0T_i} \sim \mathcal{N}(\log P_{0T_i}(\Theta), \nu^2(P_{0T_i}(\Theta), d_i)), \tag{A.16}$$

where d_i is the Macaulay duration for the bond and

$$\nu^2(p, d) = \frac{\sigma_0^2(p)[\sigma_\infty^2 d^2 b(p) + 1]}{\sigma_0^2(p) d^2 b(p) + 1}, \quad b(p) = \frac{\sigma_d^2}{\sigma_0^2(p)[\sigma_\infty - \sigma_0^2(p)]}, \tag{A.17}$$

with the error parameters

$$\sigma_0(p) = \frac{1}{3200p}, \quad \sigma_d = 0.0005, \quad \sigma_\infty = 0.001, \tag{A.18}$$

adopted in [4] for the UK bond market data between January 1992 and November 1996. As explained in [4], $\sigma_0(p)$ is based on the assumption that the published bond prices have rounding error of around 1/32 per 100 nominal price, whereas σ_d corresponds to the assumption that the difference between actual and expected yields have independent errors of the order of five basis points. Finally, σ_∞ places a limit on the magnitude of price errors for long dated bonds.

This leads to the log-likelihood function

$$L(\Theta) = -\frac{1}{2} \sum_{i=1}^{n} \Big[\log[2\pi \nu^2(P_{0T_i}(\Theta), d_i)] + \frac{(\log P_{0T_i}(\Theta) - \log \overline{P}_{0T_i})^2}{\nu^2(P_{0T_i}(\Theta), d_i)} \Big]. \tag{A.19}$$

We then use a global search procedure to find the global maximum for the log–likelihood function (A.19), that is, to avoid finding a local maximum, we repeat the procedure using 1000 different random starting points and select the best maximization result.

Having estimated the parameter vector $\widehat{\Theta}$, we denote by $y_{0T_i}(\widehat{\Theta})$ the fitted yield for maturity T_i and by $\overline{y}_{0T_i}$ the corresponding observed yield. We then define the fitting Root-Mean-Squared Percentage Error (RMSPE) as

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \Big[\frac{y_{0T_i}(\widehat{\Theta}) - \overline{y}_{0T_i}}{\overline{y}_{0T_i}} \Big]^2}. \tag{A.20}$$

We then apply the Diebold-Mariano (DM) statistics [8] based on RMSPE to compare fitting performances as is done in [18] and [29]. Here for the computation we

use the program DMARIANO[2] in the statistics package STATA, with a lag order of thirteen in both of our two data sets. The null hypothesis, which is that two models have the same fitting errors, may be rejected at 5% level if the absolute value of the DM statistics is greater than 1.96. We compare the calibration performance of the Chaos Models with a descriptive model for forward rates in Svensson form: the higher the DM statistics, the more a chaotic model outperforms the Svensson model.

Appendix B. Option calibration procedure.

For option calibration, we consider one–variable second chaos models with

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = (b_3 + b_4 s)e^{-c_2 s} \tag{B.1}$$

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s} + b_3 s e^{-c_2 s}, \quad \beta(s) = b_4 s e^{-c_3 s}, \tag{B.2}$$

factorizable second chaos models with

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = b_3 e^{-c_2 s}, \quad \gamma(s) = e^{-c_3 s} \tag{B.3}$$

and one–variable third chaos models with

$$\alpha(s) = b_1 e^{-c_1 s}, \quad \beta(s) = b_2 e^{-c_2 s}, \quad \delta(s) = b_3 e^{-c_3 s} \tag{B.4}$$

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = b_3 e^{-c_2 s}, \quad \delta(s) = b_4 e^{-c_3 s} \tag{B.5}$$

$$\alpha(s) = (b_1 + b_2 s)e^{-c_1 s}, \quad \beta(s) = (b_3 + b_4 s)e^{-c_2 s}, \quad \delta(s) = (b_5 + b_6 s)e^{-c_3 s}, \tag{B.6}$$

as well as each of the three benchmark models described in Section 4.3.

Regarding the data, zero-coupon yields are bootstrapped from the LIBOR, Future and Swap rates (see [9] for the detail of the bootstrapping technique) and interest rate option prices from ICAP (Garban Intercapital - London) and TTKL (Tullett & Tokyo Liberty - London) via the Bloomberg Database.[3] We consider the following two data sets from the UK interest rate market:

- Data between September 2000 and August 2001 at 53 dates (every Friday closing mid price) consisting of
  - zero-coupon yields with 17 maturities ranging from one month to 20 years,
  - implied volatilities for ATM caplets with 37 maturities ranging from one to 10 years,
  - implied volatilities for ATM swaptions with 7 maturities ranging from one month to 5 years and 6 tenors ranging from one to 10 years.
- Data between May 2005 and May 2006 at 53 dates (every Friday closing mid price) consisting of
  - zero-coupon yields with 22 maturities ranging from one month to 20 years.
  - implied volatilities for ATM caplets with 77 maturities ranging from one to 20 years,
  - implied volatilities for ATM swaptions with 7 maturities ranging from one month to 5 years and 6 tenors ranging from one to 10 years.

[2] http://ideas.repec.org/c/boc/bocode/s433001.html

[3] Here we would like to acknowledge helpfulness of the Bloomberg help desk staff, who have aided greatly our understanding of the actual market data. We particularly wish to extend thanks David Culshaw, from ICAP, for his assistance.

Note here that the option data corresponds to a part of the data in [29], where data was analyzed between August 1998 and January 2007. We obtain caplet implied volatilities by bootstrapping ATM cap implied volatilities observed in the market using the technique described in [12], where the ATM caplets implied volatilities maturing at six months and nine months are obtained by constant extrapolation. The extrapolation is necessary to bootstrap the other ATM caplet implied volatilities, but when we calibrate the data, the extrapolated prices give us great errors. Hence, although we follow [12] and implement the extrapolation, we do not use those two short maturities for the calibration. Moreover, we observed some obvious outliers and corrected them accordingly.

For each of the models and data sets above, we perform three distinct calibrations: first to yields and caplets, then to yields and swaptions, and finally to yields, caplets and swaptions. To define the objective function for each to these calibrations, denote the observed yields and prices of ATM caplets and ATM swaptions by $\overline{y}_{0T_i}$, $\overline{\text{Cpl}}(T_i, K_{ATM})$ and $\overline{\text{Swp}}(T_i, K_{ATM})$, their theoretical counterparties by y_{0T_i}, $\text{Cpl}(T_i, K_{ATM})$ and $\text{Swp}(T_i, K_{ATM})$, and the corresponding mean square percentage errors by

$$\text{YieldE} = \sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left[\frac{y_{0T_i} - \overline{y}_{0T_i}}{\overline{y}_{0T_i}}\right]^2}, \tag{B.7}$$

$$\text{CplE} = \sqrt{\frac{1}{n_2}\sum_{i=1}^{n_2}\left[\frac{\text{Cpl}(T_i, K_{ATM}) - \overline{\text{Cpl}}(T_i, K_{ATM})}{\overline{\text{Cpl}}(T_i, K_{ATM})}\right]^2}, \tag{B.8}$$

and

$$\text{SwpE} = \sqrt{\frac{1}{n_3}\sum_{i=1}^{n_3}\left[\frac{\text{Swp}(T_i, K_{ATM}) - \overline{\text{Swp}}(T_i, K_{ATM})}{\overline{\text{Swp}}(T_i, K_{ATM})}\right]^2}. \tag{B.9}$$

For the calibration to yields and caplets, we then minimize the objective function

$$\text{TotalE1} = \sqrt{(\text{YieldE})^2 + (\text{CplE})^2}. \tag{B.10}$$

Similarly, for the calibration to yields and swaptions, we minimize the objective function

$$\text{TotalE2} = \sqrt{(\text{YieldE})^2 + (\text{SwpE})^2}. \tag{B.11}$$

Finally, for the calibration to yields, caplets and swaptions we minimize the objective function

$$\text{TotalE3} = \sqrt{(\text{YieldE})^2 + (\text{CplE})^2 + (\text{SwpE})^2}. \tag{B.12}$$

For each of these calibrations, we test the pricing performance of the calibrated models. For example, after calibrating to yields and ATM swaptions, we use the model to price the ATM Caplets and compute the pricing error from market ATM Caplet prices. Whenever possible, for example in the Hull-White model, we calibrate initial

yield curves and options separately by minimizing their respective square errors, since these involve distinct sets of parameters. Moreover, for the LFM model, we minimize the square error in implied volatilities rather than actual prices.

To access the relative performance between the models, we use the Akaike Information Criterion (AIC) and the model selection relative frequency described in [3]. The AIC is formed with the maximized value of the likelihood function L for an estimated model and the number of parameters k in the following way:

$$\mathrm{AIC} = -2\log(L) + 2k. \tag{B.13}$$

For the least squares method under the assumed normality of residuals, this reduces to .

$$\mathrm{AIC} = n\log\left(\frac{\mathrm{RSS}}{n}\right) + 2k + n\log(2\pi) + \frac{n}{2}, \tag{B.14}$$

where RSS is the fitted residual sum of squares. Since $n\log(2\pi) + \frac{n}{2}$ is a constant, we ignore this term and conclude that

$$\mathrm{AIC} = n\log\left(\frac{\mathrm{RSS}}{n}\right) + 2k. \tag{B.15}$$

To define the model selection relative frequency, let us suppose we have two models and N calibration sets (for example the different calibration dates in our data). For a data set $j \in \{1, \cdots, N\}$ we compute AIC denoted $AIC_j^{(1)}$ for one model and $AIC_j^{(2)}$ for the other. Suppose the AIC of the first model is smaller than the AIC of the other model l times. Then the model selection relative frequencies (MSRF) for the first model and the second model are computed respectively by

$$\mathrm{MSRF}^{(1)} = \frac{l}{N} \quad \text{and} \quad \mathrm{MSRF}^{(2)} = \frac{N-l}{N}. \tag{B.16}$$

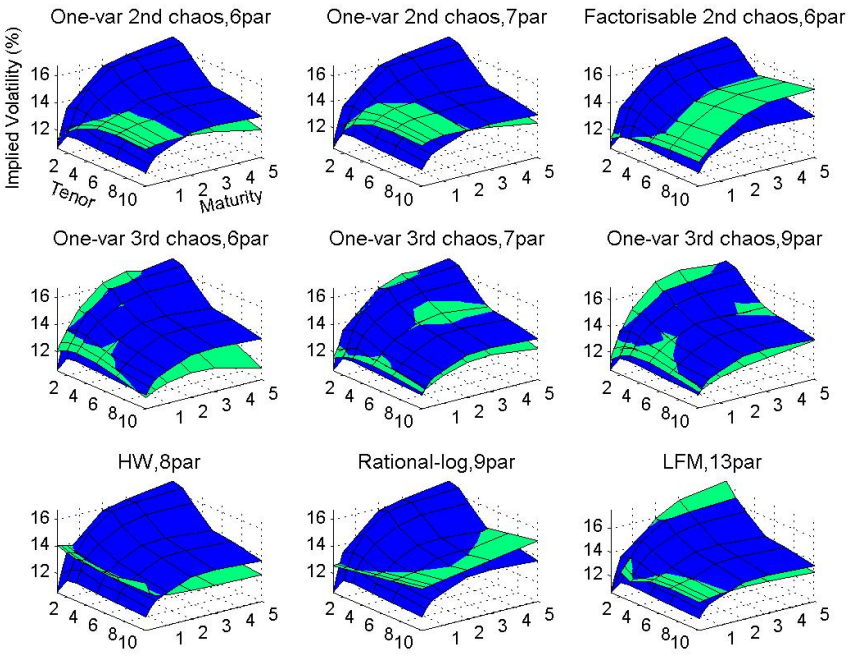

FIG. 4.1. *Implied volatility for ATM swaptions on October 20th, 2000 (Blue: Market Quotes, Green: Theoretical Values).*

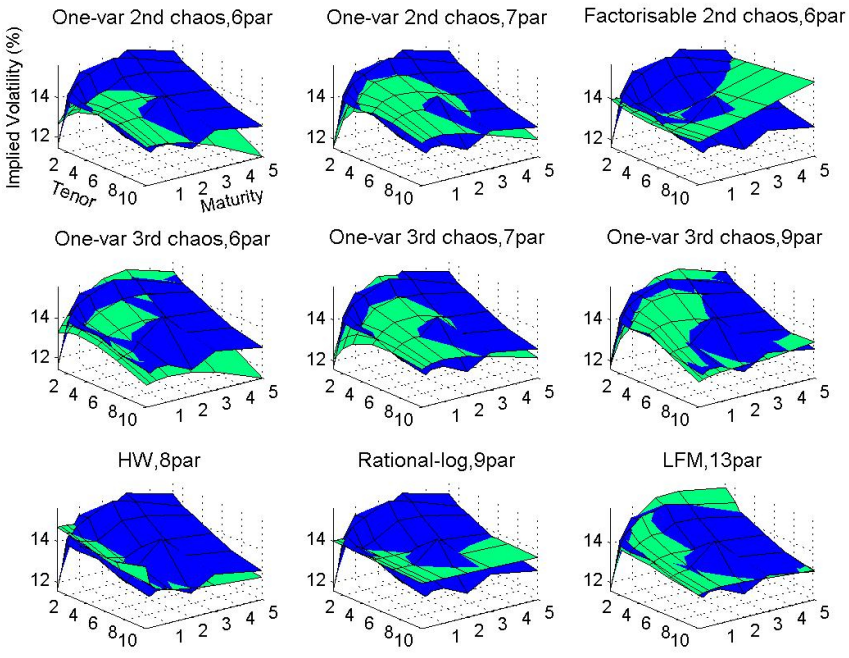

FIG. 4.2. *Implied volatility for ATM swaptions on August 12th, 2005 (Blue: Market Quotes, Green: Theoretical Values).*

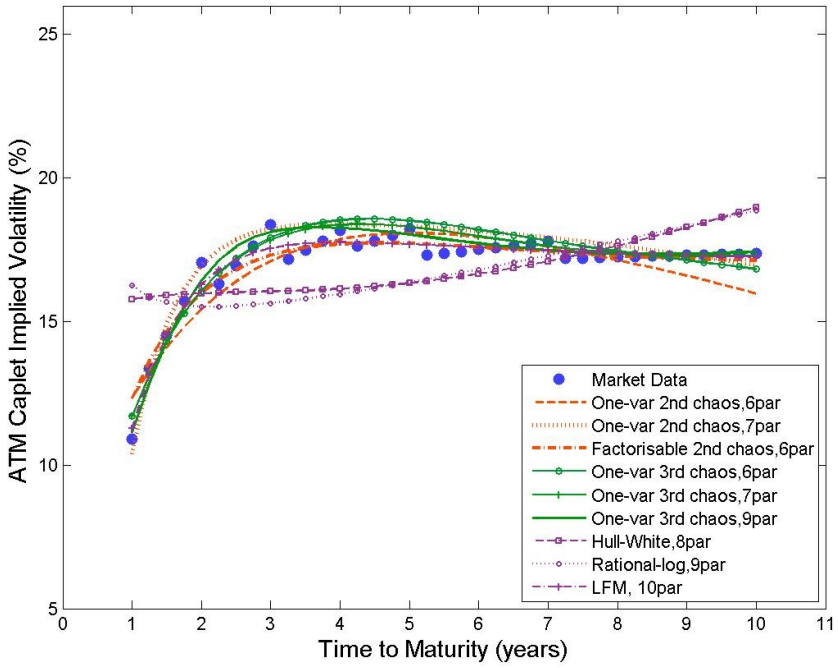

FIG. 4.3. *Implied volatility for ATM caplets on December 8th, 2000.*

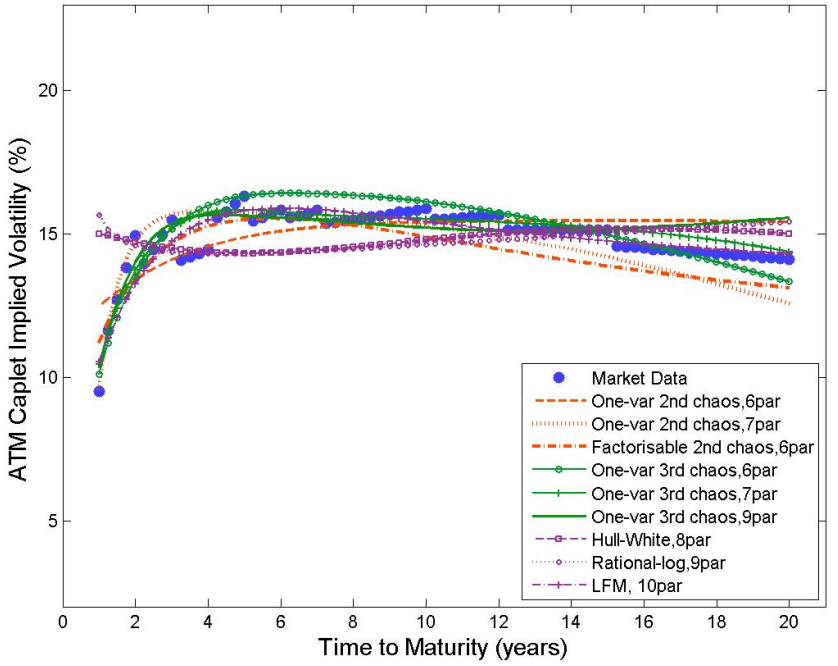

FIG. 4.4. *Implied volatility for ATM caplets on February 10th, 2006.*

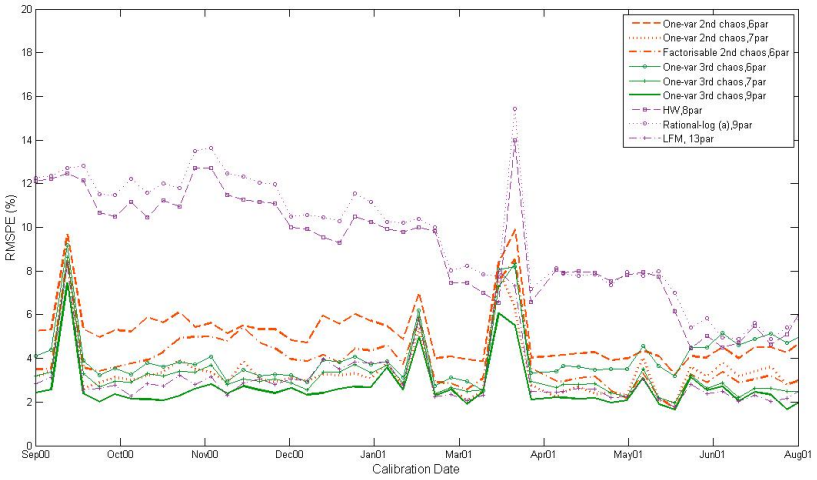

FIG. 4.5. *Total error for yields and ATM caplets in 2000–2001.*

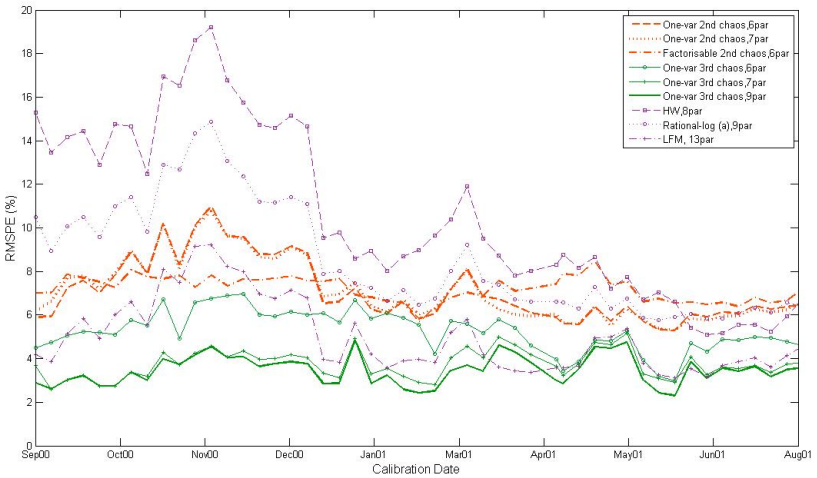

FIG. 4.6. *Total error for yields and ATM swaptions in 2000–2001.*

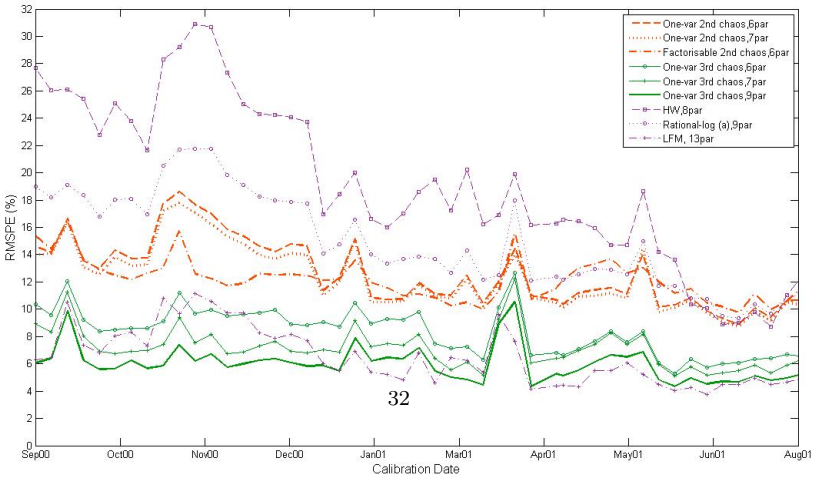

FIG. 4.7. *Total error for yields, ATM caplets and ATM swaptions in 2000–2001.*

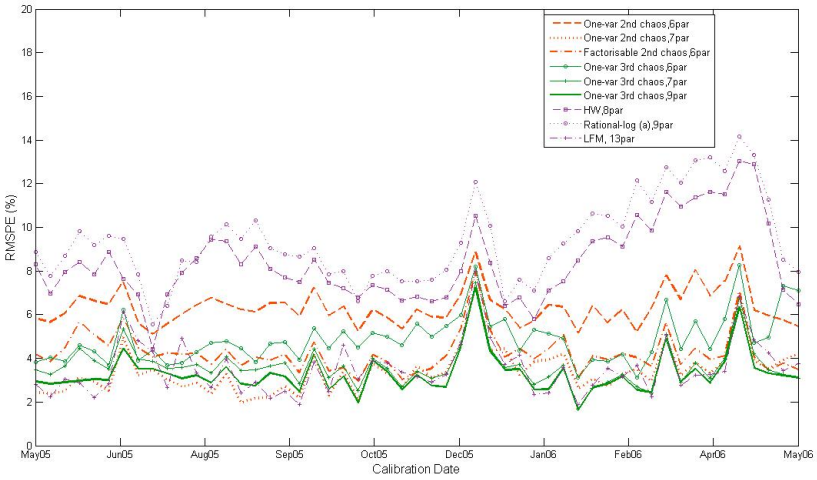

FIG. 4.8. *Total error for yields and ATM caplets in 2005–2006.*

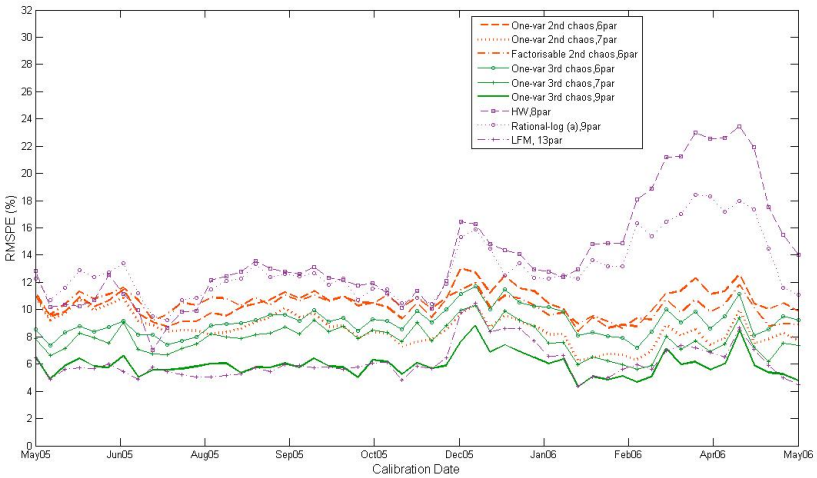

FIG. 4.9. *Total error for yields and ATM swaptions in 2005–2006.*

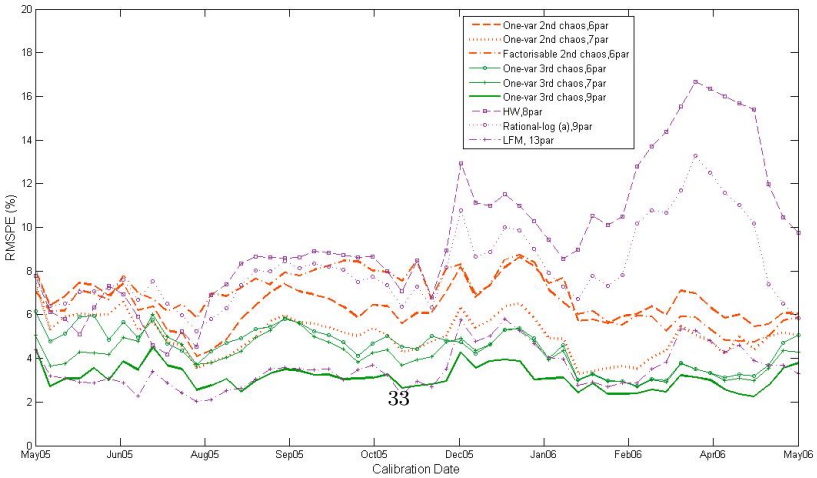

FIG. 4.10. *Total error for yields, ATM caplets and ATM swaptions in 2005–2006.*