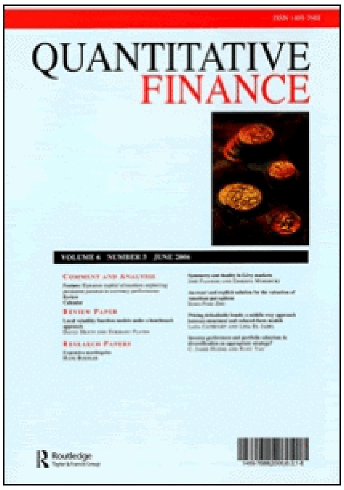

Quantitative Finance
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713665537

On the calibration of a Gaussian Heath-Jarrow-Morton model using consistent forward rate curves

A. Falcó [a]; LL. Navarro [a];J. Nave [b]
[a] Departamento de Ciencias Físicas, Matemáticas y de la Computación, Universidad CEU Cardenal Herrera, Spain [b] Departamento de Economía y Empresa, Universidad CEU Cardenal Herrera, Spain

First published on: 11 May 2010

To cite this Article Falcó, A. , Navarro, LL. andNave, J.(2010) 'On the calibration of a Gaussian Heath-Jarrow-Morton model using consistent forward rate curves', Quantitative Finance,, First published on: 11 May 2010 (iFirst)
To link to this Article: DOI: 10.1080/14697680903493565
URL: http://dx.doi.org/10.1080/14697680903493565

Quantitative Finance, 2010, 1–10, iFirst

On the calibration of a Gaussian Heath–Jarrow–Morton model using consistent forward rate curves

A. FALCÓ†, LL. NAVARRO*† and J. NAVE‡

†Departamento de Ciencias Físicas, Matemáticas y de la Computación, Universidad CEU Cardenal Herrera, San Bartolomé 55, 46115 Alfara del Patriarca (Valencia), Spain
‡Departamento de Economía y Empresa, Universidad CEU Cardenal Herrera, Luis Vives 1, 46115 Alfara del Patriarca (Valencia), Spain

(*Received 24 March 2007; in final form 25 September 2009*)

In this paper we propose a calibration algorithm, using a consistent family of yield curves, that fits a Gaussian Heath–Jarrow–Morton model jointly to the implied volatilities of caps and zero-coupon bond prices. The calibration approach is evaluated in terms of in-sample data fitting as well as stability of parameter estimates. Furthermore, the efficiency is tested against a non-consistent traditional method using simulated and market data. Also, we discuss the convergence of the algorithm by means of Monte Carlo simulations.

Keywords: Calibration of deterministic volatility; Control of stochastic systems; Derivatives pricing; Fixed income derivatives

1. Introduction

Any acceptable model that prices interest rate derivatives must fit the observed term structure. This idea, pioneered by Ho and Lee (1986), has been explored in the past by many other researchers, including Black and Karasinski (1991) and Hull and White (1990).

Contemporary models are more complex because they consider the evolution of the whole forward curve as an infinite system of stochastic differential equations (Heath *et al.* 1992) (HJM). In particular, they use a continuous forward rate curve as initial input. In reality, one only observes a discrete set composed either of bond prices or swap rates. Therefore, in practice, the usual approach is to interpolate the forward curve using splines or other parametrized families of functions.

A very plausible question arises at this point: Choose a specific parametric family, $\mathcal{G}$, of functions that represent the forward curve, and also an arbitrage-free interest rate model $\mathcal{M}$. Assume that we use an initial curve that lies within the input for model $\mathcal{M}$. Will this interest rate model evolve through forward curves that lie within the family? Motivated by this question, Björk and Christensen (1999) define the so-called consistent pairs $(\mathcal{M}, \mathcal{G})$ as those whose answer to the above question is positive. In particular, they studied the problem of consistency between the family of curves proposed by Nelson and Siegel (1987) and any HJM interest rate model with deterministic volatility, concluding that there is no such interest model consistent with it.

We remark that the Nelson and Siegel interpolating scheme is an important example of a parametric family of forward curves, because it is widely adopted by central banks (see, for instance, BIS 2005). Its forward curve shape, $G_{\mathrm{NS}}(z, \cdot)$, is given by the expression

$$G_{\mathrm{NS}}(z, x) = z_1 + z_2\, \mathrm{e}^{-z_4 x} + z_3 x\, \mathrm{e}^{-z_4 x},$$

where x denotes time to maturity and z the parameter vector

$$z = (z_1, z_2, \ldots).$$

Despite all the positive empirical features and general acceptance by the financial community, Filipovic (1999) has shown that there is no Itô process that is consistent with the Nelson–Siegel family. In a recent study, De Rossi (2004) applies consistency results to propose a consistent

*Corresponding author. Email: lluis2001@gmail.com

Quantitative Finance
ISSN 1469–7688 print/ISSN 1469–7696 online © 2010 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/14697680903493565

exponential dynamic model, and estimates it using data on LIBOR and UK swap rates. On the other hand, Buraschi and Corielli (2005) add some results to this theoretical framework, indicating that the use of inconsistent parametric families to obtain smooth interest rate curves violates the standard self-financing arguments of replicating strategies, with direct consequences in risk management procedures.

In order to illustrate this situation, we describe a very common fixed-income market procedure. In the real world, practitioners usually re-estimate the yield curve and HJM model parameters on a daily basis. This procedure consists of two steps.

- They fit the initial yield curve from discrete market data (bond prices, swap rates, short-term zero rates).
- They obtain an estimate of the parameters of the HJM model, minimizing the pricing error of some actively traded (plain vanilla) interest rate derivatives (commonly swap options or caps).

In contrast to the parsimonious assumption that model parameters are constant, an unstable HJM model parameter estimation is often observed. Perhaps this fact is not relevant for *mark to market*, but it could have practical consequences for the hedging portfolios associated with these financial instruments. It should be remembered that such dynamic strategies depend on the model assumptions. Thus, re-calibration is conceivable because practitioners are aware of *model risk*. A particular HJM model is not a perfect description of reality, and they are forced to re-estimate day-to-day model parameters in order to include new information that arrives from the market. On the other hand, unstable estimates may be the result of factors that are more theoretical, because the above-mentioned setup does not take into account that HJM model parameters are linked, in general, to the initial yield curve fit parameters. If a practitioner uses an interpolation scheme that is not consistent with the model, then the parameters will be artificially forced to change. Thus, it seems there are a plethora of motivations for the study of the empirical evidence and the practical implications that are predicted by a consistent HJM build model.

The consistency hypothesis stated by Björk implies that the zero coupon bond curve has to be determined at the same time as the parameters of the model. Angelini and Herzel (2002, 2005) propose the use of an optimization program related to the mentioned daily calibrations that is compatible with this joint estimation. The milestone of this methodology is the use of an objective function based on an error measure for just the caps portfolio. Then, the theoretical prices for the caps along the minimization of this measure can be calculated at the same time that the yield curve is fitted. This is an efficient method because consistent families of yield curves behave well in a Gaussian framework.

The purpose of this work is to extend the above strategy to a more general framework. It modifies the mentioned objective function by taking into account the error measure for the discount bonds estimation. To this we add an objective function using a convex combination of the cap and the bond error measures, by means of a fixed parameter. As a matter of fact, this approach is richer in possible outcomes. We also test the robustness of this extension using Monte-Carlo simulation.

To this end, we restrict ourselves to a particular humped volatility HJM model, proposed by Ritchken and Chuang (1999) and Mercurio and Moraleda (2000) independently. We will discuss this formalization and give some theoretical results relevant to our analysis. We chose this model because it is quite popular and analytically treatable. In particular, it provides closed formulas for Europeans caps. Moreover, it is the one-factor Gaussian model that seems better able to reproduce the humped shape of the implied volatility term structure for caps, that the normal market scenarios usually depict. Moreover, it is also the most flexible in other market conditions. We perform our study by calibrating this model, first using simulated data and second by a market data set composed of the Euro and US discount factors and the cap at-the-money flat volatilities quotes in two different periods, as shown by figure 1 for the particular case of the US market. For both Euro and US markets, the first scenario depicts a normal fixed-income market scenario, the term structure of implied volatilities for caps (hereafter TSV) have a humped shape and the term structure of interest rates (hereafter TSIR) is decreasing in the short term with a local minimum, and then it increases to mid–long-term maturities as a spoon shape. On the other hand, the second period may be considered a volatile period with a TSV monotonically decreasing, and with a higher overall implied volatility and a TSIR monotonically increasing without a local minimum. To our knowledge this is the first attempt to extend the search for empirical evidence to US market data.

The paper is organized as follows. In section 2 we give a brief overview of the model and present in this context the option valuation and the construction of the families consistent with the model. Section 3 describes the calibration procedure. Section 4 is devoted to empirical results, first comparing the consistent calibration algorithm with the non-consistent approaches with simulated data, and then presenting the results of the fitting of the different models with market data. In the last section we give some final conclusions and remarks.

2. The model

Let W be a one-dimensional Wiener–Einstein stochastic process defined in a complete probability space $(\Omega, \mathcal{F}, P)$. The single-factor Heath–Jarrow–Morton (Heath *et al.* 1992) framework is based on the dynamics of the entire forward rate curve, $\{r_t(x),\ x > 0\}$. Thus, under Musiela's (1993) parametrization it follows that the infinite-dimensional diffusion process given by

$$\begin{cases} \mathrm{d}r_t(x) = \beta(r_t, x)\,\mathrm{d}t + \sigma(r_t, x)\,\mathrm{d}W_t, \\ r_0(x) = r^*(x), \end{cases} \tag{1}$$

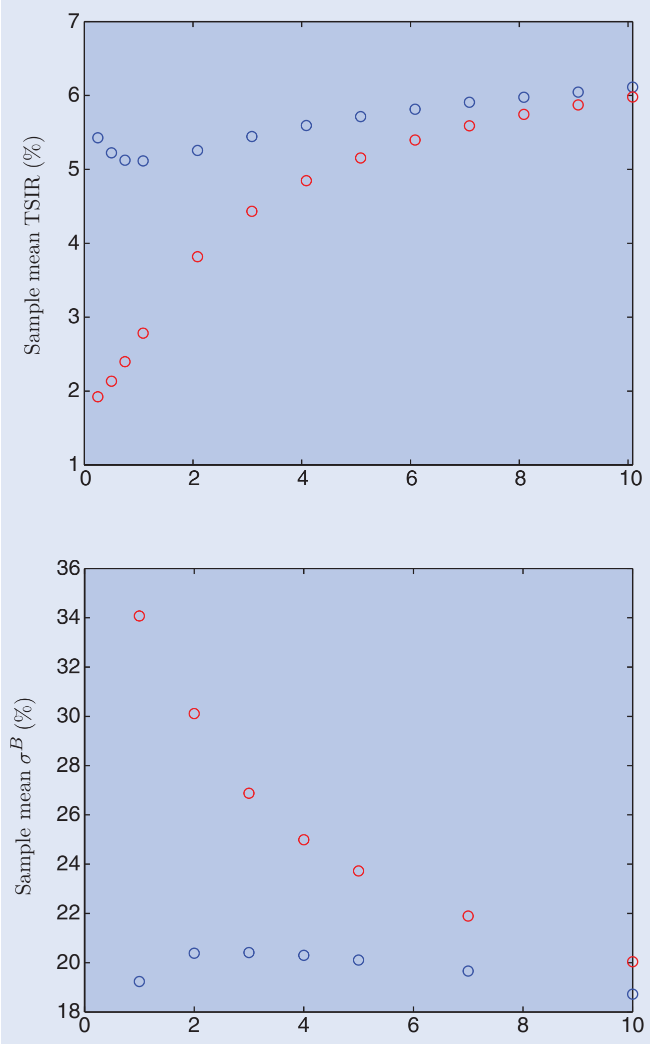

Figure 1. Market TSIR and TSV data in the two different market scenarios.

where $\{r^*(x),\ x \geq 0\}$, can be interpreted as the *observed* forward rate curve. The standard drift condition derived by Heath *et al.* (1992) can easily be transferred to the Musiela parametrization (see, for instance, Musiela 1993),

$$\beta(r_t, x) = \frac{\partial}{\partial x} r_t(x) + \sigma(r_t, x) \int_0^x \sigma(r_t, s)\,\mathrm{d}s.$$

Thus, a particular model is constructed by the choice of an explicit volatility function $\sigma(r_t, x)$. Our work is devoted to a Gaussian humped volatility model where

$$\sigma(r_t, x) = \sigma(x) = (\alpha + \beta x)\mathrm{e}^{-ax},$$

i.e. σ is a deterministic function depending only on time to maturity, and then $r_t(x)$ is a Gaussian process.

2.1. *Finite-dimensional realizations of Gaussian models*

It should also be noted that $\sigma(x)$ is a one-dimensional quasi-exponential (QE) function, because it is of the form

$$f(x) = \sum_i \mathrm{e}^{\lambda_i x} + \sum_i \mathrm{e}^{\alpha_i x}[p_i(x)\cos(\omega_i x) + q_i(x)\sin(\omega_i x)],$$

with λ_i, α_i and ω_i being real numbers and p_i and q_i real polynomials. It is well known that if $f(x)$ is an m-dimensional QE function, then it admits the following matrix representation:

$$f(x) = c\,\mathrm{e}^{Ax}B,$$

where A is an $(n \times n)$ matrix, B is an $(n \times m)$ matrix and c is an n-dimensional row vector (see lemma 2.1 of Björk (2003)). Thus, $\sigma(x)$ can be written as

$$\sigma(x) = c\,\mathrm{e}^{Ax}b, \tag{2}$$

where

$$c = [\alpha \quad \beta - a\alpha],$$
$$A = \begin{bmatrix} 0 & -a^2 \\ 1 & -2a \end{bmatrix},$$
$$b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

By means of proposition 2.1 of Björk (2001) we can write the forward rate equation (1) as

$$\mathrm{d}q_t(x) = \mathbf{F}q_t(x)\,\mathrm{d}t + \sigma(x)\,\mathrm{d}W_t, \quad q_0(x) = 0, \tag{3}$$

$$r_t(x) = q_t(x) + \delta_t(x), \tag{4}$$

where $\mathbf{F}$ is a linear operator defined by

$$\mathbf{F} = \frac{\partial}{\partial x},$$

and $\delta_t(x)$ is the deterministic process given by

$$\delta_t(x) = r^*(x+t) + \int_0^t \Sigma(x+t-s)\,\mathrm{d}s,$$

with $\Sigma(\cdot) = \sigma(\cdot)\int_0^{\cdot} \sigma(s)\,\mathrm{d}s$. Moreover, $q_t(x)$ has the concrete *finite-dimensional* realization

$$\mathrm{d}Z_t = AZ_t\,\mathrm{d}t + b\,\mathrm{d}W_t, \quad Z_0 = 0, \tag{5}$$

$$q_t(x) = C(x)Z_t, \tag{6}$$

because σ is a QE function (see, for instance, proposition 2.3 of Björk (2003)) with A and b as in (2) and $C(x) = c\,\mathrm{e}^{Ax}$. Thus, (5) is a linear SDE in the narrow sense (see Kloeden and Platen (1999) for details) with explicit solution

$$Z_t = \Phi_t \int_0^t \Phi_s^{-1} b\,\mathrm{d}W_s, \tag{7}$$

where

$$\Phi_t = \mathrm{e}^{At} = \mathrm{e}^{-at}\begin{bmatrix} 1 + at & -a^2 t \\ t & 1 - at \end{bmatrix}.$$

Now, with the definition of $S(x) = \int_0^x \sigma(u)\,\mathrm{d}u$, it is easy to obtain that

$$\int_0^t \Sigma(t+x-s)\,\mathrm{d}s = \frac{1}{2}[S^2(t+x) - S^2(x)],$$

and, therefore, combining these explicit results with decomposition (4) we arrive at the operative expression

$$r_t(x) = r^*(x+t) + \frac{1}{2}[S^2(t+x) - S^2(x)] + C(x)Z_t. \tag{8}$$

The most striking feature of the result sketched in (8) is that, starting from the initial forward curve $r^*(x)$, it allows the use of the Monte-Carlo simulation of future forward curves produced by this particular HJM model. On the other hand, as we will show later, equation (8) can also be used to build the initial forward rate curve $r^*(x)$. It must be remembered that it is consistent with the dynamics of the model.

2.2. *Interest rate option pricing*

To calibrate the model by means of real data, we actually need to determine the vector parameter $\theta = (\alpha, \beta, a)$. In order to estimate the forward rate volatility, the statistical analysis of past data is a possible approach, but practitioners usually prefer implied volatility (lying within some derivative market prices) based techniques. This strategy involves a minimization problem where the loss function can be taken as

$$l(\theta) = \sum_{i=1}^{n} (\zeta_i^* - \zeta(\theta, T_i))^2,$$

where $\zeta(\theta, T_i)$ is the ith theoretical derivative price maturing at time $t = T_i$, and $\zeta_i^* \equiv \zeta^*(T_i)$ is the ith market price. As is well known (see proposition 24.15 and pp. 364–366 of Björk 2004) the price, at $t = 0$, of the cap is given by

$$\zeta(T) = (1 + \tau K)\left(\sum_{j=0}^{n-1} \kappa D(x_j) N(-d_+) - D(x_{j+1}) N(-d_-)\right), \tag{9}$$

where

$$d_\pm = \frac{\ln[D(x_j)/\kappa D(x_{j+1})] \pm \frac{1}{2}\vartheta^2(x_j)}{\vartheta(x_j)}, \tag{10}$$

the interval $[0, T]$ is subdivided with equidistant points, i.e.

$$x_j = (j+1)\tau, \quad j = 0, 1, \ldots, n, \tag{11}$$

$D(\cdot)$ is the initial discount function, and κ equals $(1 + \tau K)^{-1}$ with K denoting the *cap rate*.

The variable ϑ in (10) is intimately related to the concrete multifactor Gaussian HJM model realization via the particular $[A, B, c]$ forward rate TSV selection:

$$\vartheta^2(x_j) = M(x_j) F(x_j) M'(x_j),$$

where $M(x_j)$ is the matrix

$$M(x_j) = cA^{-1}(\mathrm{e}^{A(x_j+\tau)} - \mathrm{e}^{Ax_j}),$$

and $F(\cdot)$ satisfies

$$F(\cdot) = \int_0^{\cdot} \mathrm{e}^{-As} BB' \mathrm{e}^{-A's}\, \mathrm{d}s.$$

Although inversion of the matrix A, the series expansion of e^{Ax}, reveals that M is not a singular matrix even for small values of parameter a. This result is also true for other Gaussian HJM models built from QE forward TSV families, because the matrix elements of A are, fortunately, polynomial functions of the model parameters. However, due to the numerical instability of the calibration process, when $a \to 0$, an asymptotically equivalent expression for ϑ must be used. Equations (9) and (10) also express the effective influence of *ab initio* yield curve estimation on cap pricing.

2.3. *Consistent curves with Gaussian models*

If we want to measure the actual impact that alternatives to the Nelson–Siegel yield curve interpolating approach produces on derivatives pricing and hedging, we need to determine consistent families for this particular model. The fundamental results can be found in Björk and Christensen (1999) in more detail. We adapt some of them to our Gaussian case study without further technical discussion for the general case.

Definition 2.1: Consider the space $\mathcal{H}$ defined as the space of all $\mathcal{C}^\infty$ functions,

$$r : \mathcal{R}_+ \to \mathcal{R},$$

satisfying the norm condition

$$\|r\|^2 = \sum_{n=0}^{\infty} 2^{-n} \int_0^\infty \left(\frac{\mathrm{d}^n r}{\mathrm{d}x^n}(x)\right)^2 \mathrm{e}^{-\gamma x}\, \mathrm{d}x < \infty,$$

where γ is a fixed positive real number.

As proved by Björk and Svensson (2001, proposition 4.2), this space $\mathcal{H}$ is a Hilbert space.

Theorem 2.2: *Consider as given the mapping*

$$G\colon \mathcal{Z} \to \mathcal{H},$$

where the parameter space $\mathcal{Z}$ is an open connected subset of R^d, $\mathcal{H}$ is a Hilbert space and the forward curve manifold $\mathcal{G} \subseteq H$ is defined as $\mathcal{G} = \mathrm{Im}(G)$. The family $\mathcal{G}$ is consistent with the one-factor model $\mathcal{M}$ with deterministic volatility function $\sigma(\cdot)$, if and only if

$$\partial_x G(z, x) + \sigma(x) \int_0^x \sigma(s)\mathrm{d}s \in \mathrm{Im}[\partial_z G(z, x)], \tag{12}$$

$$\sigma(x) \in \mathrm{Im}[\partial_z G(z, x)], \tag{13}$$

for all $z \in Z$.

Statements (12) and (13) are called, respectively, *the consistent drift* and *the consistent volatility* conditions. These are easy to apply in concrete cases as shown by Björk and Christensen (1999) and De Rossi (2004), among others. For the particular one-factor model we consider throughout this work, propositions 7.2 and 7.3 of Björk and Christensen (1999) may be directly applied to obtain the following useful result.

Proposition 2.3: *The family*

$$G_m(z, x) = (z_1 + z_2 x)\, \mathrm{e}^{-ax} + (z_3 + z_4 x + z_5 x^2)\, \mathrm{e}^{-2ax} \tag{14}$$

is the minimal dimension consistent family with the model characterized by $\sigma(x) = (\alpha + \beta x)\, \mathrm{e}^{-ax}$.

Moreover, it should also be noted that *augmented* families related to (14) can be constructed by adding to G_m an arbitrary function ϕ, that is, the map

$$G(z,x) = G_m(z,x) + \phi(z,x)$$

is also consistent with this model.

There is an alternative way to justify (14) by focusing on forward rate evolution deduced at (8), and to obtain an insight into how the Monte-Carlo procedure is implemented, we describe it next. From the definition of $S(x)$, we have $S'(x)=\sigma(x)$. Then it is easy to derive that deterministic term $\frac{1}{2}[S^2(t+x)-S^2(x)]$ is of the form

$$g_1(t)\mathrm{e}^{-2ax} + g_2(t)x\mathrm{e}^{-2ax} + g_3(t)x^2\mathrm{e}^{-2ax} + h_1(t)\mathrm{e}^{-ax} + h_2(t)x\mathrm{e}^{-2ax}.$$

On the other hand, the explicit expansion of stochastic term $C(x)Z_t$,

$$c\mathrm{e}^{Ax}\begin{bmatrix} Z_t^1 \\ Z_t^2 \end{bmatrix} = \mathrm{e}^{-ax}[\alpha\ \beta - a\alpha]\begin{bmatrix} 1+ax & -a^2x \\ x & 1-ax \end{bmatrix}\begin{bmatrix} Z_t^1 \\ Z_t^2 \end{bmatrix}$$
$$= \mathrm{e}^{-ax}\left(\alpha Z_t^1 - a\alpha Z_t^2 + \beta Z_t^2\right) + x\mathrm{e}^{-ax}\left(\beta Z_t^1 - a\beta Z_t^2\right),$$

and the forward rate evolution becomes

$$\begin{aligned} r_t(x) &= r^*(x+t) + g_1(t)\mathrm{e}^{-2ax} + g_2(t)x\mathrm{e}^{-2ax} + g_3(t)x^2\mathrm{e}^{-2ax} \\ &\quad + (h_1(t) + \alpha Z_t^1 - a\alpha Z_t^2 + \beta Z_t^2)\mathrm{e}^{-ax} \\ &\quad + (h_2(t) + \beta Z_t^1 - a\beta Z_t^2)x\mathrm{e}^{-ax}. \end{aligned} \quad (15)$$

From (15) we see that a family that is invariant under time translation is consistent with the model if and only if it contains the linear space $\{\mathrm{e}^{-ax}, x\mathrm{e}^{-ax}, \mathrm{e}^{-2ax}, x\mathrm{e}^{-2ax}, x^2\mathrm{e}^{-2ax}\}$. Consequently, to make a consistent version of a translation invariant family $\phi(z,x)$ it is sufficient to add $G_m(z,x)$.

The following concluding remarks concerning the families used throughout this work should now be clear.

- The Nelson–Siegel family (henceforth NS)
  $$G_{\mathrm{NS}}(z,x) = z_1 + z_2\mathrm{e}^{-z_4x} + z_3x\mathrm{e}^{-z_4x}$$
  is not consistent with the model.
- The family
  $$G_m(z,x) = (z_1 + z_2x)\mathrm{e}^{-ax} + (z_3 + z_4x + z_5x^2)\mathrm{e}^{-2ax}$$
  is the lowest dimension family consistent with the model (hereafter MC).
- The family
  $$G_{\mathrm{ANS}}(z,x) = z_1 + z_2\mathrm{e}^{-ax} + z_3x\mathrm{e}^{-ax} + (z_4 + z_5x + z_6x^2)\mathrm{e}^{-2ax}$$
  is the simplest adjustment based on the restricted NS family that allows model consistency (hereafter ANS).

3. Calibration to market data approaches

The calibration procedures can be described formally as follows. Let θ be the vector (α,β,a) of parameter values for the model under consideration. Assume that we have time series observations of the implied volatilities, σ_i^B, of N caps, with different ATM *strikes*, K_i, and maturities T_i with $i=1,\ldots,N$ (here $N=7$). Suppose that, at time $t=0$, we are also equipped with the discount function estimation, $D(x)$, and that the market participants translate volatility quotes into cash quotes adopting the *Black* framework. In doing so, they adopt the convention that K_i quantities must match forward swap rates of the interest rate swaps (IRS) with the same reset periods as the ith cap (these IRS start their cash flows at $t=x_0+\tau$ as the corresponding cap and have no cash value at $t=0$):

$$K_i = \frac{D(\tau) - D(T_i)}{\tau\sum_{j=1}^{n} D(x_j)}, \quad (16)$$

where τ is the length of the underlying caplets, and $x_1=2\tau,\ldots,x_n=T_i$. Derivation of formula (16) can be found, for example, in Björk (2004, proposition 20.7, p. 313). Now, by inspection, it is clear that this market convention means that K_i depends on the yield curve estimation. It allows us to denote market prices of caps by $\zeta^*(T_i, D(x), K_i(D(x)), \sigma_i^B)$. This latter expression emphasizes the explicit and implicit dependence (through ATM *strikes*) on discount function estimation even for market prices. Let $\zeta(T_i,\ D(x),\ K_i(D(x)),\ \theta)$ be the corresponding theoretical price under our particular model.

3.1. The two-step traditional method

First, we choose a non-consistent parametrized family of forward rate curves $G(z,x)$. Let $D(z,x)$ be the zero-coupon bond prices reported by $G(z,x)$. Let D_k^* be the corresponding discount factor observations on maturities x_k with $k=1,\ldots,M=11$. For each zero-coupon bond denoted by subscript k, the logarithmic pricing error† is written as follows

$$\epsilon_k(z) = \log D_k^* - \log D(z,x_k).$$

Then, we have chosen in this work the sum of squared pricing errors, *SSE*, as objective function to minimize

$$SSE_D = \min_z \|\log D^* - \log D(z,x)\|^2 = \min_z \sum_{k=1}^{M} \epsilon_k^2(z). \quad (17)$$

Now, via the least-squares estimators $\hat{z}$, an entire discount factor estimation allows us to price the caps using market practice or a HJM model. Following a similar scheme for the derivatives fitting as that used at the bond side we have

$$\varepsilon_i(\theta) = \log\zeta_i^* - \log\zeta(\theta, T_i)$$

and

$$SSE_C = \min_\theta \|\log\zeta^* - \log\zeta(\theta,T)\|^2 = \min_\theta \sum_{i=1}^{N} \varepsilon_i^2(\theta), \quad (18)$$

where we have suppressed dependencies for simplicity. Note that yield curve estimation is external to the model in the sense that there is no need to know beforehand any of the model parameters θ for solving nonlinear program (17).

†Recall that, for small ϵ_k, it is also the relative pricing error $[D_k^* - D(z,x_k)]/D(z,x_k)$.

3.2. The joint calibration to cap and bond prices

Let us now describe in detail the joint cap–bond calibration procedure that makes sense in a consistent family framework. We note that, in this situation, the parameters of the model are determined together with the initial forward rate curve. This is different from the traditional fitting of HJM models, where the two steps are separate, as discussed before. From expression (14) we note the dependency of the family on the parameter a. Let $G(z,x,a)$ be a family consistent with our model (for instance, G_m and G_{ANS}) and define least-squares estimators $\hat{z}(a)$:

$$\hat{z}(a) = \arg\min_{z} \sum_{k=1}^{M} (\log D_k^* - \log D(z, x_k, a))^2. \qquad (19)$$

From the expression

$$\log D(z, x_k, a) = -\int_0^{x_k} G(z, s, a)\, \mathrm{d}s = \sum_{j=1}^{n_p} M_{kj}(a) z_j,$$

we note that, for consistent families and for a fixed a, problem (19) is linear in z parameters (for the G_m family $n_p = 5$, and for the G_{ANS} family $n_p = 6$). Thus, $\hat{z}$ is an explicit and continuous function of a. With yield-curve estimation implemented for every fixed a, the entire discount function $D(\hat{z}(\theta), x, a)$ may be determined and it may be envisaged that the estimates $\hat{\theta}$ have to be found by solving the nonlinear program

$$SSE_C = \min_{\theta} \| \log \zeta^*[D(\hat{z}(\theta))] - \log \zeta[D(\hat{z}(\theta), \theta, T)] \|^2$$
$$= \min_{\theta} \sum_{i=1}^{N} \varepsilon_i^2(\theta). \qquad (20)$$

However, following the latter program we are not sure that the corresponding yield curve at the minimum $\hat{\theta}$, $D(\hat{z}(\hat{\theta}), x, \hat{\theta})$, was the optimal value of the sequence of yield-curve estimations implicit in the program (20). In other words, there exist reasonable doubts concerning the convergence of this algorithm because both error measures compete, in general. Now, we consider the following decomposition for the total loss function $SSE(\theta)$:

$$SSE_D(\theta) = \| \log D^* - M(\theta)\hat{z}(\theta) \|^2, \qquad (21)$$

$$SSE_C(\theta) = \| \log \zeta^*[D(\hat{z}(\theta))] - \log \zeta[D(\hat{z}(\theta)), \theta, T)] \|^2. \qquad (22)$$

Then, as a heuristic solution, we propose to modify the latter program to include pricing residuals for the discount through the convex combination

$$SSE_\lambda = \min_{\theta}((1-\lambda)\, SSE_D(\theta) + \lambda\, SSE_C(\theta)), \qquad (23)$$

for some fixed $\lambda \in [0, 1]$.

At this point, note that the program used by Angelini and Herzel (2002, 2005) in their work uses a different goal attainment,

$$SSE = \min_{\theta} SSE_C(\theta), \qquad (24)$$

where $SSE_C(\theta)$ and $\hat{z}(a)$ are defined via the identities (20) and (19). As a consequence, the program used by

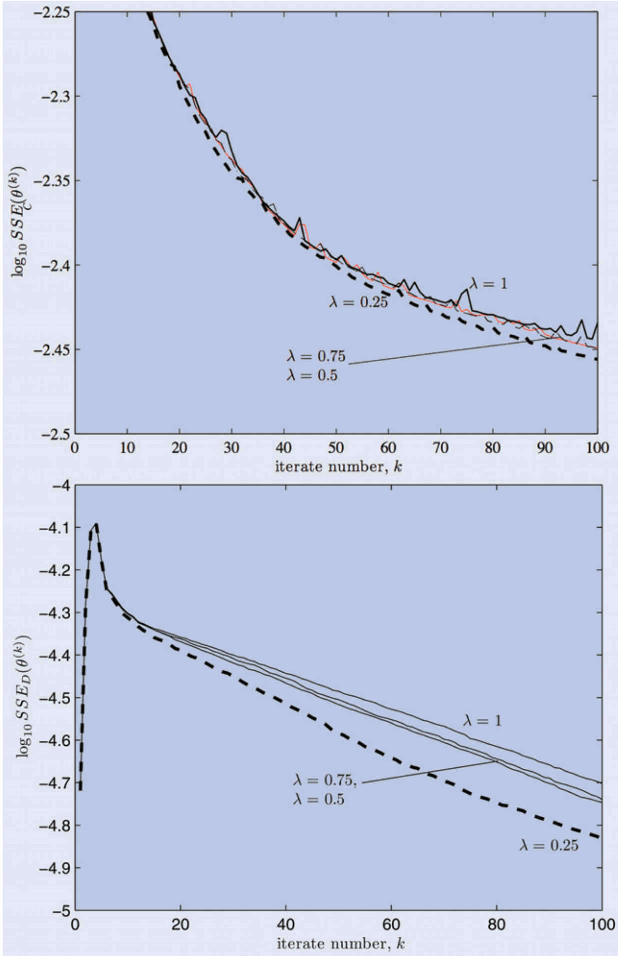

Figure 2. Convergence properties for the joint calibration algorithm for the MC family. SSE_C and SSE_D contributions to the total loss function. We present the mean of the path iterates generated by 1000 uniformly distributed choices of the initial seed, $\theta^{(0)}$.

these authors is a degenerate case of (23) with λ fixed equal to 1.

We test the robustness of this fitting algorithm for the MC family using 1000 extractions from three independent uniform distributions as initial guesses for the parameters, $\theta^{(0)}$. As representative input data, (D^*, ζ^*), we use the sample mean along the first 75 trading dates of the second (excited) period under study. Figure 2 shows the sample mean of the 1000 paths generated by the algorithm for $SSE_C(\theta^{(k)})$ and the first contribution, $SSE_D(\theta^{(k)})$, departing from simulated $\theta^{(0)}$. After the initial movements in the wrong direction, the first contribution corrects its behavior for finding its own minimum. Moreover, the second contribution exhibits a correct minimization pattern. Note the slightly better results on both sides with smaller λ. Similar results can be obtained with the ANS family and other market scenarios.

4. Empirical results

We compare three different estimations of the initial yield curve based on the Nelson–Siegel family, MC and ANS.

Table 1. Discrete data for initial yield-curve estimation.

Maturity, x	0.25	1	2	3	4	
Discount factor, $D^*(x)$	0.9886	0.9538	0.9069	0.8602	0.8142	
Maturity, x	5	6	7	8	9	10
Discount factor, $D^*(x)$	0.7693	0.7260	0.6843	0.6445	0.6066	0.5706

Our first objective is to test the stability of the implicit estimation of the model parameters θ. We consider the mean, standard deviation and coefficient of variation of parameter estimates time series. In this context the main goal is to analyse the impact that an alternative interpolation scheme has on the fitting capabilities of the model. To this end, we use as measure the mean of the daily sum of squared errors of derivatives log prices, hereafter MSE_C. The same measure is used for the zero-coupon bond prices (we denote it by MSE_D) and it is included in the analysis with the market data.

The US data set consists of 150 daily observations divided into two periods: the first period covers from 1/1/2001 to 13/4/2001 (75 trading dates) and the second starts in 15/3/2002 and finishes on 27/6/2002 (75 trading dates). The Euro denominated set used for the analysis consists of 100 daily observations from 15/2/2001 to 4/7/2001. We point out that this Euro zone database is the same as that used by Angelini and Herzel (2002, 2005). Like these authors, we divide the sample into two subperiods, Period 1 and Period 2. Period 1 runs from the beginning to 19/4/2001 (46 observations) and it is characterized by a humped implied volatility term structure. Period 2 goes from 20/4/2001 to the end (54 observations) and presents a decreasing implied volatility. The data set is composed of US and Euro discount factors for 13 maturities (3, 6 and 9 months and from 1 to 10 years) and of implied volatilities of at-the-money interest rate caps with maturities 1, 2, 3, 4, 5, 7 and 10 years. The data basis is provided by Datastream Financial Service. The simulated data was obtained from 360 extractions from the model of bond and cap prices under identical contractual features.

4.1. Simulations

We simulate, departing from alternative initial conditions $r^*(x)$, the forward curve until the time t attainable by this model. We accomplished this by working out expression (15), and writing the explicit formula for the stochastic and the deterministic coefficients, which are actually variable in time: the aforementioned $g_i(t)$ and Z_t^i and the additional value coming from the initial curve translation, $r^*(x+t)$. Now, it is possible to compute the prices of a set of zero-coupon bonds using exact integration of $r_t(x)$ over cross-sectional variable x at a fixed time t and then the prices of the seven caps with formula (9).

The fixed model parameters $\theta = (0.002,\ 0.007,\ 0.35)$ were chosen. This particular choice has a similar order of magnitude as the empirical estimations for this model reported by Angelini and Herzel (2005). As alternative initial curves, we choose MC, ANS and NS fitted to the zero-coupon bond prices shown in table 1.

Table 2. Summary statistics for calibration results with simulated data.

		MC	ANS	NS
E1:	$\epsilon_r(\alpha)$	0	0	0.23
$r_0(x) = r_m^*(x)$	$\epsilon_r(\beta)$	0	0	0.13
	$\epsilon_r(a)$	0	0	8.7×10^{-2}
	$C_v(\alpha)$	0	0	0.18
	$C_v(\beta)$	0	0	0.14
	$C_v(a)$	0	0	9.7×10^{-2}
	MSE	0	0	1.9×10^{-3}
E2:	$\epsilon_r(\alpha)$	0.25	0	0.28
$r_0(x) = r_{\mathrm{ANS}}^*(x)$	$\epsilon_r(\beta)$	0.16	0	0.16
	$\epsilon_r(a)$	0.12	0	9.5×10^{-2}
	$C_v(\alpha)$	3.8×10^{-2}	0	0.117
	$C_v(\beta)$	3.9×10^{-2}	0	9.1×10^{-2}
	$C_v(a)$	3.2×10^{-2}	0	4.8×10^{-2}
	MSE	2.6×10^{-4}	0	6.7×10^{-4}
E3:	$\epsilon_r(\alpha)$	0.313	2.7×10^{-4}	0.18
$r_0(x) = r_{\mathrm{NS}}^*(x)$	$\epsilon_r(\beta)$	0.20	2.10×10^{-4}	0.10
	$\epsilon_r(a)$	0.16	1.6×10^{-5}	6.7×10^{-2}
	$C_v(\alpha)$	2.3×10^{-2}	1.4×10^{-4}	0.17
	$C_v(\beta)$	2.6×10^{-2}	1.0×10^{-4}	0.111
	$C_v(a)$	2.2×10^{-2}	8.3×10^{-5}	6.3×10^{-2}
	MSE	3.8×10^{-4}	3.9×10^{-9}	3.5×10^{-4}

Sample statistics of the calibration on simulated data. Relative errors of the parameter estimates are expressed in absolute value. We set to 0 the table entries with value $< 10^3 \cdot \mathtt{eps}$ (variable $\mathtt{eps} \sim 10^{-16}$ measures the MATLAB internal accuracy).

Starting from the initial fitted curves, which may be denoted $r_m^*(x)$, $r_{\mathrm{ANS}}^*(x)$ and $r_{\mathrm{NS}}^*(x)$, and according to (8), the corresponding three different model evolutions are calibrated to MC, ANS and NS. In order to make the calibration results more comparable, Monte-Carlo simulations are built in from the identical random sequence $(Z_t^1,\ Z_t^2)$ in all three cases.

Following expression (8), it can be seen that there are two consistent families, G_m and G_{ANS}, for the first simulation E1, one, G_{ANS}, for the second simulation E2, and none for the last simulation E3. Table 2 shows the main consequences of the theory when the model is the *true* model. Note that perfect calibration occurs, although the model parameters are fixed *a priori*, when the family used to perform calibrations is consistent with all the future forward curves generated from the initial curve $r^*(x)$. This explains, for instance, the bad performance of the NS family even in experiment E3. Moreover, parameter instability and imprecision that produce an incorrect yield-curve selection can also be checked in figure 3.

4.2. Real data

The objective of this section is to compare the performance of the two different calibration approaches on two

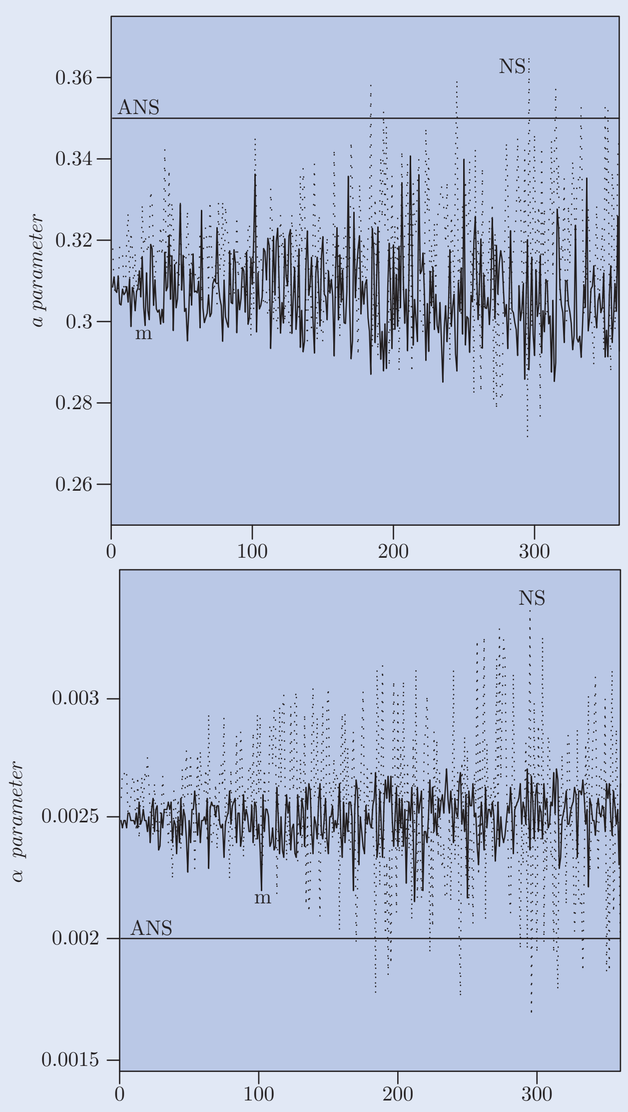

Figure 3. Daily estimates of parameters a and α for data simulated from the model with $\alpha = 0.002$ and $a = 0.35$ and starting forward curve $r_0(x) = r^*_{\text{ANS}}(x)$. The straight line corresponds to daily calibration results for the ANS family, the irregular line for the MC family and the dashed line for the NS family.

Table 3. Summary statistics for calibration results with US data for both periods.

		MC	ANS	NS
Period 1	α	0.0078	0.0079	0.0081
	β	0.0071	0.0067	0.0068
	a	0.27	0.27	0.27
	$C_v(\alpha)$	0.17	0.13	0.12
	$C_v(\beta)$	0.31	0.28	0.24
	$C_v(a)$	0.21	0.20	0.18
	MSE_C	8.2×10^{-4}	6.2×10^{-4}	5.7×10^{-4}
	MSE_D	6.4×10^{-6}	1.1×10^{-6}	1.4×10^{-5}
Period 2	α	0.0084	0.0076	0.0076
	β	0.0089	0.0117	0.0114
	a	0.30	0.39	0.37
	$C_v(\alpha)$	0.15	0.18	0.12
	$C_v(\beta)$	0.18	0.24	0.18
	$C_v(a)$	0.06	0.12	0.10
	MSE_C	0.0027	8.9×10^{-4}	2.9×10^{-4}
	MSE_D	5.8×10^{-6}	1.1×10^{-6}	2.3×10^{-5}

different periods of real trading dates. Thus, from now on we will only consider the calibration results obtained with the market data.

Concerning the US market, calibration with consistent families is carried out by setting $\lambda = 0.25$ in program (23). Table 3 shows the sample mean of the daily error fitting measures, namely MSE_C and MSE_D, and the mean and the coefficient of variation of parameter estimates. Figure 4 shows in-sample fitting time series.

The two consistent families under study report better in-sample fitting results when dealing with bond data. However, for the derivatives calibration, only the ANS family performs similar to the NS family in the two periods. This may be due to the extra factor, z_1, common for the families G_{ANS} and G_{NS}, which is independent of zero-coupon bond maturities and responsible for the better fitting observed for these families for short-term discount factors than the G_m family (note that this is not incompatible with the better summary MSE_D reported in this sample by the minimal family when compared with the Nelson–Siegel family).

Focusing on the the Euro market, we restrict ourselves to a comparison of three different estimations of the initial yield curve based on the minimal dimension family which is consistent with the model analysed in this paper. Table 4 compares the results reported by Angelini and Herzel (2002, 2005) (left column) with two of the possible extra outcomes that our extension may produce (central and right columns). Recall that the objective function of their work is a particular case of the extension presented whenever the fixed parameter λ is fixed to the value 1.

As can be seen, the results for derivatives calibration outperform those provided by the above authors in their work. For estimation of the discount function, in-sample mean statistics are marginally worse only in the second period and preserve the same order of magnitude. Thus, in both periods and for the same Euro database, we can conclude that our proposed extension clearly improves on the non-consistent methodologies that are traditionally used by practitioners.†

5. Conclusions

When calibrating a HJM model, a TSIR curve choice to fit a few market data observations is needed. In particular, it seems natural to use families of curves that do not modify their structure under the future evolution of the model, the so-called consistent families. In this work, we chose three families of curves (two consistent families and the popular Nelson–Siegel family) and we conclude that

†At this point we note that our results for the Nelson and Siegel family are omitted for brevity, but they are very close to those reported by Angelini and Herzel (2005) and are available upon request.

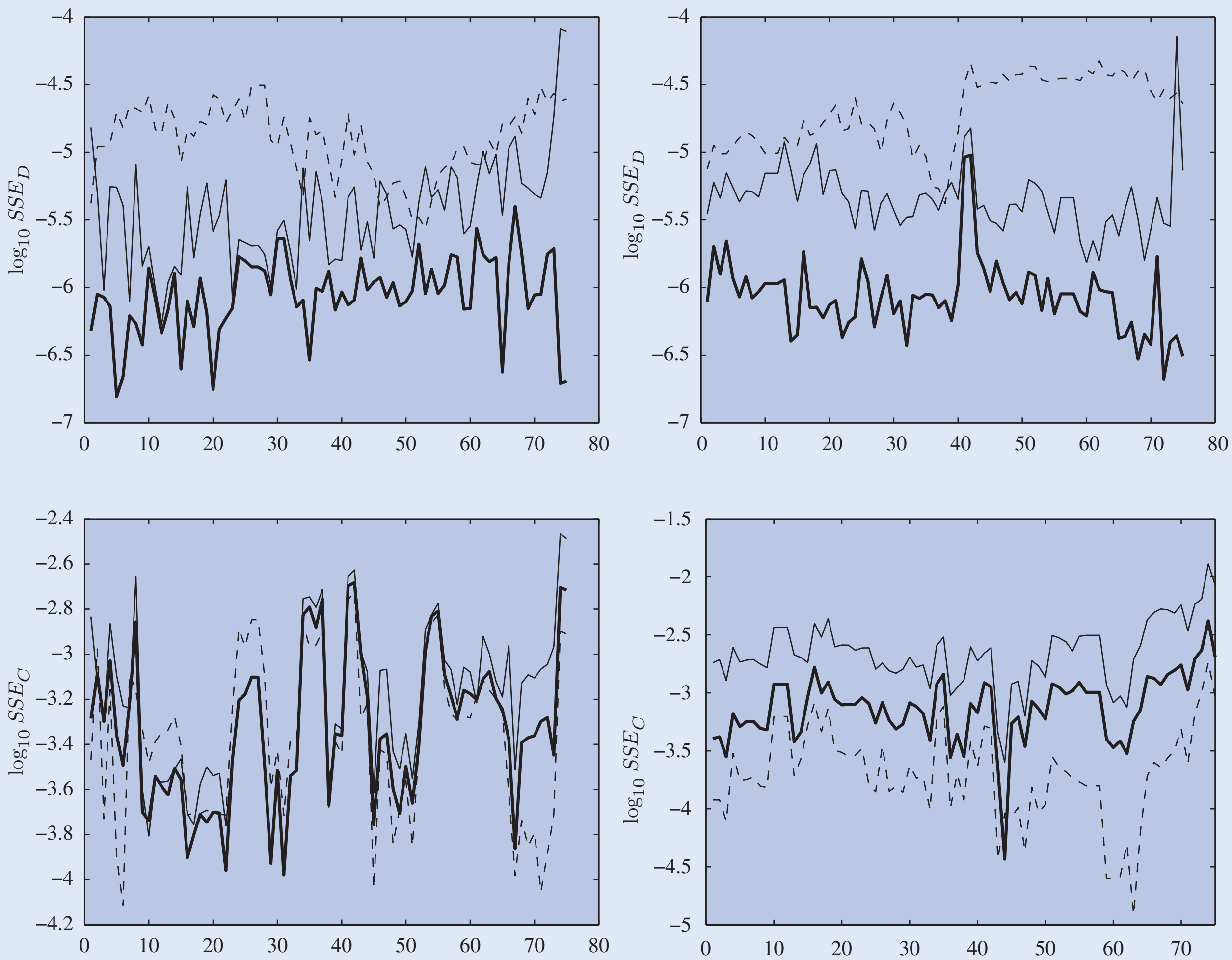

Figure 4. In-sample fitting time series for the first period (left) and the second period (right) in the US market in logarithmic terms. The thick line corresponds to the ANS family, the normal line to the MC family and the dashed line to the Nelson–Siegel family.

Table 4. In-sample mean statistics for calibration results with Euro data in both periods.

		$\lambda=1$	$\lambda=0.25$	$\lambda=0.01$
Period 1	MSE_C	2.3×10^{-4}	2.18×10^{-4}	2.19×10^{-4}
	MSE_D	8.8×10^{-7}	8.8×10^{-7}	8.4×10^{-7}
Period 2	MSE_C	3.2×10^{-4}	2.7×10^{-4}	2.7×10^{-4}
	MSE_D	6.1×10^{-7}	7.0×10^{-7}	6.8×10^{-7}

this choice has an effective impact on the quality of in-sample fitting as well as parameter estimates for both simulated and US market data.

When using simulated data it is very clear that the consistent families for experiments E1 and E2 perform much better than the non-consistent families. Moreover, the Nelson–Siegel family does not work even if it is chosen as the starting yield curve (recall experiment E3). These empirical facts constitute a nice demonstration of the theory, in the sense that even in the absence of *model risk* when only consistent families are used, perfect calibration may occur.

Translation of these consequences to real data is less clear, due to *model risk* and the quality of the data, but we can infer the following concluding remarks. In this case, the introduction of sufficiently rich *consistent families*, MC and ANS, motivated theoretically by Björk *et al.*, improves in-sample fitting capabilities on bonds. However, consistent families lead to somewhat stable parameter estimates and worse in-sample derivatives fitting results than the NS family. This may be because consistent families may exhibit undesired asymptotic features in different markets, and, in this sense, complement the empirical findings of Angelini and Herzel (2002, 2005) for different data sets like the US market data. On the other hand, note that the extension to the consistent calibration procedure presents more general features. The extension to the first consistent calibration approach is structured to allow for additional numerical outcomes. According to the results reported for the Euro database, this leads, in general, to better results also in derivatives calibration compared with non-extended consistent calibration and non-consistent methodologies.

Thus, comparative studies of the fitting of short-term zero-coupon bond capabilities and its consequences for cap pricing performance for several consistent families with a particular model and for different market bases (for instance, using different market inputs apart from US or Euro market data) should be undertaken. Moreover, we restrict our studies to a flexible one-factor Gaussian

HJM model. Future empirical research on the matter should include multi-factor models in order to more effectively capture the TSIR and TSV observed in the market. Another theoretical point regards the analytical study of the total loss function $SSE_\lambda(\theta)$ and the convergence properties of the joint calibration algorithm proposed in this work.

Acknowledgement

This work has been partially supported by grants SEJ2006-05051 and ECO2009-13616 from the Ministry of Science and Innovation of Spain, the grant PROMETEO/2008/106 from the Education Council of Generalitat Valenciana and the grant PUCH-07/08 from the CEU Cardenal Herrera University.

References

Angelini, F. and Herzel, S., Consistent initial curves for interest rate models. *J. Deriv.*, 2002, **9**(4), 8–18.

Angelini, F. and Herzel, S., Consistent calibration of HJM models to cap implied volatilities. *J. Fut. Mkts*, 2005, **25**, 1093–1120.

BIS, Zero-coupon yield curves: technical documentation. BIS Papers 25, Bank for International Settlements, Basle, 2005.

Björk, T., A geometric view of interest rate theory. In *Option Pricing, Interest Rates and Risk Management*, edited by E. Jouini, L. Cvitanic, and M. Musiela, 2001 (Cambridge University Press: Cambridge).

Björk, T., On the geometry of interest rate models. In *Paris-Princeton Lectures on Mathematical Finance 2003 (Lecture Notes in Mathematics 1847)*, edited by R.A. Carmona, E. Cinlar, I. Ekeland, E. Jouini, J.A. Scheinkman, and N. Touzi, 2003 (Springer: Berlin).

Björk, T., *Arbitrage Theory in Continuous Time*, 2nd ed., 2004 (Oxford University Press: Oxford).

Björk, T. and Christensen, B.J., Interest rate dynamics and consistent forward rate curves. *Math. Finance*, 1999, **9**, 323–348.

Björk, T. and Svensson, L., On the existence of finite dimensional realizations for nonlinear forward rate models. *Math. Finance*, 2001, **11**(2), 205–243.

Black, F. and Karasinski, P., Bond and option pricing when short rates are lognormal. *Financial Anal. J.*, 1991, **4**, 52–59.

Buraschi, A. and Corielli, F., Risk management implications of time-inconsistency: model updating and recalibration of no-arbitrage models. *J. Bank. Finance*, 2005, **29**, 2883–2907.

De Rossi, G., Kalman filtering of consistent forward rate curves: a tool to estimate and model dynamically the term structure. *J. Empir. Finance*, 2004, **11**, 277–308.

Filipovic, D., A note on the Nelson and Siegel family. *Math. Finance*, 1999, **9**(4), 349–359.

Heath, D.C., Jarrow, R. and Morton, A., Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 1992, **60**(1), 77–105.

Ho, T.S.Y. and Lee, S.B., Term structure movements and the pricing of interest rate contingent claims. *J. Financial Quant. Anal.*, 1986, **41**, 1011–1029.

Hull, J. and White, A., Pricing interest rate derivatives securities. *Rev. Financial Stud.*, 1990, **3**(4), 573–592.

Hull, J. and White, A., Single-factor interest rate models and the valuation of interest rate derivative securities. *J. Financial Quant. Anal.*, 1993, **3**(4), 573–592.

Kloeden, P.E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, 1999 (Springer: Berlin).

Mercurio, F. and Moraleda, J.M., An analytically tractable interest rate model with humped volatility. *Eur. J. Oper. Res.*, 2000, **120**, 205–214.

Musiela, M., Stochastic PDEs and term structure models. Working Paper, IGR-AFFI, La Baule, 1993.

Nelson, C.R. and Siegel, A.F., Parsimonious modelling of yield curves. *J. Business*, 1987, **60**(4), 473–489.

Ritchken, P. and Chuang, I., Interest rate option pricing with volatility humps. *Rev. Deriv. Res.*, 1999, **3**, 237–262.