# 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.

**Abstract**

In this paper we propose a calibration algorithm, by 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 by using simulated and market data. Also we discuss the convergence of the algorithm by means of Monte Carlo simulations.[1]

[1]This work has been partially supported by the grant number GV05/280 from the Generalitat Valenciana.

# 1 Introduction

Any acceptable model which prices interest rate derivatives must fit the observed term structure. This idea pioneered by Ho and Lee [16], has been explored in the past by many other researchers like Black and Karasinski [11] and Hull and White [17].

The contemporary models are more complex because they consider the evolution of the whole forward curve as an infinite system of stochastic differential equations (Heath, Jarrow and Morton [15]). 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. So, in practice, the usual approach is to interpolate the forward curve by 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 as input for model $\mathcal{M}$. Will this interest rate model evolve through forward curves that lies within the family? Motivated by this question, Björk and Christensen [8] 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 [22] and any HJM interest rate model with deterministic volatility, obtaining 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 [3]). Its forward curve shape, $G_{NS}(z, \cdot)$ is given by the expression

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

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, Filipović [14] has shown that there is no Itô process that is consistent with the Nelson-Siegel family. In a recent study De

Rossi [13] applies consistency results to propose a consistent exponential dynamic model, and estimates it using data on LIBOR and UK swap rates. On the other hand, Buraschi and Corielli [12] 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 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 with 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 on the hedging portfolios associated with these financial instruments. It must be remembered that such dynamic strategies depend on the model assumptions. Thus, re-calibration is conceivable because the 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 caused by reasons that are more theoretical, because the above mentioned set-up 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 which 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. In [1] and [2], Angelini and Herzel, propose the use of an optimization program related to the mentioned daily calibrations, which 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 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 objective function mentioned, 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 by using Monte-Carlo Simulation.

To this end, we restrict ourselves to a particular humped volatility HJM model, proposed by Mercurio and Moraleda [20] and Ritchen and Chuang [23] 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 by 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 the 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 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 –spoon-shaped. On the other hand, a 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 local minimum. To our knowledge this is the first attempt to extend the search for empirical evidence to the US–market data.

This 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 consistent families with the model. In Section 3 the calibration procedure is described. Section 4 is devoted to empirical results, first comparing the consistent calibration algorithm to the non-consistent approaches with simulated data, 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 $(\Omega, \mathcal{F}, P)$.

Single factor Heath-Jarrow-Morton [15] framework is based on the dynamics of the entire forward rate curve, $\{r_t(x),\ x > 0\}$. Thus, under Musiela's [21] parametrisation it follows that the infinite dimensional diffusion process given by

$$\begin{cases} dr_t(x) &= \beta(r_t, x)dt + \sigma(r_t, x)dW_t \\ r_0(x) &= r^*(x), \end{cases} \tag{1}$$

where $\{r^*(x),\ x \geq 0\}$, can be interpreted as the *observed* forward rate curve.

The standard drift condition derived in Heath, Jarrow and Morton [15] can easily transferred to the Musiela parametrisation (see, for instance, Musiela [21]),

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

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

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

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

## Finite Dimensional Realizations of Gaussian Models

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

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

with $\lambda_i, \alpha_i, \omega_i$ being real numbers and $p_i, q_i$ are real polynomials.

It is well-known that if $f(x)$ is a $m$-dimensional QE function, then it admits the following matrix representation

$$f(x) = ce^{Ax}B,$$

where $A$ is a $(n \times n)$-matrix, $B$ is a $(n \times m)$-matrix and $c$ is a $n$–dimensional row vector (see Lemma 2.1 in Björk [5]). Thus, $\sigma(x)$ can be written as

$$\begin{aligned} \sigma(x) &= ce^{Ax}b, \text{ where} \qquad (2) \\ c &= [\alpha \quad \beta - a\alpha], \\ A &= \begin{bmatrix} 0 & -a^2 \\ 1 & -2a \end{bmatrix}, \\ b &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \end{aligned}$$

By means of Proposition 2.1 in Björk [6] we can write the forward rate equation (1) as:

$$\begin{aligned} dq_t(x) &= \mathbf{F}q_t(x)\,dt + \sigma(x)\,dW_t, \quad q_0(x) = 0 \qquad (3) \\ r_t(x) &= q_t(x) + \delta_t(x), \qquad (4) \end{aligned}$$

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

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

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

$$\delta_t(x) = r^*(x+t) + \int_0^t \Sigma(x+t-s)\,ds,$$

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

$$\begin{aligned} dZ_t &= AZ_t\,dt + b\,dW_t, \quad Z_0 = 0, \qquad (5) \\ q_t(x) &= C(x)Z_t, \qquad (6) \end{aligned}$$

because $\sigma$ is a QE function (see, for instance, Proposition 2.3 in Björk [5]) with $A$, $b$ as in (2) and $C(x) = ce^{Ax}$. Thus, (5) is a linear SDE in the narrow sense (see Kloeden and Platen [19] for details) with explicit solution

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

where

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

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

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

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

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

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 HJM particular 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.

## 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 the 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 $i$–th theoretical derivative price maturing at time $t = T_i$, and $\zeta_i^* \equiv \zeta^*(T_i)$ is the $i$–th market price. As is well known, see

Proposition 24.15 and pages 364–366 in Björk [4], the price, at $t = 0$, of the cap is given by

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

where

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

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

$$x_j = (j+1)\tau \qquad j = 0,\ 1,\ \ldots,\ n; \qquad (11)$$

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

The variable $\vartheta$ in (10) is intimately related with 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

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

and $F(\cdot)$ satisfies

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

Although the 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 numerical instability of the calibration process, when $a \to 0$, an asymptotically equivalent expression for $\vartheta$ must be used.

The equations (9) and (10), also express the effective influence of *ab initio* yield curve estimation on cap pricing.

## 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 [8] in more detail. We adapt some of them to our Gaussian case study without further technical discussion for the general case.

**Definition 1.** *Consider the space $\mathcal{H}$ is defined as the space of all $\mathcal{C}^{\infty}$-functions,*

$$r : \mathcal{R}_{+} \rightarrow \mathcal{R}$$

*satisfying the norm condition:*

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

*where $\gamma$ is a fixed positive real number.*

As proved by Björk and Svensson [10] in Proposition 4.2, this space $\mathcal{H}$ is a Hilbert space.

**Theorem 1.** *Consider as given the mapping*

$$G : \mathcal{Z} \rightarrow \mathcal{H}$$

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

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

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

*for all $z \in Z$.*

The 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 [8] or De Rossi [13], among others. For the particular one-factor model we consider throuhgout this work, Proposition 7.2 and 7.3 in Björk and Christensen [8] may be directly applied to get the useful result:

**Proposition 1.** *The family*

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

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

Moreover, it should also be noted that *augmented* families related with (14) can be constructed by adding to $G_m$ an arbitrary function $\phi$, 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) focusing on forward rate evolution deduced at (8), and to get an insight on how the Monte-Carlo procedure is implemented, we describe it next. By the definition of $S(x)$, we have that $S'(x) = \sigma(x)$. Then it is easy to derive that deterministic term $\frac{1}{2}\left[S^2(t+x) - S^2(x)\right]$ is of the form

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

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

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

$$= e^{-ax}\left(\alpha Z_t^1 - a\alpha Z_t^2 + \beta Z_t^2\right) + xe^{-ax}\left(\beta Z_t^1 - a\beta Z_t^2\right),$$

and the forward rate evolution becomes

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

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

The following concluding remarks about the families used throughout this work should now be clear:

- The Nelson-Siegel family (henceforth NS)

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

is not consistent with the model.

- The family

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

is the lowest dimension family consistent with the model (hereafter MC).

- The family

$$G_{ANS}(z,x) = z_1 + z_2 e^{-ax} + z_3 x e^{-ax} + (z_4 + z_5 x + z_6 x^2)e^{-2ax},$$

is the simplest adjustment based on 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 $\theta$ be the vector $(\alpha, \beta, a)$ of parameter values for the model under consideration. Assume that we have time series observations of the implied volatilities, $\sigma_i^B$, 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 to cash quotes adopting *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 same reset periods that the $i$-th cap (these IRS start their cash flows at $t = x_0 + \tau$ as the corresponding cap and have no cash value at $t = 0$):

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

where $\tau$ is the length of the underlying caplets, and $x_1 = 2\tau, \ldots, x_n = T_i$. The derivation of the formula (16) can be found, for example, in Björk [4] (Proposition 20.7 on page 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 with $\zeta^*(T_i, D(x), K_i(D(x)), \sigma_i^B)$. This last expression emphasizes 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.

## 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 with subscript $k$, the logarithmic pricing error[2] 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:

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

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 to that used at the bond side we have

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

and

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

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 $\theta$ for solving non-linear program (17).

[2] Recall that, for small $\epsilon_k$, it is also the relative pricing error $\frac{D_k^* - D(z,x_k)}{D(z,x_k)}$.

## The Joint Calibration to Cap and Bond Prices

Let us now describe in detail the joint cap-bond calibration procedure which 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 we discussed before.

From expression (14), we notice the dependency of the family from 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, $\hat{z}(a)$

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

From the expression

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

we note that, for consistent families and for a fixed $a$, the 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 could be thought that the estimates $\hat{\theta}$ have to be found by solving the non-linear program

$$\begin{aligned} SSE_C =& \min_{\theta} \| \log \zeta^* \left[ D(\hat{z}(\theta)) \right] - \log \zeta \left[ D(\hat{z}(\theta), \theta, T) \right] \|^2 = \\ =& \min_{\theta} \sum_{i=1}^{N} \varepsilon_i^2(\theta). \end{aligned} \tag{20}$$

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 this program (20). In other words, there exist reasonable doubts about the convergence of this algorithm because both error measures compete in general. Now, we consider the following decomposition for the total loss function $SSE(\theta)$

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

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

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

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

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

At this point, note that the program used by Angelini and Herzel [1, 2] in their works uses a different goal attainment

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

where $SSE_C(\theta)$, and $\hat{z}(a)$ are defined via the identities (20) and (19). As a consequence, the program used by these authors is a degenerate case of (23) with $\lambda$ fixed equal to 1.

We test the robustness of this fitting algorithm for the MC family by using 1000 extractions from three independent uniform distributions as initial guesses for the parameters, $\theta^{(0)}$. As representative input data, $(D^*, \zeta^*)$, 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 first contribution, $SSE_D(\theta^{(k)})$, departing from simulated $\theta^{(0)}$. After the initial movements in the wrong direction, first contribution corrects its behaviour for finding its own minimum. Moreover, the second contribution exhibits a correct minimization pattern. Note the slightly better results on both sides with smaller $\lambda$. Similar results can be obtained with the ANS family and other market scenarios.

# 4 Empirical Results

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

Our first objective is to test the stability of the implicit estimation of the model parameters $\theta$. We consider mean, standard deviation and coefficient of variation of parameter estimates time series. In this context the main goal is to analyze the impact that an alternative interpolation scheme has on the fitting capabilities of the model. To this end, we use as a 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 with $MSE_D$) and it is included in the analysis with the market data.

The US data set consists of 150 daily observations divided into two periods: first period covers from 1/1/2001 to 13/4/2001 (75 trading dates) and the second one 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 used in Angelini and Herzel [1, 2]. 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 are composed of US and Euro discount factors for thirteen 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 1,2,3,4,5,7,10 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.

## 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 the expression (15), and writing the explicit formula for the stochastic as well as the deterministic coefficients which are actually variable in time evolution: the aforementioned $g_i(t)$, $Z_t^i$ and the additional ones coming from 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, $\theta = (0.002, 0.007, 0.35)$, have been chosen. This particular choice has a similar order of magnitude as the empirical estimations for this model reported by Angelini and Herzel [2]. As alternative initial curves, we choose MC, ANS and NS fitted to the zero coupon bond prices shown in Figure 3.

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

Following the expression (8), it is easy to observe that there are two consistent families, $G_m$ and $G_{ANS}$, for the first simulation E1, just one, $G_{ANS}$, for the second simulation E2, and none for the last simulation E3. Figure 4 shows main consequences of the theory when the model is the *true* model. Notice that perfect calibration only occurs, although model parameters are fixed *a priori*, when the family used to perform calibrations is consistent with all the future forward curves generated from initial curve $r^*(x)$. This fact explains, for instance, bad performance by the NS family even on E3 experiment. Moreover, parameter instability and imprecision that produces an incorrect yield-curve selection can also be checked in Figure 5.

## Real Data

The objective of this section is to compare the performance of the two different calibration approaches on two 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). The table on Figure 6 exhibits 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. On the other hand, Figure 7 shows in-sample fitting time series.

The two consistent families under study report better in-sample fitting results when dealing with bond data. However, on the derivatives side calibration, only the ANS family performs similar to the NS one in the two periods. This fact may be motivated by 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 these families better fitting observed short-term discount factors than $G_m$ family (note that this is not incompatible with the better summary $MSE_D$ reported in this sample by the minimal family when compared against the Nelson-Siegel ones).

Focusing on the the Euro market, we restrict ourselves to the comparison of three different estimations of the initial yield curve based on the minimal dimension family which is consistent with the model analysed in the paper. The table in Figure 8 compares the results reported by Angelini and Herzel [1, 2] (left column) with two of the possible extra outcomes that our extension may produce (central and right column). Recall that the objective function of their works is a particular case of the extension presented along the paper, whenever the fixed parameter, $\lambda$, is fixed to the value 1.

As can be seen, we stress that the results for derivatives calibration outperform those provided by the above authors in their works. As for the estimation of the discount function, in-sample mean statistics are marginally worse only in the second period and preserving the same order of magnitude. Thus, in both periods and for the same Euro database, we can conclude that our proposed extension improves clearly non consistent methodologies that are traditionally carried out by the practitioners[3].

# 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 which do not modify their structure under the future evolution of the model, the so-called consistent families.

In this work, we choose three families of curves (two consistent families and the popular Nelson-Siegel family) and we conclude that this choice has an effective impact on the quality of in-sample fitting as well as parameter estimates on both simulated and US-market data.

When using simulated data it is very clear that the consistent families for the E1 ans E2 experiments perform much better than the non-consistent ones. Moreover, Nelson-Siegel family does not work even if it is chosen as the starting yield-curve (recall E3 experiment). These empirical facts constitute a nice demonstration of the theory, in the sense that even in

[3]At this point, we must to note the reader that our results for the Nelson and Siegel family are omitted for shortness, but they are very close to the reported in [2] and available under request

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 quality of data, and we can infer the following concluding remarks. In this case, the introduction of sufficiently rich *consistent families*, MC as well as ANS, well 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 could be an insight that consistent families may exhibit undesired asymptotic features in different markets, and in this sense, complements empirical findings of Angelini and Herzel in [1, 2] to 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 more numerical outcomes. According to the results reported for the Euro database, this leads in general to better results also in derivatives calibration as compared to non-extended consistent calibration and non-consistent methodologies.

Thus, comparative studies between the fitting of short-term zero-coupon bond capabilities and its consequences on cap pricing performance for several consistent families with a particular model and on different market bases (for instance, using different market inputs apart from US or Euro market data) should be analysed deeply in the future. 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.

# References


[1] Angelini, F., & S. Herzel (2002). Consistent Initial Curves for Interest Rate Models, *Journal of Derivatives* 9(4), 8–18.

[2] Angelini, F., & S. Herzel (2005). Consistent Calibration of HJM Models to Cap Implied Volatilities, *Journal of Futures Markets* 25, 1093-1120.

[3] BIS (2005). Zero-coupon yield curves: Technical documentation, *BIS Papers* 25. Bank for International Settlements, Basle.

[4] Björk, T. (2004). *Arbitrage Theory in Continuous Time*, Second Edition, Oxford.

[5] Björk, T. (2003). On the Geometry of Interest Rate Models, en: Carmona, R.A., Cinlar, E., Ekeland, I., Jouini, E., Scheinkman, J.A., Touzi, N., *Paris-Princeton Lectures on Mathematical Finance 2003*, Lecture Notes in Mathematics, 1847, Springer.

[6] Björk, T. (2001). A Geometric View of Interest Rate Theory, en: Jouini, E., Cvitanic, J. & M. Musiela, *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press.

[7] Björk, T., & A. Gombani (1999). Minimal Realizations of interest rate models. *Finance and Stochastics* 3(4), 413-432.

[8] Björk, T., & B.J. Christensen (1999). Interest Rate Dynamics and Consistent Forward Rate Curves, *Mathematical Finance* 9, 323-348.

[9] Björk, T., & C. Landen (2002). On the construction of finite dimensional realizations for nonlinear forward rate models, *Finance and Stochastics* 6(3), 303-331.

[10] Björk, T., & L. Svensson (2001). On the existence of finite dimensional realizations for nonlinear forward rate models, *Mathematical Finance* 11(2), 205-243.

[11] Black, F., & P. Karasinski (1991). Bond and Option Pricing when Short Rates are Lognormal, *Financial Analysts Journal* 4, 52-59. 127–155.

[12] Buraschi, A., & F. Corielli (2005). Risk Management Implications of Time-Inconsistency: Model Updating and Recalibration of No-Arbitrage Models. *Journal of Banking & Finance* 29, 2883-2907.

[13] De Rossi, G., (2004). Kalman Filtering of Consistent Forward Rate Curves: a Tool to Estimate and Model Dynamically the Term Structure, *Journal of Empirical Finance* 11,277–308.

[14] Filipovic D., (1999). A Note on the Nelson and Siegel Family. *Mathematical Finance* 9(4), 349–359.

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

[16] Ho, T.S.Y, & S.B. Lee (1986). Term Structure Movements and the Pricing of Interest Rate Contingent Claims. *Journal of Financial and Quantitative Analysis* 411011-1029.

[17] Hull, J., & A. White (1990). Pricing Interest Rate Derivatives Securities. *The Review of Financial Studies* 3(4), 573–592.

[18] Hull, J., & A. White (1993). Single-factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis* 3(4), 573–592.

[19] Kloeden, P. E., & E. Platen (1999). *Numerical solution of stochastic differential equations*, Springer.

[20] Mercurio, F., & J.M. Moraleda (2000). An analytically tractable interest rate model with humped volatility. *European Journal of Operational Research* 120, 205–214.

[21] Musiela, M., (1993). Stochastic PDEs and term structure models. Working Paper, J. Intern. Finance, IGR-AFFI, La Baule.

[22] Nelson, C. R., and A. F. Siegel (1987). Parsimonious Modelling of Yield Curves. *Journal of Business* 60(4), 473–489.

[23] Ritchken, P., & I. Chuang (1999). Interest rate option pricing with volatility humps. *Review of Derivatives Research* 3, 237–262.

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

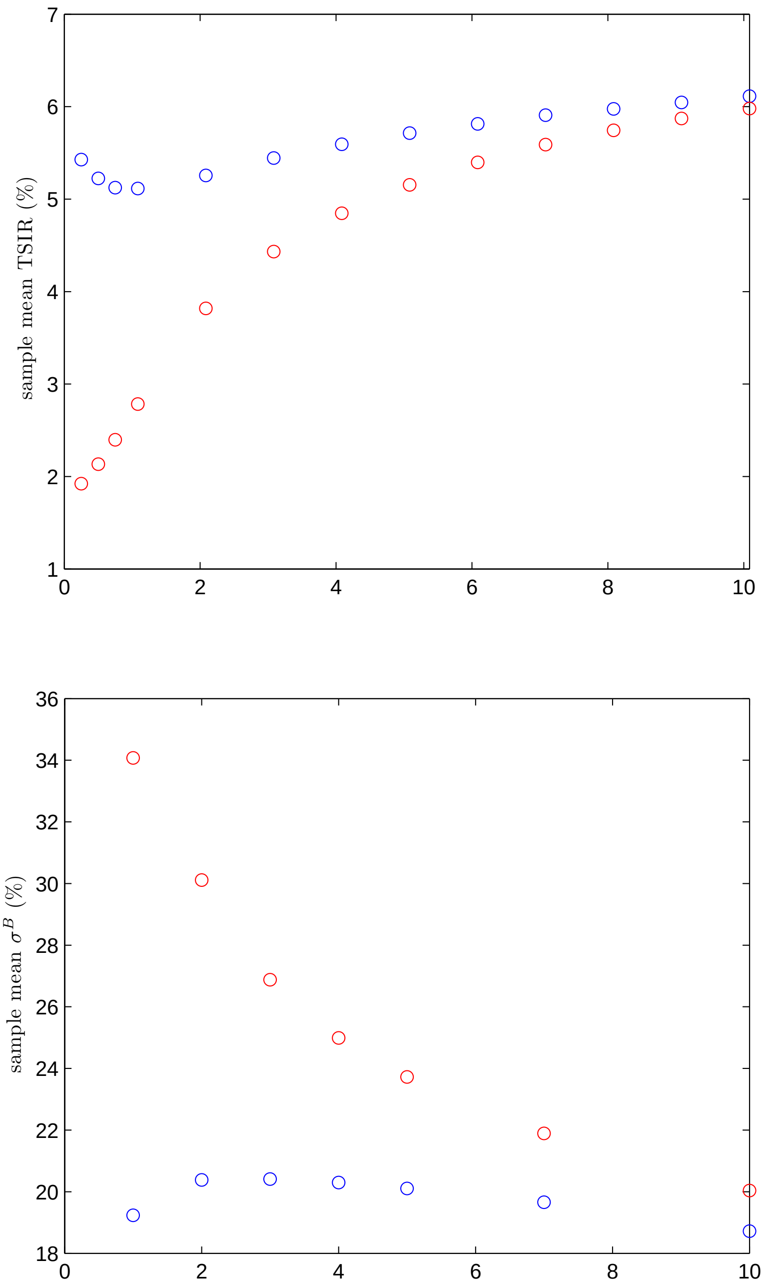
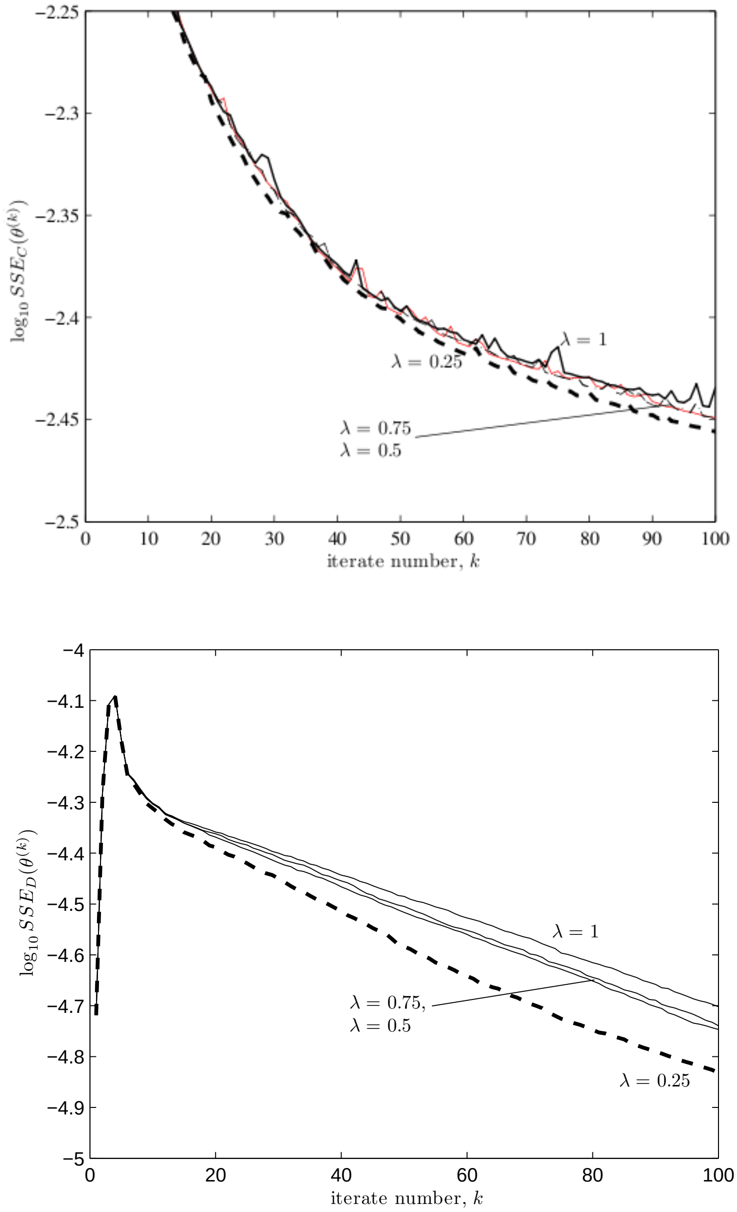


Figure 2: Convergence properties for the joint calibration algorithm with MC family. $SSE_C$ contribution to total loss function on top $SSE_D$ contribution, from zero-coupon bond pricing errors on the bottom.




We represent the mean of the path iterates generated by the 1000 uniform distributed choices of the initial seed, $\theta^{(0)}$. 1

Figure 3: Discrete data for initial yield-curve estimation.

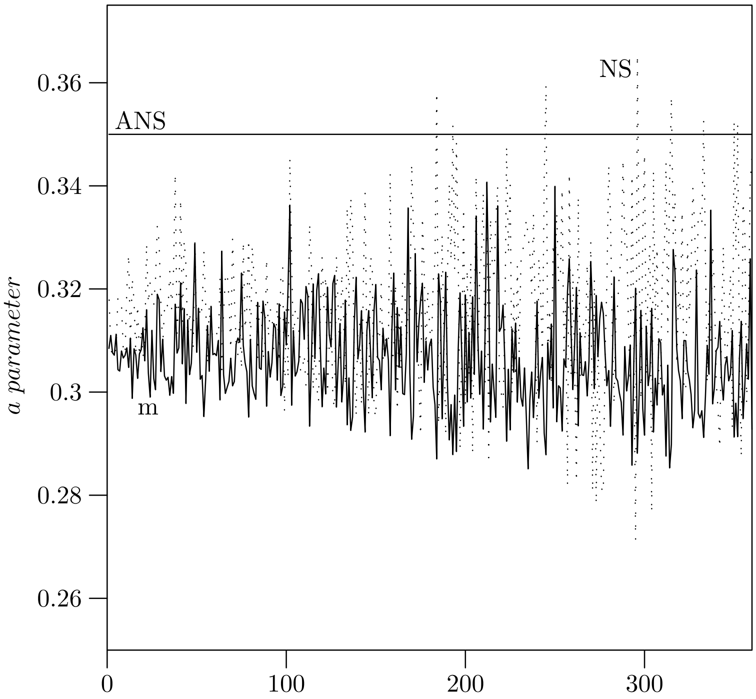
| | | | | | | |
|---|---|---|---|---|---|---|
| Maturity, $x$ | 0.25 | 1 | 2 | 3 | 4 | |
| Discount Factor, $D^*(x)$ | 0.9886 | 0.9538 | 0.9069 | 0.8602 | 0.8142 | |
| Maturity, $x$ | 5 | 6 | 7 | 8 | 9 | 10 |
| Discount Factor, $D^*(x)$ | 0.7693 | 0.7260 | 0.6843 | 0.6445 | 0.6066 | 0.5706 |

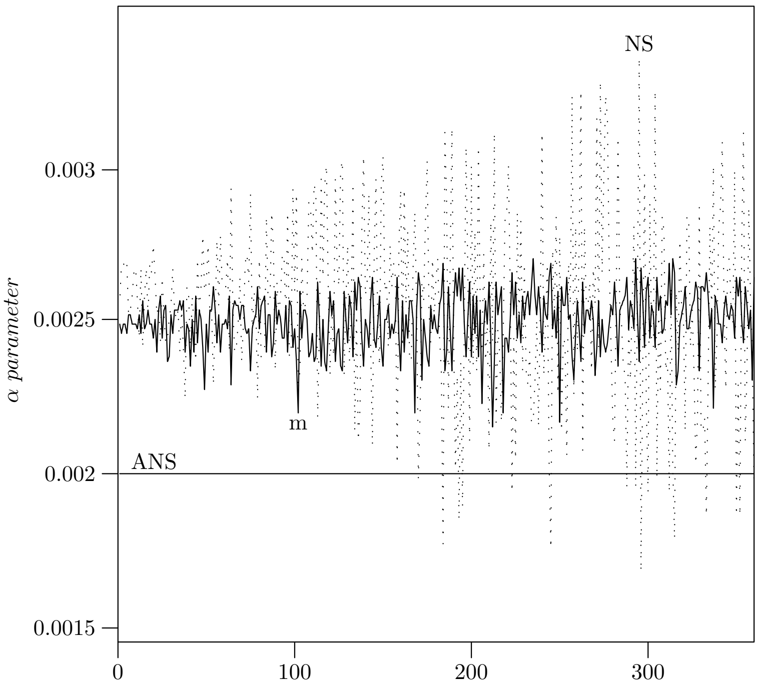
Figure 4: Summary statistics for calibration results with simulated data.

| | | MC | ANS | NS |
|---|---|---|---|---|
| | $\varepsilon_r(\alpha)$ | 0 | 0 | 0.23 |
| | $\varepsilon_r(\beta)$ | 0 | 0 | 0.13 |
| | $\varepsilon_r(a)$ | 0 | 0 | $8.7\ 10^{-2}$ |
| E1: | $C_v(\alpha)$ | 0 | 0 | 0.18 |
| $r_0(x) = r^*_m(x)$ | $C_v(\beta)$ | 0 | 0 | 0.14 |
| | $C_v(a)$ | 0 | 0 | $9.7\ 10^{-2}$ |
| | MSE | 0 | 0 | $1.9\ 10^{-3}$ |
| | $\varepsilon_r(\alpha)$ | 0.25 | 0 | 0.28 |
| | $\varepsilon_r(\beta)$ | 0.16 | 0 | 0.16 |
| | $\varepsilon_r(a)$ | 0.12 | 0 | $9.5\ 10^{-2}$ |
| E2: | $C_v(\alpha)$ | $3.8\ 10^{-2}$ | 0 | 0.117 |
| $r_0(x) = r^*_{ANS}(x)$ | $C_v(\beta)$ | $3.9\ 10^{-2}$ | 0 | $9.1\ 10^{-2}$ |
| | $C_v(a)$ | $3.2\ 10^{-2}$ | 0 | $4.8\ 10^{-2}$ |
| | MSE | $2.6\ 10^{-4}$ | 0 | $6.7\ 10^{-4}$ |
| | $\varepsilon_r(\alpha)$ | 0.313 | $2.7\ 10^{-4}$ | 0.18 |
| | $\varepsilon_r(\beta)$ | 0.20 | $2.10\ 10^{-4}$ | 0.10 |
| | $\varepsilon_r(a)$ | 0.16 | $1.6\ 10^{-5}$ | $6.7\ 10^{-2}$ |
| E3: | $C_v(\alpha)$ | $2.3\ 10^{-2}$ | $1.4\ 10^{-4}$ | 0.17 |
| $r_0(x) = r^*_{NS}(x)$ | $C_v(\beta)$ | $2.6\ 10^{-2}$ | $1.0\ 10^{-4}$ | 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 parameters estimates are expressed in absolute value. We set to 0 table entries with value $< 10^3 \cdot$`eps` (variable `eps` $\sim 10^{-16}$ measures MATLAB internal accuracy).

Figure 5: Daily estimates of parameters $a$ and $\alpha$ for data simulated from the model with $\alpha = 0.002$ and $a = 0.35$ and starting forward curve $r_0(x) = r^*_{ANS}(x)$.



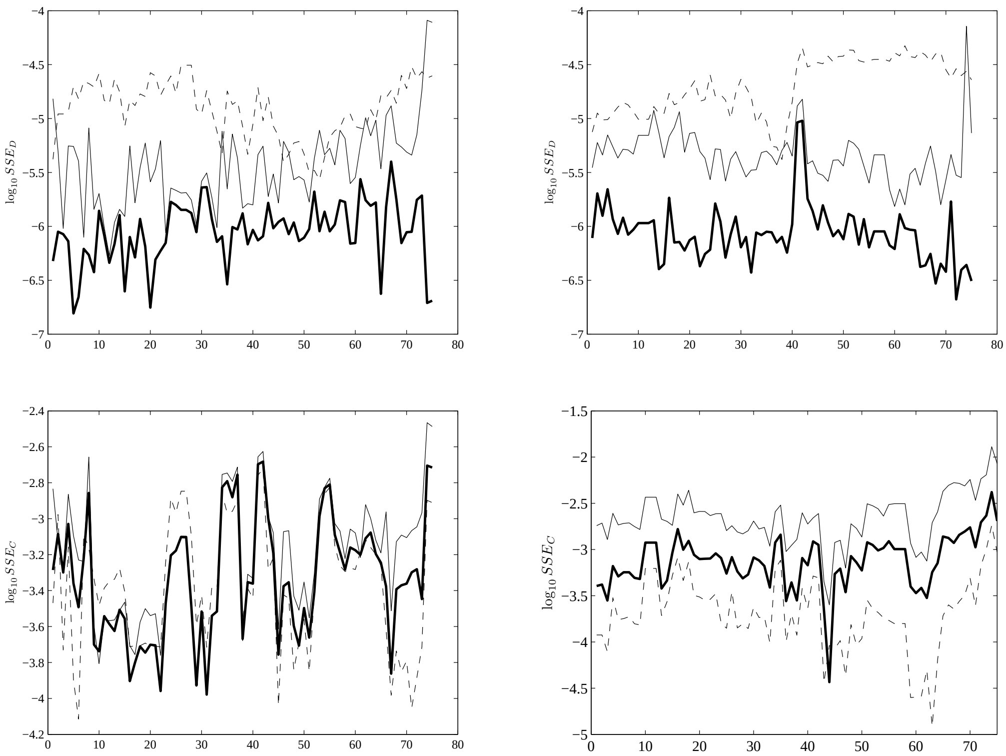




The straight line corresponds to daily calibration results belonging ANS family, normal but irregular line to MC family and the dashed one to the NS family.

Figure 6: Summary statistics for calibration results with US data on both periods.

| | | MC | ANS | NS |
|---|---|---|---|---|
| Period 1 | $\alpha$ | 0.0078 | 0.0079 | 0.0081 |
| | $\beta$ | 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 | $\alpha$ | 0.0084 | 0.0076 | 0.0076 |
| | $\beta$ | 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 | $8.9\ 10^{-4}$ | $2.9\ 10^{-4}$ |
| | $MSE_D$ | $5.8\ 10^{-6}$ | $1.1\ 10^{-6}$ | $2.3\ 10^{-5}$ |

Figure 7: In-sample fitting time series for the first period (left) and the second period (right) in the US market in logarithmic terms.




Thick line corresponds to ANS family, normal line to MC family and the dashed one to the Nelson-Siegel family.

Figure 8: In-sample mean statistics for calibration results with Euro data on 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 | $MSE_C$ | $3.2\ 10^{-4}$ | $2.7\ 10^{-4}$ | $2.7\ 10^{-4}$ |
| | $MSE_D$ | $6.1\ 10^{-7}$ | $7.0\ 10^{-7}$ | $6.8\ 10^{-7}$ |