The Hull-White Model and Multiobjective Calibration with Consistent Curves: Empirical Evidence

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

We present a new methodology for the calibration of the Hull-White model to US market prices with consistent curves. It falls into the general class of nonlinear multicriteria optimization problems and we show how this algorithm is able to build a set of dicrete Pareto points of the implied trade-off curve. We also evaluate its fitting capabilities against non-consistent traditional methods with very promising results.
Keywords: Consistent Forward Rate Curves, Multiobjective Calibration, Interest Rate Models.
AMS 2000 Subject Classification: 91B28, 62P05.

1 Introduction

Any acceptable model which prices interest rate derivatives must fit the observed term structure. This idea pioneered by Ho and Lee [19], has been explored in the past by many other researchers like Black and Karasinski [12] and Hull and White [20].

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 [18]). In particular, they use as initial input, a continuous forward rate curve. In reality, we just observe a discrete set composed either by 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 lay within as input for model $\mathcal{M}$. Will this interest rate model evolve through forward curves that lay within the family? Motivated by this question, Björk and Christensen [8] define the so–called consistent pairs ($\mathcal{M}$, $\mathcal{G}$) as ones whose answer to the above question is positive. In particular, they studied the problem of consistency the family of curves proposed by Nelson and Siegel [23] 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

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

Despite all the positive empirical features and general acceptance by the financial community, Filipović [17] has shown that there is no Itô process that is consistent with the Nelson-Siegel family. In a recent study De Rossi [15] 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 [13] add results to 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, the practitioners usually reestimate 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), then
- They obtain an estimation 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)

We remark that this pure cross-sectional procedure itself ought to take into account that it will be repeated. If $\mathcal{M}$ and $\mathcal{G}$ are inconsistent, then the interest rate model will produce forward curves outside the family used in the calibration step, and this will force the analyst to change model parameters all time, and not because the model is not the *truth model*, but simply because the family does not go well with the model. Put into applied terms, if we want to maintain recalibration as a coherent practice for incorporating newly arrived information from the market, then our family $\mathcal{G}$ of forward rate curves should be chosen to be consistent with the model $\mathcal{M}$.

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 [1] propose the use of a 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 with 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 have a good behaviour 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 scope, we construct the 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 rigorous approach is richer in possible outcomes.

To this end, we restrict ourselves to the one-factor extended Vasicek [24] model originally introduced by Hull and White [20], calibrated on a US data set consisting of US term structures of interest rates (TSIR, from now) and cap quotes between 12/09/2001 and 23/08/2002 (see Figure 1).

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 US–market data. In the last section we give some concluding remarks.

2 The Model

Let W be a one dimensional Wiener stochastic process defined in a complete probability space $(\Omega, \mathcal{F}, P)$.

Single factor Heath-Jarrow-Morton [18] framework is based on the dynamics of the entire forward rate curve, $\{r_t(x),\ x > 0\}$. Thus, under Musiela's [22] parameterization it follows that the infinite dimensional diffusion process given by

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

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 [18] can easily transferred to the Musiela parametrization (see, for instance, Musiela [22]),

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

Thus, a particular model is constructed by the choice of an explicit volatility function $\sigma(r_t, x)$.

Recall that our work is devoted to the Hull-White model that falls into the class of HJM models with

$$\sigma(r_t, x) = \sigma(x) = \sigma e^{-ax}.$$

The Hull-White model improve Ho-Lee model incorporating mean-reversion and providing closed formulas for liquid options like interes rate caps. This model is one of the simplest Gaussian HJM models which preserves the Markov property, allowing very efficient numerical methods for the pricing of exotic options. On the negative side, it does not capture humped shapes of the term structure of volatilities (TSV hereafter). However, it exhibits a relative good performance when it is chosen as a parsimonious solution for bussiness cycles with monotonically decreasing TSV, as it is shown by [2].

2.1 Forward Representation of the Model

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

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

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

$$dq_t(x) = \mathbf{F}q_t(x)\,dt + \sigma(x)\,dW_t, \quad q_0(x) = 0 \qquad (3)$$

$$r_t(x) = q_t(x) + \delta_t(x), \qquad (4)$$

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

$$\Sigma(x) = \sigma(x) \int_0^x \sigma(s)\,ds.$$

Moreover, $q_t(x)$ has the concrete *finite dimensional* realization

$$dZ_t = -aZ_t\,dt + \sigma\,dW_t, \quad Z_0 = 0, \qquad (5)$$

$$q_t(x) = e^{-ax}Z_t, \qquad (6)$$

see, for instance, Proposition 2.3 in Björk [5]. Thus, (5) is a linear SDE in the narrow sense (see Kloeden and Platen [21] for details) with explicit solution

$$Z_t = \sigma e^{-at} \int_0^t e^{as}\,dW_s, \qquad (7)$$

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 to the forward rate dynamics

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

Equation (8) allows to perform the Monte Carlo simulation of future forward curves produced by the model and may be useful for risk managing purposes. As we show next, this last expression may be used to build initial forward rate curves $r^*(x)$ time-consistent with the model.

2.2 Consistent Curves with the Model

If we want to measure the actual impact that alternative choices 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,*

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

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 γ is a fixed positive real number.

As proved by Björk and Landen [9], this space $\mathcal{H}$ is a Hilbert space.

Theorem 1 *Consider as given the mapping*

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

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

$$\begin{aligned} \partial_x G(z,x) + \sigma(x)\int_0^x \sigma(s)ds \quad &\in \quad \mathrm{Im}\left[\partial_z G(z,x)\right], &(9)\\ \sigma(x) \quad &\in \quad \mathrm{Im}\left[\partial_z G(z,x)\right], &(10)\end{aligned}$$

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 Björk and Christensen [8] or De Rossi [15], among others.

For the particular one-factor model we consider along 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*

$$G_m(z,x) = z_1 e^{-ax} + z_2 e^{-2ax}, \tag{11}$$

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

There is a way to justify (11) focusing on forward rate evolution deduced at (8), and to get an insight on how the simulations may be implemented for risk management purposes, 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

$$g(t)e^{-ax} + h(t)e^{-2ax}.$$

Thus, the forward rate evolution becomes

$$r_t(x) = r^*(x+t) + \left(g(t) + Z_t\right) e^{-ax} + h(t)e^{-2ax}. \tag{12}$$

From (12) 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 $\{e^{-ax}, e^{-2ax}\}$. It should be also noted that the map

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

where $\phi(\cdot)$, is an arbitrary function, is also consistent with this model.

The following concluding remarks about the families used along 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

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

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

- The family

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

  is the simplest adjustment based on restricted NS family that allows model consistency (hereafter ANS).

2.3 Interest Rate Option Pricing

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

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

where $C_i(\theta)$ are the i–th theoretical derivative price and $C_i^* \equiv C^*(T_i)$ is the i–th market price one. As it is well known, see Propositions 24.12 and 24.13 on pages 361–362 in Björk [4], the price, at $t = 0$, of the cap is given by

$$C(T) = (1 + \tau K) \left(\sum_{j=0}^{n-1} \kappa D(x_j) N(-d_+) - D(x_{j+1}) N(-d_-) \right), \tag{13}$$

where

$$d_\pm = \frac{\ln \frac{D(x_j)}{\kappa D(x_{j+1})} \pm \frac{1}{2}\vartheta^2(x_j)}{\vartheta(x_j)}. \tag{14}$$

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

$$x_j = (j + 1)\tau \qquad j = 0,\ 1,\ \ldots,\ n; \tag{15}$$

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

$$\vartheta(x_j) = \frac{\sigma}{a}\left(1 - e^{-a\tau}\right)\sqrt{\frac{1 - e^{-2ax_j}}{2a}}.$$

The equations (13) and (14), also express the effective influence of *ab initio* yield curve estimation on cap pricing.

3 Calibration to Market Data Approaches

The calibration procedures can be described formally as follows. Let θ be the vector (σ, a) of parameter values for the model under consideration. Assume that we have time series observations of the implicit 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 we are also equipped with the discount function estimation, $D(x)$, at time $t = 0$.

Market participants translate volatility quotes to cash quotes adopting Black model [11]. In addition, they make the well-known convention that K_i quantities must be equal to

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

where $\tau = x_{j+1} - x_j$ is the length of the underlying caplets. The derivation of the formula (16) can be found, for example, in Björk [4] (Proposition 20.7 on pages 312–313). Now, by inspection, it is clear that this market convention makes that K_i depends on the yield-curve estimation. It allows to us to denote market prices of caps with $C^*(T_i, D(x), K_i(D(x)), \sigma_i^B)$. This last expression emphasizes explicit and implicit dependence (through ATM *strikes*) on discount function estimation even for market prices. Let $C(T_i, D(x), K_i(D(x)), \theta)$ 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 $k = 1, \ldots, M = 10$. For each zero-coupon bond denoted with subscript k, the logarithmic pricing error[1] 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 logarithmic pricing errors, l_D, as the objective loss function to minimize:

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

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

$$\eta_i(\theta) = \log C_i^* - \log C(\theta, T_i).$$

and

$$l_C = \min_\theta \| \log C^* - \log C(\theta, T) \|^2 = \min_\theta \sum_{i=1}^{N} \eta_i^2(\theta), \tag{18}$$

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

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

3.2 The Joint Calibration to Cap and Bond Prices

Let us now describe in detail the joint cap-bond calibration procedure which has 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 the Hull-White model, where the two steps are separate, as we discussed before. From the expression (11), we notice the dependency of the family from the parameter a. Let $G(z, x, a)$ be a family consistent with our gaussian model (for instance, G_m and G_{ANS}) and define least-squares estimators, $\hat{z}(a)$

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

From the expression

$$\log D(z, a, x_k) = -\int_0^{x_k} G(z, a, s)\, ds = \sum_{j=1}^{n_p} M_{kj}(a) z_j, \tag{20}$$

we note that, for consistent families and for a fixed a the problem (19) is linear in z-parameters (for the G_m family $n_p = 2$, and for the G_{ANS} family $n_p = 4$). Thus, $\hat{z}$ is an explicit and continuous function of a. Strictly speaking, joint calibration must be formalized as a multicriteria optimization problem (MOP):

$$\min_{(\sigma,a)\in S} \boldsymbol{l}(\sigma, a) = \begin{bmatrix} l_D(a) \\ l_C(\sigma, a) \end{bmatrix}, \qquad \boldsymbol{l} : \mathcal{R}^2 \to \mathcal{R}^2$$

where

$$S = \{a : h(a) = 0\}$$

with

$$h(a) = z - R(a) Q^{-1}(a) \log D^*$$

being Q, R matrices of the reduced QR decomposition of $M(a)$ which is defined by the relation (20) and *partial loss functions*, $l_i(\sigma, a)$, defined as

$$\begin{aligned} l_C(\sigma, a) &= \| \log C^* \left(D(z, \sigma, a) \right) - \log \left(D(z, \theta), \sigma, a \right) \|^2 \\ l_D(a) &= \| \log D^* - M(a) z \|^2 \end{aligned}$$

Note that it is highly probable that these objectives would both be conflicting, in general, and since no single $(\hat{\sigma}, \hat{a})$ would generally minimize every l_i simultaneously, we are dealing with Pareto optimality. A popular and acceptable method for finding a discrete set of Pareto optimal points requires to build a convex combination of the objectives into a single objective function and minimize the single objective over various values of the control parameter used to combine the objectives, see for instance [16]:

$$\min_{(\sigma,a)\in S} \boldsymbol{\lambda}^T \boldsymbol{l}(\sigma, a) = \lambda_D l_D(a) + \lambda_C l_C(\sigma, a) \tag{21}$$

with $\boldsymbol{\lambda} \in (\mathcal{R}_+ \cup \{0\})^2$ and $\lambda_C + \lambda_D = 1$.

This algorithm provides a discrete collection of Pareto optimal points representative of the entire spectrum of efficient solutions as noted in [14]. Thus ideally, consistent calibration carried out with consistent families involves the entire Pareto optimal set, in contrast to the uniqueness for the solution that appears in the two-step scalar problem.

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

$$l = \min_{(\sigma,a)\in S} l_C(\sigma, a) \tag{22}$$

where $l_C(\sigma, a)$, and $\hat{z}(a)$ are defined trough the identities (18) and (19). As a consequence, the program used by these authors is a degenerate case of (21) with λ_C fixed equal to 1, so it just allows to obtain one point of the implied trade-off curve.

4 Empirical Results

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 daily (on average) relative pricing errors, hereafter RPE_C:

$$RPE_C = \frac{1}{N} \sum_{i=1}^{N} \frac{|C_i^* - C(\hat{\sigma}, \hat{a}, T_i)|}{C_i^*}$$

The same kind of measure is used for the zero-coupon bond prices and we denote it with RPE_D:

$$RPE_D = \frac{1}{M} \sum_{k=1}^{M} \frac{|D_k^* - D(\hat{z}(\hat{a}), \hat{a}, x_k)|}{D_k^*}$$

We perform such analysis focusing on US market. The real date consists of 248 daily observations, between 12/09/2001 and 23/08/2002 . The data set is composed of US discount factors for ten maturities (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. This database is provided by Datastream Financial Service

As it have been explored before, daily joint calibration of caps and bonds with consistent families must be properly carried out as a constrained vector optimization problem. Figure 2 shows the in-sample fitting results reported by the MC family for all sample under analysis. We remark that we have divided them into two graphs just for the ease of visual inspection. The method of convex combinations was run for every date in sample with several weight vectors $\boldsymbol{\lambda}$. In doing so, we assume the same 10 logarithmically spread values

$$\lambda_C = 10^x \qquad \text{with } x = -2 + j\frac{2}{9} \qquad \text{and } j = 0, 1, 2, \ldots, 9$$

as the second vector component for all trading dates. Observe that efficient frontiers with regular shapes appear all the days nicely revealing the intrinsic multiobjective nature of the consistent calibration. Moreover, note that it can be found very different topologies for this frontiers depending on the date. Some days the objectives are conflicting and the better we fit the zero coupon bonds the worse we calibrate caps portfolio. However, another days we can achieve better results for both components of vector objective without a trade-off (there exists what is called a utopia point for the implied Pareto curve). The tables on Figure 3 show, as a numerical example, the two different situations explained before restricting to the minimal parametrized family. If we look on both tables, it must be also noted that for a fixed trading date the best cap fit results may occur with $\lambda_C \neq 1$, even if the objectives are competing. In figure 4, we analize more deeply the latter fact this time for both, MC and ANS, consistent families. We plot the second component of weight vectors, λ_C, which is responsible of the best calibration for caps on top graph. Then, we repeat the same exercise on the bottom graph, searching, in this case, for the ones which produce the best fit of the corresponding discount curve. As it is shown, most of the days the weight vector $(\lambda_D = 0, \lambda_C = 1)$ produces the best cap calibration results but there is a non-negligible number of bussiness dates where another weights produce better goals than it. On the zero curve side, in most of the cases, we find the best fit results when weights are fixed to $(\lambda_D = 0.99, \lambda_C = 0.01)$, but again, in some dates another weight choices achieve a better yield curve estimation.

For the shake of simplicity, from now on we will only consider the calibration results obtained with daily weights choices that produce the best calibration for the caps on every trading date. For instance, this rational approximation to the problem may be followed by a market participant which pursues a good risk management or pricing tool restricted to the

OTC derivatives market. On the opposite direction, note that the most desirable behaviour for a regulator (like Federal Reserve or ECB) may be consider the weights which allow the best fits for the zero rates. Following the first rational approach to Pareto point selection, in Figure 5, we compare summary statistics of the parameter estimates and the in-sample fit measures reported by NS, MC and ANS families. In addition, Figure 6 shows the comparison of in-sample fitting results in time series. The two consistent families under study report better RPE results when we restrict the analysis to cap data. For RPE on bonds, only the ANS family outperforms NS in the sample. Recall that this fact is acceptable since MC family is a family with less number of parameters than the other ones proposed. Moreover, on caps, note that the MC family appears to give better results than its consistent counterpart, ANS. Now, this behaviour can be explained because the major of dates considered, market conditions make the objective functions l_D and l_C to conflict.

5 Conclusions

When calibrating the Hull-White model, a TSIR curve choice to fit a few market data observations is needed. In particular it seems to be 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 have an effective impact on the quality of in-sample fitting for US-market data. Moreover, this paper extends the seminal calibration algorithm proposed in Angelini and Herzel [1].

In a consistent approach the parameters of parameters of model are estimated jointly with the esmation of initial discount function. Thus, from a rigorous point of view, joint calibration of caps and bonds must be viewed as a constrained vector optimization problem. Although the main objective of the algorithm is to minimize the relative differences of cap prices too, note that the vector extension of the consistent calibration presents more general features. Such extension is structured to allow more numerical outcomes and we observe that it allows to better fit results for both, caps and bonds, than the above mentioned. In particular, it is possible to find better cap calibration outcomes with $\lambda_C \neq 1$, and this is definitively different from what worked Angelini and Herzel [1] on Hull-White model, where only the fixed $\lambda_C = 1$ seems to be considered for all consistent families. The empirical findings of this paper show that, in general, consistent calibration on every date must to be carried out by analyzing the entire shape of the Pareto curve.

In this sense, this work confirms and complements the shown by Angelini and Herzel [1, 2] restricte to a Euro data set. We restrict possible outcomes on every date, by choosing the Pareto points which are responsible of better fit results on caps. Then the minimal consistent family gives the best performance in terms of caps pricing errors and becomes a good candidate for the calibration of the Hull-White model. The ANS consistent family performs very close to the Nelson-Siegel family, though it seems to be the best solution for estimating the discount function. Now, this could be explained in the context of vector optimization. We show empirically the usual competing behaviour followed by the objectives through the sample considered. Then, the minimal parameterized consistent family relax the performance on the estimation of the discount function, allowing minor relative pricing errors on caps.

Future empirical research on the matter should include multi-factor models for capturing more general TSIR and TSV observed in the market.

Another technical point regards the adaptation of the *Normal Boundary Intersection* (NBI) method to use it in the calibration problems that usually appear in the private and public financial institutions. As is mentioned by Das and Dennis in [14], NBI method surpass in flexibility as well as efficiency the popular method of minimizing weighted combinations of objective functions.

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 Future Markets* 25,1093–1120

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

[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-Verlag.

[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. The Pricing of Commodity Contracts, *Journal of Financial Economics* 3, 167–179.

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

[13] 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.

[14] Das I., & J.E. Dennis (1998). Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems, *SIAM Journal on Optimization* 8(3), 631–657.

[15] 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.

[16] Eschenauer, H., Koski J. & A. Osyczka (1990). *Multicriteria Design Optimization*, Springer.

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

[18] 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.

[19] 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* 41, 1011-1029.

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

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

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

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

[24] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5, 177-188.

Figure 1: Average of the US market TSIR and TSV with 99% confidence levels.

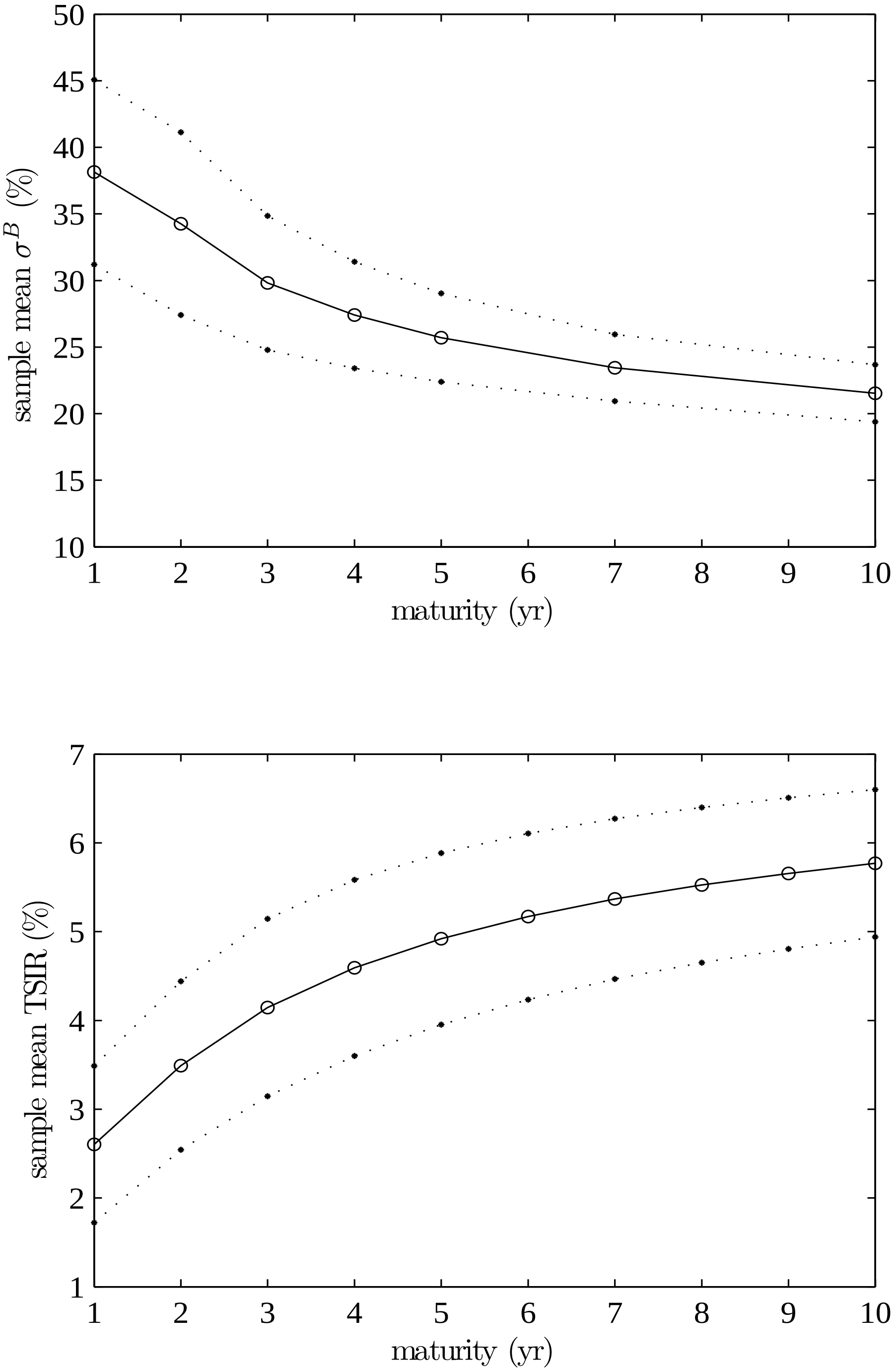
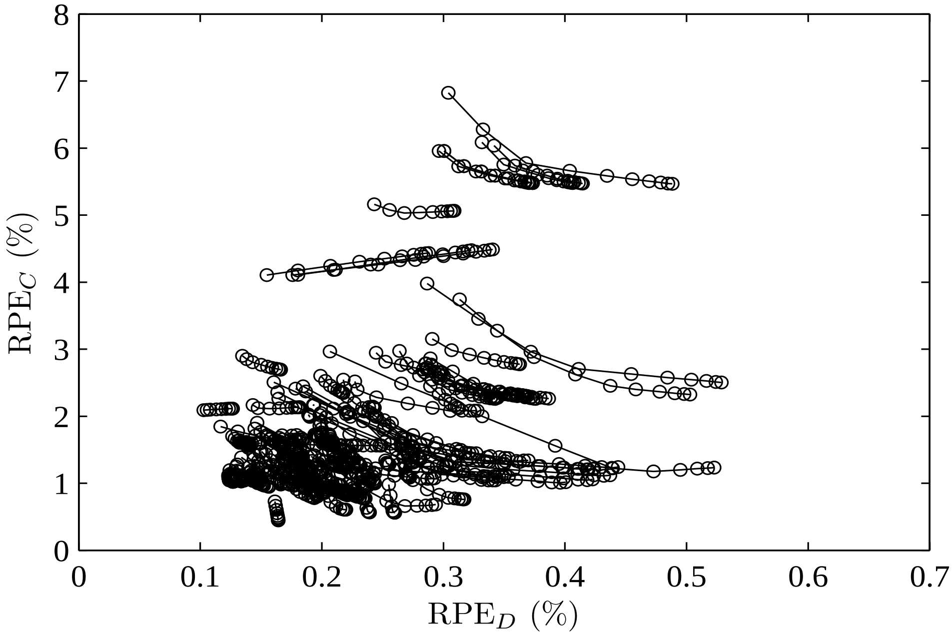


Figure 2: Daily calibration results for the minimal consistent family. The sample is divided into two periods for ease of visualization.



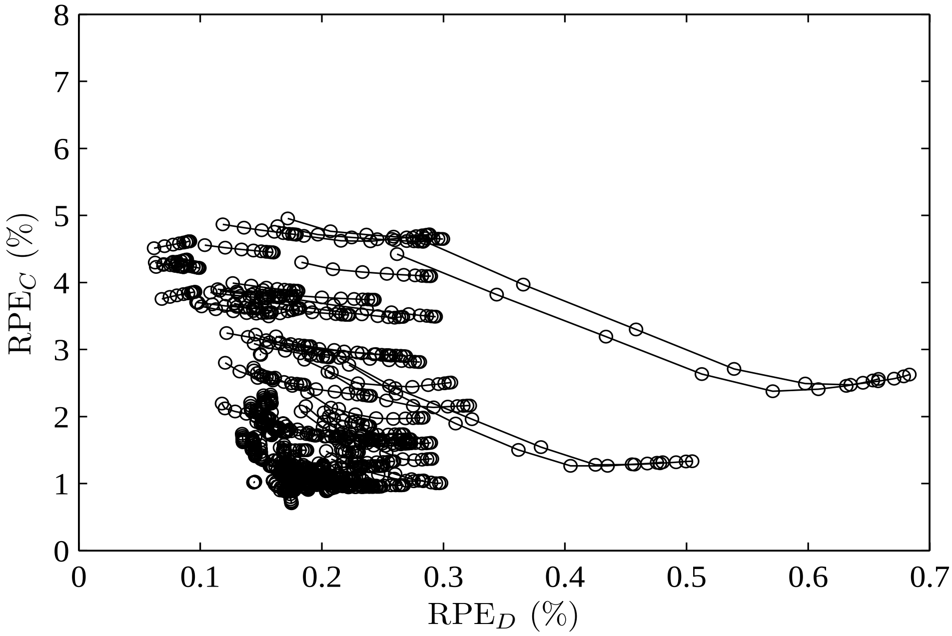


Figure 3: Efficient points in the $RPE_D - RPE_C$ space using the method of convex combinations for two different days in sample. The partial objectives, l_C and l_D are cooperative, for the Day 1 (top). In contrast, the latter ones are conflicting for the Day 2 (bottom).

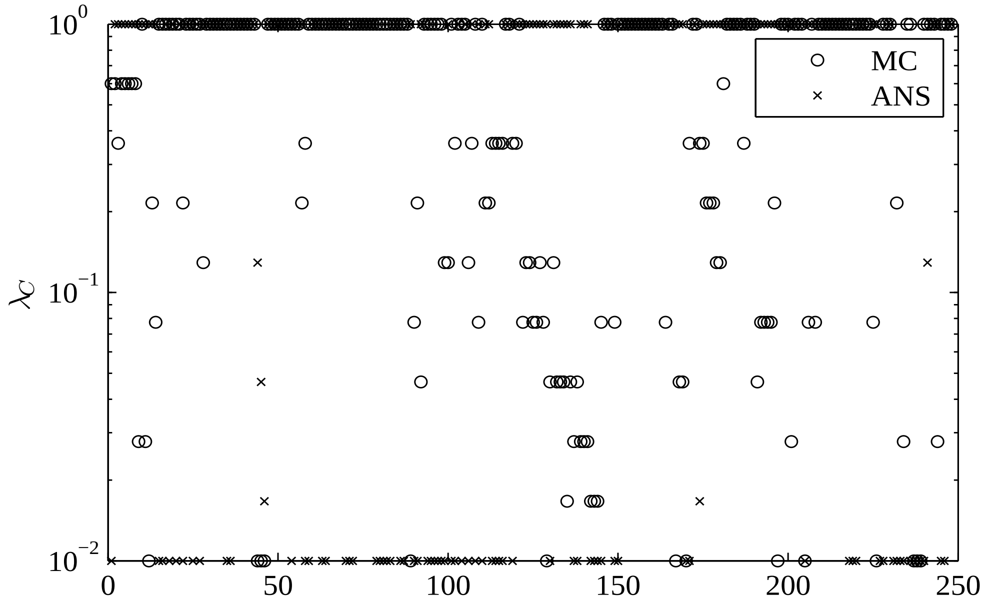
Day 1

λ_D	λ_C	RPE_D (%)	RPE_C (%)
0.99	**0.01**	**0.1695**	**0.8851**
0.98	0.02	0.1705	0.8865
0.97	0.03	0.1714	0.8880
0.95	0.05	0.1722	0.8895
0.92	0.08	0.1728	0.8906
0.87	0.13	0.1733	0.8915
0.78	0.22	0.1736	0.8921
0.64	0.36	0.1738	0.8925
0.40	0.60	0.1739	0.8928
0.00	1.00	0.1740	0.8929

Day 2

λ_D	λ_C	RPE_D (%)	RPE_C (%)
0.99	**0.01**	**0.1321**	1.6436
0.98	0.02	0.1347	1.6103
0.97	0.03	0.1372	1.5969
0.95	0.05	0.1393	1.5963
0.92	**0.08**	0.1423	**1.5962**
0.87	0.13	0.1452	1.5964
0.78	0.22	0.1472	1.5966
0.64	0.36	0.1484	1.5968
0.40	0.60	0.1492	1.5969
0.00	1.00	0.1497	1.5970

Figure 4: On top, daily weights of the multiobjective program with the best RPE_C for both consistent families. On the bottom, we choose the daily values which are responsible for the best RPE_D.



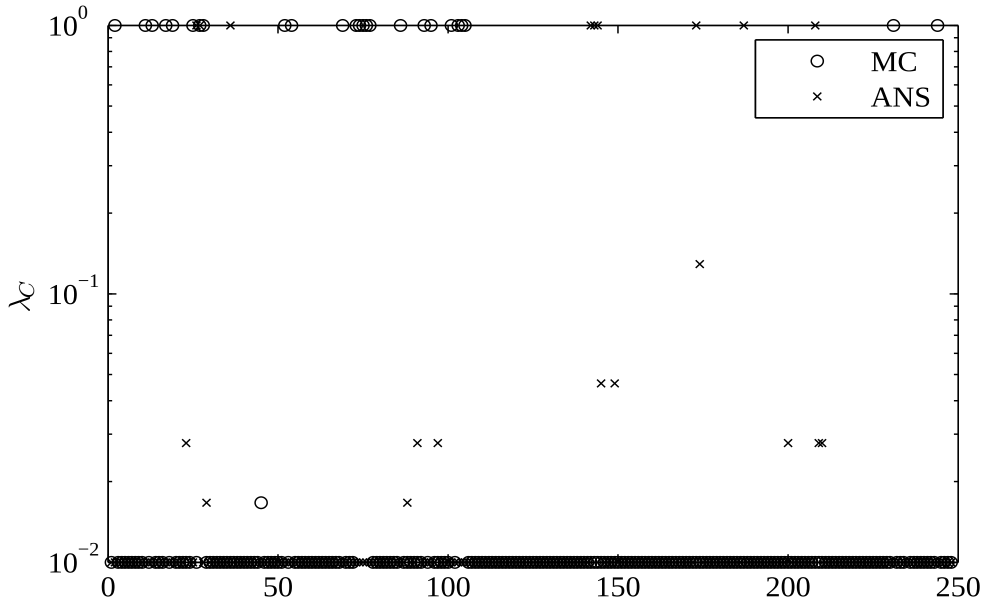


Figure 5: Summary statistics for the calibration results. In-sample descriptive statistics are carried out using the daily Pareto points with best derivative fit capabilities

SUMMARY STATISTICS

	MC	ANS	NS
σ	0.0186	0.0221	0.0218
a	0.0838	0.1911	0.1796
$C_v(\sigma)$	0.0934	0.1453	0.1406
$C_v(a)$	0.2245	0.3922	0.3821
RPE_C (%)	1.8059	2.4123	2.5997
RPE_D (%)	0.2278	0.0467	0.0567

Figure 6: Time Series Comparison.