A Multiobjective Approach using Consistent Rate Curves to the Calibration of a Heath-Jarrow-Morton Model

A. Falcó, Ll. Navarro and J. Nave

CEU Universidad Cardenal Herrera

Third Conference on Numerical Methods in Finance

Single factor HJM framework

Single factor Heath-Jarrow-Morton framework is based on the dynamics of the entire forward rate curve,

$$\{r(t,x),\ x = T - t \geq 0\}.$$

Under Musiela's parameterization

$$\begin{cases} dr(t,x) &= \beta(t,x,\boldsymbol{\theta})dt + \gamma(t,x,\boldsymbol{\theta})\,dW_t \\ r(0,x) &= G(\mathbf{z},x), \textit{observed}\text{ discount factors} \end{cases} \tag{1}$$

where

$$\beta(t,x,\boldsymbol{\theta}) = \frac{\partial}{\partial x} r(t,x) + \gamma(t,x,\boldsymbol{\theta}) \int_0^x \gamma(t,u,\boldsymbol{\theta})du$$

Motivation: Two step calibration method

1. Estimation of the initial curve $r(t = 0, x)$ from market data (e.g. bond, futures, swaps) by means a parametrized yield curve obtained by:
   $\min_{\mathbf{z}} \|(\log D_{T_i})_{T_i} - (\int_0^{T_i} G(\mathbf{z}, s)ds)_{T_i}\|^2$
2. Calibration of the parameters of the equation,

$$dr(t, x) = \beta(t, x, \boldsymbol{\theta})dt + \gamma(t, x, \boldsymbol{\theta})\, dW_t$$

   minimizing the pricing error of some interest rate derivatives (Caps, Floors, Swaption), using the initial curve computed in step 1.

The procedure is repeated every day

Consistency: Björk and Christensen (1999)

- The concept of consistency between an interest rate model $\mathcal{M}$ and a family G of forward curves: $\mathcal{M}$ and G are consistent if all forward curves produced by $\mathcal{M}$ belong to G.
- If one uses, in the calibration process of step 1, a family which is inconsistent with the model, this will produce forward curves outside the family, forcing the parameters to change.

Some initial curves for $\gamma(t, x, \sigma, a) = \sigma e^{-ax}$

- The Nelson-Siegel family (NS)

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

- The lowest dimension family consistent (MC)

$$G_{\min}(z_1, z_2, a, x) = z_1 e^{-ax} + z_2 e^{-2ax},$$

- The Adjusted NS consistent family (ANS)

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

The Hull-White Model

$$\begin{cases} dr(t,x) = \beta(t,x,a,\sigma)dt + \sigma e^{-ax}\, dW_t \\ r(0,x) = G_*(z_1,z_2,z_3,z_4,a,x), \end{cases} \tag{2}$$

- Model Parameters: $\boldsymbol{\theta} = (a, \sigma)$
- Curve Parameters: $(z_1, z_2, z_3, z_4, a) = (\mathbf{z}, a)$ where

Step 1 and Step 2 are coupled by the a-parameter

Joint Calibration with Caps and Bonds

- D_k^* be the corresponding discount factor on maturity x_k, here $k = 1, 2, \ldots, d$.
- $D(z_1, z_2, z_3, z_4, a, x) = D(\mathbf{z}, a, x)$ the discount factor obtained by using $G(\mathbf{z}, a, x)$.
- C_i^* a finite sequence of market prices of caps, here $i = 1, 2, \ldots, N$.
- $C_i(a, \sigma, D(\mathbf{z}, a, x))$ be the corresponding theoretical price under a given model $\mathcal{M}$.

A log-linear model

- Note that

$$\begin{aligned}\log D(z_1, z_2, z_3, z_4, a, x_k) &= -\int_0^{x_k} G(z_1, z_2, z_3, z_4, a, s)ds \\ &= \sum_{j=1}^4 M_{k,j}(a) z_j\end{aligned}$$

- We can write

$$\log \mathbf{D} = M(a)\mathbf{z},$$

and to obtain $r(0, x)$ we need to solve

$$\min_{(\mathbf{z},a)} \| \log \mathbf{D}^* - M(a)\mathbf{z}\|^2$$

where $\log \mathbf{D}^* = [\log D_1^* \cdots \log D_d^*]^{\mathrm{T}}$.

Two step calibration procedure

- Let

$$\widehat{\mathbf{z}}(a) \in \arg\min_{\mathbf{z}} \| \log \mathbf{D}^* - M(a)\mathbf{z}\|^2$$

- Solve

$$\min_{(a,\sigma)} \sum_{i=1}^{N} \left(\log C_i^* - \log C_i(a,\sigma,D(\widehat{\mathbf{z}}(a),a,x))\right)^2 .$$

- See, for example, Angelini and Herzel, *Journal of Derivatives* 9(4), 8–18 (2002).

A Multicriteria Optimization Problem

- Let
$$\widehat{\mathbf{z}}(a) \in \arg\min_{\mathbf{z}} \| \log \mathbf{D}^* - M(a)\mathbf{z}\|^2$$

- The joint calibration can be formalized by the following multicriteria nonlinear optimization problem:

$$\min_{(a,\sigma)} \begin{bmatrix} \| \log \mathbf{D}^* - M(a)\widehat{\mathbf{z}}(a)\|^2 = f_1(a) \\ \\ \sum_{i=1}^{N} (\log C_i^* - \log C_i(a, \sigma, D(\widehat{\mathbf{z}}(a), a, x)))^2 = f_2(a, \sigma) \end{bmatrix}$$

Solving Nonlinear MOP

- Since no single (a^*, σ^*) would generally minimize every f_i simultaneously, a concept of optimality which is useful in the multiobjective framework is that of Pareto optimality, as explained below.
- A point (a^*, σ^*) is said to be *globally Pareto optimal or a globally efficient point for (MOP)* if and only if there does not exist (a, σ) satisfying that $f_i(a, \sigma) \leq f_i(a^*, \sigma^*)$ for all $i \in \{1, 2\}$ and $f_j(a, \sigma) < f_j(a^*, \sigma^*)$ for some $j \in \{1, 2\}$.
- If the above definition holds for some open neighborhood of (a^*, σ^*) then we say that it is *a locally Pareto optimal or a locally effcient point for (MOP)*. **Pareto optimality will henceforth refer to local Pareto optimality unless qualified explicitly.**

Weighted Convex Combinations Method

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

$$\min_{(\sigma,a)} \left\{\lambda f_1(a) + (1-\lambda) f_2(a,\sigma)\right\}, \quad (3)$$

with $0 \le \lambda \le 1$.

- We perform (3) by means the Matlab Optimization Toolbox function `lsqnonlin`.
- We point out that `lsqnonlin` might only give local solutions.

Pareto optimal points

- This algorithm provides a discrete collection of Pareto optimal points representative of the entire spectrum of efficient solutions:

$$\lambda \mapsto (\widehat{a}(\lambda), \widehat{\sigma}(\lambda)) \in \arg\min_{(a,\sigma)} \{\lambda f_1(a) + (1-\lambda) f_2(a,\sigma)\}$$

- In particular, $\lambda = 0$ gives us the two step calibration procedure.
- Indraneel Das,*Nonlinear Multicriteria Optimization and Robust Optimality*. Ph. D. Thesis (Rice U.) April 1997.

Empirical Measures

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:

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

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

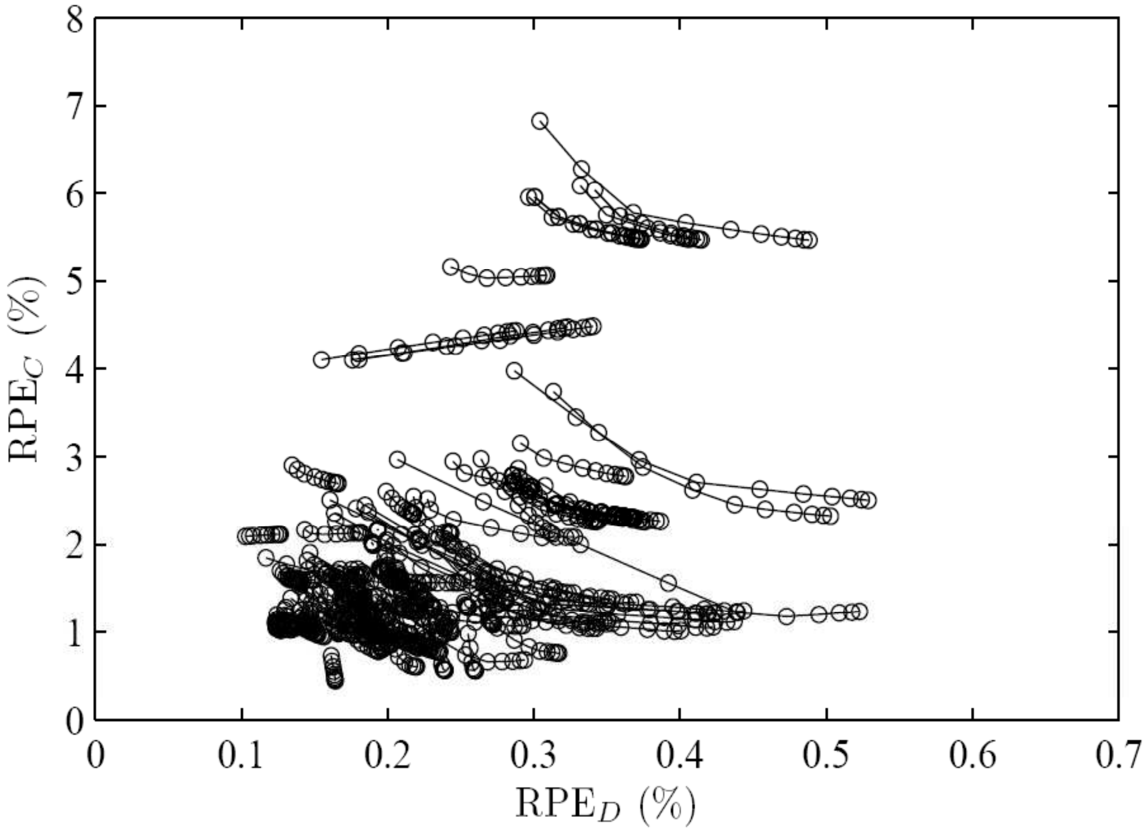
$$RPE_D(\text{day}, \lambda) = \frac{1}{d} \sum_{k=1}^{d} \frac{|D_k^* - D(\hat{z}(\hat{a}), \hat{a}, x_k)|}{D_k^*}$$

Data Basis Set

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 was provided by Datastream Financial Service.

The map $(RPE_D(\text{day}, \lambda), RPE_C(\text{day}, \lambda))$



$$RPE_*(\%) \times 10^{-2} = RPE_*$$

Empirical Results: Day 1

λ	$1-\lambda$	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

Empirical Results: Day 2

λ	$1-\lambda$	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

Summary Statistics

SUMMARY STATISTICS

	MC	ANS	NS
$\mathrm{mean}(\sigma)$	0.0186	0.0221	0.0218
$\mathrm{mean}(a)$	0.0838	0.1911	0.1796
$\mathrm{C}_v(\sigma)$	0.0934	0.1453	0.1406
$\mathrm{C}_v(a)$	0.2245	0.3922	0.3821
$\mathrm{mean}(RPE_C$ (%))	1.8059	2.4123	2.5997
$\mathrm{mean}(RPE_D$ (%))	0.2278	0.0467	0.0567

In-sample descriptive statistics are carried out using the daily Pareto points with best fit capability in Caps.

The End

THANK YOU FOR YOUR ATTENTION