UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA

Jury Falini

Pricing caps with HJM models: the benefits of humped volatility

n. 563 – Agosto 2009

Abstract - In this paper we compare different multifactor HJM models with humped volatility structures, to each other and to models with strictly decreasing volatility. All the models are estimated on Euribor and swap rates panel data. We develop the analysis in two steps: first we study the in-sample properties of the estimated models, then we study the pricing performance on caps. We find the humped volatility specification to greatly improve the model estimation and to provide sufficiently accurate cap prices, although the models has been calibrated on interest rates data and not on cap prices. Moreover we find the two factor humped volatility model to outperform the three factor models in pricing caps.

Keywords: Finance, interest rates, humped volatility, Kalman filter, cap and floor pricing
JEL classification: E43, G12, G13

Jury Falini, Dipartimento di Economia Politica, Università degli Studi di Siena, Italy - e-mail: falini3@unisi.it

1 Introduction

Interest rate option markets are the most liquid and important in the financial community, both from the point of view of trading volume (caps/floors and swaptions are the most traded) and variety of securities, and these derivatives are used both for speculative purpose as well as to hedge against term structure curve movements. This huge quantity of derivatives prompted a push to academic research, developing a large number of theoretical papers, even if the popolarity among practitioners has been hampered due to model complexity.

The most important empirical result is that humped volatility improves the model specification, both in terms of likelihood score, analysis of the yield errors and caps pricing performance. Moreover, the two factor model outperforms the three factor models in terms of pricing accuracy and this result is due to the combination of different types of volatility functions: the humped shape volatility mixed with the strictly decreasing one.

The literature on interest rate modelling can be coarsely divided into two different approaches: spot rate models, like Vasicek (1977) and Cox, Ingersoll and Ross (1985), based on model formulation that assumes the spot rate process as the single state variable which determines the yield curve movements, and the Heath, Jarrow and Morton (1992) methodology which describes the term structure dynamics in terms of infinite set of forward rate processes. The HJM framework has allowed to extend the class of models used to study the yield curve, including most of the more popular short rate models, but the higher flexibility of this approach came along with the higher complexity of the procedures applied to the parameter estimation and derivative securities pricing (hedging).

Most of the empirical papers about HJM models study the deterministic volatility specification because it is simpler to implement, and Brace and Musiela (1994) provide closed formulas for pricing interest rate sensitive contingent claims. When estimating stochastic volatility HJM model, ad hoc volatility specifications have been used (such as strictly decreasing volatility with respect to the maturity) with the purpose of simplifying the model even if they do not fit well market data. Morever the strictly decreasing volatility leads to a single factor model misspecification (Amin and Morton, 1994, Driessen et al., 2003, Mercurio and Moraleda, 2000).

Empirical studies have pointed out two very important issues: the first one is that interest rates volatility can depend on the level of the interest rates themselves (Chan et al., 1992 and Amin and Morton, 1994), moreover the volatility function is increasing in the short end of the curve, and decreasing in the long end, with an humped type movement (Amin and Morton, 1994, Moraleda and Vorst, 1997 and Mercurio and Moraleda, 2000).

This paper deals with the estimation of multifactor HJM models with stochastic humped volatility, and under specific conditions on the volatility structure, we express the forward rate process as an affine function of a finite set of state variables which are jointly Markovian. In fact, if the term structure of interest rates satisfies the Markov property, the numerical procedures for estimation and simulation are faster.

Our sample data consists in Euribor rates and swap rates with maturity which ranges from three months to ten years, and ATM cap volatilities with maturities 1, 2, 3, 4, 5, 7, 10 years. For estimation we use Kalman filter which is a natural way to approach panel data estimation of term structure models. In this context the state variables are treated as unobservable variables to be filtered from the observed interest rate data using the Kalman filter. Moreover, the parameters of the model can be estimated using the quasi-likelihood function obtained from the filter.

Our in-sample analysis compares the goodness of the model estimation by means of the Likelihood Ratio Test for nested models, and by means of information criteria for non-nested ones. In the second part of this paper the estimated models are compared analysing their ability in pricing interest rate derivative products. It is remarkable to note that our pricing results have been obtained without calibrating the model to observed option prices, which is instead the usual approach in the literature (Driessen et al., 2003). Option-based estimation has much appeal on practitioners which, for example, calibrate the Hull and White (1990) model on Bermudan swaptions, Libor Market Model on caps and Swap Market Model on swaptions; unfortunately it is known that these models are incompatible among them, making impossible to manage the interest rate risk of non homogeneous interest rate derivative portfolios.

This paper is organized as follows: section 2 reviews the existing literature on humped volatility; section 3 briefly discusses the HJM framework and the volatility specifications of the estimated models; in section 4 we explain the estimation method; section 5 contains the in-sample results and section 6 shows the caps pricing performance; section 7 concludes.

2 Literature review

Amin and Morton (1994) study six different forward volatility specifications, estimating the parameters on the Eurodollar future options; one of these specifications is $\sigma(t,T) = \gamma e^{-k(T-t)}$ that leads to the Gaussian model. They find an estimated value of the parameter k that is negative on average making strictly rising the volatility function, and they conclude that the volatility is humped.

Moraleda and Vorst (1997) and Mercurio and Moraleda (2000) analyse the single factor Gaussian model using cap and floor prices, and they also find a negative estimate of the exponential parameter of the volatility function, deducing that the volatility can be humped. This phenomenon has always been difficult to handle mathematically, especially with stochastic volatility, Moraleda and Vorst (1997), Ritchken and Chuang (2000) and Fan, Gupta and Ritchken (2001).

Among the first attempts to implement an HJM model with humped volatility is Mercurio and Moraleda (2000), where the humped volatility function is given by

$$\sigma(t,T) = (\sigma + \gamma(T-t))\, e^{-\lambda(T-t)} \tag{1}$$

This specification is stationary and leads to analytical formula for pricing European options on discount bonds, but it is a non-stochastic single factor model. In alternative, Moraleda and Vorst (1997) specify the volatility structure within the Ritchken and Sankarasubramanian (1995) class of models

$$\sigma(t,T) = \sigma \frac{1+\gamma T}{1+\gamma t} e^{-\lambda(T-t)} \tag{2}$$

so that recombining trees can be used for pricing derivative securities; unfortunately this model is not stationary, deterministic and single factor only. Successively Ritchken and Chuang (2000) use the Nelson and Siegel (1987) family function

$$\sigma(t,T) = (\sigma + \gamma(T-t))\, e^{-\lambda(T-t)} + \theta \tag{3}$$

Also in this case, they derive closed formula for pricing European options on discount bonds and the model is described in terms of three Markovian state variables, but it is a single factor model with deterministic volatility.

Recently, Fan, Gupta e Ritchken (2001) have estimated a two-factor model with stochastic volatility where the humped volatility function is defined by

$$\sigma(t,T) = \left[(\sigma + \gamma(T-t))\, e^{-\lambda(T-t)} + \theta\right] f(t,T)^{\alpha} \tag{4}$$

When α is positive the forward volatility depends on the level of rates, so computational problems arise. Moreover, they do not provide the term structure dynamics.

In this paper we analyse multifactor HJM models combining special cases of the following forward volatility function

$$\sigma\left(t,T\right)=\sqrt{f\left(t,T_{1}\right)}\left[\left(c_{1}+c_{2}\left(T-t\right)\right)\exp\left\{-c_{3}\left(T-t\right)\right\}+c_{4}\right] \tag{5}$$

3 The model

In this section we present the HJM methodology (Heath et al., 1992) and the conditions which allow to get a Markovian model. We consider a time interval $[0,\tau]$, for a fixed $\tau>0$, and suppose $(\Omega,F,\{\mathcal{F}_{t}:t\in[0,\tau]\},P)$ is a filter probability space satisfying the usual conditions where the flow of information evolves according the filtration $\mathcal{F}_{t}$ generated by $n\geq 1$ independent Brownian motions. The HJM framework is based on the specification of the entire forward rate curve dynamics under the natural measure P

$$f\left(t,T\right)=f\left(0,T\right)+\int_{0}^{t}\widetilde{\alpha}\left(u,T\right)du+\sum_{i=1}^{n}\int_{0}^{t}\sigma_{i}\left(u,T\right)d\widetilde{W}_{i}\left(u\right) \tag{6}$$

where $\alpha\left(u,T\right)$ e $\sigma_{i}\left(u,T\right)$ are $\mathcal{F}_{t}-$adapted processes and $f\left(0,T\right)$ is the initial forward rate curve. Under the risk-neutral measure Q, non arbitrage condition implies that the drift process must be related with the volatility structure by

$$\alpha\left(t,T\right)=\sum_{i=1}^{n}\sigma_{i}\left(t,T\right)\int_{t}^{T}\sigma_{i}\left(t,u\right)du=\sum_{i=1}^{n}\sigma_{i}^{*}\left(t,T\right) \tag{7}$$

so we can always specify the model under the equivalent martingale measure Q

$$f\left(t,T\right)=f\left(0,T\right)+\sum_{i=1}^{n}\int_{0}^{t}\sigma_{i}^{*}\left(u,T\right)du+\sum_{i=1}^{n}\int_{0}^{t}\sigma_{i}\left(u,T\right)dW_{i}\left(u\right) \tag{8}$$

In general the forward rate is not a Markov process and consequently also the spot rate is not Markovian; the main input of an HJM model is the forward volatility structure $\sigma_{i}\left(t,T\right)$, so the analysis is focused on the conditions on the forward volatility that allow to get an affine and Markovian model. Carverhill (1994) shows which conditions on the volatility structure allow to get the spot rate a Markov process, only in the determistic case; successively Jeffrey (1995) analyses the stochastic volatility case, within the single factor models. This paper deals with multifactor HJM models that satisfy the Markov property, not as regards the spot rate, but in general with respect to non observable state variables.

An HJM model is representable in terms of a $d-$dimensional Markovian system if there exists a $d-$dimensional Markovian process $Z(t)$ and a deterministic function Φ such that

$$f(t,T) = \Phi(t,T,Z(t)) \tag{9}$$

$$dZ(t) = \mu_z(t,Z)\,dt + \sigma_z(t,Z)\,dW(t) \tag{10}$$

Moreover, the model is affine if the function Φ is affine in $Z(t)$

$$f(t,T) = h_0(t,T) + h_1(t,T)\,Z(t) \tag{11}$$

The vector $Z(t)$ is the state variable vector. Inui and Kijima (1998), Chiarella and Kwon (2001,2003), deal with the conditions on the volatility structure which lead to the representation (9)-(10) and (11). Suppose that for each $1 \le i \le n$, there exists m_i such that the forward volatility $\sigma_i(t,T)$ can be expressed in the following form

$$\sigma_i(t,T) = \sum_{j=1}^{m_i} c_{ij}(t)\,\sigma_{ij}(T) \tag{12}$$

where $c_{ij}(t)$ are stochastic processes and $\sigma_{ij}(T)$ are deterministic functions. This property with the following theorem leads to the Markov system representation.

Theorem 1 (Chiarella and Kwon, 2003) *Let $\sigma_i(t,T)$, for $1 \le i \le n$, satisfy (12), then the corresponding HJM model admits a finite dimensional affine realisation*

$$f(t,T) = f(0,T) + \sum_{i=1}^{n}\sum_{j=1}^{m_i} \sigma_{ij}(T)\,\psi_j^i(t)$$
$$+ \sum_{i=1}^{n}\sum_{\substack{j,k=1\\ j\le k}} \left[\sigma_{ij}(T)\,\overline{\sigma}_{ik}(T) + \epsilon_{jk}\sigma_{ik}(T)\,\overline{\sigma}_{ij}(T)\right]\rho_{jk}^i(t) \tag{13}$$

where

$$
\begin{aligned}
\overline{\sigma}_{ij}\left(T\right) &= \int_0^T \sigma_{ij}\left(s\right)ds \\
\psi_j^i\left(t\right) &= \int_0^t c_{ij}\left(s\right)dW_i\left(s\right) - \sum_{k=1}^{m_i}\int_0^t c_{ij}\left(s\right)c_{ik}\left(s\right)\overline{\sigma}_{ik}\left(s\right)ds \\
\rho_{jk}^i\left(t\right) &= \int_0^t c_{ij}\left(s\right)c_{ik}\left(s\right)ds \\
\epsilon_{jk} &= \begin{cases} 1 & if\ j \neq k \\ 0 & if\ j = k \end{cases}
\end{aligned}
$$

In this formulation the variables describing the model are: the deterministic functions

$$\overline{\sigma}_{ij}\left(t\right) = \int_0^t \sigma_{ij}\left(u\right)du \tag{14}$$

and the stochastic processes

$$\rho_{jk}^i\left(t\right) = \int_0^t c_{ij}\left(u\right)c_{ik}\left(u\right)du \tag{15}$$

$$\psi_j^i\left(t\right) = \int_0^t c_{ij}\left(u\right)dW_i\left(u\right) - \sum_{k=1}^{m_i}\int_0^t c_{ij}\left(u\right)c_{ik}\left(u\right)\overline{\sigma}_{ik}\left(u\right)du \tag{16}$$

Let $X\left(t\right) = \left\{\rho_{jk}^i\left(t\right), \psi_j^i\left(t\right) :\ i \leq j \leq n,\ 1 \leq j,k \leq m_i\right\}$ the state variable vector and denoting with d its dimension, then (13) can be rewritten as

$$f\left(t,T\right) = f\left(0,T\right) + \sum_{i=1}^{d} a_i\left(T\right)X_i\left(t\right) \tag{17}$$

Consequently the bond price can be represented by the exponentially affine form

$$P\left(t,T\right) = \frac{P\left(0,T\right)}{P\left(0,t\right)}\exp\left\{\sum_{i=1}^{d}\overline{a}_i\left(T\right)X_i\left(t\right)\right\} \tag{18}$$

We can proof that the state variables form $d-$dimensional Markovian system which dynamics can be set in following affine form

$$dX\left(t\right) = \left[a\left(t\right) + A\left(t\right)X\left(t\right)\right]dt + U\left(t\right)dW\left(t\right) \tag{19}$$

$$U\left(t\right) = C\left(t\right)\sqrt{b\left(t\right) + B\left(t\right)X\left(t\right)} \tag{20}$$

Most of empirical papers implementing multifactor HJM models, use a strictly decreasing volatility specification given by

$$\sigma_i(t,T) = \sqrt{\sigma_i^2 r(t) + a_i^2} e^{-c_i(T-t)} \; i = 1, ..., n \tag{21}$$

This volatilty structure is an affine function of the spot rate process, and it includes as a special case, $\sigma_i = 0$, the Gaussian model (or Vasicek type). In this case, it is possible to perform a statistical test to value the parameter σ_i using the Likelihood Ratio Test for nested models. The volatility (21) satisfies the condition (12), so the term structure can be expressed by a Markovian system with $2n$ state variables. The affine specification (21) has been studied by De Jong and Santa-Clara (1999): they set the initial forward rate curve constant, an assumption which implies a stationary model. This assumption is useful to write the conditional moments of the state variables in closed form when it is necessary to discretise the dynamics of the state variables in order to implement the Kalman filter; moreover, it is a two factor model. In this paper the initial forward curve is a function of the maturity, and a three-factor model is implemented.

Humped shape volatility is more sticking to market data and it has already been studied in several papers, Moraleda and Vorst (1997), Mercurio and Moraleda (2000), Fan, Gupta and Ritchken (2001) and Angelini and Herzel (2005) but none of them studies models being at same time multifactor, stochastic, affine and Markovian. Different types of volatility specifications will be implemented, combining special cases of (5). Table 1 summarizes the estimated models.

4 Estimation

Since the relationship that links yield to maturity with the state variables is linear, as shown by (18), and the dynamics of the state variables is affine and Markovian (that is, the drift and variance of the process are linear functions with respect to the variables), we use the Kalman

Table 1 presents the volatility specifications of the estimated models.

Volatility Function		Label
Three factor (mixed) model: $\begin{cases} \sigma_1(t,T) = a_1\sqrt{r(t)} \\ \sigma_2(t,T) = \sqrt{r(t)}\left(a_2 e^{-a_3(T-t)} + a_4\right) \\ \sigma_3(t,T) = \sqrt{r(t)}\left[(a_5 + a_6(T-t))\, e^{-a_7(T-t)} + a_8\right] \end{cases}$		
All parameters		3-MIXED1
$a_8 = 0$		3-MIXED2
$a_8 = a_4 = 0$		3-MIXED3
Two factor (mixed) model: $\begin{cases} \sigma_1(t,T) = \sqrt{f(t,T_1)}\left(a_2 e^{-a_3(T-t)} + a_4\right) \\ \sigma_2(t,T) = \sqrt{f(t,T_2)}\left[(a_5 + a_6(T-t))\, e^{-a_7(T-t)} + a_8\right] \end{cases}$		
$a_8 = a_4 = 0, T_1 = T_2 = t$		2-MIXED4
$a_8 = a_4 = 0$		2-MIXED5
$a_8 = 0, T_1 = T_2 = t$		2-MIXED6
Models with same volatility function for all factors:		
$\sigma_i(t,T) = a_i e^{-c_i(T-t)}$	$i = 1, 2, 3$	i-GAUSSIAN
$\sigma_i(t,T) = \sqrt{\sigma_i^2 r(t) + a_i^2}\, e^{-c_i(T-t)}$	$i = 1, 2, 3$	i-SQRT
$\sigma_i(t,T) = (a_i + b_i(T-t))\, e^{-c_i(T-t)}$	$i = 1, 2, 3$	i-HUMP
$\sigma_i(t,T) = \sqrt{r(t)}\,(a_i + b_i(T-t))\, e^{-c_i(T-t)}$	$i = 1, 2, 3$	i-HUMP-STOC

filter to estimate the proposed models. The Kalman filter is a tractable and reasonably accurate estimation method even though the exact likelihood function is not known (Duffee and Stanton, 2001).

In the literature, some empirical works deal with the calibration of HJM models with the Kalman Filter (De Jong and Santa Clara, 2001 and Chiarella, Hung and To, 2009), while some authors estimate affine spot rate models (Jegadeesh and Pennacchi, 1996, Duan and Simonato, 1999, Geyer and Pichler, 1999, Babbs and Nowman,1999 and Chen and Scott, 2003).

Suppose that in the bond market M zero-coupon bonds with different maturities are traded, and denote $y_k(\tau_i) = -\ln(P(t_k, t_k + \tau_i))/\tau_i$ the yield observed at time t_k $k = 1..T$ with time to maturity τ_i, then using (18) the measurement equation can be written in the following form

$$y_k = d_k(\theta) + A_k(\theta) X_k + \epsilon_k \tag{22}$$

where $X_k \in R^d$ is the vector of the state variables, $y_k = (y_k(\tau_1), ..., y_k(\tau_M))$ is the yield vector observed at time t_k and θ is the parameter vector. It is necessary to introduce a measurement error, Normally distribuited, with zero mean and covariance matrix $H(\theta)$.

The differential equation (19)-(20) is defined in continuous time, while the yields are observed at fixed time intervals, so if we discretise following the Euler scheme, we obtain the "transition equation" which represents the evolution of the process between t_k and t_{k+1}

$$\begin{aligned} X_{k+1} &= a_k \Delta t + (I_d + A_k \Delta t) X_k + U_k \sqrt{\Delta t} \eta_k \\ X_{k+1} &= c_k(\theta) + M_k(\theta) X_k + Q_k(\theta) \eta_k \\ \eta_k &\sim N(0,1) \end{aligned} \tag{23}$$

In the Gaussian case the innovations η_k are Normally distribuited and then the assumptions of the Kalman filter are satisfied, allowing to obtain the parameter estimation by maximum likelihood. In general case we are not able to compute the innovation probability distributions, so we cannot exploit the maximum likelihood estimation. However we suppose that, for high-frequency data, the innovations are Normally distribuited obtaining the quasi-likelihood estimation.

Under the hypothesis of Normality, in order to derive the discrete time exact representation of (19)-(20) we would be having to compute the conditional moments at time t_k of $X(t_k + \Delta t)$, that is $E_k[X(t_k + \Delta t)]$ and $V_k[X(t_k + \Delta t)]$ (Fackler, 2000). In this paper, Δt is fixed equal one day and this short time interval allows Euler scheme to get the same results, in terms of likelihood function, as the exact representation (Lund, 1997).

The Kalman filter is a set of recursive equations; let z_k the optimal estimator of X_k and P_k the associated MSE matrix of z_k, the algorithm consists in two set of equations, the prediction equations

$$\begin{aligned} z_{k,k-1} &= M_{k-1} z_{k-1} + c_{k-1} \\ P_{k,k-1} &= M_{k-1} P_{k-1} M'_{k-1} + Q_{k-1} Q'_{k-1} \end{aligned} \tag{24}$$

and the updating equations

$$
\begin{aligned}
z_k &= z_{k,k-1} + P_{k,k-1} A_k' R_k^{-1} v_k \\
P_k &= P_{k,k-1} - P_{k,k-1} A_k' R_k^{-1} A_k P_{k,k-1} \\
v_k &= y_k - d_k - A_k z_{k,k-1} \\
R_k &= A_k P_{k,k-1} A_k' + H
\end{aligned} \tag{25}
$$

where v_k and R_k are the prediction error and its MSE matrix.

The value z_k is the optimal estimate of X_k given the available information at time t_k, and it is referred as the filtered estimate of X_k; since the predictor error v_k is Gaussian, in order to estimate the parameter vector θ we can maximize the log-likelihood function

$$
\log L\left(y|\,\theta\right) = -\frac{dT}{2}\log\left(2\pi\right) - \frac{1}{2}\sum_{k=1}^{T}\log\left(\left|R_k\left(\theta\right)\right|\right) - \frac{1}{2}\sum_{k=1}^{T} v_k'\left(\theta\right) R_k^{-1}\left(\theta\right) v_k\left(\theta\right) \tag{26}
$$

4.1 Monte Carlo study

In this section we analyse the finite sample properties of the Kalman filter using Monte Carlo simulations: we simulate a time-series of 700 daily observations, and for each realization a set of 13 zero-coupon yields for maturities of 3, 6, 9 months and from 1 to 10 years. We repeat the procedure 400 times. To perform Monte Carlo experiments, the unobserved state variables of the model must first be simulated; we discretise the dynamics of the state variables using the Euler scheme. Monte Carlo analysis is applied to three models: three-factor Gaussian model (3-GAUSSIAN), three-factor square root model (3-SQRT) and two-factor humped volatility model (2-MIXED4). In order to simulate the paths of the state variables, the true parameter values are taken from the estimates reported in next section, setting to zero the market price of risk. The covariance matrix H is fixed to $h^2 I_{13}$, and the initial yield curve $y\left(0,t\right) = b_0 + \left(b_1 + b_2 t\right)\exp\left(-b_3 t\right)$. In tables 2 and 3 the results are shown.

Table 2 shows the Monte Carlo results for the maximum likelihood estimation. We simulate 30000 sample paths at daily frequency, for maturities of 3, 6 and 9 months, and from 1 to 10 years. Standard deviation is computed both in-sample and with the average of the Hessian matrix. To simulate the state variable dynamics we use the Euler scheme.

	three-factor Gaussian model				three-factor square root model			
	generated	mean	std	std(Hessian)	generated	mean	std	std(Hessian)
k_1	0.109	0.0109	0.00307	0.00427	0.109	0.0109	0.00307	0.00407
k_2	0.486	0.486	0.00899	0.0211	0.486	0.486	0.00899	0.0196
k_3	1.74	1.75	0.0316	0.0541	1.74	1.75	0.0316	0.0544
σ_1	0	0	0	0	1.81e-7	1.83e-7	6.71e-6	2.42e-8
σ_2	0	0	0	0	1.89e-7	1.91e-7	9.91e-6	2.52e-8
σ_3	0	0	0	0	5.46e-7	5.58e-7	4.23e-6	1.25e-7
γ_1	0.0111	0.0111	3.91e-4	4.18e-4	0.0111	0.0111	3.91e-4	4.36e-4
γ_2	0.0111	0.0111	4.83e-4	6.38e-4	0.0111	0.0111	4.83e-4	6.19e-4
γ_3	0.0121	0.0121	5.15e-4	6.11e-4	0.0121	0.0121	5.15e-4	5.7e-4
h	2.7e-4	2.7e-4	2.28e-6	2.27e-6	2.7e-4	2.7e-4	2.28e-6	2.26e-6
b_0	0.0519	0.0519	1.97e-4	1.8e-4	0.0519	0.0519	1.97e-4	1.87e-4
b_1	-0.02	-0.02	2.86e-4	2.21e-4	-0.02	-0.02	2.86e-4	2.24e-4
b_2	-0.0044	-0.0044	8.67e-5	7.31e-5	-0.0044	-0.0044	8.67e-5	6.83e-5
b_3	0.185	0.185	0.00235	0.00233	0.185	0.185	0.00235	0.00239

Table 2 shows the Monte Carlo results for the maximum likelihood estimation. We simulate 30000 sample paths at daily frequency, for maturities of 3, 6 and 9 months, and from 1 to 10 years. Standard deviation is computed both in-sample and with the average of the Hessian matrix. To simulate the state variable dynamics we use the Euler scheme.

	two-factor humped volatility model			
	generated	mean	std	std(Hessian)
a_2	0.0204	0.0204	8.77e-4	0.00124
a_3	0.387	0.415	0.00939	0.08560
a_5	-0.00939	-0.00951	0.00164	0.00185
a_6	0.055	0.0551	0.00267	0.00216
a_7	0.387	0.0387	0.00203	0.00301
h	5.27e-4	5.27e-4	4.13e-6	4.07e-6
b_0	0.0553	0.0553	1.93e-4	1.71e-4
b_1	-0.0223	-0.0223	3.69e-4	2.87e-4
b_2	-0.0067	-0.00666	1.46e-4	8.77e-5
b_3	0.198	0.198	0.00198	0.00168

Table 4. Summary statistics for yield to maturity.

	3m	6m	9m	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
mean	3.86	3.88	3.91	3.99	4.2	4.39	4.56	4.71	4.86	4.99	5.11	5.21	5.28
std	0.81	0.78	0.77	0.78	0.74	0.7	0.67	0.63	0.6	0.57	0.54	0.51	0.49
min	2.56	2.57	2.63	2.63	2.77	2.95	3.16	3.37	3.53	3.66	3.80	3.92	4.02
max	5.11	5.14	5.15	5.29	5.43	5.51	5.56	5.61	5.69	5.75	5.82	5.89	5.96

5 Empirical Results

The sample data consists in 13 time-series of Euribor rates (3, 6, 9 months) and swap rates (from 1 to 10 years) and it ranges from January 1999 to December 2001 for a total of 777 daily observations. We obtain the yield curve bootstrapping the swap rates.

The estimation procedure requires, besides the volatility function, the initial yield curve as a function of the maturity given by

$$y\left(0,t\right) = b_0 + (b_1 + b_2 t)e^{-b_3 t} \tag{27}$$

and it is estimated simultaneously with the other parameters of the model, differently from Driessen et al. (2003) where the initial curve is estimated separately and then used to determine the other parameters. Figure 1 shows the estimated curves for one and two factor models.

We compare nested models with the usual Likelihood Ratio test, otherwise we use the information criteria BIC, AIC and HQC defined as

$$\begin{aligned} AIC &= -2\log\left(L\right) + 2p \\ BIC &= -2\log\left(L\right) + p\log\left(N\right) \\ HQC &= -2\log\left(L\right) + 2\log\left(\log\left(N\right)\right) \end{aligned} \tag{28}$$

where p is the number of parameters, N is the sample dimension and L is maximum value of the likelihood function; in order to value the fitting on market data, we analyse the yield errors expressed in basis points. Table 10 reports the yield errors of the estimated models.

When we build the Kalman filter algorithm, in the measurement equation we add a disturbance ϵ_k whose covariance matrix is $H(\theta)$; for all models, the matrix $H(\theta)$ will be $H = h_0^2 I_M$. This choice has been made after several unsuccessful attempts to specify a suitable functional form of the error covariance matrix $H(\theta)$. Even if some papers, Geyer e Pichkler (1999) e Brandt e He (2002), have shown the opportunity of defining the error covariance matrix as a function of the maturity, this choice is due to computational requirements to avoid adding a parameter for each maturity.

It is important to remark that the Kalman filter requires an initial state vector X_0 from which the algorithm starts, and this can be a random variable. From equations (**??**)-(16) and (18) we note that for $t = 0$ the model must provide the initial term structure of interest rates, this implies that $X(0) = 0$ and this is the value to inizialise the filter, so the initial state vector is not a random variable.

To estimate the models using market data, we need to specify in (19)-(20) the market price of risk since data are observed under the objective probability measure; the market price of risk is fixed proportional to the diffusion coefficient in order to maintain the affine property of state variable dinamics

$$dX(t) = [a(t) + A(t)X(t)]\,dt + C(t)\sqrt{b(t) + B(t)X(t)}\,dW(t) \tag{29}$$

$$dW(t) = d\widetilde{W}(t) - \lambda(t)\,dt \tag{30}$$

$$\lambda(t) = \theta C(t)\sqrt{b(t) + B(t)X(t)} \tag{31}$$

In our estimations we set $\theta = 0$ because it does not provide any improvement in terms of likelihood score. This is in line with Jegadeeshand Pennacchi (1996), De Jong and Santa Clara (1999), Babbs and Nowman (1999), Duan and Simonato (1999) and Chen and Scott (2003).

5.1 One factor models

Table 5 shows the estimated parameters for one factor models. For the strictly decreasing volatility models (1-GAUSSIAN/1-SQRT), the parameter c is estimated to be negative and the initial yield curve does not fit the observed data; if we instead introduce the humped specification (1-HUMP/1-HUMP-STOC) the parameter c is positive and the initial yield curve fits well the market data. The Likelihood Ratio test between the models 1-HUMP and 1-GAUSSIAN, is equal to 4220. Comparing the humped stochastic volatility model 1-HUMP-STOC to the corresponding nested model 1-SQRT with $a = 0$ (setting $b = 0$ and $a = \sigma$), the test rejects the null hypothesis . Analysing the yield errors, we clearly note the improvement yielded by the humped volatility specification: the errors have gone down of 30-40% with respect to the strictly decreasing volatility model. This is in line with the misspecification of strictly decreasing volatility evidenced by Driessen et al. (2003), Amin and Morton (1994) and Mercurio and Moraleda (2000).

If we further include the parameter σ in the model estimation, the volatility function is stochastic and the model 1-SQRT includes the determistic model 1-GAUSSIAN; even if the Likelihood Ratio test with respect to the Gaussian model is equal to 45.38, the parameter c is still estimated to be negative.

Overall, the fitting of the single factor model is satisfactory compared to other papers that use the Kalman filter on panel data (Babbs and Nowman, 1999, Geyer and Pichkler, 1999 and Brandt and He, 2002): the yield errors are less than 30 basis point (1-GAUSSIAN/1-SQRT), while in Brandt and He (2002) the single factor model provides yield errors from 10 to 65 bps. De Jong and Santa Clara (1999) estimate a one factor model with the volatility function (21) but under the constant initial yield curve assumption.

Table 5 presents the parameter estimation for one factor models where L is the log-likelihood score and p is the number of parameters. In the last row we report the BIC criterion for non-nested models. Estimated standard errors are reported in parentheses.

	DETERMINISTIC				STOCHASTIC					
	1-GAUSSIAN		1-HUMP		1-SQRT		1-SQRT ($a = 0$)		1-HUMP-STOC	
c	-0.0902	(0.00138)	0.622	(0.0034)	-0.0901	(0.00135)	-0.0901	(0.00135)	0.621	(0.0034)
b	0		0.0159	(7.62e-4)	0		0		0.0847	(0.00402)
a	0.00627	(2.86e-4)	-0.00324	(1.7e-4)	5.64e-9	(2.96e-5)	0		-0.0173	(9.02e-4)
σ	0				0.032	(0.0015)	0.032	(0.0015)		
h	0.00177	(1.34e-5)	0.00117	(8.82e-6)	0.00177	(1.34e-5)	0.00177	(1.34e-5)	0.00117	(8.83e-6)
b_0	0.0268	(2.76e-4)	0.0531	(2.26e-4)	0.0267	(2.75e-4)	0.0267	(2.75e-4)	0.053	(2.22e-4)
b_1	0.0192	(5.50e-4)	-0.0214	(3.86e-4)	0.0197	(5.50e-4)	0.0197	(5.50e-4)	-0.0214	(3.81e-4)
b_2	-0.0397	(4.92e-4)	-0.00879	(3.76e-4)	-0.0401	(4.87e-4)	-0.0401	(4.87e-4)	-0.00885	(3.7e-4)
b_3	1.172	(0.00817)	0.279	(0.00396)	1.18	(0.00817)	1.18	(0.00818)	0.279	(0.00388)
L	52779		56599.26		52801.69		52801.69		56584.69	
p	7		8		8		7		8	
BIC	-105512		-113146		-105551		-105558		-113117	

5.2 Two factor models

Many papers show that increasing the number of the factors produces a tangible improvement of the model estimation (Babbs and Nowman, 1999, Geyer and Pichkler, 1999, Brandt and He, 2002 and De Jong and Santa Clara, 1999), and this also occurs in our analysis: if we add a risk factor, the model is more flexible and it is able to represent the term structure movements. Indeed, the estimate of parameters c_i in the models 2-GAUSSIAN/2-SQRT is now positive for both factors. Yield errors are considerably lower, being smaller than 10 bps for all maturities; the measurement error standard deviation also reduces considerably, from 1.77e-3 of the model 1-GAUSSIAN to 5.64e-4 of the model 2-GAUSSIAN. Babbs and Nowman (1999) specify the covariance error matrix using a parameter for each maturity, they obtain yield errors that vary from 1.4e-4 to 2.8e-3 using eight time series of interest rates.

Here as well, the humped volatility specification provides better estimation results, both with deterministic and stochastic volatilities. The model 2-HUMP includes as a special case

Table 6 presents the parameter estimation for two factor models where L is the log-likelihood score and p is the number of parameters. In the last row we report the BIC criterion for non-nested models. Estimated standard errors are reported in parentheses.

	DETERMINISTIC				STOCHASTIC					
	2-GAUSSIAN		2-HUMP		2-SQRT		2-SQRT ($a_i = 0$)		2-HUMP-STOC	
c_1	0.218	(0.00325)	0.834	(0.00898)	0.218	(0.00324)	0.217	(0.00323)	0.259	(0.00306)
b_1	0		0.0132	(4.66e-4)	0		0		0.0319	(0.00157)
a_1	0.0119	(4.77e-4)	0.00105	(5.87e-5)	0.0119	(4.77e-4)	0		-0.0191	(8.83e-4)
σ_1	0		0		8.52e-7	(2.74e-4)	0.0654	(0.00263)	0	
c_2	0.678	(0.00967)	0.257	(0.00311)	0.678	(0.00967)	0.681	(0.00964)	0.839	(0.009)
b_2	0		0.00551	(2.64e-4)	0		0		0.0679	(0.00242)
a_2	0.0121	(5.02e-4)	-0.00341	(1.54e-4)	0.0121	(5.02e-4)	0		0.00547	(3.05e-4)
σ_2	0		0		2.06e-6	(2.84e-4)	0.0661	(0.0027)	0	
h	5.65e-4	(4.56e-6)	3.76e-4	(2.98e-6)	5.65e-4	(4.56e-6)	5.66e-4	(4.58e-6)	3.76e-4	(2.98e-6)
b_0	0.0538	(1.85e-4)	0.0548	(2.7e-4)	0.0539	(1.85e-4)	0.0538	(1.85e-4)	0.0548	(2.67e-4)
b_1	-0.0209	(4.00e-4)	-0.0229	(3.14e-4)	-0.0209	(4.00e-4)	-0.0209	(4.0e-4)	-0.0229	(3.11e-4)
b_2	-0.00676	(1.32e-4)	-0.00443	(8.81e-5)	-0.00676	(1.32e-4)	-0.00675	(1.32e-4)	-0.00443	(8.84e-5)
b_3	0.211	(0.00213)	0.165	(0.00233)	0.211	(0.00213)	0.211	(0.00214)	0.165	(0.0023)
L	62276.22		65974.19		62276.22		62214.92		65945.51	
p	9		11		11		9		11	
BIC	-124493		-131876		-124480		-124371		-131819	

the model 2-GAUSSIAN, and comparing them with the Likelihood Ratio test we reject the null hypothesis of $b_i = 0$. As shown in Figure 1, the initial yield curve of the humped volatility model fits the market data better than the strictly decreasing volatility model. Considering stochastic volatility, the model 2-HUMP-STOC includes as a special case the model 2-SQRT with $a_i = 0$ (by setting $b_i = 0$ and $a_i = \sigma_i$): also in this case, we reject the null hypothesis. Contrary to the single factor model, the parameters σ_i are not significant.

Generally, most of empirical papers which deal with multifactor model, use the same volatility specification for all factors, and the functional form commonly used is the strictly decreasing volatility with respect to the maturity. From the principal component analysis (Litterman and Scheinkman, 1991 and Rebonato, 1998), we argue that is necessary to use different types of functions. We implement the following volatility structure

$$\sigma_1(t,T) = \sqrt{f(t,T_1)}\left[a_2 e^{-a_3(T-t)} + a_4\right] \tag{32}$$

$$\sigma_2(t,T) = \sqrt{f(t,T_2)}\left[(a_5 + a_6(T-t))\, e^{-a_7(T-t)} + a_8\right] \tag{33}$$

The model (32)-(33) improves the model estimation in comparison with the strictly decreasing volatility (2-GAUSSIAN/SQRT), but it results worse than humped volatility (2-HUMP/HUMP-STOC). However, this ranking will be reverted on the pricing performance. Table 7 shows the results: the model 2-MIXED4 includes the model 2-SQRT with $a_i = 0$, and applying the Likelihood Ratio test we reject the null hypothesis of fixing $a_6 = 0$ and we can assess that combining different functional forms of the forward volatility improves the model estimation. Regarding the initial yield curve, the estimated parameters of model 2-MIXED4 are similar to the values reported for the model 2-SQRT with $a_i = 0$; valuing the yield errors, the model 2-MIXED4 is slightly better than the model 2-SQRT with $a_i = 0$.

Moreover, the model (32)-(33) allows us to estimate which maturity to choose in the volatility function. The model 2-MIXED5 shows that it is significant estimating the forward rate maturity of the volatility functions, the Likelihood Ratio test is equal to 24.64 and we reject the null hypothesis of $T_1 = T_2 = t$. We remark that the humped volatility depends on a long-term forward rate, about 22 years, while the other volatility is fitted on a shorter maturity (about 4 months): this is not a surprise, because the principal component analysis asserts that the first component represents the average level ad the second one accounts for the slope of the yield curve. The forward rate with 22-years maturity is a long bond yield, i.e. the average level, while the short maturity determines the slope of the yield curve by the spread with the long maturity. Comparing the mixed model with the humped one, the Likelihood Ratio test clearly rejects the null hypothesis, $b_1 = 0$.

Many papers use forward volatilities which approach to zero as the maturity approaches to

Table 7 presents the parameter estimation for two factor (mixed) models where L is the log-likelihood score and p is the number of parameters. In the last row we report the BIC criterion for non-nested models. Estimated standard errors are reported in parentheses.

	2-MIXED4		2-MIXED5		2-MIXED6	
a_4	0		0		-0.0314	(0.00145)
a_3	0.387	(0.00939)	0.388	(0.00944)	0.767	(0.0194)
a_2	0.0204	(8.77e-4)	0.0199	(0.00102)	0.052	(0.00233)
T_1	t		0.346	(0.407)	t	
a_8	0		0		0	
a_7	0.387	(0.00203)	0.388	(0.00203)	0.717	(0.00628)
a_6	0.055	(0.00267)	0.0384	(0.00318)	0.0655	(0.00226)
a_5	-0.00957	(0.00164)	-0.00633	(0.00124)	0.00464	(2.76e-4)
T_2	t		22.2	(18.4)	t	
h	5.27e-4	(4.13e-6)	5.28e-4	(4.12e-6)	3.84e-4	(3.04e-6)
b_0	0.0553	(1.93e-4)	0.0553	(1.93e-4)	0.0476	(1.99e-4)
b_1	-0.0223	(3.69e-4)	-0.0223	(3.65e-4)	-0.0151	(3.25e-4)
b_2	-0.0067	(1.46e-4)	-0.00673	(1.45e-4)	-0.0047	(1.04e-4)
b_3	0.198	(0.00198)	0.198	(0.00198)	0.233	(0.00246)
L	63210.36		63222.68		65790.33	
p	10		12		11	
BIC	-126355		-126367		-131509	

infinity; the specification (32)-(33) allows us to check this hypothesis, estimating the parameters a_4 and a_8. In our in-sample analysis, the best choice is to estimate a_4 and fixing $a_8 = 0$: the model 2-MIXED6 reports the results. The improvement in the likelihood score clearly rejects the null hypothesis; also the yield errors improves.

Summarizing, using humped functions to specify the forward volatility leads to better estimation: the initial yield curve fit well observed data, the yield errors are smaller and applying the Likelihood Ratio test, we reject the hypothesis of strictly decreasing volatility.

5.3 Three factor models

Also for the three factor models, the humped volatility performs better than the strictly decreasing one. Comparing the model 3-HUMP[1] with the model 3-GAUSSIAN, it is significant to estimate the parameters b_i which allow us to model the humped shape. Valuing the BIC index and the yield errors, the model 3-HUMP outperforms the model 3-HUMP-STOC. We can't use the Likelihood Ratio test because they are not nested models. Considering the stochastic volatility, we value the benefit of the humped specification comparing the model 3-HUMP-STOC with the model 3-SQRT with $a_i = 0$: the Likelihood Ratio test rejects the null hypothesis, so the humped volatility specification improves the model estimation.

Adding a risk factor is clearly significant and it improves the fitting of the model: the yield errors are halved in comparison to the two factor models, less than 5 basis point for all maturities. The parameters σ_i is not significant, the test applied to the models 3-SQRT and 3-GAUSSIAN gives a value less than 1e-2, so we do not reject the null hypothesis.

This section ends with the mixed models: we use the following forward volatility specifications

$$\sigma_1(t,T) = a_1\sqrt{r(t)} \tag{34}$$

$$\sigma_2(t,T) = \sqrt{r(t)}\left[a_2 \exp\{-a_3(T-t)\} + a_4\right] \tag{35}$$

$$\sigma_3(t,T) = \sqrt{r(t)}\left[(a_5 + a_6(T-t))\exp\{-a_7(T-t)\} + a_8\right] \tag{36}$$

The estimates shown in table 9 must be compared to the those in table 8, using the BIC index and the yield errors. Setting $a_4 = a_8 = 0$ (3-MIXED3), the volatility functions approach to zero as the maturity approaches to infinity; we relax this hypothesis in the other models. These models can be compare with the Likelihood Ratio test because they are nested. The results

[1] We experienced difficulties in fitting three humped volatilities. For this reasons, one factor has been replaced with a strictly decreasing volatility.

Table 8 presents the parameter estimation for two factor models where L is the log-likelihood score and p is the number of parameters. In the last row we report the BIC criterion for non-nested models. Estimated standard errors are reported in parentheses.

	DETERMINISTIC				STOCHASTIC					
	3-GAUSSIAN		3-HUMP		3-SQRT		3-SQRT ($a_i = 0$)		3-HUMP-STOC	
c_1	0.109	(0.00307)	-0.0042	(1.71e-4)	0.109	(0.00307)	0.106	(0.00307)	0.252	(0.00253)
b_1	0		0.00457	(1.95e-4)	0		0		0.028	(9.35e-4)
a_1	0.0111	(3.91e-4)	0.214	(0.00342)	0.0111	(3.91e-4)	0		-0.00557	(0.00103)
σ_1	0		0		1.81e-7	(6.71e-6)	0.0598	(0.0021)	0	
c_2	0.486	(0.00899)	0.00193	(9.24e-5)	0.486	(0.00899)	0.491	(0.00914)	1.01	(0.0126)
b_2	0		0.0112	(3.8e-4)	0		0		0.0781	(0.00439)
a_2	0.0111	(4.83e-4)	0.716	(0.0109)	0.0111	(4.83e-4)	0		-0.0179	(0.00161)
σ_2	0		0		1.89e-7	(9.91e-6)	0.0583	(0.00257)	0	
c_3	1.74	(0.0316)	-0.00781	(2.97e-4)	1.74	(0.0316)	1.73	(0.0319)	0.252	(0.0133)
b_3	0		0.0335	(0.0013)	0		0		0	
a_3	0.0121	(5.15e-4)	3.73	(0.0358)	0.0121	(5.15e-4)	0		0.0197	(7.09e-4)
σ_3	0		0		5.46e-7	(4.23e-6)	0.0641	(0.00274)	0	
h	2.7e-4	(2.28e-6)	2.36e-4	(1.92e-6)	2.7e-4	(2.28e-6)	2.71e-4	(2.29e-6)	2.66e-4	(2.15e-6)
b_0	0.0519	(1.97e-4)	0.0522	(2.13e-4)	0.0519	(1.97e-4)	0.0518	(1.97e-4)	0.0536	(2.03e-4)
b_1	-0.02	(2.86e-4)	-0.0202	(2.62e-4)	-0.02	(2.86e-4)	-0.0199	(2.86e-4)	-0.0217	(2.72e-4)
b_2	-0.0044	(8.67e-5)	-0.00447	(5.94e-5)	-0.0044	(8.67e-5)	-0.00441	(8.68e-5)	-0.00433	(7.85e-5)
b_3	0.185	(0.00235)	0.183	(0.00208)	0.185	(0.00235)	0.186	(0.00237)	0.17	(0.0021)
L	67819.84		69420.05		67819.84		67759.29		68542.24	
p	11		14		14		11		13	
BIC	-135568		-138748		-135548		-135447		136999	

assert that it is significant to estimate both parameters, but a_4 implies a larger improvement in the likelihood score; valuing the yield errors, we do not see any difference among the models. Comparing the mixed models with the models of the table 8, they fit well the market data, and using the BIC criterion, we can assert that the mixed models are better than the strictly decreasing volatility models and they are worse than the humped volatility ones.

Table 9 presents the parameter estimation for two factor (mixed) models where L is the log-likelihood score and p is the number of parameters. In the last row we report the BIC criterion for non-nested models. Estimated standard errors are reported in parentheses.

	3-MIXED1		3-MIXED2		3-MIXED3	
a_1	0.0158	(6.31e-4)	0.0159	(6.4e-4)	0.0261	(9.75e-4)
a_4	-0.0319	(0.00133)	-0.0327	(0.00139)	0	
a_3	0.839	(0.0293)	0.818	(0.0284)	0.779	(0.0261)
a_2	0.0465	(0.00165)	0.047	(0.00171)	0.0322	(0.00125)
a_8	-2.22e-8	(3.32e-9)	0		0	
a_7	0.806	(0.00642)	0.789	(0.0591)	0.779	(0.00601)
a_6	0.0729	(0.00386)	0.0719	(0.00377)	0.0883	(0.00441)
a_5	-0.00866	(0.00161)	-0.00863	(0.00159)	-0.0263	(0.0021)
h	2.81e-4	(2.27e-6)	2.82e-4	(2.28e-6)	2.87e-4	(2.36e-6)
b_0	0.0466	(1.43e-4)	0.0468	(1.45e-4)	0.0471	(1.38e-4)
b_1	-0.0139	(2.51e-4)	-0.0144	(2.49e-4)	-0.0148	(2.49e-4)
b_2	-0.00565	(1.39e-4)	-0.00519	(1.23e-4)	-0.00506	(1.28e-4)
b_3	0.264	(0.00277)	0.252	(0.00247)	0.247	(0.00257)
L	68098.8		68076.72		67742.09	
p	13		12		11	
BIC	-136112		-136075		-135412	

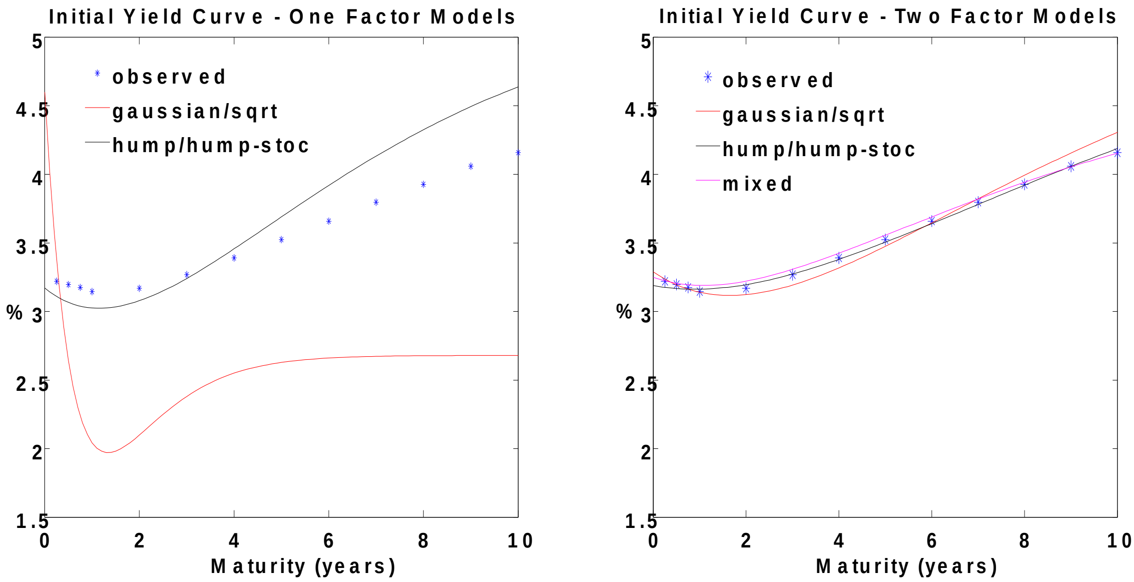

Figure 1. Comparison of the estimated initial yield curve for one and two factor models.

Table 10 reports the mean absolute errors (yield errors in basis points) of the estimated models.

	3m	6m	9m	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y	avg
1-GAUSSIAN	19.0	15.1	16.8	17.8	16.2	13.8	10.3	8.46	8.32	10.1	13.4	16.1	18.8	14.2
1-HUMP	14.4	13.1	13.2	12.1	9.32	7.5	5.66	4.46	5.08	6.76	8.54	9.81	10.9	9.29
1-SQRT	19.1	15.0	16.7	17.8	16.1	13.8	10.2	8.44	8.31	10.0	13.3	16.0	18.7	14.1
1-SQRT($a = 0$)	19.1	15.0	16.7	17.8	16.1	13.8	10.2	8.44	8.31	10.0	13.3	16.0	18.7	14.1
1-HUMP-STOC	14.4	13.1	13.2	12.1	9.31	7.5	5.67	4.48	5.09	6.76	8.53	9.8	10.9	9.3
2-GAUSSIAN	7.45	3.17	4.21	5.82	6.08	4.98	4.32	3.92	3.71	4.02	4.86	5.59	6.34	4.96
2-HUMP	6.79	3.53	3.52	3.89	4.51	4.12	3.71	3.48	3.34	3.27	3.47	3.52	3.88	3.93
2-SQRT	7.45	3.17	4.21	5.82	6.08	4.98	4.32	3.92	3.71	4.02	4.86	5.59	6.34	4.96
2-SQRT($a_i = 0$)	7.46	3.17	4.24	5.83	6.08	4.99	4.32	3.92	3.71	4.03	4.87	5.6	6.34	4.97
2-HUMP-STOC	6.81	3.52	3.52	3.89	4.51	4.11	3.7	3.49	3.34	3.26	3.46	3.52	3.87	3.92
2-MIXED4	7.5	3.36	4.26	5.94	5.83	4.58	4.08	3.91	3.63	3.67	4.33	4.93	5.83	4.76
2-MIXED5	7.51	3.36	4.27	5.95	5.86	4.6	4.09	3.91	3.63	3.68	4.34	4.94	5.86	4.77
2-MIXED6	6.56	3.69	3.6	3.98	4.46	4.32	3.88	3.51	3.29	3.25	3.48	3.62	3.98	3.97
3-GAUSSIAN	3.18	2.77	3.43	3.18	3.42	3.54	3.54	3.54	3.38	3.19	3.36	3.5	3.64	3.36
3-HUMP	2.1	2.68	2.69	3.37	3.46	3.73	3.68	3.48	3.3	3.17	3.3	3.37	3.57	3.22
3-SQRT	3.18	2.77	3.43	3.18	3.42	3.54	3.54	3.54	3.38	3.19	3.36	3.5	3.64	3.36
3-SQRT($a_i = 0$)	3.18	2.77	3.44	3.19	3.42	3.55	3.55	3.54	3.38	3.19	3.36	3.5	3.64	3.36
3-HUMP-STOC	3.41	2.91	3.48	3.21	3.39	3.52	3.5	3.57	3.44	3.19	3.32	3.45	3.61	3.38
3-MIXED1	4.01	2.99	3.66	3.49	3.34	3.57	3.51	3.51	3.43	3.2	3.24	3.4	3.6	3.46
3-MIXED2	3.89	2.89	3.74	3.59	3.47	3.65	3.66	3.58	3.44	3.25	3.31	3.47	3.67	3.51
3-MIXED3	4.04	2.98	3.65	3.53	3.34	3.58	3.51	3.5	3.42	3.23	3.26	3.4	3.58	3.46

6 Pricing

In this section the estimated models will be compared in terms of pricing errors. The data set consists in ATM cap quoted volatilities with maturities 1,2,3,4,5,7,10 years and we use the Black (1976) formula to convert volatilities in prices. Our procedure consists in the following steps: we estimate the parameters of the model using a rolling window of 550 daily observations for a total of 100 trading days; successively, we price the caps quoted in the last trading day of the sample. Closed formula are used for Gaussian models, while for stochastic models cap prices are computed using Monte Carlo simulations. Specifically, we simulate 30000 sample paths and we divide the time to maturity in trading intervals of length $\Delta t = 9/252$ years. Table 11 reports the results[2]; to compare the pricing performance we use the MAPE index given by

$$MAPE_K = \frac{100}{N} \sum_{i=1}^{N} \frac{\left| Cap_K^M(i) - Cap_K^B(i) \right|}{Cap_K^B(i)} \tag{37}$$

where $N = 100$ is the number of daily estimates, $Cap_K^M(i)$ is the model price of the cap with maturity K and $Cap_K^B(i)$ is the market price given by the Black formula.

Comparing the pricing errors, the humped volatility models outperform the other models. The in-sample analysis reveals that the mixed models do not improve the estimation with respect the humped volatility models; valuing the pricing accuracy the mixed models 2-MIXED4 and 2-MIXED6, one factor with humped volatility and the other with strictly decreasing volatility, improve the MAPE error and we can assert that the combination of two different volatility functions is the optimal choice. Moreover, is not clear the benefit of adding a risk factor. Among the strictly decreasing volatility models, the one factor model outperforms the multi-factor models. But, in the case of humped volatility, the two factor model provides better pricing performance than the three factor model. Interest rate volatility can depend on the level of the interest rates themselves (Chan et al., 1992 and Amin and Morton, 1994), and the stochastic volatility specification leads to accurate pricing performance: the models i-HUMP-STOC has a lower pricing errors than the models i-HUMP.

As reported in the previous section, the estimate of parameters a_4 (model 2-MIXED6) improves the estimation, valuing both the likelihood score and the yield errors. In terms of pricing performance, we have the same improvement, the pricing errors have gone down of 30%, from 8.2% to 5.8%.

We want to underline that the MAPE error of the model 2-MIXED6 it is relevant because our model is calibrated using only the interest rate time series. Two separete reasons influence this performance: the first one concerns the volatility specification, the humped volatility fits better the observed data and it allows to price accurately interest rate derivatives. The second

[2] We could not get reliable results for the mixed model 3-MIXED1.

Table 11 shows the pricing errors (MAPE) for caps with maturities 1,2,3,4,5,7 and 10 years. Results are based on 100 trading days using a rolling window of 550 daily observations.

	CAP MATURITY							
model	1Y	2Y	3Y	4Y	5Y	7Y	10Y	Avg
1-GAUSSIAN	87.3	35.5	8.9	23.0	32.3	36.4	20.2	34.8
2-GAUSSIAN	166.7	118.0	87.5	72.3	59.3	40.1	26.0	81.4
3-GAUSSIAN	128.9	89.4	72.9	65.9	57.8	43.4	31.9	70.0
1-SQRT1	91.0	41.2	7.3	14.4	23.1	26.5	13.4	31.0
2-SQRT1	168	119.3	89.0	74.0	61.1	42.1	27.1	82.9
3-SQRT1	129	90.6	74.8	68.4	60.6	46.5	35.4	72.2
1-HUMP	45.1	43.1	32	20.9	15.0	13.3	17.2	26.7
2-HUMP	25.6	20.8	18.1	12.5	10.1	9.2	8.5	15.0
3-HUMP	34.8	24.5	21.0	14.4	10.6	8.4	7.2	17.3
1HUMP-STOC	39.2	32.2	21.3	10.6	5.6	4.4	6.4	17.1
2-HUMP-STOC	18.7	12.0	8.5	6.7	6.1	4.9	6.4	9.0
3-HUMP-STOC	29.5	19.4	10.7	3.7	4.0	6.3	8.5	11.7
2-MIXED4	12.9	11.0	11.9	8.0	5.8	4.6	3.3	8.2
2-MIXED6	11.8	8.1	6.0	4.5	4.2	3.2	3.0	5.8
3-MIXED2	21.1	11.1	11.3	15.0	15.4	11.8	8.5	13.4
3-MIXED3	24.6	15.2	7.3	4.6	6.6	8.8	11.3	11.2

one is the estimation methods, the Kalman filter, which exploits both cross section and time series data. A huge quantity of interest rate models have been implemented using option-based estimation with the purpose of pricing and hedging a specific interest rate derivative, unfortunately, these models are often incompatible, such as the Libor Market Model for caps and the Swap Market Model for swaptions. As a consequence, it is very complicated to manage the interest rate risk when the portfolio includes different types of interest rate derivatives.

Finally, we analise the three factor models based on three different forward volatility functions, see (34)-(36). Also in this case, adding the third factor does not provide improvement in terms of pricing performance, therefore we can assert that, in our sample, in order to pricing interest rate derivatives, two factor models are sufficient.

7 Conclusions

In this work we show the benefits of humped volatility specification in multifactor HJM models; in literature, the empirical applications of multifactor models has been limited by the complexity of the procedures applied to the parameter estimation and derivative pricing (hedging). This work analyses the humped volatility specification within the stochastic multifactor HJM models which satisfy the Markov property. We develop the empirical application in two steps: in-sample analisys and pricing performance.

We show that the humped volatility improves the parameter estimation and it allows us to avoid unrealistic (implying diverging volatility) parameter estimates induced by misspecification; moreover, the humped shape allows us to correctly estimate the initial yield curve. The improvement is clear both in terms of likelihood score and yield errors. Mixed models, i.e. with different specifications of volatility functions, outperform the strictly decreasing volatility models, but they do not perform better than the humped volatility ones; the assumption of level dependent volatility does not improve the model estimation.

With respect to cap pricing, the humped volatility specification allows us to reduce heavily the pricing errors, obtaining good results despite the interest rate-based estimation. We show that the two factor model with different types of forward volatility functions, one with humped shape and the other strictly decreasing, outperforms the other models. About the level dependent volatility assumption, the results show that it improves the pricing performance.

An important direction for future research is the analisys of the benefits of the humped volatility in hedging interest rate derivatives; moreover, the analysis should be extended to more complex derivative products, for instance swaptions or Bermudan swaptions.

8 References

Angelini, F., Herzel, S. (2005): Consistent calibration of HJM models to cap implied volatilities. *Journal of Futures Markets* 25(10), 1093-1120.

Amin, K.I., Morton, A.J. (1994): Implied volatility functions in arbitrage-free term structure models. *Journal of Financial Economics* 35, 141-180.

Babbs, S.H., Nowman, K.B. (1999): Kalman filtering of generalized Vasicek term structure models. *The Journal of Financial and Quantitative Analysis* 34(1), 115-130.

Black, F. (1976): The pricing of commodity contracts. *Journal of Financial Economics* 3, 167-179.

Brace, A., Gatarek, D., Musiela, M. (1997): The market model of interest rate dynamics. *Mathematical Finance* 7(2), 124-147.

Brace, A., Musiela, M. (1994): A multifactor Gauss Markov implementation of Heath, Jarrow and Morton. *Mathematical Finance* 4(3), 259-283.

Brandt, M.W., He, P. (2002): Simulated likelihood estimation of affine term structure models from panel data. The Wharton School and Economics Department University of Pennsylvania, Manuscript

Carverhill, A. (1994): When is the short rate Markovian? *Mathematical Finance* 4(4), 305-312.

Chan, K.C., Karolyi, G.A., Longstaff F.A., Sanders, A.B. (1992): An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47(3), 1209-1227.

Chen, R., Scott, L. (2003): Multifactor Cox-Ingersoll-Ross models of the term structure: estimates and tests from a Kalman filter model. *Journal of Real Estate Finance and Economics* 27(2), 143-172.

Chiarella, C., Hung, H., To, T. (2009): The volatility structure of the fixed income market under the HJM framework: a nonlinear filtering approach. *Computational Statistics and Data Analysis* 53(6), 2075-2088.

Chiarella, C., Kwon, O.K. (2001): Classes of interest rate models under the HJM framework. *Asia-Pacific Financial Markets* 8(1), 1-22.

Chiarella, C., Kwon, O.K. (2001): Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model. *Finance and Stochastics* 5, 237-257.

Chiarella, C., Kwon, O.K. (2001): State variables and the affine nature of Markovian HJM term structure models. School of Finance and Economics, University of Technology Sydney, Working Paper.

Chiarella, C., Kwon, O.K. (2003): Finite dimensional affine realisations of HJM models in terms of forward rates and yields. *Review of Derivatives Research* 6, 129-155.

Cox, J.C., Ingersoll J.E., Ross, S.A. (1985): A theory of the term structure of interest rates. *Econometrica* 53(2), 385-407.

De Jong, F. (2000): Time-series and cross-section information in affine term structure models. *Journal of Business & Economic Statistics* 18(3), 300-314.

De Jong, F., Santa-Clara P. (1999): The dynamics of the forward interest rate curve: a formulation with state variables. *The Journal of Financial and Quantitative Analysis* 34(1), 131-157.

Driessen, J., Klaassen, P., Melenberg, B. (2003): The performance of multi-factor term structure models for pricing and hedging cpas and swaptions. *Journal of Financial and Quantitative Analysis* 38(3), 635-672

Duan, J., Simonato, J. (1999): Estimating and testing Exponential-affine term structure models by Kalman filter. *Review of Quantitative Finance and Accounting* 13, 111-135.

Duffee, G., Stanton R. (2001): Estimation of dynamic term structure models. Berkeley University of California, Working Paper.

Fackler, P. (2000): Moments of affine diffusions. North Carolina State University, Working Papers.

Fan, R., Gupta, A., Ritchken, P. (2001): On pricing and hedging in the swaption market: how many factors, really? Case Western Reserve, Working Paper.

Geyer, A.L.J., Pichler, S. (1999): A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure. *The Journal of Financial Research* 22(1), 107-130.

Gupta, A., Subrahmanyam, M.G. (2005): Pricing and hedging interest rate options: evidence from cap-floor markets. *Journal of Banking&Finance* 29, 701-733.

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

Hull, J., White, A. (1990): Pricing interest-rate-derivative secutities. *Review of Financial Studies* 3(4), 573-592.

Hull, J., White, A. (1990): One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of Financial and Quantitative Analysis* 28(2), 235-254.

Inui, K., Kijima, M. (1998): A Markovian framework in multi-factor Heath-Jarrow-Morton models. *Journal of Financial and Quantitative Analysis* 33(3), 429-440.

Jagannathan, R., Kaplin, A., Sun, S. (2003): An evaluation of multi-factor CIR models using Libor, swap rates, and cap and swaption prices. *Journal of Econometrics* 116, 113-146.

Jamshidian, F. (1997): Libor and swap market models and measures. *Finance and Stochastics* 1, 293-330.

Jeffrey, A. (1995): Single factor Heath-Jarrow-Morton term structure models based on markov spot interest rate. *Journal of Financial and Quantitative Analysis* 30(4), 619-642.

Jegadeesh, N., Pennacchi, G. (1996): The behavior of interest rates implied by the term structure of eurodollar futures. *Journal of Money, Credit and banking* 28(3), 426-446.

Longstaff, F., Santa-Clara, P., Schwartz, E. (2001): The ralative valuation of caps and swaptions: theory and empirical evidence. *Journal of Finance* 56(6), 2067-2109.

Longstaff, F., Schwartz, E.S. (1992): Interest rate volatility and the term structure: a two-factor general equilibrium model. *Journal of Finance* 47, 1259-1282.

Litterman, R., Scheinkman, J. (1991): Common factors affecting bond returns. *Journal of Fixed Income* 1, 54-61.

Lund, J. (1997): Econometric Analysis of continuos-time arbitrage-free models of the term structure of interest rates. Aarhus School of Business, Working Paper.

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

Moraleda, J.M., Vorst, T.C.F. (1997): Pricing American interest rate claims with humped volatility models. *Journal of Banking & Finance* 21, 1131-1157.

Nelson, C., Siegel, A. (1987): Parsimonious Modelling of the yield curves. *Journal of Business* 60, 473-489.

Rebonato, R. (1996): *Interest-rate option models.* John Wiley&Sons

Ritchken, P., Chuang, I. (1999): Interest rate option pricing with volatility humps. *Review of Derivatives Research* 3, 237-262.

Ritchken, P., Sankarasubramanian, L. (1995): Volatility structure of forward rates and the dynamics of the term structure. *Mathematical Finance* 5(1), 55-72.

Vasicek, O. (1977): An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177-188.