Bayesian extensions to diebold-li term structure model

Article · January 2008

Source: RePEc

2 authors:

Márcio Poletti Laurini
University of São Paulo

Luiz K. Hotta
University of Campinas

BAYESIAN EXTENSIONS TO DIEBOLD-LI TERM STRUCTURE MODEL

MÁRCIO POLETTI LAURINI
LUIZ KOODI HOTTA

Abstract. In this article we propose a statistical model to adjust, interpolate and forecast the term structure of interest rates. This model is based on extensions for the term structure model of interest rates proposed by [Diebold & Li, 2006], through a Bayesian estimation using Markov Chain Monte Carlo. The proposed extensions involve the use of a more flexible parametric form for the yield curve, making all parameters time-varying using a structure of latent factors, and adding a stochastic volatility structure to control the presence of conditional heteroscedasticity observed in the interest rates.
The Bayesian estimation enables the exact distribution of estimators in finite samples, and as a sub product, the estimation enables obtaining the distribution of forecasts for the term structure of interest rates. The methodology developed does not need a pre-interpolation of the yield curve as it happens in some econometric models of term structure. We do an empirical exercise of this methodology in which we adjust daily data of the term structure of interest rates implicit in Swap DI-PRÉ contracts traded in the Mercantile and Futures Exchange (BM&F)
Keywords: Term Structure, Bayesian Inference, Markov Chain Monte Carlo.
JEL Codes: G1,C22.

First author : Ibmec São Paulo and PhD student in Statistics at the Statistics Department, IMECC-UNICAMP. Second author - Statistics Department, IMECC-UNICAMP. Email contact marciopl@isp.edu.br - First Version - April/2007.

1. Introduction

The term structure of interest rates may be defined as a collection of interest rates, indexed in two dimensions: maturity and time. The first index shows the relation between rates with different maturities for contracts of the same nature in a determined period. The second index shows the time evolution of the rates of contracts with the same maturities. The term structure of interest rates shows the dynamics of the yield curve, linking a functional structure of observations in cross-section (evolution of the rates over maturity) and the evolution of the yield curve over time. As such, the term structure may be represented by a multivariate stochastic process.

Modeling the term structure of interest rate is highly important to the financial market. Between January and July of 2006 the Brazilian market presented about 88 million DI contracts, which encompassed a net asset value of R$ 172 bi. According to data provided by the Special System of Liquidation and Custody between January and July, 2006, the operation with public bonds issued by the Brazilian government corresponded to a movement involving R$ 2.3 trillion and about 310,000 trades of effective purchases and sales of those bonds.

Operations of financial instruments of fixed income are an essential part of the financial institution's portfolio. This is because the financial institutions need to control their exposure to the risk of interest rate's oscillation by balancing their positions on public debt bonds or future contracts that have their value linked to interest rates. In addition, the strategies of portfolio immunization need the forecast of movements of the term structure of interest rates

There is a wealth of literature regarding models of the term structure. To simplify, we can classify this literature by 3 classes of models. The first class encompasses the equilibrium models, such as [Vasicek, 1977], [Brennan & Schwartz, 1979], [Cox *et al.* , 1985] and [Duffie & Kan, 1996]. In these models, the evolution of the term structure of interest rates is given by the specification of a generating process of short-term interest rates, generally in the form of a diffusion process and a discount function that yield the relation between longer maturity rates based on short-term rates. The econometric estimation of the equilibrium models involves estimating the parameters of the diffusion process of the short-term interest rate. An example of equilibrium models estimation lies in the seminal article of [Chan *et al.* , 1992].

The second classification of models of term structure of interest rates is based on the arbitrage-free models, of which [Heath *et al.* , 1992] is a representative model. In this class the fitting of the yield curve is executed so that there are no arbitrage conditions between rates. Nor is there a direct estimation of parameters subjacent to the generator process of interest rates. The observed yield curve is perfectly fit for each day (generally using binomial or trinomial trees); however, there is no dynamic structure in the short-term rates and the problem does not involve parameters estimation. The purpose of these models is not to forecast the yield curve but rather to price financial instruments that use the interest curve observed in the market.

The third relevant literature is the use of statistical models without a structural interpretation, that is, models which synthesize data patterns and to allow for the forecasting of the curve without necessarily representing the theoretical models fit under equilibrium and free-arbitrage conditions. Examples of this kind of model include the methodology of principal components ([Litterman & Scheinkman, 1991]), curve interpolation models such as splines ([Shea, 1984]), smoothing splines ([Shea, 1984]), kernel regression ([Linton. *et al.* , 2001]), and parametric models for curve fitting such as [Nelson & Siegel, 1987] and [Svensson, 1994]. The dynamic extension of the Nelson-Siegel model, presented in [Diebold & Li, 2006] and the basis of the procedure studied in this article, is an example of an statistical model with great success in forecasting the term structure of interest rates.

Despite its basis in theoretical models for interest rates, the structural models based on equilibrium conditions have low forecasting power for the term structure. The calibration models based on no-arbitrage don´t permits a direct forecasting of the yield curve. Statistical models are generally used in the fit and forecasting of term structure of interest rates because of their superior adjustment to equilibrium based econometric models and for their greater simplicity.

The article is structured as follows: section 2 reviews the Diebold-Li ([Diebold & Li, 2006]) model that represents the basic framework of our article; section 3 presents the proposed extensions to Diebold-Li model; the formal representation of the proposed model is shown in section 4; the methodology of estimation is in section 5, and the application of the model for the term-structure of Brazilian SWAP DIxPRÉ operations in on section 6; and the final conclusions are in section 7.

2. Diebold-Li Model

Among the statistical models for interest rate, the influential model designed by Diebold-Li [Diebold & Li, 2006] is widely used in market applications. This model is a dynamic extension of the Nelson-Siegel model ([Nelson & Siegel, 1987]) for the cross-section fit for the yield curve. The Nelson-Siegel model corresponds to fitting the following equation for the yield curve observed in the market on a specific date:

$$y_{it}(m_{it}) = \beta_{1t} + \beta_{2t}\frac{1-e^{-m_{it}/\tau_t}}{m_{it}/\tau_i} + \beta_{3t}\left[\frac{1-e^{-m_{it}/\tau_t}}{m_{it}/\tau_t} - e^{-m_{it}/\tau_t}\right] + \epsilon_{it} \tag{1}$$

where $y_{it}(m_{it})$ are the observed rates on a given date i and maturity t, and $\beta_{1t}, \beta_{2t}, \beta_{3t}$ and τ_t are parameters. The Nelson-Siegel model is a parsimonious way of fitting the yield curve while managing to capture a part of the stylized facts in interest rate process, such as the exponential formats present in the yield curves. The parameters β_{it} have economic interpretations, where β_{1t} presents a long-term level interpretation, β_{2t} short-term components, and β_{3t} medium-term components. It may also be interpreted as decompositions of level, slope and curvature of the yield curve, according to the terminology developed by [Litterman & Scheinkman, 1991]. These components may be used directly in the immunization process of interest rate portfolios. A extension of this model is to use the formulation proposed by [Svensson, 1994] to fit the interest cross-sections. This formulation considers adding an additional term to the formulation proposed by [Nelson & Siegel, 1987], thus corresponding to:

$$y_{it}(m_{it}) = \beta_{1t}+\beta_{2t}\frac{1-e^{-m_{it}/\tau_{1t}}}{m_{it}/\tau_{2t}}+\beta_{3t}\left[\frac{1-e^{-m_{it}/\tau_{1t}}}{m_{it}/\tau_{1t}} - e^{-m_{it}/\tau_{1t}}\right]+\beta_{4t}\left[\frac{1-e^{-m_{it}/\tau_{2t}}}{m_{it}/\tau_{2t}} - e^{-m_{it}/\tau_{2t}}\right]+\epsilon_{it} \tag{2}$$

allowing a more flexible fit for the yield curve and enabling the capture of multiple changes in the yield curve slope. The purpose of these models is to allow fitting, and subsequent interpolations and extrapolations of the yield curve based on a parametric structure, which concurs with other non-parametric fitting models such as smoothing-splines. Besides the parsimonious estimation, the [Nelson & Siegel, 1987] model has two additional advantages over non-parametric models.

The first advantage is that the extrapolation of the curve has a better performance due to the exponential nature of this model. The second advantage is that the parameters β_{1t}, β_{2t} and β_{3t} have interpretation of level, slope and curvature compatible with the interpretation of three factors proposed by [Litterman & Scheinkman, 1991], a benchmark in literature. This makes the interpretation and comparison of the results obtained in the curve fitting easier.

The extension formulated by [Diebold & Li, 2006] renders the [Nelson & Siegel, 1987] model dynamic (adjusting the several days observed for the yield curve) by means of a procedure in 3 stages:

(1) The Nelson-Siegel model (with τ fixed, thus making the model linear in the parameters) is fitted by Ordinary Least Squares for each date, estimating the parameters β_{1t}, β_{2t}, β_{3t}.
(2) The dynamics of the system is modelled by a vector autoregressive (VAR) model for the parameters β_{1t},β_{2t} and β_{3t}, estimated in the first stage.
(3) Forecasts for these parameters are made through the VAR model estimated for vectors β_{1t},β_{2t} and β_{3t}. By substituting the forecasted parameters in Nelson-Siegel model given by equation 2 it is possible to forecast future interest rate curves.

According to [Diebold & Li, 2006], this dynamic formulation has the purpose of capturing the set of existing stylized facts in the term structure of interest rates, such as the fact that while the yield curve is crescent and concave, it may also assume inverted shapes like decreasing curves and slope changes. Other stylized facts captured by [Diebold & Li, 2006] models are the high persistence in the temporal dynamics (rates with same maturity are highly dependent on the past), and the fact that persistence in the long-term rates is higher than in the short-term rates. Although the Diebold-Li model is simple to implement and has a superior predictive potential in comparison with other related models in the literature, some problems still arise when it is used. The three main objections to this model are:

(1) To consider the τ fixed (linearization imposed in the model) may be troublesome for the more unstable yield curves, such as those of emerging countries.
(2) The functional form adapted from the [Nelson & Siegel, 1987] model does not allow for capturing more complicated yield curves, such as when there are multiple changes in the slope and the curvature.
(3) No econometric property of the estimation method has been presented. Consider that it is an two-step estimation, where the VAR is estimated on the basis of a estimated vector of Beta parameters. The main problem is the construction of confidence intervals in finite samples for the forecasts from this model. These intervals should take into account the uncertainty in the estimation of the vectors of hyperparameters β_{1t},β_{2t} and β_{3t}.
(4) Because the forecasted curve may be contaminated by arbitrage situations, there is significant resistance to the use of models that are not based on no-arbitrage conditions.

There are some proposed solutions to these problems. Problem 1 may be addressed by estimating the full [Nelson & Siegel, 1987] models without fixing parameter τ, generally using non-linear least squares. Yet, considering the limited number of observations in the yield curve, the problem of minimizing the non-linear least squares may be complicated, presenting more than one local minimum, a possibility that may lead to an inappropriate fit of the yield curve. This is one of the justifications to keep the parameter τ fixed to avoid the numeric optimization problems involved in the estimation of non-linear models with a restricted number of observations.

The simultaneous estimation of Betas may be performed through the state-space formulation using the Kalman filter, but τ is kept fixed in the sample due to the need for linearity in the use of the linear Kalman filter. Some statistic properties of a model derived from the [Diebold & Li, 2006] formulation were derived in [Huse, 2007], in which a form similar to the Nelson-Siegel model is used with the incorporation of spatial dependences and macroeconomic variables. The estimation is performed on two steps, but some properties of the estimation method in finite samples are studied through the Monte Carlo simulation.

3. Proposed Extensions

To overcome these problems, we propose an extended version of the [Diebold & Li, 2006] model using Bayesian methods. The Bayesian methods based on Markov Chain Monte Carlo are proposed as alternatives to Maximum Likelihood Estimation in a large number of situations where the Maximum Likelihood methods are complicated or unfeasible to apply. Examples of estimation procedures using MCMC include: estimation of continuous-time diffusion processes for term structure of interest rates; option pricing, stochastic volatility and regime switching models, as summarized in [Johannes & Polson, 2007].

The advantages of the Bayesian formulation are that it enables us to treat both parameters and state vectors as latent variables. This is performed through the dynamic linear model formulation for the time evolution of those parameters. In the Bayesian formulation it is not necessary to assume linearity, which is why it is not necessary to fix parameter τ as it is done in the Diebold-Li method. Notice that the construction of posterior distribution of parameters is performed by simulation; hence the various local minima that affect the estimation based on non-linear least squares of equations 1 and 2 do not constitute a problem.

The first Bayesian formulation of the model of [Diebold & Li, 2006], proposed by [Migon & Abanto-Valle, 2007], corresponds to an analogous specification of the original model, using Nelson-Siegel 1 equation with parameter τ kept fixed, but estimated simultaneously with the other parameters of the model. We propose some extensions to the Bayesian formulation of the [Diebold & Li, 2006] model proposed by [Migon & Abanto-Valle, 2007]. The first is to use of the Svensson model (Equation 2), rather than the original Nelson-Siegel formula (Equation 1)), which makes the curve format more flexible. The second extension is to make parameters τ_1 and τ_2 time-varying, adding two latent factors to these components. The third extension is that the formulation of our model allows of different number of observations for each day, which avoids the first stage of curve interpolation in order to obtain a set of observations for the same maturities as it was originally performed in the article by [Diebold & Li, 2006], and which may introduce distortions in the yield curves used in the estimation.

The last extension introduced is to add a stochastic volatility structure to the model. This addition is of fundamental importance since one of the stylized facts in interest rates is the presence of conditional heteroscedasticity, generally captured in no-arbitrage and equilibrium models by the addition of factors which specifically control the stochastic evolution of the variance. Examples of this kind of formulation include the [Hull & White, 1990] and [Scott, 1996] models, and a detailed discussion may be found in [Fouque *et al.* , 2000].

An analysis of the importance of modeling volatility effects is found in [Chan *et al.* , 1992], which studies a wide range of equilibrium models for short-term interest rate. Among the studied models, those assuming constant volatility were statistically rejected. This fact indicates the need to include a stochastic volatility effect in the modeling of the term structure for interest rates.

The advantages of the Bayesian formulation are that the properties of the estimators are obtained in the exact form for finite samples, which allows calculating confidence intervals for the hyperparameters, and forecasting the term structure for interest rates considering the uncertainty in the parameter estimation. The model structure will be described in the next section, and in section 5 we will demonstrate the estimation mechanism of this model using the hybrid Markov Chain Monte Carlo algorithm, using Gibbs algorithm and Metropolis-Hastings algorithm simultaneously.

4. Model Description

We can describe the extensions proposed in this article by the following set of equations:

$$y_{it}(m_{it}) = \beta_{1t}+\beta_{2t}\frac{1-e^{-m_{it}/\tau_{1t}}}{m_{it}/\tau_{2t}}+\beta_3\left[\frac{1-e^{-m_{it}/\tau_{1t}}}{m_{it}/\tau_{1t}}-e^{-m_{it}/\tau_{1t}}\right]+\beta_{4t}\left[\frac{1-e^{-m_{it}/\tau_{2t}}}{m_{it}/\tau_{2t}}-e^{-m_{it}/\tau_{2t}}\right]+e^{\sigma_t\eta_t} \tag{3}$$

$$\begin{bmatrix}\beta_{1t}\\ \beta_{2t}\\ \beta_{3t}\\ \beta_{4t}\\ \tau_{1t}\\ \tau_{2t}\end{bmatrix} = \begin{bmatrix}\mu\beta_1\\ \mu\beta_2\\ \mu\beta_3\\ \mu\beta_4\\ \mu\tau_1\\ \mu\tau_2\end{bmatrix} + \Phi\begin{bmatrix}\beta_{1t-1}\\ \beta_{2t-1}\\ \beta_{3t-1}\\ \beta_{4t-1}\\ \tau_{1t-1}\\ \tau_{2t-1}\end{bmatrix} + \epsilon_t \tag{4}$$

$$ln\sigma_t^2 = \phi_0 + \phi_1 ln\sigma_{t-1}^2 + \upsilon_t \tag{5}$$

$$\eta_t \sim IID(0,1)\ e\ \eta_t \perp \eta_s \veebar\ t \neq s \tag{6}$$

$$\sum_{\eta,\epsilon,\upsilon} = \begin{bmatrix}\sigma_\eta & 0 & 0\\ 0 & \Omega_\epsilon & 0\\ 0 & 0 & \sigma_\upsilon\end{bmatrix}$$

In this specification, which may be seen as a non-linear state-space model, Equation 3 corresponds to a measurement equation, connecting the observed rates y_{it} that describe the interest rate as functions of maturities i at time t. The formulation of this equation follows the specification of Svensson model, but with the addition of latent factors β_{jt} and τ_{ht} , $j = 1, 2, 3, 4$ and $h = 1, 2$ time-varying rather than fixed in time. Matrix $\sum_{\varepsilon,\epsilon,\upsilon}$ denotes the expanded variance-covariance matrix, where σ_η^2 is a scalar, the variance in the measurement equation, Ω_ϵ is the variance-covariance matrix between the latent factors, and σ_υ^2 is a scalar, the variance in the stochastic volatility equation. We assume that the matrix is diagonal, except for the sub-matrix of components Ω_ϵ , which may be correlated.

The evolution of the latent factors is given by Equation 4, which describes an first order autoregressive model for these components with a parameter matrix given by Φ, containing the coefficients of autoregressive estimation. We adopted a specification of first order for the autoregressive model, though noting that there is no theoretical limitation to a superior order. A possibility is to implement a restricted vector autoregressive structure, working with only one autoregressive structure for each parameter. Although this may be imposed a priori, a possible alternative is the use of informative priors in the estimation of vector autoregressive models, as advocated by [Doan *et al.* , 1984].

Finally Equation 5 describes the stochastic volatility components for errors in the measurement equation. The formulation used is that of an autoregressive model for the unobserved stochastic volatility component, according to the original specification of the stochastic volatility model introduced by [Taylor, 1986]. The addition of the stochastic volatility model represents a relevant extension since the presence of conditional heteroscedasticity is a stylized fact in modeling the series of interest rates. An ample class of models is based on the construction of multi-factors models for the term structure for interest rates with specific factors describing the evolution of volatility as in [Hull & White, 1990] and [Lund & Andersen, 1997]. See [Fouque *et al.* , 2000] for a revision of these models under the perspective of derivative pricing. We noted that the addition of stochastic volatility components is especially important at the moments of changes in the shape of the yield curve, especially because these moments are linked to greater uncertainties about future interest rates and expectations about the ways the monetary and fiscal policy will assume. A relevant stylized fact is that the volatility of the interest rates is greater in emerging economies, thus the component of stochastic volatility is especially relevant to the set of data used in this study.

5. Markov Chain Monte Carlo Estimation

Note that the model specification given by the equation system 3,4 and 5 corresponds to a non-linear state-space model, and thus cannot be treated by methods such as the Linear Kalman Filter. A way to perform the simultaneous estimation is through methods of Bayesian Inference using the Markov Chain Monte Carlo.

When we use methods of Bayesian inference, the objective is to find the so called posterior distribution of parameters of interest, conditioned to the observed sample, denoted by $p(\Theta|y)$. To find the distribution of parameters conditioned to the sample, the following relation is used:

$$p(\Theta|y) = p(\Theta, y)/p(y) = p(y|\Theta)p(\Theta)/p(y) \tag{7}$$

where $p(y|\Theta)$ is the likelihood of the model, $p(\Theta)$ denotes the assumed prior distribution of the parameter and $p(y)$ is the marginal distribution of the sample, which needs to be known up to a constant of integration:

$$p(\Theta|y) = p(\Theta, y)/p(\Theta) = p(y|\Theta)p(\Theta)/c \tag{8}$$

that is, the posterior distribution is proportional to the product of likelihood and the prior distribution:

$$p(\Theta|y) \propto p(y|\Theta)p(\Theta) \tag{9}$$

After obtaining the posterior distribution, the results may be summarized by calculating the expected values and the variance of the posterior distribution of each parameter:

$$E(\theta_k|y) = \int \theta_k p(\Theta|y) d\theta \tag{10}$$

$$Var(\theta_k|y) = \int \theta_k^2 p(\Theta|y) d\theta - [E(\theta_k|y)]^2 \tag{11}$$

and the marginal density of parameter θ_j, can be evaluated by:

$$p(\theta_j|y) = \int p(\Theta|y) d\theta_1 d\theta_2 ... d\theta_d \tag{12}$$

Only in some specific cases can the analytical forms for these expressions be obtained. Integration techniques that use Monte Carlo methods in Bayesian Inference, especially the Markov Chain Monte Carlo (MCMC) methodology are useful for this case. The idea of the MCMC method is to simulate a Markov chain whose stationary distribution converges to the distribution $p(\Theta|y)$. MCMC methodology simplifies the calculus, factoring this distribution in a set of conditional distributions of inferior dimensions that can make the simulation easier. The main idea to obtain estimators for the set of parameters $\Theta_1, \Theta_2, ..., \Theta_n$ is to use the set of conditional distributions:

$$\begin{gathered} p(\Theta_1|\Theta_2, \Theta_3, ..., \Theta_n, y) \\ p(\Theta_2|\Theta_1, \Theta_3, ..., \Theta_n, y) \\ \vdots \\ p(\Theta_n|\Theta_2, \Theta_3., ..., \Theta_{n-1}, y) \end{gathered} \tag{13}$$

The Hammersley-Clifford Theorem (see [Robert & Casella, 2004] for a derivation of this result) ascertains that under certain conditions this set of conditional distributions will uniquely characterize the posterior distribution $p(\Theta|y)$, and the MCMC methodology is based on obtaining random samples of the conditional distributions given by 13, where a Markov Chain structure is used. The MCMC methodology may be summed up as a way to evaluate the integral related to the posterior distribution by using simulations of conditional distributions through Markov chains. An evident advantage of this method is that it does not involve any methodology of numerical maximization, thus avoiding the numerical problems involved in the non-linear maximization of functions such as those found in our problem. The validity of the methodology is verified through methods that check the convergence of the Markov chains for its stationary distribution.

The methodology of Bayes Hierarchical estimators is a convenient way to address the problem when the model to be estimated can be placed in a state space formulation. Following the example given in [Lehmann & Casella, 1998] a form to represent these models is:

$$X|\theta \sim f(x|\theta)$$
$$\Theta|\gamma \sim \pi(\theta|\gamma)$$
$$\Gamma \sim \psi(\gamma)$$

thus we place a hierarchy structure among the priors distributions. This formulation is especially useful in state space models since the hierarchical specification allows estimating the hyperparameters related to the latent factors using the disponible data, specifying the dynamics for the latent factors. For example the local level model is formulated as:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \mu_t &= \mu_{t-1} + \nu_t \end{aligned} \tag{14}$$

where we use as a priori distribution of the latent factor μ_t the value of μ_{t-1} and then $\mu_t \sim p(\mu_{t-1})$, which corresponds to the idea of state equation in the state space formulation. The specification of the latent factors uses a generalized formulation $\xi_t \sim p(\xi_{t-1})$ 4, where ξ denotes the set of latent factors in our model given by β_{it}, τ_{it} and σ_i^2. This methodology is also known as empirical Bayes estimators ([Lehmann & Casella, 1998]).

Given the complexity of conditional distributions it is not possible to take direct samples from all conditional distributions involved. A simple form of MCMC algorithm is the Gibbs sampler, where the estimation is done by sampling the conditional distributions directly. A limitation is that all the conditional distributions need to be analytically known. When it not be possible to sample the analytical conditional distribution, an idea is to use the Metropolis-Hastings algorithm, which may be seen as a generalization of the method of acceptance-rejection simulation of random variables for the samples of conditional distributions.

In our problem we cannot directly sample all the conditional distributions given the non-linear forms involved. Thus a Hybrid Markov Chain Monte Carlo will be used, when we simultaneous use of the Gibbs algorithm and the Metropolis- Hastings algorithm, a methodology initially proposed in [Tierney, 1994]. A hybrid MCMC algorithm ([Robert & Casella, 2004]) may be seen through as iterations in the following stages: θ

`For i=1,...,p ,and given` $(\theta_1^{(t+1)}, ..., \theta_{i-1}^{(t+1)}, \theta_i^{(t)}, ...\theta_p^{(t)})$

`1 - Simulate`

$$\widetilde{\theta}_i \sim q_i(\theta|\theta_1^{(t+1)}, ..., \theta_{i-1}^{(t+1)}, \theta_i^{(t)}, ...\theta_p^{(t)})$$

`2 - Accept`

$$\theta_i^{(t+1)}, = \begin{cases} \theta_i^{(t)} & with\, probability\, 1-\rho \\ \widetilde{\theta}_i & with\, probability\, \rho \end{cases}$$

`where`

$$\rho \bigwedge \begin{cases} \left(\frac{g_{ii}(\widetilde{\theta}_i|\theta_1^{(t+1)},...,\theta_{i-1}^{(t+1)},,\theta_i^{(t)},...\theta_p^{(t)})}{q_{ii}(\widetilde{\theta}_i|\theta_1^{(t+1)},...,\theta_{i-1}^{(t+1)},\theta_i^{(t)},\theta_i^{(t)},...\theta_p^{(t)})} \right) \\ \left(\frac{g_{ii}(\theta_i^{(t)}|\theta_1^{(t+1)},...,\theta_{i-1}^{(t+1)},,\theta_i^{(t)},...\theta_p^{(t)})}{q_{ii}(\theta_i^{(t)}|\theta_1^{(t+1)},...,\theta_{i-1}^{(t+1)},\theta_i^{(t)},\theta_i^{(t)},...\theta_p^{(t)})} \right) \end{cases}$$

where q is the so called tentative distribution (we assume a normal multivariate as tentative distribution) and g is the conditional distribution. A simpler form is to use the Metropolized Gibbs sampler, where the previous algorithm is simplified to:

`Given` $\theta^{(t)}$

`1- Simulate` $z_i \neq \theta^{(t)}$ `with probability`

$$\frac{g_i(z_i|\theta_j^{(t)}, j \neq i)}{1 - g_i(\theta_i^{(t)}|\theta_j^{(t)}, j \neq i)}$$

`2 - Accept` $\theta^{(t+1)} = z_i$ `with probability`

$$\frac{1 - g_i(\theta_i^{(t)}|\theta_j^{(t)}, j \neq i)}{1 - g_i(z_i|\theta_j^{(t)}, j \neq i)} \bigwedge 1$$

Iterating steps 1 and 2 up to convergence, we obtain our estimators for the parameters and hyperparameters related to the unobserved latent factors in our model.

To completely characterize our model, the prior distributions are the normal-gamma pair inverse for β_{it} and τ_{it}, using the hierarchical characterization with the mean given by the vector autoregressive structure. For the parameters Φ of the autoregressive vector, we assume a normal multivariate structure with variance matrix given by a Wishart distribution; for the latent factor of stochastic volatility we assume $\sigma_t^2 \sim LogNormal(\phi_0 + \phi_{1_t}\sigma_{t-1}^2, \tau_{\sigma^2})$, with a gamma distribution for τ_{σ^2}, normal for ϕ_0 and finally $\phi_1 \sim Beta$.

For the parameters β_{it}, ϕ_0, parameters of Wishart distribution and the gamma distributions, we use a Gibbs sampling step; for τ_{it} we use Metropolis-Hastings; and for parameter ϕ_1 we use the algorithm known as Slice Sampler [Neal, 2003]).

6. Application

We show an application of this model for the fitting of term structure implicit in the Swap DI-PRÉ curves provided by the BM&F. This yield curve is notoriously difficult to adjust by conventional methods. We use BM&F data on yield curves implicit in the SWAP operations for time interval from 01/12/2004 to 12/12/2006, a sample of 722 yield-curve days. Figure 1 shows the evolution of the yield curves over time.

The interesting fact is that the curves in our study presents several slope and curvature changes, going from the usual crescent shape to inverted curve several times throughout the period. In the same interval, it also presents several days where the yield curves have two slope changes. This fact cannot be adequately captured by the model of [Diebold & Li, 2006], because the [Nelson & Siegel, 1987] formulation does not allow more than one slope and curvature change. Another point of importance is that the yield curve in Brazil has intense oscillation, both in terms of curve level and format, which reinforces the necessity to make the parameters time-varying, and challenges the maintenance of parameter τ fixed as assumed by the [Diebold & Li, 2006] model.

Another important point is that the yield curve lengthens and retracts in the mentioned period, that is, the maximal maturities observed in the Swap contracts alter in the analyzed sample, varying between 1800 and 2400 days. Note that our study does not carry out a pre-interpolation and extrapolation on the data; the methodology permits to work with distinct maturities in each day. This fact must be highlighted because the interpolation stage may distort the data, and the estimated model may be used to interpolate and extrapolate the curve if necessary.

To estimate the model, we used 10,000 iterations of the MCMC algorithm described in section 5, discarding the first 5,000 iterations (Burn-In period) and using the other 5,000 in the construction

Figure 1. Swap Di-PRÉ Term Structure(12/01/2004 - 05/12/2006)

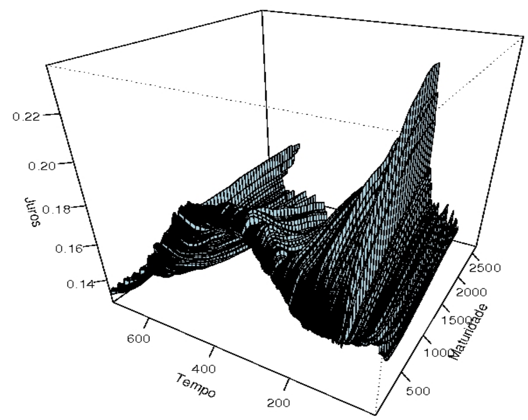


Figure 2. Adjusted Curve and Residuals

of posterior distributions. The Gelman-Rubin convergence diagnoses indicate that the Markov chains converge to the stationary distributions, thus validating the estimation methodology used.

Figure 2 shows the model fit inside the sample and the residuals in relation to the observed curves. It is possible to see that the model is able to reproduce the variations observed in the term structure of the interest rates, and that the relative magnitude of the residues is very low, not existing any systematic pattern in the model residuals. Note that with the estimated parameters, it is possible to do any interpolation and extrapolation of the yield curves, since we just replace the value for maturity to be interpolated in equation 3.

Figures 3, 4, 5 and 6 show the evolution of the latent factors $\beta_{1t}\beta_{2t}, \beta_{3t},\ \beta_{4t}$ obtained as medians of posterior distributions. The evolution of β_{1t} clearly shows the level interpretation of this parameter, following the evolution of the mean yield curve over time. The evolution of the other hyperparameters also captures in adequate form the evolution of the slope and curvature components of the term structure observed in the interest rates.

FIGURE 3. Beta 1

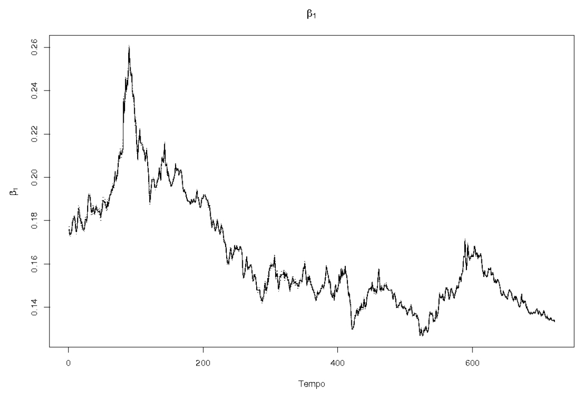


FIGURE 4. Beta 2

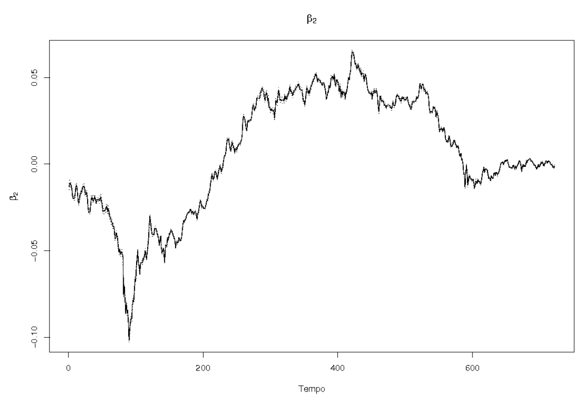
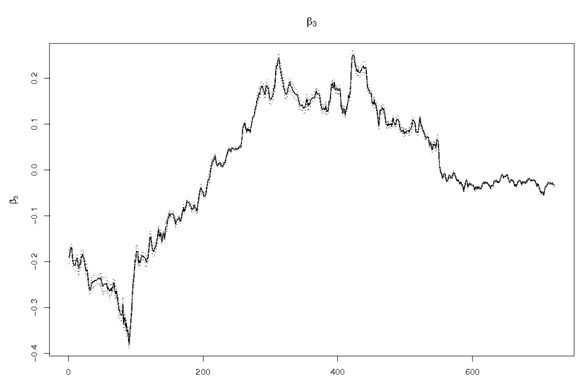


FIGURE 5. Beta 3



Figures 7 and 8 are of special importance because they show that the a priori fixing of parameter τ assumed in [Diebold & Li, 2006] model is not a valid restriction as it becomes evident by the great temporal variation observed in parameters τ_1 and τ_2. This gives evidence to the necessity

FIGURE 6. Beta 4

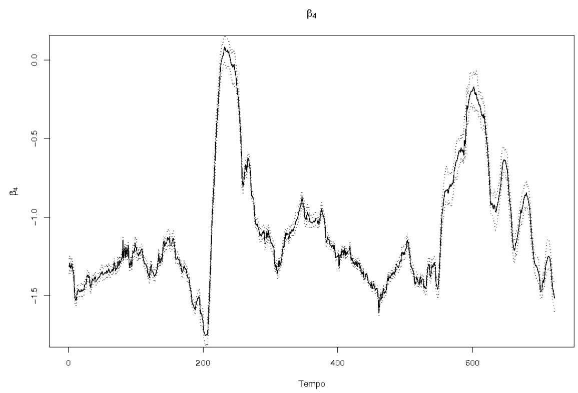


FIGURE 7. Tau1

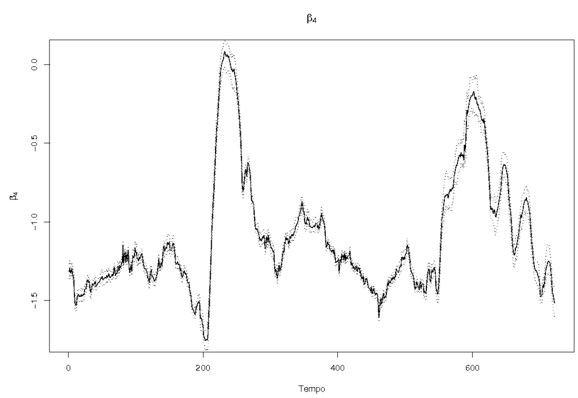
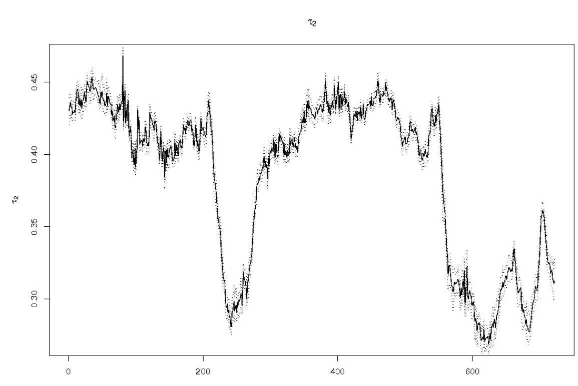


FIGURE 8. Tau 2



to incorporate variation in these parameters for the yield curves with great variation of format as observed in emerging countries.

FIGURE 9. Stochastic Volatility

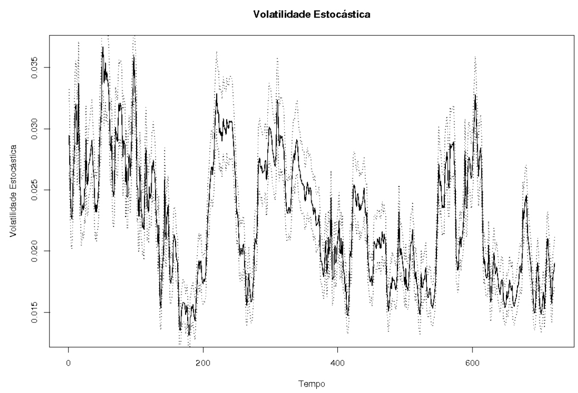


TABLE 1. Credibility Intervals 95% - Φ

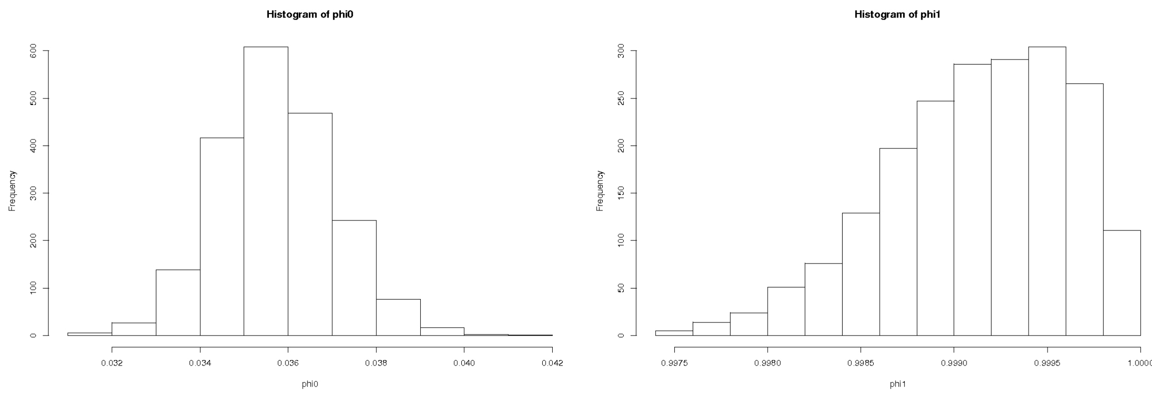
	μ	β_{1t-1}	β_{2t-1}	β_{3t-1}	β_{4t-1}	β_{5t-1}	β_{6t-1}
$\Phi_{\beta 1}$ (.025)	.2310	1.262	.5203	-.1122	-.4625	-.0152	-.1183
$\Phi_{\beta 1}$(.50)	.2407	1.281	.5504	-.1077	-.4432	-.0136	-.1115
$\Phi_{\beta 1}$ (.975)	.2532	1.293	.5721	-.1033	-.4266	-.0139	-.0999
$\Phi_{\beta 2}$ (.025)	-.2582	-.2877	.4215	.1051	.4337	.0135	.1005
$\Phi_{\beta 2}$(.50)	-.2453	-.2755	.4437	.1102	.4490	.0141	.1126
$\Phi_{\beta 2}$ (.975)	-.2316	-.2572	.4746	.1149	.4702	.0154	.1189
$\Phi_{\beta 3}$ (.025)	-.4741	-.4378	-.9778	1.141	.5951	.0183	.1253
$\Phi_{\beta 3}$(.50)	-.4020	-.3385	-.8204	1.171	.7172	.0224	.1619
$\Phi_{\beta 3}$ (.975)	-.3331	-.2330	-.6642	1.201	.8445	.0270	.1998
$\Phi_{\beta 4}$ (.025)	.1570	.1605	.3217	-.0822	.1397	-.0115	-.0896
$\Phi_{\beta 4}$(.50)	.1764	.1918	.3705	-.0728	.1782	-.0100	-.0783
$\Phi_{\beta 4}$ (.975)	.1984	.2205	.4234	-.0639	.2124	-.0087	-.0659
$\Phi_{\beta 5}$(.025)	.3965	.4719	.9298	-.5165	-2.2090	.9181	-.6076
$\Phi_{\beta 5}$(.50)	.7881	1.1180	1.8360	-.3394	-1.5045	.9456	-.3833
$\Phi_{\beta 5}$(.975)	1.1680	1.8060	2.787	-.1688	-.7929	.9707	-.1819
$\Phi_{\beta 6}$ (.025)	.2788	.3752	.6515	-.1911	-.8212	-.0282	.7723
$\Phi_{\beta 6}$(.50)	.3651	.5022	.8621	-.1567	-.6838	-.0232	.8116
$\Phi_{\beta 6}$ (.975)	.4360	.6283	1.0570	-.1191	-.5263	-.0173	.8538

The estimated Stochastic Volatility component (Figure 9) shows the capacity of the model to capture the stylized fact of the presence of conditional heteroscedasticity in interest rates. The structure of conditional volatility captures the uncertainty existing in the periods of change in the yield curves shapes, since we can notice the correlation between increase in volatility and periods of inversion in the curve format.

Figure 10 shows another fact captured by the Stochastic Volatility structure - the high persistence of shocks in the volatility. This is noticeable because parameter ϕ_1 is concentrated on values close to 1. The high persistence of shocks in the volatility of interest rates and financial series is in one of the most important stylized facts in financial econometrics.

Table 1 shows the credibility intervals calculated for the matrix of coefficients Φ. To verify the stationarity of the process, we calculated the eigenvalues of matrix Φ for the upper and lower limits of this matrix. The highest eigenvalues for the higher limit is 1.0029, and for the lower limit is

Figure 10. Posterior Distribution - Phi0 and Phi1



.9783, indicating that the region of nonstationarity is included in the credibility intervals. Although this is not a problem in Bayesian inference, this result attests the high-persistence observed in the level of interest rates in Brazil.

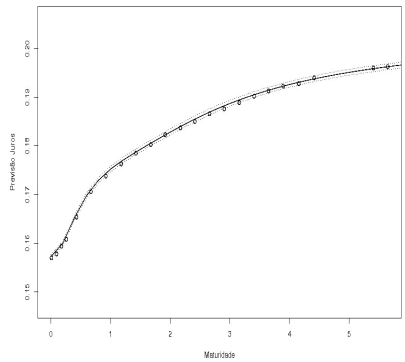
To demonstrate the predictive potential of the model, we show the forecasts for some specific days, characterized by distinct shapes of the yield curve. We also show the forecasts and one-step-ahead forecast errors for all the days observed in the sample.

Figure 11 shows one-step-ahead forecasts obtained by the extended Diebold-Li model, with confidence intervals at the 2.5% and 97.5% limits, for 4 days observed in the yield curve. The first sub-figure shows the prediction for July 20, 2004, with the format generally observed in the interest rates, with a positive trend in maturity. The second curve, predicted for Feb. 01, 2005, shows a curve with slope change, normally associated with expected changes in the long-term interest rates. The curve predicted for June 27, 2006 shows an opposite situation, with a decreasing curve at medium-term maturities, and increasing at long-term maturities. The sub-figure d shows a one-step-ahead forecast for the last observation in the sample, the observation referring to Dec 06, 2006.

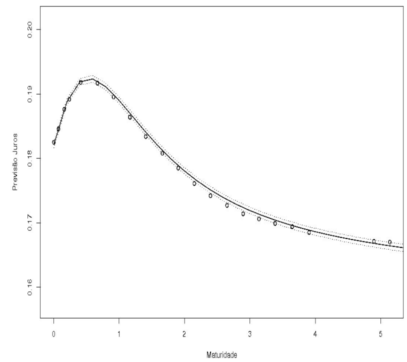
The one-step-ahead forecasts for all sample, along with the inclusion of extrapolations for unobserved maturities and associated prediction errors, are shown in Figure 12. The forecast errors have relatively low magnitude. It is noticeable that the bigger errors are concentrated at the moments of change in the shape of the yield curve.

We also carried out a comparative forecast analysis between the extended Diebold-Li model and the original formulation of Diebold-Li with τ fixed, the time-varying specification of the Diebold-Li model, and a modification in the Diebold-Li model using Svensson Equation 2 to replace the original formulation based on the Nelson-Siegel specification, with parameters τ_1 and τ_2 fixed and time-varying. The estimation of these reference models for time-varying τ is based on non-linear least squares while the linearized forms are based on the estimation by ordinary least squares. Table 2 presents the Root Mean Square Error using the one-step-ahead forecast errors for the five compared models. Parameters $\tau, \tau_1\, and\, \tau_2$ are fixed by the mean value of the corresponding time-varying parameters.

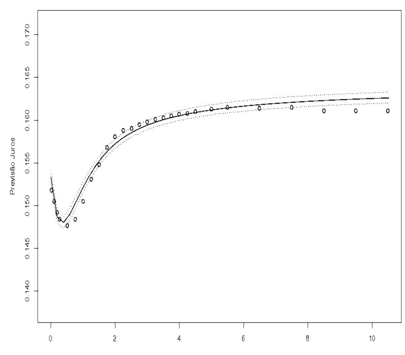
FIGURE 11. One Step Ahead Forecasts and Forecast Errors - Specific Days



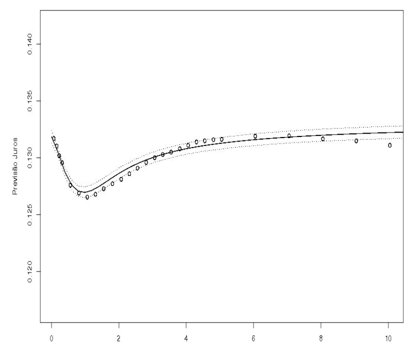
(a) 20/07/2004



(b) 01-02-2005

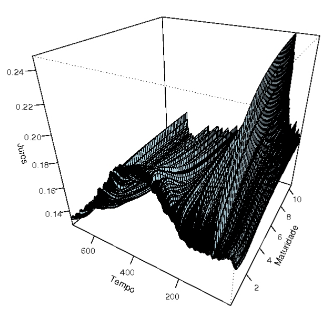


(c) 27/06/2006

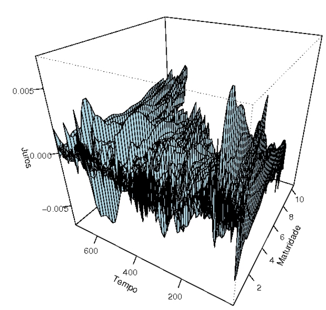


(d) 06/12/2006

FIGURE 12. One step Ahead Forecasts and Forecast Errors



(a) One step Ahead Forecasts



(b) Forecast Errors

The results of this comparative analysis show that the Diebold-Li model with the proposed extensions has a superior forecast performance when compared with the other models, as it is shown in Table 2. The original Diebold-Li model with parameter τ fixed, using the Nelson-Siegel

Table 2. Root Mean Squared Forecast Error

Model	Root Mean Squared Error
Extended Diebold-Li	1.1987
Original Diebold-Li τ fixed	36.61
Original Diebold-Li τ varying	23.92
Diebold-Li-Svensson τ_1 and τ_2 fixed	17.05
Diebold-Li-Svensson τ_1 and τ_2 varying	204.6

specification, is not a valid specification because it substantially reduces the predictive power of the model when compared with the varying-parameter version of the same model. In the case of the Diebold-Li model using Svensson specification, fixed parameters result in better predictive power than the estimation with free parameters. This result may be explained by the difficulty in estimating Svensson specification, since on many days the estimation does not converge due to non-linearity, which makes the model fitting inadequate, thus raising the mean quadratic error value with the presence of high forecasting errors for all the maturities observed on those days. This problem also contaminates the autoregressive vector estimation, which compromises the curve forecasting for the following day. The use of Bayesian estimation with informative priors allows us to use the more flexible Svensson specification, but without being affected by the instability problems in the non-linear estimation that occur in the classical estimation that use non-linear least squares.

7. Conclusions

In this article we implemented some extensions for the [Diebold & Li, 2006] model, including the use of Bayesian estimation methods using MCMC for the parameters and latent factors of this model. The proposed extensions for the model were changes in the functional form, making it more flexible with the use of the [Svensson, 1994] specification, including latent factors to make the model parameters time-varying, and enabling the use of different observations on each day and the inclusion of a stochastic volatility structure. The more flexible form adopted allows capturing changes in the shapes associated with the yield curve in emerging countries. This flexibility is reflected in the low forecasting and fitting errors observed in this model. The use of the Bayesian estimation associated with informative priors in the latent factors specification avoids the procedure of linearized estimation in two stages used in the Diebold-Li model, which results in more precise fits and forecasting. The parameter specification as modeled latent factors through a Bayesian hierarchical structure allows obtaining the distribution in finite samples of parameters and model predictions, thus enabling the quantification of the uncertainty present in the estimation of the term structure of interest rates.

We showed that making the τ parameters time-varying is important to fit the term structure, in special to yield curves in emerging countries with constant modifications in the shape of the yield curve, as observed by the behavior of the latent factors τ_{1t} and τ_{2t}. The latent factors, due to their informative priors structure, allow us to overcome the common problem of numerical instability associated with the non-linear estimation of the Nelson-Siegel and Svensson models in the presence of a restricted number of observations.

The proposed methodology does not need pre-interpolation at the yield curve, which avoids the distortions that may be introduced in this process. The stochastic volatility component has two objectives: the first is to capture one of the most important stylized facts in the interest rates, the presence of conditional heteroscedasticity and the high persistence of shocks in volatility; the second is to allow the quantification of the uncertainty associated with the yield curve on each day - which proves to be of special importance at the moments of changes in the shape of interest rate curve.

The Bayesian estimation methodology through the Markov Chain Monte Carlo algorithms does the estimation simultaneously and allows us to avoid the linearization of the model and the estimation in two stages used in [Diebold & Li, 2006]. The specification of the model is based on a standard set of priors, and the estimation algorithm, based on a mixture of Gibbs and Metropolis-Hastings, is widely used and its properties extensively studied, making the estimation of the model simple and trustworthy.

References

[Brennan & Schwartz, 1979] Brennan, M. J., & Schwartz, E. J. 1979. A Continuos Time Approach to the Pricing of Bonds. *Journal of Banking and Finance*, **3**, 133–155.

[Chan *et al.* , 1992] Chan, K. G., Karolyi, G.A., Longstaff, F.A., & Sanders, A. B. 1992. An Empirical Comparasion of Alternative Models of Term Structure of Interest Rates. *Journal of Finance*, **47**, 1209–1227.

[Cox *et al.* , 1985] Cox, J. C., Ingersoll, J . E., & Ross, S. A. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica*, **53**, 385–408.

[Diebold & Li, 2006] Diebold, F.X., & Li, C. 2006. Forecasting the Term Structure of Government Bond Yields. *Journal Of Econometrics*, **130**, 337–364.

[Doan *et al.* , 1984] Doan, T., Litterman, R., & Sims, C. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, **3**, 1–100.

[Duffie & Kan, 1996] Duffie, D., & Kan, R. 1996. A yield-factor model of interest rates. *Mathematical Finance*, 6, 379–406.

[Fouque *et al.* , 2000] Fouque, J-P., Papanicolaou, G., & Sircar, K. R. 2000. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press.

[Heath *et al.* , 1992] Heath, David, Jarrow, Robert, & Morton, Andrew. 1992. Bond Pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, **60**(1).

[Hull & White, 1990] Hull, John, & White, Alan. 1990. Pricing interest rate derivative securities. *Review of Financial Studies*, **3**(4), 573–94.

[Huse, 2007] Huse, C. 2007. *Term Structure Modelling with Observable State Variables*. Unpublished Working Paper - FMG - LSE.

[Johannes & Polson, 2007] Johannes, M., & Polson, N. 2007. *Handbook of Financial Econometrics*. Chap. MCMC Methods for Continuos Time Financial Econometrics.

[Lehmann & Casella, 1998] Lehmann, E., & Casella, G. 1998. *Theory of Point Estimation (2nd Edition)*. Springer.

[Linton. *et al.* , 2001] Linton., O., Mammen, E., Nielsen, J., & Tanggard, C. 2001. Estimating Yield Curves By Kernel Smoothing Methods. *Journal Of Econometrics*, **105:1**, 185–223.

[Litterman & Scheinkman, 1991] Litterman, R., & Scheinkman, J. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income*, **1**, 54–61.

[Lund & Andersen, 1997] Lund, J., & Andersen, T. 1997. Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate. *Journal Of Econometrics*, **77**, 343–377.

[Migon & Abanto-Valle, 2007] Migon, H., & Abanto-Valle, C. 2007. A Bayesian Term Structure Modelling. *Pages 200–203 of:* Fernandes, C., Schimidli, H., & Kolev, N. (eds), *Proceedings of the Third Brazilian Conference on Statistical Modelling in Insurance and Finance*. IME-USP.

[Neal, 2003] Neal, R. 2003. Slice Sampling (with discussions). *Annals of Statistics*, **31**, 705–767.

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

[Robert & Casella, 2004] Robert, C., & Casella, G. 2004. *Monte Carlo Statistical Methods*. Springer.

[Scott, 1996] Scott, L. O. 1996. Simulating a multi-factor term structure model over relatively long discrete time periods. *In: Procedings of the IAFE First Annual Computacional Finance Conference*.

[Shea, 1984] Shea, G. 1984. Pitfalls in Smoothing Interest Rate Structure Data: Equilibrium Models and Spline Approximation. *Journal of Financial and Quantitative Analysis*, **19**, 253–269.

[Svensson, 1994] Svensson, L. E. O. 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. *NBER Working Paper*.

[Taylor, 1986] Taylor, S. J. 1986. *Modelling Financial Time Series*. John Wiley& Sons.

[Tierney, 1994] Tierney, L. 1994. Markov Chains for Exploring Posterior Distributions (with Discussion). *Annals of Statistics*, **22**, 1701–1786.

[Vasicek, 1977] Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics*, **5**, 177–88.