Efficient Yield Curve Estimation and Forecasting in Brazil

João F. Caldeira
Department of Economics, Federal University of Rio Grande do Sul and Banco SICREDI, Brazil

Guilherme V. Moura
Department of Econometrics, Vrije Universiteit Amsterdam, Netherlands

Marcelo S. Portugal
Department of Economics, Federal University of Rio Grande do Sul and CNPq, Brazil

Abstract
Modeling the term structure of interest rate is very important to macroeconomists and financial market practitioners in general. In this paper, we used the Diebold-Li interpretation to the Nelson Siegel model in order to fit and forecast the Brazilian yield curve. The data consisted of daily observations of the most liquid future ID yields traded in the BM&F from January 2006 to February 2009. Differently from the literature on the Brazilian yield curve, where the Diebold-Li model is estimated through the two-step method, the model herein is put in the state-space form, and the parameters are simultaneously and efficiently estimated using the Kalman filter. The results obtained for the fit and for the forecast showed that the Kalman filter is the most suitable method for the estimation of the model, generating better forecast for all maturities when we consider the forecasting horizons of one and three months.

Keywords: Term Structure of the Interest Rate, Yield Curve, State-Space Model, Kalman Filter

JEL Classification: C53, E43, G17

Resumo
Modelar a estrutura a termo da taxa de juros é extremamente importante para macroeconomistas e participantes do mercado financeiro em geral. Neste artigo é empregada a formulação de Diebold-Li para ajustar e fazer previsões da estrutura a termo da taxa de juros brasileira. São empregados dados diários referentes às taxas dos contratos de DI Futuro negociados na BM&F que apresentaram maior liquidez para o período de Janeiro de 2006 a Fevereiro de 2009. Diferentemente da maior parte da literatura sobre curva de juros para dados brasileiros, em que o modelo de Diebold-Li é estimado pelo método de dois passos, neste trabalho o modelo é colocado no formado de estado espaço, e

os parâmetros são estimados simultaneamente, de forma eficiente, pelo Filtro de Kalman. Os resultados obtidos tanto para o ajuste, mas principalmente no que diz respeito à previsão, mostram que a estimação do modelo através do Filtro de Kalman é a mais adequada, gerando melhores previsões para todas as maturidades quando é considerado horizontes de previsão de um mês, três meses e seis meses.

1. Introduction

Understanding the behavior of the term structure of interest rate is important to macroeconomists, financial economists and fixed income managers, and such understanding has prompted remarkable improvement in theoretical modeling and in the estimation of this type of process in the past few decades. The major models developed during this period can be classified as follows: no-arbitrage models; equilibrium models; and statistical or parametric models. No-arbitrage models focus on the perfect adjustment of the term structure in a given time period, warranting that arbitrage possibilities will not occur, which is important for derivatives pricing. Examples of these models include Hull and White (1990), and Heath et al. (1992). Equilibrium models place emphasis on the modeling of the instantaneous yields, typically through affine models; then the yields of other maturities can be derived under several hypotheses about the risk premium. Models of this type were developed by Vasicek (1977), Cox et al. (1985) and by Duffie and Kan (1996).

Statistical or parametric models consist of principal component models, factor models or latent variables, and also interpolation models. According to Matzner-Lober and Villa (2004), most of the intuition about the dynamics of bond and bonus profitability arises from models belonging to this class, as in Robert and Scheinkman (1991) and in Pearson and Sun (1994). Among factor models, the model developed by Nelson and Siegel (1987) and its variants, are the most popular amidst fixed income managers and central banks. The attractiveness of factor models of the Nelson Siegel type is due to its parsimony and good empirical performance. Models of this type can capture most of the behavior of the term structure of interest rate by means of only three factors. Models with a larger number of factors were used by Svensson (1994), Almeida and Vicente (2008), Laurini and Hotta (2008), among others.

Interpolation models were developed, for instance, by McCulloch (1971, 1975), who interpolated the discount function rather than the yields or the asset prices in a direct manner; and by Vasicek and Fong (1982), who adjusted exponential splines to the discount curve, obtaining smoother adjustments for the longest section of the curve.

⋆ Submitted in November 2009, accepted in December 2009.
E-mail address: joao@ppge.ufrgs.br

Diebold and Li (2006) argue that, despite major improvements in theoretical modeling of the term structure of interest rate, little attention has been paid to the forecast of the term structure. No-arbitrage models place emphasis on fit to a given time period and say too little about out-of-sample dynamics or forecast. Conversely, equilibrium models have some dynamic implications in view of a certain risk premium, which allows drawing some conclusions about out-of-sample forecasts. However, according to Diebold and Li, most studies on equilibrium models focused on in-sample performance. Exceptions include Duffee (2002), who demonstrates that arbitrage-free models exhibit poor performance in out-of-sample forecasts; and Egorov et al. (2010), who show that affine models with stochastic volatility can predict the conditional joint distribution of bonus profitability.

Having good yields forecasts is essential to calculate the market value of an asset portfolio, to assess fixed income derivatives, to build investment strategies and to develop monetary policies.

Following a different line of research, Ang and Piazzesi (2003), Hördahl et al. (2006), and Wu (2001) analyzed models with macroeconomic variables and showed that these variables contribute towards improving the forecast of the yield curve dynamics. Diebold et al. (2006) (hereinafter referred to as DRA), used a model with latent factors for the yield curve and also included macroeconomic variables. Unlike previous models which considered a unidirectional relationship of macroeconomic variables towards the yield curve, or of the yield curve towards the macroeconomic variables, DRA assessed the possibility of a bidirectional relationship and observed that the inclusion of macroeconomic variables improved the predictability of the model, mainly for six-month and one-year-ahead forecasting horizons, for the medium-term maturities of the yield curve analyzed.

Vicente and Tabak (2007) compared the Gaussian affine model with Diebold and Li model for Brazilian data and concluded that the latter model is slightly superior in terms of yield curve forecasts. Almeida et al. (2007) obtained better forecasting results than those from Diebold and Li model using a dynamic version of Svensson's model. Vargas (2007) uses Brazilian data for future ID contracts to replicate the results obtained by Diebold and Li. Laurini and Hotta (2008), on the other hand, estimate an extended Svensson's model, where decay parameters vary over time and stochastic volatility is added to the measurement equation of the state-space system.

In this paper, we use the three-factor model for the term structure as proposed by Nelson and Siegel (1987), but we reinterpret the factors as level, slope and curvature of the yield curve just as in Diebold and Li, in order to make out-of-sample forecasts. To estimate the models and perform the forecasting exercise, we use the state-space approach introduced in this context by DRA, which allows simultaneously to fit the yield curve in each time period and estimating the dynamics of the underlying factors using maximum likelihood. This procedure obviates the a priori selection of the decay parameter and permits obtaining optimal smoothed estimates of the factors. The database consists of daily spot rate series of future ID contracts traded in the BM&F, precluding the use of swap rates, which often do not represent actual trading rates.

The contribution of the paper is twofold: we estimate Diebold and Li yield curve model efficiently in a single step by means of the Kalman filter, avoiding the *a priori* selection of the decay parameter λ; and we use a new data set based on future ID rates for maturities with higher liquidity than the often used data on swap rates. Besides this Introduction, the paper is organized as follows. Section 2 introduces the structure of Diebold-Li model for the yield curve and its state-space form. Section 3 presents the data used in the estimation, and in addition to analyzing the adjustment of the model, it performs a forecasting exercise to verify whether the model can produce good out-of-sample forecasts. Section 4 concludes and suggests avenues for further investigation.

2. The Factor Model of the Term Structure

This section introduces the factor model for the term structure of interest rates. The version herein follows the three-factor model devised by Diebold and Li (2006), and represents a reinterpretation of the yield curve that appears in Nelson and Siegel (1987), where the three factors are interpreted as level, slope and curvature factors.

2.1. *Discount function, forward curve and yield curve*

Before describing the structure of the model, it is necessary to define discount curve, forward curve and yield curve, as well as their interrelations. The term structure of interest rates is represented by a set spot rates for different maturities. Each point corresponds to an yield $y_i(\tau)$ associated with maturity τ, obtained from a security traded on the market (Varga, 2008).

At any point of time t, there will be a collection of zero-coupon bonds that differ only in terms of maturity. However, in a given moment, there may not be a bond available to all desired maturities as bonds are not negotiated for all possible maturities.

One of the most basic constructions describing the term structure of the interest rate, from which other curves are often derived, is the discount function. Let $P_t(\tau)$ be the price of a zero-coupon bond at time t, which pays \$1 at maturity τ. Supposedly, every zero-coupon bond is default-free and has strictly positive prices. Thus, the discount function is defined by:

$$P_t(\tau) = e^{-\tau y_t(\tau)} \tag{1}$$

The yield $y_t(\tau)$ at which the bond is discounted is the internal rate of return of the zero-coupon bond, at time t, and with maturity τ, expressed as:

$$y_t(\tau) = \frac{-\ln(P_t(\tau))}{\tau} \tag{2}$$

The forward rate at time t applied to the time interval between τ_1 and τ_2 , is defined by:

$$f_t(\tau_1, \tau_2) = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} y_t(x)dx \tag{3}$$

The same argument applies to forward rates for k-periods. The forward rate can be interpreted as the marginal rate of return necessary to maintain a bond for an additional period. The limit of expression (3) when τ_1 draws closer to τ_2, denoted by $f_t(\tau)$, is the instantaneous forward rate:

$$f_t(\tau) = \frac{-P_t(\tau)}{P_t(\tau)} \tag{4}$$

The instantaneous forward rate curve, $f_t(\tau)$, provides the decay rate of discount function $P_t(\tau)$ in each point τ. The yield curve $y_t(\tau)$ is the average decay rate for the interval between 0 and τ, expressed by:

$$y_t(\tau) = \frac{1}{\tau} \int_0^{\tau} f_t(x)dx \tag{5}$$

The function $f_t(\tau)$ of forward rates describes the (instantaneous) rate of return of an investment that is maintained for a very short time interval. The instantaneous forward rate curve is a very important theoretical construct, even though its value for a single maturity is of little practical interest, due to the high transaction cost associated with a contract between two points in the future if these two points are too close to one another. Only the mean of $f_t(\tau)$ for a future time interval is of practical interest.

In any point at time t, there will be a set of bonds with different maturities, τ, and different payment flows, which may be used to estimate the yield curves, discount curves and forward curves, which are not observable in practice. There are different approaches to the construction of yield curves. McCulloch (1971, 1975) and Vasicek and Fong (1982) build yield curves using estimated smooth discount curves and converting them into rates at relevant maturities. The method put forward by McCulloch (1971, 1975) employs a cubic spline discount function interpolation. The advantage of this method is that the estimation model only has linear parameters, its disadvantage is the resulting erratic curves for longer maturities, i.e., the adjusted discount curve diverges for longer maturities instead of converging to zero. Vasicek and Fong's approach suggests the use of exponential splines to adjust the discount function, which would eliminate the divergence problem for longer maturities.

Statistical models were also used to estimate the term structure of interest rate as in Nelson and Siegel (1987), Svensson (1994), among others. These models proved quite useful in the analysis and pricing of fixed income securities, and special attention should be paid to the work carried out by Nelson and Siegel (1987), which was given a new interpretation in Diebold and Li (2006), wherein short-, medium- and long-term factors began to be interpreted as slope, curvature and level factors.

Fama and Bliss (1987) proposed a method for the construction of the term structure using forward rates estimated for the observed maturities. The method consists in sequentially building the forward rates necessary to successively price bonds with longer maturities, known as the unsmoothed forward rates proposed by Fama and Bliss. The yield curve resulting from this procedure is a (discontinuous) function with jumps relative to the maturity of the bond being traded.

2.2. *Diebold and Li yield curve model*

The classic problem with the term structure requires the estimation of a smooth yield curve based on bond prices observed. In recent years, the method has consisted in computing the implicit forward rates in order to successively price bonds with longer maturities in the observed maturities, known as unsmoothed forward rates. Then a smooth forward rate curve is obtained by fitting a parametric functional form using unsmoothed rates. One of the parametric functional forms most widely used in the estimation of the yield curve was proposed by Nelson and Siegel (1987), who developed a sufficiently flexible model that could represent curves of different shapes. In this model, the parameters are associated with the long-term, medium-term and short-term interest rates. Basically, this form describes the yield curve through three factors, which are interpreted as level, slope and curvature, and another factor that represents a time scale.

Nelson and Siegel (1987) suggest to fit the forward rate curve at a given date with a mathematical class of approximating functions. The functional form they advocate uses Laguerre functions which consist of the product between a polynomial and an exponential decay term. The resulting Nelson-Siegel approximating forward curve can be assumed to be the solution to a second order differential equation with equal roots for spot rates:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\lambda_t \tau} + \beta_{3,t} \lambda_t e^{-\lambda_t \tau} \tag{6}$$

The parameters $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ are determined by initial conditions and λ_t is a constant associated with the equation. Recently, Diebold and Li (2006) reinterpreted the exponential model proposed by Nelson and Siegel (1987), considering a parametric form for the behavior of the term structure over time, in which coefficients are treated as level, slope and curvature. By averaging over forward rates, as in (5), we obtain the corresponding yield curve:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1 - e^{\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3,t}\left(\frac{1 - e^{\lambda_t \tau}}{\lambda_t \tau} - e^{\lambda_t \tau}\right) \tag{7}$$

Nelson and Siegel yield curve also corresponds to the discount function, assuming value 1 at maturity zero, and drawing close to zero when maturity tends to infinity.

The shape of the yield curve is determined by the three factors and by the factor loadings associated with them. Parameter λ, kept fixed in Diebold and Li (2006), governs the exponential decay rate, small (large) values of λ are associated with a slow (quick) decay and fit better to long (short) maturities. The factor loading of the first component is 1 (constant) and is interpreted as the level of the yield curve, which equally influences the short- and long-term rates. The factor loading of the second component $\left(\frac{1-e^{\lambda_t \tau}}{\lambda_t \tau}\right)$ begins at 1 and converges to zero monotonically and quickly, being interpreted as slope. This factor has a major influence over short-term interest rates. The factor loading of the third component, $\left(\frac{1-e^{\lambda_t \tau}}{\lambda_t \tau} - e^{\lambda_t \tau}\right)$, is a concave function, assuming value zero for maturity zero, increasing, and converging monotonically to zero at longer maturities. Thus, this factor is associated with medium-term interest rates, and is treated as the curvature of the yield curve.

Since the factor loading of the first component is the only one that is equal to 1 when the maturity draws close to infinity, $\beta_{1,t}$ is associated with the long-term interest rate. The slope of the yield curve is usually defined as $y_t(\infty) - y_t(0)$, in which case the slope converges to $\beta_{2,t}$. The curvature is defined as $2y_t(\tau^*) - y_t(\infty) - y_t(0)$, where τ^* represents a medium-term maturity. Note that the curvature is virtually $-\beta_{3,t}$.[1]

Given that bonds with different maturities are observed in each time period, one has a set of yields with maturities $(\tau_1, \tau_2, \ldots, \tau_N)$ for every t. Therefore, Diebold and Li (2006) propose to fix λ to an *a priori* value and estimate equation (7) by means of an ordinary least squares regression for every time period t, from which the time series of the factors $\beta_{j,t}$ are obtained. The $\beta_{j,t}$ are obtained by estimating the following regression for a given t:

$$y_t(\tau_i) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{\lambda\tau_i}}{\lambda\tau_i}\right) + \beta_{3,t}\left(\frac{1-e^{\lambda\tau_i}}{\lambda\tau_i} - e^{\lambda\tau_i}\right) + \epsilon_{i,t} \tag{8}$$

where the errors $(\epsilon_{1,t}, \epsilon_{2,t}, \ldots, \epsilon_{N,t})$ are assumed to be independent, with zero mean and constant variance σ_t^2 for a given time t.

In general, there are several specifications that can be used to fit the data. Nevertheless, most of the existing literature basically relies on two specifications for the fitting of the model. In one of the cases, it is assumed that the three state variables follow an independent and first-order autoregressive process, used, for instance, in Diebold and Li (2006). In the other case, the three nonobservable factors in the state-space model, are modeled by a first-order vector autoregressive process, VAR (1), as in Diebold et al. (2006) and Koopman et al. (2007). Both in Diebold and Li (2006), and in Diebold et al. (2006), the factor loadings depend upon a single decay parameter, and to allow the estimation of time-varying latent factors in a linear fashion, the factor loadings are kept constant over time for each maturity.

[1] Diebold and Li (2006) define the slope as $y_t(120) - y_t(3) = -0.78\beta_{2,t} + 0.06\beta_{3,t}$, and the curvature as $2y_t(24) - y_t(120) - y_t(3) = 0.000053\beta_{2,t} + 0.37\beta_{3,t}$.

In the case of VAR(1) the transition equation, which governs the dynamics of the state vector, is defined by:

$$\begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} = \begin{pmatrix} \mu_{\beta_1} \\ \mu_{\beta_2} \\ \mu_{\beta_3} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix} \quad (9)$$

In the case in which the model is adjusted by a first-order autoregressive process, the matrix A with elements $a_{i,j}$ above is diagonal. The measurement equation, which associates the interest rates of N maturities with the three unobserved components, is given by:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_t \tau_1}}{\lambda_t \tau_1} & \frac{1-e^{-\lambda_t \tau_1}}{\lambda_t \tau_1} - e^{-\lambda_t \tau_1} \\ 1 & \frac{1-e^{-\lambda_t \tau_2}}{\lambda_t \tau_2} & \frac{1-e^{-\lambda_t \tau_2}}{\lambda_t \tau_2} - e^{-\lambda_t \tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t \tau_N}}{\lambda_t \tau_N} & \frac{1-e^{-\lambda_t \tau_N}}{\lambda_t \tau_N} - e^{-\lambda_t \tau_N} \end{pmatrix} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{N,t} \end{pmatrix} \quad (10)$$

The system comprising the transition equation and the measurement equation can be written using a matrix notation:

$$\beta_{t+1} = \mu + \Phi \beta_t + \eta_t \quad (11)$$

$$y_t = \Lambda(\lambda)\beta_t + \epsilon_t \quad (12)$$

where $\Lambda(\lambda)$ is an $N \times 3$ matrix of factor loadings, which will be time-varying only if the decay parameter is variable.

Measurement equation (12) defines the vector of yields $N \times 1$ for N maturities at time t, as the sum of factors multiplied by their factor loadings, with a normally distributed and independent error vector across maturities. The 3×1 vector β_t represents the factors.

If the purpose is only to fit the yield curve, the measurement equation suffices. However, to make forecasts of the term structure, it is necessary to model the dynamics of factors as well. Following Diebold et al. (2006) and Koopman et al. (2007), the dynamics of factors is specified as a first-order vector autoregressive process. Finally, the errors of the measurement and state equations are assumed to be orthogonal to each other and to the vector of initial states, β_0, and are distributed as:

$$\begin{pmatrix} \eta_t \\ \epsilon_t \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \begin{pmatrix} \Sigma_\eta & \mathbf{0} \\ \mathbf{0} & \Sigma_\epsilon \end{pmatrix} \right]$$

In addition, the errors of the transition and measurement equations are assumed to be orthogonal to the initial state:

$$E[\beta_0 \eta'] = \mathbf{0}$$
$$E[\beta_0 \epsilon'] = \mathbf{0}$$

The covariance matrix of the measurement errors ϵ_t constitutes a $(N \times N)$ diagonal matrix Σ_ϵ. The assumption that matrix Σ_ϵ is diagonal implies that the deviations of the interest rates to different maturities are uncorrelated. This assumption facilitates the estimation of the model by reducing the number of parameters, and is quite common in the literature. On the other hand, the assumption that matrix Σ_η is unrestricted allows shocks on the three factors to be correlated.

2.2.1. *Estimation and forecasting using the state-space form*

When the state-space form is used, two approaches can be employed to estimate the latent factors and the parameters. The initial approach proposed by Diebold and Li (2006) is based on two steps and, therefore, is inefficient, disregarding the uncertainty that is inherent to the first-step estimates in the subsequent step. In the first stage, the measurement equation is estimated using cross-sectional data, in which the estimates for the factors are obtained for each time period. Assuming that the decay parameter is constant, the measurement equation becomes linear and can be estimated by ordinary least squares. In the second stage, the time dynamics of the parameters is specified and estimated as an AR(1) or VAR(1) process.

Diebold et al. (2006) showed that it is possible to estimate this model by maximum likelihood in a single step by using the Kalman filter, providing efficient estimates for the parameters and smoothed estimates for the unobservable factors. This approach is not only adopted in Diebold et al. (2006), but also in Koopman et al. (2007), among others. The procedure utilizes the Kalman filter to build the likelihood function, which is then maximized in order to obtain parameters estimates. We consider the Nelson-Siegel model 11 and 12 as a linear Gaussian state space model. The vector of unobserved states β_t can be estimated conditional on the past and current observations $y_1, \ldots, y_t$ via Kalman filter. Define $\beta_{t|t-1}$ as the expectation of β_t given $y_1, \ldots, y_{t-1}$ with mean square error (MSE) matrix $P_{t|t}$. For given values of $\beta_{t|t-1}$ and $P_{t|t-1}$, when observation y_{t-1} is available, the predition error can be calculated $v_t = y_t - \Lambda(\lambda)\beta_{t|t-1}$. Thus, after observing y_t, a more acurate inference can be made of β_t and $P_{t|t}$:

$$\beta_{t|t} = \mu + P_{t|t-1} + P_{t|t-1}\Lambda(\lambda)' F_t^{-1} v_t$$
$$P_{t|t} = \beta_{t|t-1} - P_{t|t-1}\Lambda(\lambda)' F_t^{-1}\Lambda(\lambda) P_{t|t-1}$$

where $F_t = \Lambda(\lambda)P_{t|t-1}\Lambda(\lambda)' + \Sigma_\varepsilon$ is the predition error covariance matrix. The estimator of the state vector for the next period $t+1$, conditional on $y_1, \ldots, y_t$, is given by the predition step

$$\beta_{t+1|t} = \mu + \Phi\beta_{t|t}$$
$$P_{t+1|t} = \Phi P_{t|t}\Phi' + \Sigma_\eta$$

For a given time series $y_1, \ldots, y_T$, the Kalman filter computations are carried out recursively for $t = 1, \ldots, T$ with initializations $\beta_{1|0}$ and $P_{1|0}$. The parameters in the VAR coefficient matrix Φ, variance matrices Σ_η and Σ_ε together with μ and λ are treated as unknown coefficients which are collected in the parameter vector ψ. Estimation of ψ is based on the numerical maximization of the loglikelihood function that is constructed via the predition error decomposition and given by

$$l(\psi) = -\frac{NT}{2}log2\pi - \frac{1}{2}\sum_{t=1}^{T}log|F_t| - \frac{1}{2}\sum_{t=1}^{T}v_t' log F_t^{-1} v_t \tag{13}$$

As a result, $l(\psi)$ can be evaluated by Kalman filter for a given value of ψ (for details about Kalman filter estimation see Durbin and Koopman 2001; Anderson and Moore 1979; Simon 2006).

The maximum likelihood estimator obtained thereby is preferable to the two-step method, as in the later the estimation of parameters in the second stage does not take into consideration the uncertainty over the values of the estimated factors in the first stage, producing inefficient parameter estimates. The joint estimation of the measurement and state equations on the other hand does not have such problem and yields efficient estimates for the parameters. Another advantage of likelihood estimation is the joint estimation of the decay parameter which, in the two-step method, has to be calibrated according to some measure. Almeida et al. (2007) show that different rules for the calibration of the decay parameter yield different results for the out-of-sample forecast of the term structure of interest rate, indicating that the two-step estimation method lacks robustness. Moreover, the Kalman smoother allows obtaining smoothed estimates for the latent variables, which take the whole sample information into account in order to infer on the time series of the unobserved factors.

3. Data and Analysis of the Results

The future interbank deposit (future ID) contract with maturity τ is a future contract of which the basic asset is the interest rate[2] accrued on a daily basis (ID), capitalized between trading period t, and τ. The contract value is set by its value at maturity, $R\$100,000.00$ discounted according to the accrued interest rate, negotiated between the seller and the buyer.

When buying a future ID contract for the ID price at time t and keeping it until maturity τ, the gain or loss is given by:

[2] The ID rate is the average daily rate of interbank deposits (borowing/lending), calculated by the Clearinghouse for Custody and Settlements (CETIP) for all business days. The ID rate, which is published on a daily basis, is expressed in annually compounded terms, based on 252 business days.

$$100.000\left(\frac{\prod_{i=1}^{\zeta(t,\tau)}(1+y_i)^{\frac{1}{252}}}{(1+DI^*)^{\frac{\zeta(t,\tau)}{252}}}-1\right)$$

where y_i denotes the ID rate, $(i-1)$ days after the trading day. The function $\zeta(t,\tau)$ represents the number of working days between t and τ.

The ID contract is quite similar to the zero-coupon bond, except for the daily payment of marginal adjustments. Every day the cash flow is the difference between the adjustment price of the current day and the adjustment price of the previous day, indexed by the ID rate of the previous day.

Future ID contracts are negotiated in the BM&F, which determines the number of maturities with authorized contracts. In general, there are around 20 maturities with authorized contracts every day, but not all of them have liquidity. Approximately 10 maturities have contracts with greater liquidity. There exist contracts with monthly maturities for the months at the beginning of each quarter (January, April, July and October). In addition, there are contracts with maturities for the four months that follow the current month. The maturity date is the first working day of the month in which the contract is due.

3.1. *Data*

The data used in this paper consist of daily observations of yields of future ID contracts, closing prices. In practice, contracts with all maturities are not observed on a daily basis. Therefore, based on the rates observed daily for the available maturities, the data were converted to fixed maturities of 1, 2, 3, 4, 6, 9, 12, 15, 18, 24, 27, 29, 31 and 33 months, by means of interpolations using cubic splines. The data were observed between January 2006 and February 2009, and represent the most liquid ID contracts negotiated during the analyzed period. Table 1 presents statistics for Brazilian yield curve.

Only the data referring to the adjustments of future ID contracts were used, thus excluding swap rates. According to the BM&F (Brazilian Mecantile Exchange – Bolsa de Mercadorias e Futuros) selection criteria, the closing data on PRE ID swap rates are obtained from data on the adjustment of future ID contracts negotiated in the BM&F, thus not corresponding to data on actually processed trades in the swap modality. Therefore, as swap data are obtained from the future ID contract or by its interpolation, we consider that by using only the future ID data the model will be free of distortions arising from the use of published swap rates as if they were information about actually processed transactions. Thus, the interpolation for obtaining fixed maturity rates used in the model will rely on the data on the adjustment of future ID contracts as source of information, as these data reflect the rates of actually processed transactions, avoiding an interpolation of data that result from a previous interpolation.

The yield curve for the analyzed period has several shapes, with many changes in slope and curvature, often assuming ascending and inverted shapes throughout

Table 1
Descriptive statistics, yield curves (Jan-06 to Feb-09)

Maturity (months)	Mean	Std. Dev.	Minimum	Maximum	$\widehat{\rho}(21)$	$\widehat{\rho}(63)$
1	13.13	1.73	11.05	17.71	0.894	0.672
2	13.08	1.68	11.04	17.60	0.890	0.662
3	13.05	1.62	11.02	17.40	0.885	0.650
4	13.03	1.57	11.02	17.21	0.879	0.635
6	13.03	1.52	10.90	16.72	0.866	0.597
9	13.06	1.53	10.78	16.54	0.858	0.575
12	13.12	1.57	10.68	16.47	0.851	0.560
15	13.19	1.60	10.60	16.91	0.849	0.553
18	13.25	1.61	10.46	17.15	0.845	0.550
24	13.29	1.62	10.37	17.36	0.840	0.543
27	13.31	1.63	10.28	17.57	0.836	0.536
29	13.32	1.63	10.21	17.72	0.831	0.527
31	13.36	1.63	10.15	17.83	0.828	0.521
33	13.33	1.63	10.13	17.87	0.827	0.517

Note: We present descriptive statistics for daily yields at different maturities. The last three columns contain sample autocorrelations at displacements of 1 and 3 months.

the period. Figure 1 shows the 3D graph of the analyzed curve.

Note that there is a large amount of time changes in the level of the curve. The analyzed period was characterized by several changes in the Brazilian monetary policy conduct. These changes in monetary policy conduct influence the yields utilized on the market of public and private bonds, causing the Brazilian term structure of interest rate to take different shapes throughout the period. Quite often, the term structure of interest rate demonstrates changes not only in the curvature pattern, but also in the slope pattern. This way, the analyzed period seems quite appropriate for checking the predictability of Nelson and Siegel (1987), extended by Diebold and Li (2006).

3.2. *Empirical assessment of the model for future ID data*

In Section 2, the model of Diebold and Li (2006) was laid out in state-space form, with a VAR(1) for the transition equation, which models the dynamics of the factors, and a linear measurement equation that relates the observed yields to the state vector. The parameters were estimated simultaneously by maximum

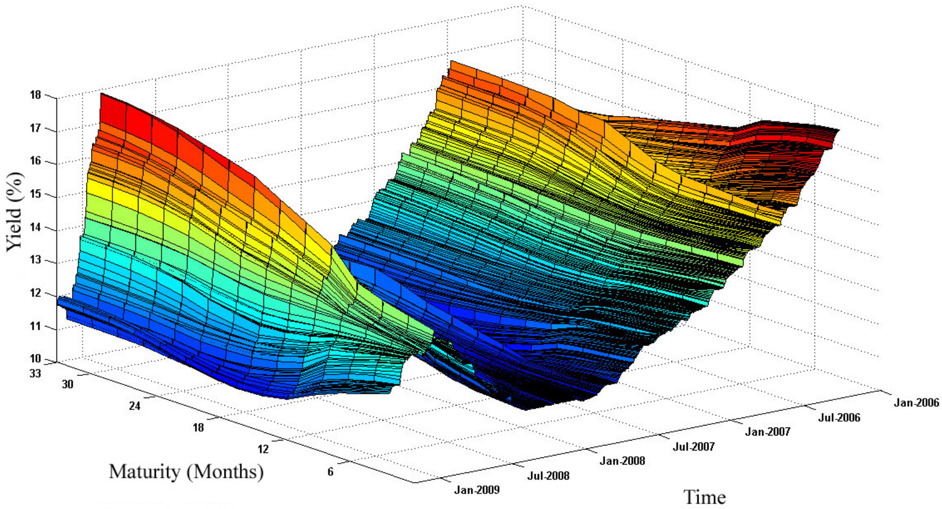

Fig. 1. Term structure of the interest rate (Jan 2006 to Feb 2009)

likelihood using the Kalman filter, which is an efficient estimator and also eliminates the problem related to how to calibrate the decay parameter. The yields used consist of daily data on future ID rates between January 2006 and February 2009, totaling 772 daily observations for each of the 14 maturities; of these observations the latter 252 (one business year) were used for the out-of-sample analysis.

The minimization of the negative of the logarithm of the likelihood function was obtained by the quasi-Newton method with updates of the inverse Hessian matrix using the BFGS method. More specifically, we used the csminwel algorithm, developed by Christopher Sims to be robust to certain pathologies common to likelihood functions such as hyperplane discontinuities, which is available at `http://sims.princeton.edu/yftp/optimize/`. The algorithm was configured to cease iterations when it is no longer possible to increase the function value by at least 1.0e-05.

Unlike the two-step method, in the Kalman filter estimation the parameters are estimated in a single step. Besides, λ which governs the decay rate of factor loadings of both the level and curvature, was estimated jointly with other parameters and not determined a priori. The initial values of the parameters for Kalman filter initialization were obtained from the estimation as in Diebold and Li (2006) using the two-step method. Figure 2 shows the factor loadings for level, slope and curvature, obtained from the estimate of λ.

With an estimated λ equal to 0.1047, the factor loading relative to the curvature assumes maximum value for maturities between 13 and 18 months.

The main argument in favor of Diebold-Li three-factor model is its capacity to yield good forecasts. Although it is not the best model when the fit of the term structure of interest rate is the major goal, the model put forward by Diebold-Li

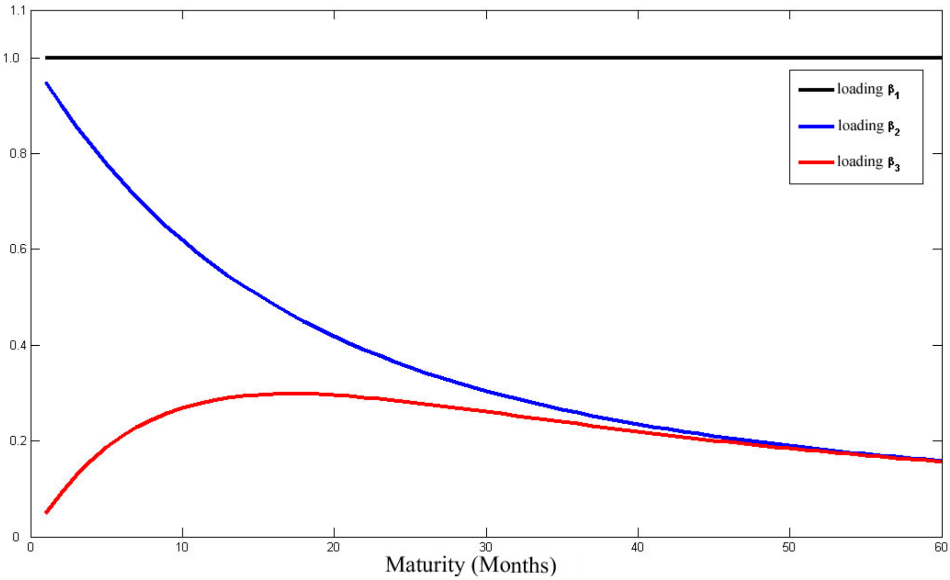

Fig. 2. Factor loadings ($\lambda = 0.1047$)

can replicate the several shapes taken by the yield curves. Figure 3 shows the real data on the yield curve for some days and the curve adjusted by the parameters of the estimated factor model.

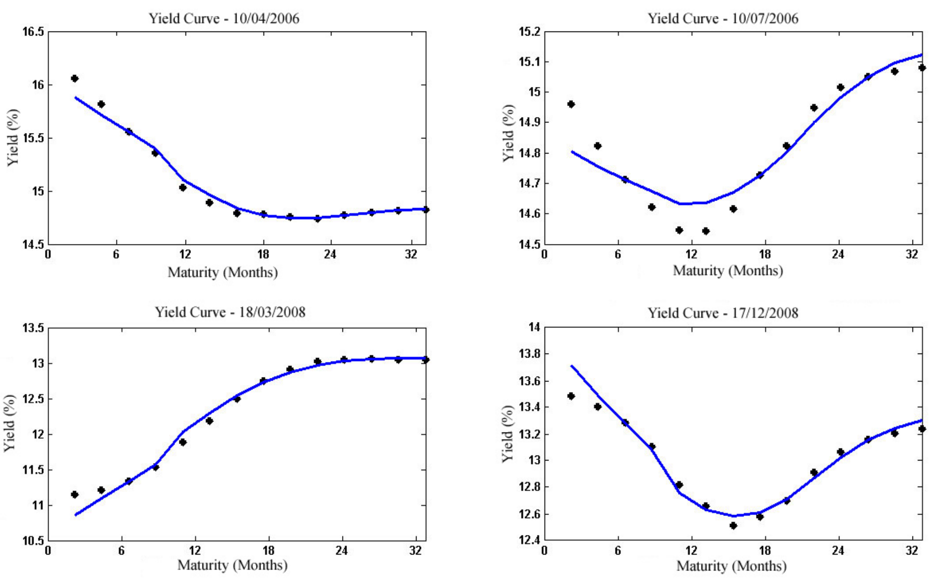

Fig. 3. Yield curve fit for different points in time

Note that the model estimated with three factors fits well to a wide variety of shapes of the yield curve: positively sloped, negatively sloped and with different curvature shapes. Figure 4 plots the daily residuals of the yield curve obtained from the fitted model. Observe that the residuals do not have a systematic behavior and are of small magnitude, indicating that the model can replicate the patterns exhibited by the yield curve for the period. The graph shows that residuals have greater volatility to shorter maturities. This situation is regarded as a stylized fact when it comes to yield curves – shorter maturities are more volatile than the rates of longer maturities. One of the possible explanations is that shorter maturities are more susceptible to fluctuations of the benchmark interest rate (Selic rate) – a monetary policy instrument.

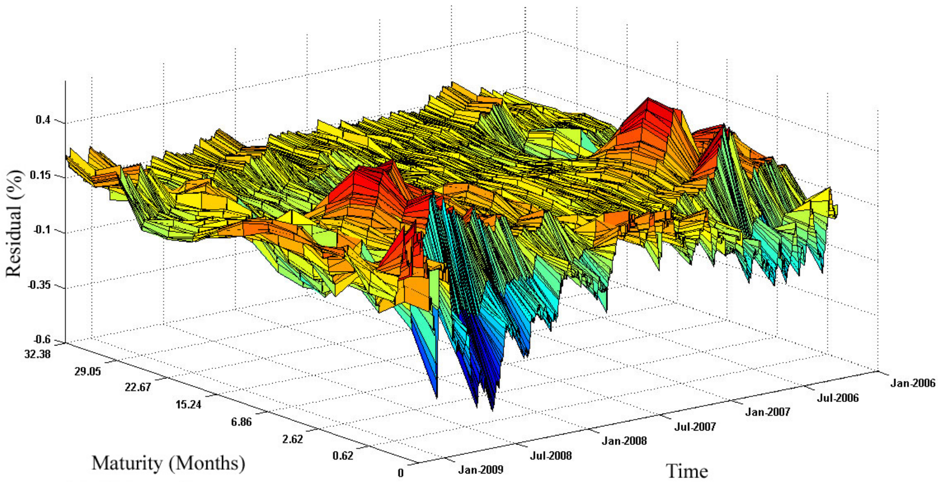

Fig. 4. Residuals of the estimated yield curve (Jan 2006 to Feb 2009)

The estimated level of the yield curve, β_1, has a statistically significant mean of 13.72%, with high persistence. Note that the level of the yield curve exhibited a more volatile behavior after August 2008, when the financial crisis had a stronger impact on the assets traded in the Brazilian market. Also during this period, the level of the yield curve had its highest value (19.47). It should also be observed that there is a sudden change of behavior in the slope and curvature factors. Figures 5.1 through 5.3 clearly show that not only the level, but also the slope and curvature oscillate during this period, and that autocorrelations reveal the high persistence of these two factors.

The assessment of the predictability of the model is made by splitting the sample into two parts. One of these parts is used to estimate the model and includes 520 observations, with data obtained from January 2006 to January 2008. The second part is used to assess the performance of forecasts produced by the model, with data from February 2008 to February 2009, totaling 252 observations. Forecasts

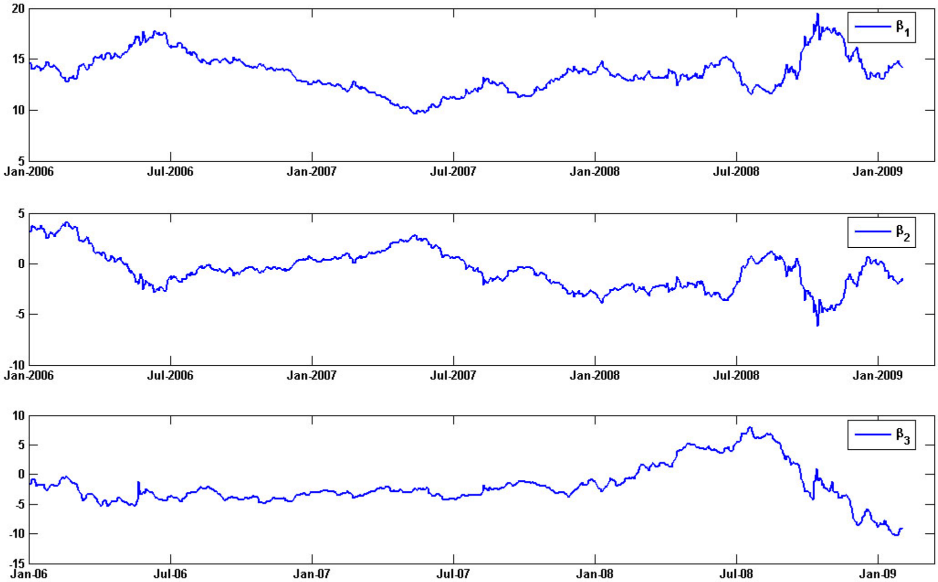

Fig. 5. Smoothed estimates of the factors

for one day, one month, three months and six months ahead are analyzed based on a rolling window scheme with growing window size. Note that the period used to asses the out-of-sample performance of the model includes the present crisis. To complete the forecasting exercise, we obtained forecasts from a random walk model for the yields and from the Diebold-Li model estimated by the two-step method(RW and 2S). [3]

Tables 2 and 3 show the estimated VAR parameters and the covariance matrix of the estimated factors. Most covariances estimated are statistically significant, indicating that the VAR is the most suitable structure to capture the dynamics of the factor.

[3] In the two-step estimation λ was set to 0.1182, which is the value that minimizes the RMSE. This value maximizes the loading of the curvature factor for the average maturity of 13 to 15 months.

Table 2
Estimated VAR Parameters

	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	μ
$\beta_{1,t}$	0.991^{***}	0.003^{*}	0.001	13.72^{***}
	(0.002)	(0.002)	(0.002)	(1.01)
$\beta_{2,t}$	0.008^{***}	0.993^{***}	0.002	$-.72$
	(0.003)	(0.002)	(0.001)	(1.35)
$\beta_{3,t}$	-0.003	$-.0012^{***}$	1.000^{***}	1.46
	(.002)	(.003)	(0.002)	(1.66)

Note: Standard errors in parenthesis are calculated as the square root of the diagonal elements of the inverse Hessian.
*,**, and *** indicates statistical significance t̂he 10%, 5% and 1% significance level respectively.

Table 3
Covariance matrix of the estimated VAR

	$\eta_{1,t}$	$\eta_{2,t}$	$\eta_{3,t}$
$\eta_{1,t}$	0.25^{***}	-0.07^{***}	-0.04^{***}
	(0.001)	(0.001)	(0.002)
$\eta_{2,t}$		0.26^{***}	0.03
		(0.001)	(0.002)
$\eta_{3,t}$			0.40^{***}
			(0.003)

Standard errors in parenthesis are calculated as the square root of the diagonal elements of the inverse Hessian.
*,**, and *** indicates statistical significance at the 10%, 5% and 1% significance level respectively.

The analysis of VAR parameters indicates high persistence in the dynamics of the latent factors. The statistically significant cross-effects for the dynamics of the factors are observed from $\beta_{1,t-1}$ in $\beta_{2,t}$, and in $\beta_{3,t}$, reinforce the importance of modelling the fators as a VAR.

The approach to forecast the yield curve using the Diebold Li model consists in predicting the factors and then using the forecasted factors to adjust the predicted yield curve. Forecasts at time t, for $t+h$, of yield with maturity τ , are given by:

$$\widehat{y}_{t+h|t}(\tau) = \widehat{\beta}_{1,t+h|t} + \widehat{\beta}_{2,t+h|t}\left(\frac{1-e^{-\widehat{\lambda}\tau}}{\widehat{\lambda}\tau}\right) + \widehat{\beta}_{3,t+h|t}\left(\frac{1-e^{-\widehat{\lambda}\tau}}{\widehat{\lambda}\tau} - e^{-\widehat{\lambda}\tau}\right)$$

The forecasts of the factors are obtained using the estimated parameters of the transition equation (1):

$$\widehat{\beta}_{t+h|t} = \widehat{\mu} + \widehat{A} \cdot \widehat{\beta}_t$$

Tables 4 and 5 show the RMSE for the out-of-sample forecasts made with the model estimated by the Kalman filter (KF), for horizons of one day, one month, three months and six months ahead. We also present the RMSE for the same horizons, obtained by random walk and by Diebold-Li model estimated by the two-step method, for comparison of the results.

Table 4
RMSE for out-of-sample forecasts (Feb 2008 to Feb 2009)

Maturity	One day ahead			One month ahead		
(months)	RW	2S	FK	RW	2S	FK
1	4.13	10.81	26.06	38.55	48.30	47.67
2	4.57	5.96	11.61	41.25	51.41	37.42
3	5.44	10.74	5.10	45.33	56.01	32.90
4	6.94	14.02	9.71	52.54	62.94	36.23
7	10.49	16.72	16.02	70.18	78.20	51.58
9	13.01	17.36	16.59	80.78	88.53	63.01
12	15.76	18.39	16.29	91.76	99.35	75.86
15	18.02	20.18	17.92	98.02	106.21	84.61
19	19.18	21.15	19.32	105.00	110.52	90.86
23	20.65	21.79	20.36	111.21	113.80	95.91
26	21.53	22.74	21.22	117.03	117.72	101.20
29	22.20	24.59	22.06	121.28	121.86	105.86
31	22.66	26.36	22.99	123.64	124.93	109.09
32	22.84	27.70	23.73	124.79	126.89	111.01

Note: RMSE expressed in basis points, and maturities in months (RW =Random Walk Model, 2S=Two-Step Estimation, KF=Kalman Filter Estimation).

For the one-day-ahead forecasting horizon, the model estimated by the Kalman filter outperforms the random walk only in the case of maturities of 3, 15, 23 and 26 months. The remaining maturities, even though they are often better than the forecasts obtained when the model is estimated by the two-step method, are worse than the random walk. Note that the worst performance is observed for the shortest maturity which, as previously mentioned, is more susceptible to Selic rate fluctuations. However, the quality of the forecasts improves substantially when the horizon is broadened. For one-month-ahead forecasts, the model estimated by the Kalman filter outperforms the model estimated in two steps in all maturities, and underperforms the random walk only for the shortest one. For medium-term maturities, between 3 and 24 months, the forecasts obtained by the KF have an RMSE on average 15 basis points lower than the RW and 2S. When the forecasts for the 3-month horizon are analyzed, we note that the KF consistently outperforms its counterparts for all maturities, showing an RMSE on average 35 basis points lower than those obtained by 2S. For six-months-ahead forecasts, the KF outperforms the RW in all maturities, showing an RMSE on average 40 basis points lower than those obtained by RW, but is able to outperform the 2S only for shorter maturities.

Table 5
RMSE for out-of-sample forecasts (Feb 2008 to Feb 2009)

Maturity	Three months ahead			Six months ahead		
(months)	RW	2S	FK	RW	2S	FK
1	88.44	100.69	71.49	141.17	133.81	96.20
2	94.41	106.67	63.03	144.59	133.30	91.99
3	103.89	112.69	62.97	151.28	132.77	95.78
4	117.66	121.39	71.31	161.65	134.04	108.59
7	146.13	140.99	97.11	189.94	144.11	144.43
9	161.48	153.07	114.38	205.41	150.81	166.10
12	176.11	164.42	131.42	222.08	158.63	186.12
15	184.61	170.51	140.98	230.55	162.08	196.50
19	186.01	172.97	144.58	229.31	161.33	197.64
23	187.03	173.43	144.46	223.25	157.71	192.37
26	189.83	173.79	144.18	219.10	154.77	186.64
29	192.88	175.16	144.63	214.69	153.46	182.12
31	194.89	176.22	145.20	212.30	152.78	179.17
32	196.01	177.05	145.72	211.16	152.52	177.48

Note: RMSE expressed in basis points, and maturities in months (RW =Random Walk Model, 2S=Two-Step Estimation, KF=Kalman Filter Estimation).

In the case of forecasts for the 6-month horizon, the KF outperforms the RW in all maturities, but it is outclassed by DP for longer maturities. The smallest liquidity of contracts with longer maturities may deteriorate the quality of forecasts, as pointed out by Bali et al. (2006). According to their work, liquidity plays an important role in the predictability of yields.

To confirm whether the differences among the out-of-sample forecasts generated by the model using Kalman filter are statistically significant, we applied the Diebold and Mariano (DM) test to compare forecasts (see Diebold and Mariano 1995; McCracken 2007). We compared the paired forecasts generated with Kalman filter, using the two step and random walk methodologies. Let $\{d_i\}_{i=1}^{n}$ be a function of the difference of square forecast errors produced by two models. We can write d_i as:

$$d_i = \left(\widehat{y}_{t+h,i}(\tau) - y_{t+h}(\tau)\right)^2 - \left(\widehat{y}_{t+h,j}(\tau) - y_{t+h}(\tau)\right)^2 \qquad (14)$$

where $i = FK$ or $2S$ and $j = 2S$ or RW. The variables $\widehat{y}_{t+h,i}$ are h-step ahead forecast at time t of FK, 2S and random walk (RW) models, respectively. DM propose a test to check whether the average loss differential $\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ is statistically different from zero, which is given by:

$$DM = \frac{\overline{d}}{\sqrt{\frac{\widehat{\delta}}{n}}} \longrightarrow^{d} N(0,1) \qquad (15)$$

where $\widehat{\delta}$ is an estimate of the long run covariance matriz of the d_i. We employ the Newey and West (1987) estimate for $\widehat{\delta}$, which allows controlling for serial correlation in the forecasting errors.

Tables 6 and 7 show Diebold and Mariano statistics for a quadratic loss function. Negative values indicate the superiority of the first method of the pair. High absolute values for DM statistics indicate large probability to reject the null hypothesis (differences between the mean quadratic errors are negligible). Absolute values higher than 1.96 indicate rejection of the null hypothesis at 95% confidence. Diebold and Mariano statistics support most of the conclusions above, such as, the good results of the model estimated through Kalman filter for forecasts one and three months ahead, overperforming the two step method for one day forecast and showing a slightly inferior performance for the 6-month horizon.

Table 6
Diebold and Mariano test for out-of-sample forecasts

Maturity	One day ahead			One month ahead		
(months)	FK×2S	FK×RW	2S×RW	FK×2S	FK×RW	2S×RW
1	2.66	2.94	4.05	0.58	1.99	4.58
2	3.30	4.22	4.39	-0.15	-1.59	0.90
3	-3.50	-0.58	3.46	-0.90	-3.18	-0.09
4	1.00	3.28	3.38	-1.51	-3.68	-0.11
7	1.05	2.96	5.27	-2.24	-3.17	-0.16
9	-1.35	3.31	6.15	-2.46	-2.52	-0.03
12	-7.63	1.45	8.22	-2.29	-1.63	0.21
15	-5.83	-0.66	6.20	-2.19	-1.73	0.27
19	-5.23	0.61	4.95	-1.56	-1.82	-0.21
23	-5.04	-0.79	4.68	-1.41	-2.10	-1.01
26	-4.27	-1.25	3.99	-1.36	-2.20	-1.36
29	-3.17	-0.60	2.77	-1.32	-2.13	-1.48
31	-2.30	1.13	2.19	-1.31	-2.13	-1.52
32	-1.63	2.22	2.07	-1.23	-2.13	-1.56

Note: Diebold and Mariano statistics for one day and one month ahead, comparing the methods by pairs. Negative values indicate superiority of the forecast for the first method of the pair.

Table 7
Diebold and Mariano test for out-of-sample forecasts

Maturity	Three months ahead			Six months ahead		
(months)	FK×2S	FK×RW	2S×RW	FK×2S	FL×RW	2S×RW
1	-1.95	-1.30	-0.93	-0.18	-1.55	-1.17
2	-8.35	-1.89	-1.04	-0.17	-17.91	-7.18
3	-1.31	-2.24	-1.24	-0.09	-23.38	-5.30
4	-1.14	-2.76	-1.28	0.09	-9.88	-7.51
7	-1.03	-3.53	-1.40	2.44	-9.16	-2.15
9	-1.23	-3.45	-1.39	0.78	-21.02	-1.43
12	-6.85	-3.25	-1.38	0.73	-12.93	-1.58
15	-5.99	-3.32	-1.51	0.60	-32.44	-1.81
19	-2.16	-3.52	-1.83	0.52	-1.77	-1.94
23	-0.77	-3.07	-2.23	0.42	-1.68	-1.99
26	-0.78	-2.71	-2.50	0.35	-1.64	-1.92
29	-0.82	-2.45	-2.48	0.86	-1.62	-1.89
31	-0.82	-2.37	-2.31	9.67	-1.62	-1.90
32	-0.84	-2.34	-2.18	4.91	-1.64	-1.92

Note: Diebold and Mariano statistics for one day and one month ahead, comparing the methods by pairs. Negative values indicate superiority of the forecast for the first method of the pair.

4. Conclusion

In the present paper, Diebold and Li model, usually estimated for Brazilian data by the inefficient two-step method, was put in the state-space form and efficiently estimated by maximum likelihood using the Kalman filter. The maximum likelihood estimation allows for the joint estimation of all parameters of the model, preventing the a priori selection of the decay parameter. Smoothed estimates of the parameters, which contemplate the whole information of the sample in order to infer on the time series of the factors, were obtained by the Kalman smoother and used for the out-of-sample forecast. The results indicate that the model estimated by maximum likelihood yields better out-of-sample forecasts than the model estimated by the two-step method for all forecasting horizons. In addition, the forecasts based on the model estimated by maximum likelihood are better than those of the random walk model for all maturities when horizons of one month, three months and six months are considered. The yield curve factor models are the most widely used by central banks worldwide and by most market participants to adjust and forecast the term structure of interest rate. The results obtained herein show the flexibility and capacity of the model to adjust itself to a wide variety of yield curve shapes, and that the estimation by the Kalman filter is better than its counterparts estimated by the two-step method. A possible suggestion for further investigation is the estimation of the model using four factors as proposed by Cochrane and Piazzesi (2005), which include an additional curvature that improves the predictability in markets with more volatile curves, as occurs in emerging markets.

References

Almeida, C., Gomes, R., Leite, A., & Vicente, J. (2007). Does curvature enhance forecasting? Working Papers Series 155, Central Bank of Brazil, Research Department.

Almeida, C. & Vicente, J. (2008). The role of no-arbitrage on forecasting: Lessons from a parametric term structure model. *Journal of Banking & Finance*, 32(12):2695–2705.

Anderson, B. & Moore, J. (1979). *Optimal Filtering*. Englewood Cliffs: Prentice Hall, New York.

Ang, A. & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787.

Bali, T., Heidari, M., & Wu, L. (2006). Predictability of interest rates and interest-rate portfolios by Turan Bali, Massoud Heidari, Liuren Wu.

Cochrane, J. H. & Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, 95(1):138–160.

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

Diebold, F. X. & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364.

Diebold, F. X. & Mariano, R. (1995). Comparing preditive acuracy. *Journal of Business and Economic Statistics*, 13:253–263.

Diebold, F. X., Rudebusch, G. D., & Borag[caron]an Aruoba, S. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, 131(1-2):309–338.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):405–443.

Duffie, D. & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406.

Durbin, J. & Koopman, S. (2001). *Time Series Analysis by State Space Methods*. Oxford, New York.

Egorov, A. V., Li, H., & Ng, D. (2010). A tale of two yield curves: Modeling the joint term structure of dollar and euro interest rates. *Journal of Econometrics*.

Forthcoming.

Fama, E. F. & Bliss, R. R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77(4):680–92.

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.

Hördahl, P., Tristani, O., & Vestin, D. (2006). A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics*, 131(1-2):405–444.

Koopman, S. J., Mallee, M. I., & van der Wel, M. (2007). Analyzing the term structure of interest rates using the dynamic nelson-siegel model with time-varying parameters. Tinbergen institute discussion papers, Tinbergen Institute.

Laurini, M. P. & Hotta, L. K. (2008). Bayesian extensions to diebold-li term structure model. Ibmec Working Papers, Ibmec Working Paper, Ibmec São Paulo.

Matzner-Lober, E. & Villa, C. (2004). Functional principal component analysis of the yield curve. In *21th International Conference AFFI*. Association Française de Finance.

McCracken, Michael, W. (2007). Asymtoptics for out of sample tests of granger causality. *Journal of Econometrics*, 140:719–752.

McCulloch, J. H. (1971). Measuring the term structure of interest rates. *Journal of Business*, 44(1):19–31.

McCulloch, J. H. (1975). The tax-adjusted yield curve. *Journal of Finance*, 30(3):811–30.

Nelson, C. R. & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, 60(4):473–489.

Newey, W. & West, K. (1987). A simple, positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.

Pearson, N. D. & Sun, T.-S. (1994). Exploiting the conditional density in estimating the term structure: An application to the cox, ingersoll, and ross model. *Journal of Finance*, 49(4):1279–1304.

Robert, L. & Scheinkman, J. (1991). Common factors affecting bond returns. *Journal of Fixed Income*, 1(1):54–61.

Simon, D. (2006). *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, New Jersey.

Svensson, L. E. O. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994. Technical Report 94-114, IMF Working Papers.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.

Vasicek, O. A. & Fong, H. G. (1982). Term structure modeling using exponential splines. *Journal of Finance*, 37(2):339–48.

Vicente, J. & Tabak, B. M. (2007). Forecasting bonds yields in the Brazilian fixed income market. Working Papers Series 141, Central Bank of Brazil, Research

Department.

Wu, T. (2001). Monetary policy and the slope factor in empirical term structure estimations. Working Papers in Applied Economic Theory 2002-07, Federal Reserve Bank of San Francisco.