EFFICIENT INTEREST RATE CURVE ESTIMATION AND FORECASTING IN BRAZIL

João Frois Caldeira[1]
Guilherme Valle Moura[2]
Marcelo Savino Portugal[3]

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 approach of the Nelson Siegel model in order to adjust and forecast the Brazilian yield curve. The data consisted of daily observations of future ID yields traded in the BM&F which presented more liquidity 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 adjustment, but mainly 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, three and six months.

Keywords: Term structure, interest rate, yield curve, state-space model, Kalman filter

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.

Palavras-chave: Estrutura a termo, taxa de juros, curva de juros, formato estado espaço, Filtro de Kalman

Área 7 - Microeconomia, métodos quantitativos e finanças
JEL: C53, E43, G17

[1] PhD student (UFRGS) and Derivatives Analyst (Banco SICREDI S.A.).
[2] PhD student in Econmics (Christian-Albrechts-Universität zu Kiel).
[3] Professor of Economics (UFRGS) and CNPq researcher.

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, Jarrow and Morton (1992). Equilibrium models place emphasis on the modeling of the instantaneous rate, typically through affine models; then the rates of other maturities can be derived under several hypotheses about the risk premium. Models of this type were developed by Vasicek (1977), Cox, Ingersoll and Ross (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 and Villa (2004), most of the intuition about the dynamics of bond and bonus profitability arises from models belonging to this class, as in Litterman 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 *et al*. (1998), Laurini and Hotta (2007), among others.

Interpolation models were developed, for instance, by McCulloch (1971, 1975), who interpolated the discount function rather than the interest rate 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.

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 adjustment 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 Ergorov *et al*. (2006), who show that affine models with stochastic volatility can predict the conditional joint distribution of bonus profitability.

Having good interest rate 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*.(2002), and Wu (2002) analyzed models with macroeconomic variables and showed that these variables contribute towards improving the forecast of the interest rate curve dynamics. Diebold, Rudebush and Aruoba (2006) (hereinafter referred to as DRA), used a model with latent factors for the interest rate curve and also included macroeconomic variables. Unlike previous models which considered a unidirectional relationship of macroeconomic variables towards the interest rate curve, or of the interest rate 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 interest rate 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 interest rate curve forecasts. Almeida *et al*. (2007a) obtained better forecasting results than those from Diebold and Li model using a dynamic version of Svensson's model (1994). Vargas (2007) uses Brazilian data for future ID contracts to replicate the results obtained by Diebold and Li. Laurini and Horta (2008), on the other hand, estimate an extended Svensson's model (1994), where decay parameters vary over time and the stochastic volatility is added to the equation used to measure 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 interest rate 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 adjusting the interest rate 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 smoothed estimates of the factors, which are later utilized in the forecasting exercise. 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.

Our paper adds to the literature because we estimate Diebold and Li model in a single step by means of the Kalman filter using data on future ID rates for maturities with higher liquidity. Besides this introduction, the paper is organized as follows. Section 2 introduces the structure of Diebold-Li model for the interest rate 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 interest rate curve that appears in Nelson and Siegel (1987), where the three factors of the Nelson and Siegel curve are interpreted as level, slope and curvature factors.

2.1 Discount Function, Forward Curve and Interest Rate Curve

Before describing the structure of the model, it is necessary to define discount curve, forward curve and interest rate 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 interest rate $y_i(t)$ associated with maturity t, obtained from a security traded on the market (Varga, 2008).

In any point at 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, in the financial market, 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:

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

The interest rate $y_t(\tau)$, 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:

$$y_t(\tau) = \frac{-\ln\left(P_t(\tau)\right)}{\tau} \quad (2)$$

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

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

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 (2) when τ_1 draws closer to τ_2, denoted by $f_t(\tau)$, is the instantaneous forward rate:

$$f_t(\tau) = \frac{-P_t^{'}(\tau)}{P_t(\tau)} \quad (4)$$

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

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

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 τ 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 interest rate curves, discount curves and forward curves, which are not observable in practice. There are some approaches to the construction of interest rate curves. McCoulloch (1971, 1975) and Vasicek and Fong (1977) build interest rate curves using estimated smooth discount curves and converting them into rates at relevant maturities. The method put forward by McCoulloch (1971, 1975) employs a cubic spline discount function interpolation. The advantage of this method is that the estimation model only has linear parameters. A disadvantage of this method is that it produces erratic curves for longer maturities, i.e., the adjusted discount curve differs for longer maturities instead of converging to zero. Vasicek and Fong's approach (1977) suggests the use of exponential splines to adjust the discount function, which would eliminate the divergence problem for longer maturities.

Statistical nonlinear models were also used to estimate the term structure of interest rate, rather than the discount function as in Nelson and Siegel (1971), Svenson (1984), Bolder and Gusba (2002), 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 (1971), 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 Interest Rate Curve Model

The classic problem with the term structure requires the estimation of a smooth interest rate curve based on the 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 adjusting a parametric functional form using unsmoothed rates. One of the parametric functional forms most widely used in the estimation of the interest rate 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 interest rate curve through three factors, which are interpreted as level, slope and curvature, and another factor that represents a time scale. If the instantaneous forward rate for a maturity τ is given by the solution of a second-order differential equation with real and different roots, it can be expressed as:

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

Recently, Diebold and Li (2003) 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. The corresponding interest rate curve is:

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

Nelson and Siegel interest rate 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 interest rate 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 slight (quick) decay and adjust best to long (short) maturities. The factor loading of the first component is 1 (constant) and is interpreted as level of the interest rate 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 curvature of the interest rate 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 interest rate curve is usually defined as $y_t(\infty)-y_t(0)$, in this case, the slope converges to $-\beta_{2,t}$[4]. The curvature is defined as $2y_t(\tau^*)-y_t(\infty)-y_t(0)$, where τ^* represents an medium-term maturity, usually a maturity of 24 or 30 months. Note that the curvature is virtually $-\beta_{3,t}$.

Given that bonds with different maturities are observed in each time period, one has a set of interest rates with maturities $(\tau_1,\tau_2,\ldots,\tau_N)$ for every *t*. Therefore, equation (7) can be estimated by ordinary least squares for every *t*, from which the time series of $\beta_{j,t}$ are obtained. The $\beta_{j,t}$ are obtained by estimating the following regression for every *t*:

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

Where the errors $(\varepsilon_{1,t},\varepsilon_{2,t},\ldots,\varepsilon_{N,t})$ are, by supposition, 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 adjust the data. Nevertheless, most of the extant literature basically relies on two specifications for adjustment 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 factors that were not observed in the model, in state-space form, are modeled by a first-order vector autoregressive process, VAR (1), as in Diebold, Rudebush and Aruoba (2006) and Koopan, Mallee and Well (2007). Both in Diebold and Li (2006), and in Diebold, Rudebush and Aruoba (2006), the factor loadings depend upon a

---

[4] Diebold and Li (2003) define the slope as $y_t(120)-y_t(3)=-0.78\beta_2+0.06\beta_{3,t}$, and the curvature as $2y_t(24)-y_t(120)-y_t(3)=0.00053\beta_{2,t}+0.37\beta_{3,t}$.

single decay parameter, and to permit the estimation of time-varying latent factors in a linear fashion, the factor loadings are kept constant over time for each maturity.

If the dynamics of the factors follows an autoregressive vector – VAR, the model can be put in the state-space form. In the case of VAR(1), the transition equation, which governs the dynamics of the state vector, is defined by:

$$\begin{pmatrix} \beta_{1,t}-\mu_L \\ \beta_{2,t}-\mu_S \\ \beta_{3,t}-\mu_C \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1}-\mu_L \\ \beta_{2,t-1}-\mu_S \\ \beta_{3,t-1}-\mu_C \end{pmatrix} + \begin{pmatrix} \eta_t(L) \\ \eta_t(S) \\ \eta_t(C) \end{pmatrix} \qquad (9)$$

For $t=1,\ldots,T$. In the case in which the model is adjusted by a first-order autoregressive process, matrix *A*, above, is diagonal. The measurement equation, which associates the interest rates of *N* maturities with the three unobserved components, is given by:

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

For $t=1,\ldots,T$. The system, comprising the transition equation and the measurement equation, can be written using a matrix notation:

$$y_t = \Lambda(\lambda)\beta_t + \varepsilon_t \qquad (11)$$

Where $\Lambda(\lambda)$ is an $N\times 3$ matrix of factor loadings, which will be time-varying only if the decay parameter is variable, whose elements (i,j) are defined by:

$$\Lambda_{ij}(\lambda) = \begin{cases} 1, & j=1, \\ \left(1-e^{-\lambda\tau_i}\right)/\lambda\tau_i, & j=2, \\ \left(1-e^{-\lambda\tau_i}-\lambda\tau_i e^{-\lambda\tau_i}\right)/\lambda\tau_i, & j=3. \end{cases}$$

$$\beta_t = (I-\Phi)\mu + \Phi\beta_{t-1} + \eta_t \qquad (12)$$

Measurement equation (11) defines the interest rate vector ($T\times N$) for *N* maturities, as the sum of factors multiplied by their factor loadings, with a normally distributed and independent error vector across maturities. Vector β_t, with size (3×1), represents the factors and Φ is the matrix containing the VAR parameters, that determine the dynamics of the states.

If the purpose is only to adjust the term structure of interest rate 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, Rudebush and Aruoba (2006) and Koopman, Mallee and Well (2007), the dynamics of factors is specified as a first-order autoregressive process. Equation (12) describes the dynamics of factors as a multivariate process, VAR (1). The vector of means μ is sized as (3×1), whereas matrix Φ is sized as (3×3) and will be diagonal or complete depending on the specification of the model, AR(1) or VAR(1) following this order. 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 usually distributed as:

$$\begin{pmatrix}\eta_t\\ \varepsilon_t\end{pmatrix} \sim \left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\Sigma_\eta & 0\\ 0 & \Sigma_\varepsilon\end{pmatrix}\right]$$

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

$$E\left(\beta_0\eta_t^{'}\right)=0,$$
$$E\left(\beta_0\varepsilon_t^{'}\right)=0.$$

The error variance for each maturity constitutes diagonal matrix Σ_ε, with size $N\times N$. The assumption that matrix Σ_ε is diagonal implies that the deviations of the interest rates to different maturities are uncorrelated. This supposition 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.

2.3 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 DL is based on two stages and, therefore, it is inefficient, disregarding the uncertainty that is inherent to the first-stage estimates in the subsequent stage. In the first stage, the measurement equation is estimated using cross-sectional data, in which the estimators for the parameters 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.[5] In the second stage, the time dynamics of the parameters is specified and estimated as an AR(1) or VAR(1) process.

Diebold, Rudebush and Aruoba (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, Rudebush and Aruoba (2006), but also in Pooter

---

[5] The choice of the decay parameter is not obvious. Usually, the criterion consists in building a grid of values for the parameter and choosing the one that minimizes the RMSE.

(2007) and Koopman, Mallee and Well (2007), among others. The procedure utilizes the Kalman filter to build the likelihood function, which is then maximized in order to obtain the estimates for the parameters. The maximum likelihood estimator obtained thereby is preferable to the two-step method, as the estimation of parameters in the second stage does not take into consideration the uncertainty over the values of the parameters estimated in the first stage, producing inefficient estimators. 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*. (2007b) 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 into account in order to infer on the time series of the factors, which is then used in the forecasting exercise.

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[6] 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:

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

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

[6] The ID rate is the average daily rate of interbank deposits (borrowing/lending), calculated by the Clearinghouse for Custody and Settlement (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.

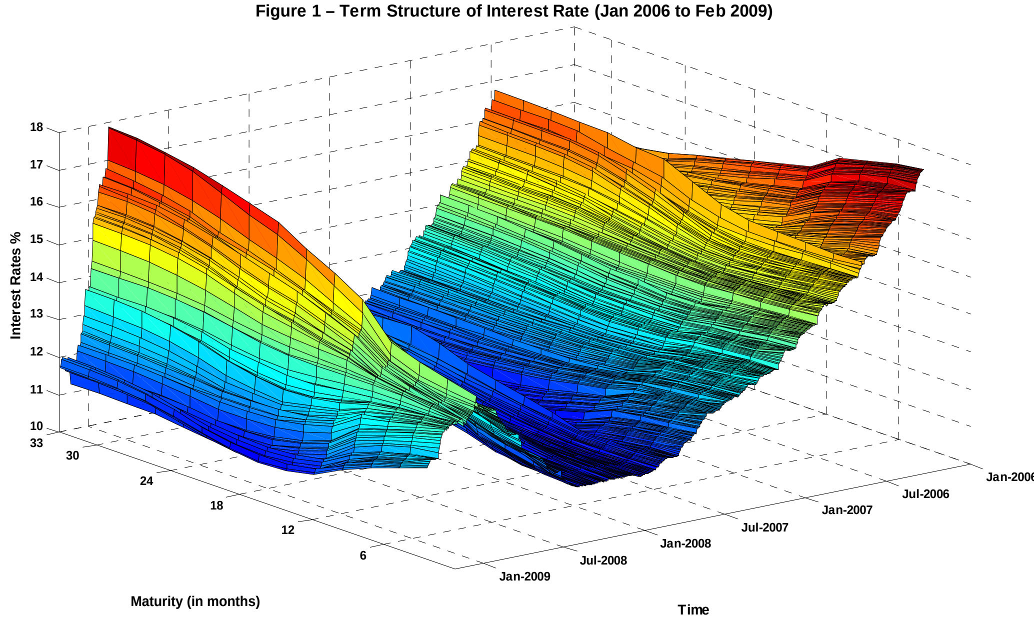
3.1 Data

The data used in this paper consist of daily observations of interest rates 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 the exponential spline method. The data were observed between January 2006 and February 2008, and represent the most liquid ID contracts negotiated during the analyzed period.

Only the data referring to the adjustments of future ID contracts were used, thus excluding swap rates. According to the BM&F 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 tradings 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 interest rate curve for the analyzed period has several shapes, with many changes in slope and curvature, often assuming ascending and inverted shapes throughout the period. Figure 1 shows the 3D graph of the analyzed curve.

Figure 1 – Term Structure of Interest Rate (Jan 2006 to Feb 2009)



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 interest rates 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 factor model (1987), extended by Diebold-Li (2006).

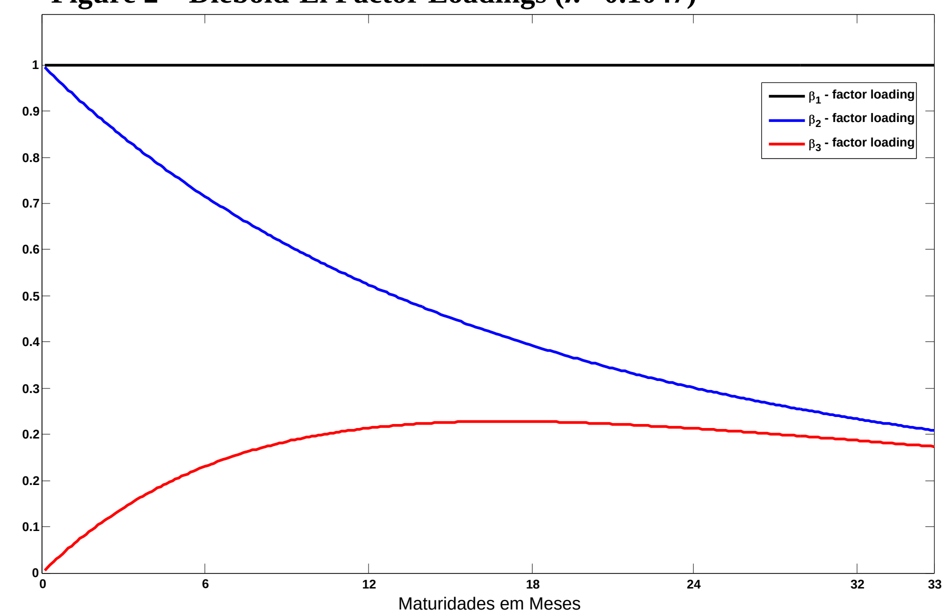
3.2 Empirical Assessment of the Model for Future ID Data

In Section 2, Diebold-Li model (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 interest rates to the state vector. The parameters were estimated simultaneously by maximum likelihood using the Kalman filter, which is an efficient estimator and also eliminates the problem related to how to calibrate the decay parameter. The interest rates 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 maximization of the likelihood function logarithm 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. The algorithm was configured to eliminate iterations when it is no longer possible to increase the function value by at least 1.0e-05.

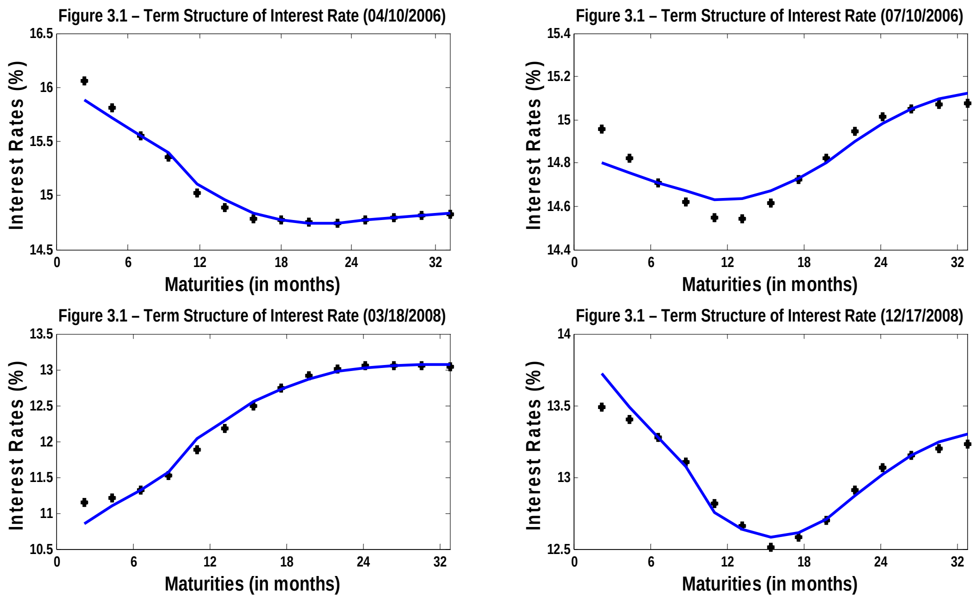
Unlike the two-step method, in the Kalman filter estimation the parameters are estimated in a single step. The lambda parameter governs the decay rate of factor loadings of both the level and curvature, estimated together with other parameters, and not determined a priori. The initial values of the parameters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

Figure 2 – Diebold-Li Factor Loadings (λ= 0.1047)



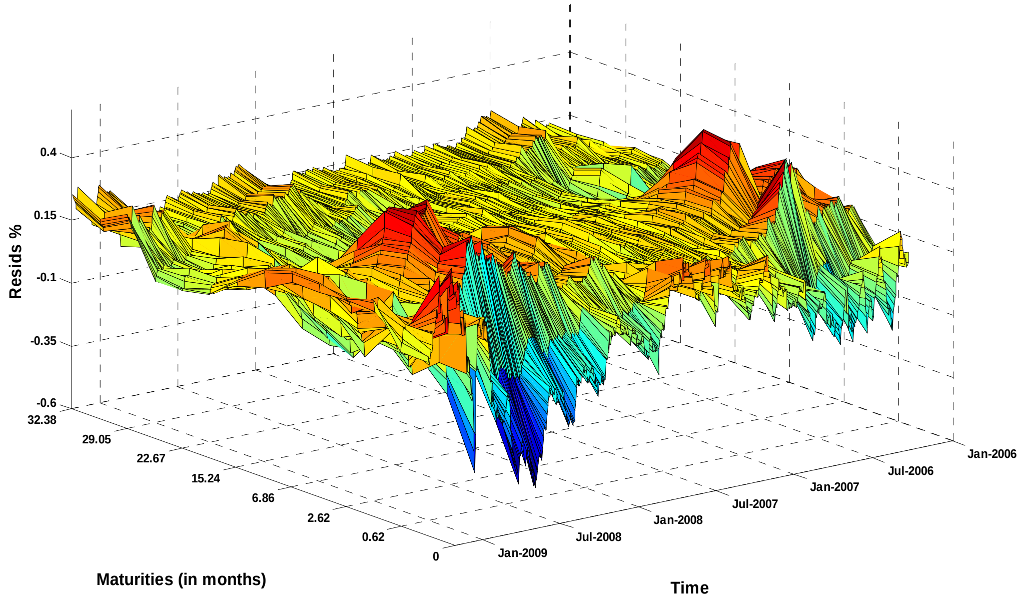
With an estimated λ of 0.1047, the factor loading relative to the curvature assumes maximum value for maturities between 13 and 18 months.

The main argument in favor of Diebold-Li three-factor model (2006) is its capacity to yield good forecasts. Although it is not the best model when the adjustment of the term structure of interest rate is the major goal, the model put forward by Diebold-Li can replicate the several shapes taken by the interest rate curves. Figure 3 shows the real data on the interest rate curve for some days and the curve adjusted by the parameters of the estimated factor model.



Note that the model estimated with three factors fits well to a wide variety of shapes of the interest rate curve: positively sloped, negatively sloped and with different curvature shapes. Figure 4 plots the daily residuals of the interest rate 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 interest rate 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 interest rate 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.

Figure 4 – Residuals of the Term Structure of Interest Rate (Jan 2006 to Feb 2009)



Optimal smoothed estimates were obtained for the three latent factors of the interest rate curve (level, slope and curvature) using the Kalman[7] smoother. Figure 5 shows the time series associated with the factors. Observe that the level is the most stable factor, driving slightly away from the mean, except in the period around October 2008. The curvature is the factor with highest instability, assuming values between -8.39 and 9.35, with a standard deviation of 3.40.

---

[7] For further details on the Kalman filter, see Durbin and Koopman (2001), Anderson and Moore (1979), Hamilton (1994) or Harvey (1989).

Figure 5.1 – Estimated Level of Interest Rate Curve (β_1)

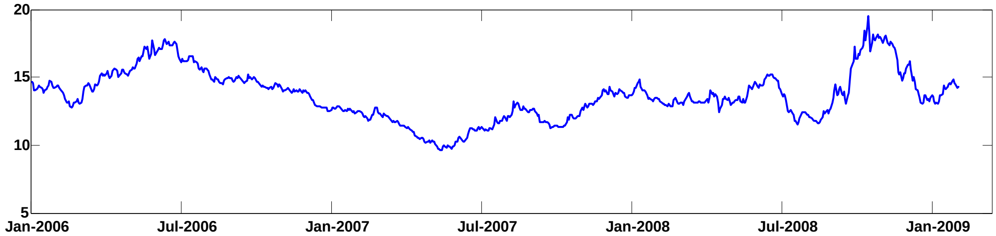


Figure 5.2 – Estimated Slope of Interest Rate Curve (β_2)

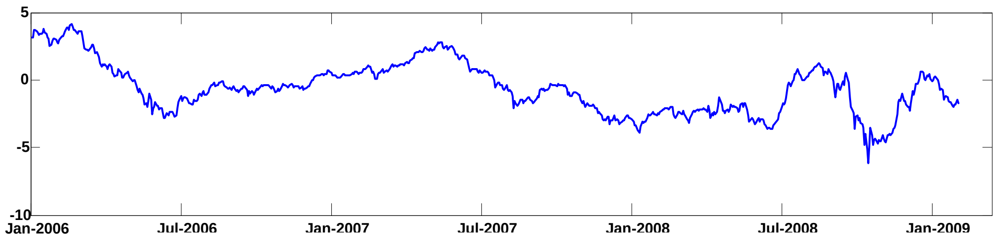
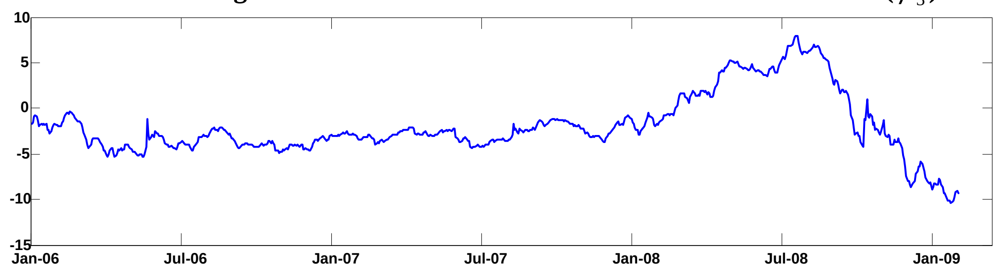


Figure 5.3 – Estimated Curvature of Interest Rate Curve (β_3)



The estimated level of the interest rate curve, β_1, has a statistically significant mean of 13.72%, with high persistence. Note that the level of the interest rate 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 interest rate curve had its highest value (19.47). It should also be observed that there is a sudden change in behavior in the slope and curvature latent 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.

Table 1 – Statistics of the Estimated Latent Factors

	Mean	SD	Max	Min	ρ_1	ρ_{21}	ρ_{63}
β_1	13.72	1.84	19.47	9.60	0.99	0.74	0.33
β_2	-0.72	1.89	4.11	-6.18	0.99	0.67	0.15
β_3	1.46	3.40	9.35	-8.89	0.99	0.79	0.34

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 for four horizons are analyzed: one day, one month, three months and six months. To complete the forecasting exercise, we obtained forecasts via random walk and from the Diebold-Li model estimated by the two-step method[8] (RW and DP).

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

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	-0.72
	(0.003)	(0.002)	(0.001)	(1.35)
$\beta_{3,t}$	-0.003	**-0.012**	**1.000**	1.46
	(0.002)	(0.003)	(0.002)	(1.66)

Standard deviation in parentheses

Table 3 – Covariance Matrix of Estimated Factors

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

Standard deviation in parentheses

The analysis of VAR parameters indicates high persistence of the dynamics of the three 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 $\beta_{2,t-1}$ in $\beta_{3,t}$.

The approach to forecast the interest rate 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 interest rate with maturity τ, are given by:

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

The forecasts of the factors are obtained by the estimated VAR(1) parameters to model the dynamics of the state vector:

$$\hat{\beta}_{t+h/t} = \hat{\mu} + \hat{\mathrm{B}}\hat{\beta}_t$$

[8] The estimation of the model by the two-step method was used with $\lambda = 0.1182$, which maximizes the factor loading of the curvature for the mean maturity of 13 to 15 months.

Tables 5 and 6 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.

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 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 its counterparts in all maturities, except 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 DP.

When the forecasts for the 3 and 6-month horizons are analyzed, we note that the KF consistently outperforms its counterparts. For three-month-ahead forecasts, the KF outperforms the RW and DP in all maturities, showing an RMSE on average 35 basis points lower than those obtained by DP.

Table 4 – RMSE for Out-of-Sample Forecasts (Feb 2008 to Feb 2009)

	One Day Ahead			One Month Ahead		
Maturity	**RW**	**DP**	**FK**	**RW**	**DP**	**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	37.22
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	98.82	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

RMSE expressed in basis points (RW = Random Walk, DP = Two-Step Estimation, KF = Kalman Filter)
Maturity in months

Table 5 - RMSE for Out-of-Sample Forecasts (Feb 2008 to Feb 2009)

	Three Months Ahead			Six Months Ahead		
Maturity	**RW**	**DP**	**FK**	**RW**	**DP**	**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	188.94	143.11	144.43
9	161.48	153.07	114.38	205.41	150.81	166.10
12	176.11	164.42	131.42	222.08	158.53	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

RMSE expressed in basis points (RW = Random Walk, DP = Two-Step Estimation, KF = Kalman Filter)
Maturity in months

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*. (2007). According to their work, liquidity plays an important role in the predictability of interest rates.

4 Conclusion

In the present paper, Diebold and Li model, usually estimated 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 interest rate 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 interest rate 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.

5 References

Almeida C. and J. Vicente (2008). Does Curvature Enhance Forecasting? ***forthcoming* in *International Journal of Theoretical and Applied Finance,*** (2009).

Almeida C. and J. Vicente (2008). The Role of No-arbitrage on Forecasting: Lessons from a Parametric Term Structure Model. ***Journal of Banking and Finance***, 32, 12, 2695-2705.

Almeida C., Vicente J., Gomes R., Leite A (2007). Movimentos da Estrutura a Termo e critérios de Minimização do Erro de Previsão em um Modelo Paramétrico Exponencial. ***Revista Brasileira de Economia***, 62, 4, (2008), 497-510.

Ang. A. and M. Piazzesi (2003). A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. ***Journal of Monetary Economics***, 50, 745-787.

Bali T., M. Heidari, and L. Wu (2007). Predictability of Interest Rates and Interest Rate Portfolios. ***Working Paper***, Baruch College.

Bliss, R. R. "Testing term structure estimation methods". ***Federal Reserve Bank of Atlanta, Working Paper*** 96-12a, November 1996.

Bowsher C. and R. Meeks (2008). The Dynamics of Economic Functions: Modelling and Forecasting the Yield Curve. ***forthcoming in Journal of the American Statistical Association***.

Cajueiro, O. D., Divino, J. A. e Tabak B. M. (2007). Forecasting the Yield Curve for Brazil. ***Working Paper Banco Central do Brasil.***

Cochrane J. and M. Piazzesi (2005). Bond Risk Premia. ***American Economic Review***, 95, 1, 138-160.

Cochrane J. and M. Piazzesi (2006). Decomposing the Yield Curve. ***Working paper, Chicago Graduate School of Business***.

Dai Q. and Singleton K. (2002). Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure. ***Journal of Financial Economics***, 63, 415-441.

Diebold F.X. and C. Li (2006). Forecasting the Term Structure of Government Bond Yields. ***Journal of Econometrics***, 130, 337-364.

Diebold F.X., G.D. Rudebusch, and B. Aruoba (2006). The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. ***Journal of Econometrics***, 131, 309-338.

Duffee G. R. (2002). Term Premia and Interest Rates Forecasts in Affine Models. ***Journal of Finance***, 57, 405-443.

Duffie D. and Kan R. (1996). A Yield Factor Model of Interest Rates. ***Mathematical Finance***, 6, 4, 379-406.

Egorov, A. V., Li, H., David, Ng. (2008). A Tale of Two Yield Curves: Modeling the Joint Term Structure of Dollar and Euro Interest Rates. ***Forthcoming, Journal of Econometrics***.

Fama, E. F. and Bliss R.R. (1987). The information in long-maturity forward rates. ***American Economic Review,*** 77, 4, 680-692.

Fontaine J.S. and R. Garcia (2007). Bond Liquidity Premia. ***Working Paper, Department of Economics, Université de Montréa***l.

Heath D., R. Jarrow and A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. ***Econometrica***, 60, 1, 77-105.

Koopman S.J., M.I.P. Mallee, and M. van der Wel (2008). Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters. ***forthcoming in Journal of Business and Economic Statistics***.

Litterman R. and Scheinkman J. A. (1991). Common Factors Affecting Bond Returns. ***Journal of Fixed Income***, 1, 54-61.

McCulloch J.H. (1971). Measuring the Term Structure of interest Rates. ***Journal of Business***, 44, 19-31.

McCulloch, J. H. "The tax-adjusted yield curve". ***Journal of Finance***, 30, pp. 811-830, June 1975.

Nelson C. and A. Siegel (1987). Parsimonious Modeling of Yield Curves. ***Journal of Business***, 60, 4, 473-489.

Varga, G. (2008). Teste de Modelos Estatísticos para a Estrutura a Termo no Brasil. ***aceito para publicação na Revista Brasileira de Economia***

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

Vasicek, O. A., Fong, H. G. "Term structure modeling using exponential Splines", ***The Journal of Finance***, 37, 339-348, 1982.