Journal of Forecasting, *J. Forecast.* **32**, 111–123 (2013)
Published online 20 November 2011 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/for.1256

Modeling and Forecasting the Yield Curve by an Extended Nelson-Siegel Class of Models: A Quantile Autoregression Approach

RAFAEL B. DE REZENDE[1]* AND MAURO S. FERREIRA[2]
[1] *Department of Finance, Stockholm School of Economics, Sweden*
[2] *Department of Economics, Universidade Federal de Minas Gerais – CEDEPLAR, Belo Horizonte, Brazil*

ABSTRACT
This paper compares the in-sample fitting and the out-of-sample forecasting performances of four distinct Nelson–Siegel class models: Nelson–Siegel, Bliss, Svensson, and a five-factor model we propose in order to enhance the fitting flexibility. The introduction of the fifth factor resulted in superior adjustment to the data. For the forecasting exercise the paper contrasts the performances of the term structure models in association with the following econometric methods: quantile autoregression evaluated at the median, VAR, AR, and a random walk. As a pattern, the quantile procedure delivered the best results for longer forecasting horizons. Copyright © 2011 John Wiley & Sons, Ltd.

KEY WORDS yield curve; in-sample fitting; out-of-sample forecasts; Nelson–Siegel; quantile autoregression

INTRODUCTION

The yield curve plays a central role in macroeconomics and finance. For macroeconomists, it carries information on a variety of variables, such as expected inflation and future gross domestic product (Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1996; , 1998). In finance, the yield curve allows marking to market, pricing derivatives, and hedging, among other uses. Thus it is not surprising that substantial research effort has been applied to estimating and forecasting the yield curve.

The three most popular estimation approaches are: the affine equilibrium models (Vasicek, 1977; Cox *et al.*, 1985; Duffie and Kan, 1996), no-arbitrage models (Hull and White, 1990; Heath *et al.*, 1992), and statistical and parametric models (McCulloch, 1971, 1975; Vasicek and Fong, 1982; Bliss, 1997; Nelson and Siegel, 1987; Svensson, 1994; Fisher *et al.*, 1995).

The first two have not been very successful in terms of forecasting. According to Duffee (2002), equilibrium models only pay attention to instantaneous short rates, resulting in poor yield curve predictions.[1] The same is true for arbitrage-free models, as they are specialized in fitting yield curves at a particular point in time, leaving aside their dynamics.[2]

The Nelson–Siegel class of parametric models, on the other hand, has delivered good estimation and forecasting properties, which has increased its popularity among users. The Bank for International Settlements (2005) reports that nine of the 13 main central banks of the world rely on Nelson and Siegel (1987) model (NS, henceforth) and/or on Svensson (1994) model (SV, henceforth) for their yield curve construction, while Gürkaynak *et al.* (2007) suggest the use of the SV by the Federal Reserve Board. For forecasting purposes, Diebold and Li (2006) show that a dynamic version of the NS model predicts the US yield curve more accurately than other competing models, especially at longer horizons. Moench (2008) partially confirms these results, while Bolder (2006) shows the superiority of the NS for estimation and forecasting purposes when compared to affine equilibrium and other parametric models.

Motivated by the practical use of the Nelson–Siegel models, our paper compares the in-sample fitting and the out-of-sample predictive power of some of the models we propose: NS, Bliss (1987) (BL, henceforth), SV and a five-factor model (FF, henceforth). The decision to include a fifth factor as an extension to the SV model is an attempt to enhance their flexibility in order to improve the in-sample fitting and the predictive power.

* Correspondence to: Rafael B. De Rezende, Department of Finance, Stockholm School of Economics, SE-113 83 Stockholm, Sweden. E-mail: rafael.rezende@hhs.se

[1] However, Ang and Piazzesi (2003) show that the imposition of no-arbitrage restrictions and the insertion of macroeconomic variables improve the forecasts of equilibrium models. Since then, many articles have been written on no-arbitrage macro-finance equilibrium models (Rudebusch and Wu, 2008; Hördahl *et al.*, 2005; Wu, 2006; Moench, 2008).

[2] In a recent paper, however, Christensen *et al.* (2007) introduced the no-arbitrage restrictions to the Nelson–Siegel model, improving its predictability.

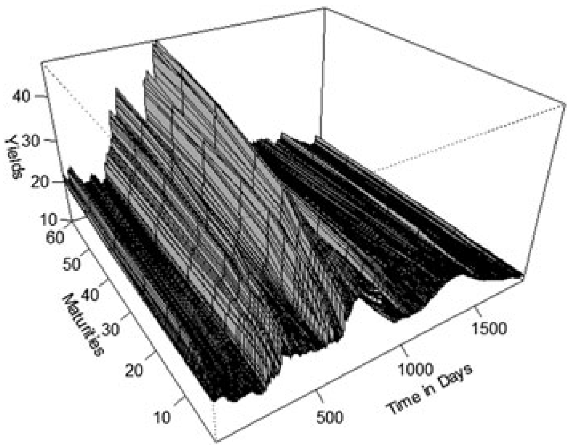

Figure 1. ID × PRE swap yield curves. The sample ranges from 16 March 2000 to 15 October 2007. We consider the following maturities: (in months): 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 16, 24, 36, 48, and 60

The paper also innovates when comparing the forecasting performance of a first-order quantile autoregressive model (QAR),[3] estimated at the median, to other standard procedures: AR(1), VAR(1), and a random walk (RW) applied directly on the term structure yields. The use of a robust estimation method, such as QAR, reduces estimation bias provoked by outliers which normally affects out-of-sample forecasting performance.[4] This robustness seems particularly attractive when dealing with financial variables that normally suffer from wide oscillations, even more in emerging markets.

We conduct the analysis for the Brazilian zero-coupon data, whose characteristics are described in the next section. Although most of the term structure literature focuses on data for the richest countries, the international investors' increasing interest in the high yields of emerging economies, which has been accentuated due to their recent conquest of macroeconomic and institutional stability, naturally requires a better comprehension of such financial markets. This paper takes a step in this direction when analyzing the Brazilian case—one of the most popular destinies for international investors seeking high returns in the emerging world.

In the third section we present three standard term structure models before introducing the FF model. The fourth section describes the in-sample and the out-of-sample estimation procedures, while the fifth section discusses the results. The sixth and last section brings our concluding remarks.

We anticipate the main results of our study: (i) the FF shows superior in-sample fitting; (ii) for the out of sample exercise, both the SV and BL are more accurate; (iii) among the econometric forecasting procedures, QAR is superior to AR and to VAR; (iv) the random walk beats QAR for 1-day-ahead prediction, but (v) QAR shows superior performance for 1- and 3-month horizons.

THE DATA[5]

The work uses daily data on the implicit yield curves extracted from swap operations between the Brazilian Interbank Deposit rate and the Predetermined Interest, which is a fixed coupon rate. The Interbank Deposit rate is a weighted average of daily rates on the interbank lending operations. This swap, referenced by ID-PRE, is probably the most important and liquid instrument of the Brazilian fixed-income market. The contract has the same characteristics of a risk-free zero-coupon since BM&FBOVESPA provides full assurance.

Our sample ranges from 16 March 2000 to 15 October 2007, constituting 1883 business days. Because very long maturity contracts have only recently become more liquid, we decided to work with the following 15 vertices: 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 18, 24, 36, 48 and 60 months (in running days). Figure 1 provides a three-dimensional picture of the data.

Table I reports descriptive statistics for the yields and for the empirical factors: level, slope and curvature.[6] Data reveal that the average yield curve is positively sloped and volatility increases with maturities. For all maturities, the yields are highly persistent over time. They also seem to depart from a normal distribution, as suggested by the

[3] For details regarding the quantile autoregression model see Koenker and Xiao (2002, 2004, 2006).
[4] Ledolter (1989) and Hotta (1993) discuss that outliers lead to bias parameter estimates and inflate the estimated variance of the series, thus harming the quality of out-of-sample forecasts.
[5] The data are provided by BM&FBOVESPA—Securities, Commodities and Futures Exchange and can be found in http://www.bmfbovespa.com.br/en-us/home.aspx?idioma=en-us
[6] The level is the 60-month yield; the slope is the 60-month yield less 1-month yield; and the curvature is defined as twice the 1-year yield less the sum of 1-month and 60-month yields.

Table I. Interest rate descriptive statistics

Maturities	Mean	SD	Skew	Kurt	Min.	Max.	JB p	ρ_1	ρ_{21}	ρ_{63}
1	17.771	3.525	0.547	40.556	11.050	26.950	0.000	0.997	0.967	0.781
2	17.873	3.642	0.486	19.785	11.040	27.390	0.000	0.997	0.963	0.782
3	17.983	3.764	0.437	1.637	11.010	27.770	0.000	0.996	0.960	0.784
4	18.105	3.900	0.399	−14.370	11.020	28.170	0.000	0.996	0.959	0.784
5	18.186	4.018	0.399	−20.813	11.030	28.490	0.000	0.996	0.960	0.784
6	18.259	4.136	0.426	−21.699	10.990	28.990	0.000	0.996	0.959	0.783
7	18.330	4.248	0.454	−21.522	10.910	29.450	0.000	0.996	0.959	0.782
8	18.405	4.366	0.494	−18.014	10.850	30.130	0.000	0.997	0.958	0.779
9	18.473	4.481	0.538	−12.578	10.810	30.650	0.000	0.997	0.958	0.778
12	18.670	4.818	0.679	8.476	10.710	32.690	0.000	0.997	0.957	0.776
16	19.098	5.497	0.893	40.142	10.490	36.900	0.000	0.997	0.957	0.773
24	19.442	6.043	1.025	61.567	10.350	39.280	0.000	0.997	0.958	0.775
36	20.094	6.998	1.144	73.944	10.090	43.380	0.000	0.998	0.958	0.777
48	20.582	7.672	1.192	80.481	9.970	45.980	0.000	0.998	0.958	0.782
60 (Level)	20.944	8.130	1.199	80.924	9.840	47.430	0.000	0.998	0.958	0.787
Slope	3.174	6.171	1.429	163.078	−4.670	27.690	0.000	0.997	0.897	0.708
Curvature	−1.374	2.468	−0.781	15.961	−10.020	4.550	0.000	0.983	0.795	0.588

Note: The table shows summary statistics for the Brazilian swap ID × PRE yields. The sample ranges from 16 March 2000 to 15 October 2007. We report the mean, standard deviation, skewness, excess kurtosis, minimum, maximum, the p-value of the Jarque-Bera (JB) test for normality and the 1st, 21st and 63rd sample autocorrelations.

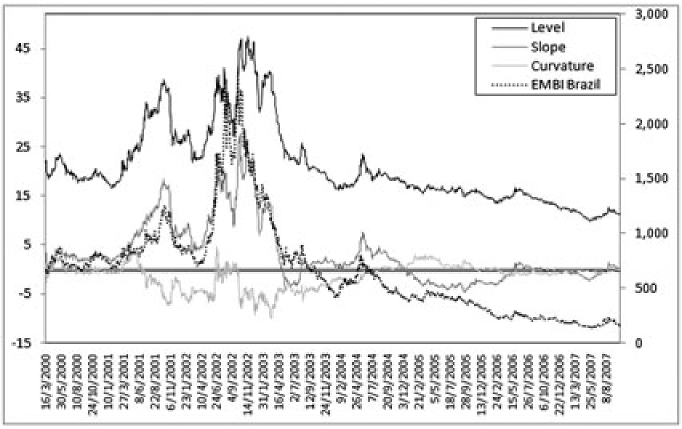

Figure 2. Emperical factors and EMBI: Brazil. Empirical factors are left scaled and the EMBI Brazil is right scaled. The level is defined as the 60-month yield, the slope as the 60-month less 1-month yield and the curvature as twice the 1-year yield less the sum of the 1-month and 60-month yields

positive skewness, either high or low excess kurtosis, and also by the low p-values from the Jarque–Bera normality test. This last pattern is also observed for all empirical factors. The level, however, presents higher persistency and more volatility than the slope and the curvature.

In order to familiarize the readers with the data, we contrast, in Figure 2, the time series of the empirical factors (left scale) and the Brazilian risk premium (right scale), measured in basis points by the Emerging Markets Bond Index for Brazil (EMBI Brazil), which is computed by JP Morgan.

It is straightforward to visualize the close co-movements between these series, reflecting how the Brazilian interest rate policy was strongly affected by the country risk perception. This pattern, observed in several emerging markets, has been studied by Uribe and Yue (2006). Following an adverse shock (normally a contagion, a default or even an increase in the political uncertainty), investors take their money out of the country, provoking a huge currency depreciation that passes through to the local inflation indices. Central banks following an inflation targeting regime, which is the Brazilian situation, react by increasing the targeting interest rate, which ends up oscillating according to the country risk perception.

This is the explanation why the empirical factors and the EMBI Brazil varied so abruptly from the beginning of 2001 until the middle of 2003. During this period Brazil suffered contagion from Argentine's sovereign default in December of 2001 and from the increase in the international risk aversion following the 9/11 attack. The country also had to deal with a local political uncertainty during the 2002 presidential election in October, since the leading candidate, Lula, and his Labor Party had always announced that they would deviate from orthodox economic policies.

The analysis of Uribe and Yue (2006) also explains the coincident decline in the risk premium and the empirical factors. As Lula positively surprised by honoring debt contracts, increasing budget surplus and maintaining the inflation targeting framework, the risk premium declined, attracting foreign investors back to Brazil and appreciating the local currency, which helped in lowering inflation. As a result, the term structure level declined with the perspective that the Central Bank would reduce further the interest rate.

NELSON–SIEGEL CLASS MODELS

Basic definitions

The term structure of interest rates can be described in terms of the spot (or zero-coupon) rate, the discount rate and the forward rate. A forward rate $f(\tau^*, \overline{\tau})$ is the interest rate of a forward contract on an investment which will be initiated $\overline{\tau}$ periods in the future and will mature τ^* periods beyond the start date of the contract. The instantaneous forward rate $f(\overline{\tau})$ is defined as $\lim_{\tau^* \to 0} f(\tau^*, \overline{\tau}) = f(\overline{\tau})$ from which one obtains the forward curve, $(f\tau)$.

The spot rate, $y(\overline{\tau})$, implicit in a zero-coupon bond with maturity $\overline{\tau}$, is defined as $y(\overline{\tau}) = \frac{1}{\tau}\int_0^{\tau} f(x)\mathrm{d}x$, from which one gets the spot curve (or zero-coupon curve), $y(\tau)$.

The discount curve $d(\tau)$ is formed by the present value of a zero-coupon bond paying the nominal value of \$1.00 after τ periods and is obtained as follows: $d(\tau) = e^{-y(\tau)\tau}$. It is then straightforward to move from one curve to another using the relations $d(\tau) = \exp\left[-\int_0^{\tau} f(x)\mathrm{d}x\right]$ and $f(\tau) = -\frac{d'(\tau)}{d(\tau)}$.

Three-factor Nelson–Siegel (1987) model

Nelson and Siegel (1987) fitted a zero-coupon curve at a particular point in time using the following model:

$$y(\tau) = \beta_1 + \beta_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \varepsilon^{\tau} \tag{1}$$

Later, Diebold and Li (2006) introduced dynamics to this model by regressing, period by period, the yields on the exponential components of (1), which resulted in time series for its coefficients: $\beta_{1t}, \beta_{2t}, \beta_{3t}$ and λ_t. Following Litterman and Scheinkman (1991), Diebold and Li (2006) interpreted them as latent factors of level (β_{1t}), slope (β_{2t}) and curvature (β_{3t}), while the exponential terms (inside parentheses of equation (1)) were interpreted as factor loadings whose shapes depend on the decaying parameter (λ_t). For matters of simplicity, Diebold and Li (2006) fixed λ_t at $\lambda = 0.00609$ for every t, which allowed (1) to be estimated by ordinary least squares (OLS). In Figure 3(a) we plotted some specific shapes of these exponential components.

Despite its importance, the Nelson–Siegel model has gone through several modifications to enhance flexibility in order to capture a wider variety of curve shapes. These improvements were mainly done by incorporating additional factors and decaying parameters.

Bliss (1997) three-factor model

Bliss (1997) tried to improve the term structure fitting by incorporating two different decaying parameters: λ_1 and λ_2. The dynamic spot curve is then given by

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

The factor loading of equation (2) is shown in Figure 3(b).

Svensson (1994) four-factor model

Svensson (1994) included another exponential term, similar to the third, but with a different decaying parameter. The SV dynamic four factor model is written as

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \beta_{4t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) + \varepsilon_t^{\tau} \tag{3}$$

The fourth component can be interpreted as a second curvature. The factor loadings of equation (3) are shown in Figure 3(c).

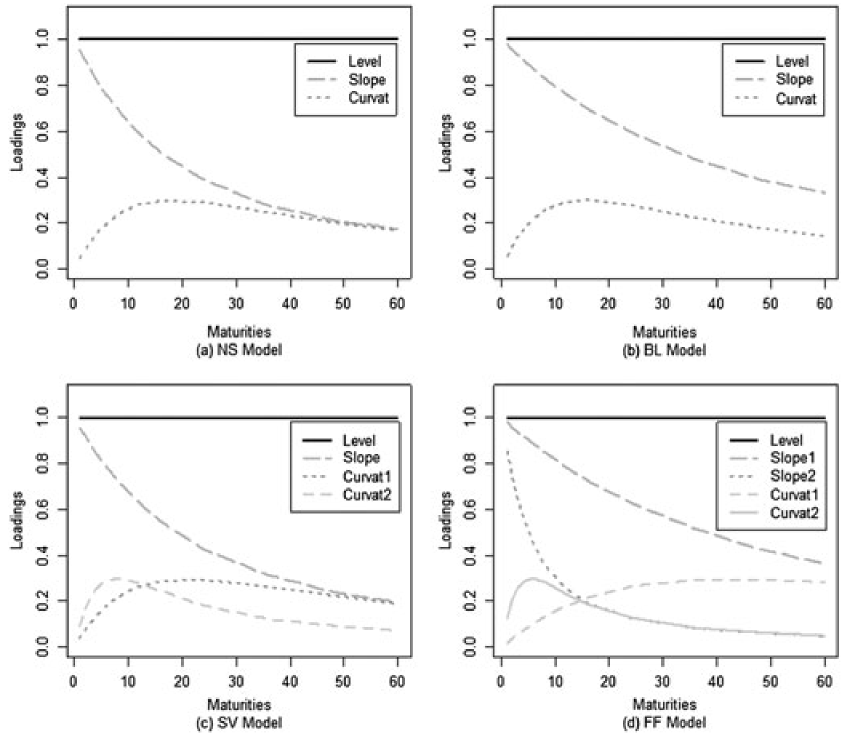

Figure 3. Loadings of the NS class models. The factor loadings of the NS, BL, SV and FF models are plotted considering $\lambda = 0.097$, $\lambda_1 = 0.048$ and $\lambda_2 = 0.114$, $\lambda_1 = 0.084$ and $\lambda_2 = 0.222$, $\lambda_1 = 0.042$ and $\lambda_2 = 0.32$, respectively

Five-factor model

We propose a five-factor (FF) model, as a natural extension of the SV, by including a term resembling the second loading, but with a different decaying parameter:

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau}\right) + \beta_{4t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \beta_{5t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) + \varepsilon_t^{\tau} \quad (4)$$

The third factor loading, shown in Figure 3(d), can be interpreted as a second slope. We expect the FF model to enhance the fitting of more complex and twisted curves.

ESTIMATION METHODS

The in-sample fitting

The models presented in the previous section can be nested in the following representation:

$$Y_t = X_t\beta_t + \varepsilon_t \quad (5)$$

where $Y_t = [y_t(\tau_1), y_t(\tau_2), \ldots, y_t(\tau_N)]'$ is the $N \times 1$ column vector representing the term structure of all N interest rates at period t; X_t is the $N \times F$ matrix of factor loadings, F being the number of factors in each model; β_t is the $F\times 1$ vector of latent factors; and ε_t is the $N\times 1$ vector of errors with typical element, ε_{nt}, for $n = 1,..., N$, satisfying $\varepsilon_{nt,} \sim$ i.i.d.$N(0, \sigma_n^2)$.

Fitting the term structure according to (5) requires, for each period t, estimating the vector β_t and the decaying parameters λ_t, $\lambda_{1,t}$, and $\lambda_{2,t}$. Following Nelson and Siegel (1987) and Diebold and Li (2006), we keep these last parameters constant for all t, i.e. $\lambda_t = \lambda, \lambda_{1,t} = \lambda_1$, and $\lambda_{2,t} = \lambda_2$.

Diebold and Li (2006) decided to fix λ at 30 months maturity,[7] but we adopted a less arbitrary strategy. For choosing the decaying parameter for the NS model, we initially constrained the range of λ between 0.03 and 0.42,

[7] Diebold and Li (2006) chose λ that maximized the curvature of the factor loadings at the average of 2- and 3-year yields, i.e. the 30th-month maturity. They argue that, historically, the curvature has been linked to changes in the medium-term yields, which represents a value between 2- and 3-year yields.

Table II. Optimal parameters λ

Models	In-sample fitting	Out-of-sample forecasting
NS	0.097	0.095
BL	0.048; 0.114	0.048; 0.112
SV	0.084; 0.222	0.084; 0.229
FF	0.042; 0.320	0.046; 0.259

because these values correspond to a maximum of the curvature loadings at 5 years and 0.05 years (0.6 months), respectively.[8] Given these boundaries, we constructed the set $\Omega = \{0.029 + 0.001l\}_{l=1}^{291}$. Given $\lambda \in \Omega$ and the correspondent matrix of factor loadings X_t, the vector β_t is estimated by OLS at each period t. We chose the optimal decaying factor $\hat{\lambda} \in \Omega$ that minimizes the average of the root mean squared error (RMSE) computed at each period t. More specifically, $\hat{\lambda}$ is the solution of the following problem:

$$\hat{\lambda} = \underset{\lambda \in \Omega}{\arg\min} \left\{ \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(y_t(\tau_n) - \hat{y}_t(\tau_n, \lambda, \hat{\beta}_t) \right)^2} \right\} \tag{6}$$

where T is the number of yield curves in the sample.

In the case of the BL, SV and FF we solved a similar problem, but now we had to find the values $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$ from the set $\Lambda = \left\{ \left(\lambda_1^l, \lambda_2^l\right) | \lambda_1 \in \Omega, \lambda_2 \in \Omega \right\}$, i.e. the Cartesian product $\Omega \times \Omega$. Thus $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$ solves the following problem:

$$\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = \underset{\left(\hat{\lambda}_1, \hat{\lambda}_2\right) \in \Lambda}{\arg\min} \left\{ \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(y_t(\tau_n) - \hat{y}_t(\tau_n, \lambda_1, \lambda_2, \hat{\beta}_t) \right)^2} \right\} \tag{7}$$

Table II shows the values of optimal decaying parameters for each model.

The out-of-sample forecasting

We use the first T^* yield curves to set $\hat{\lambda}$ and $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$ based on the procedure described in the previous subsection.[9] As the algorithm selects the optimal values for the decaying parameters, it simultaneously defines a time series of estimated factors $\hat{\beta}_{f,t}$ for $t = 1, \ldots, T^*$ and $f = 1, \ldots, F$. Given this time series we estimate the equation

$$\hat{\beta}_{f,t} = \mu_f + \phi_f \hat{\beta}_{f,t-h} + \nu_{f,t} \tag{8}$$

where h is the forecasting horizon, μ_f and ϕ_f are parameters and $\nu_{f,t}$ is the residual whose restrictions we will comment on later.

Let $m = 0, \ldots, (T - T^* - h)$. Given $\hat{\lambda}$ and $\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$, equation (5) and (8) can then be used to forecast each factor in the following iterative manner: (i) use the data available up to $T^* + m$ and estimate the parameters $\hat{\mu}_f^{(T^*+m)}$ and $\hat{\phi}_f^{(T*+m)}$ according to (8), where the indexed superscript indicates the last observation used in the regression; (ii) obtain $\tilde{\beta}_{f,T^*+h+m|T^*+m} = \hat{\mu}_f^{(T^*+m)} + \hat{\phi}_f^{(T*+m)} \hat{\beta}_{f,T^*+m}$, where $\tilde{\beta}_{f,T^*+h+m|T^*+m}$ is the h-step-ahead forecasted value of $\hat{\beta}_{f,T^*+h+m}$, given the information available at the moment $T^* + m$; (iii) forecast the vector of yields $\hat{Y}_{T^*+h+m|T^*+m}$ using $\hat{Y}_{T^*+h+m|T^*+m} = \hat{X}\left(\hat{\lambda}\right) \tilde{\beta}_{T^*+h+m|T^*+m}$ or $\hat{Y}_{T^*+h+m|T^*+m} = \hat{X}\left(\hat{\lambda}_1, \hat{\lambda}_2\right) \tilde{\beta}_{T^*+h+m|T^*+m}$.

We are particularly interested in contrasting the forecasting power of the following autoregressive models: AR(1), VAR(1), and QAR(1). When equation (8) is estimated according to AR(1), $\nu_{f,t} \sim$ i.i.d.$N\left(0, \sigma_{\nu_f}^2\right)$, for each f. In the case of QAR(1), we only assume independency of $\nu_{f,t}$ over t. In both cases, we assume independency of the errors

[8] There is no reason to search for optimal values outside this interval, since the greatest time to maturity in our database is 5 years, and the shortest is 1 month. Given a value for λ, the solution of the equation $1 + \lambda\tau + (\lambda\tau)^2 = e^{\lambda\tau}$ for τ gives the time to maturity where the curvature loading achieves its maximum.

[9] Laurini and Hotta (2007) and Gimeno and Nave (2006) argue that the estimation of the NS models by nonlinear least squares, especially the SV, generates extremely unstable time series for β, showing that the term structure forecasts are more accurate when λ_t is fixed. In this paper, all models were estimated considering a fixed value for λ_t.

across regressions, i.e. $\nu_{f_1,t} \perp \nu_{f_2,t}$, for any $f_1 \neq f_2$. This last assumption is dropped in the case of the VAR(1) estimates, in which case (8) needs to be rewritten so that, instead of estimating μ_f, β_f, and $\sigma^2_{\nu f}$, we respectively estimate the matrices M, Φ, and Ω. The VAR(1) is represented by

$$\hat{\beta}_t = \mathrm{M} + \Phi\hat{\beta}_{t-h} + \nu_t \tag{9}$$

where $\hat{\beta}_t$ is an $F \times 1$ vector of all F factors; M is an $F \times 1$ vector representing the intercept of each equation; Φ is an $F \times F$ matrix; ν_t is an $F \times 1$ vector of serially uncorrelated residuals; and Ω is the $F \times F$ variance–covariance matrix where it is allowed that $\mathrm{cov}(\nu_f, \nu_g) \neq o$ for any $\nu_f \neq \nu_g$.

We compare the forecasting performance of each combination of *term structure model* / *econometric method* against each other and also against the predictions of a random walk applied directly on the yields: $\hat{Y}_{\tau_n,T^*+h+m|T^*+m} = Y_{\tau_n,T^*+m}$.

EMPIRICAL RESULTS

The forecasting exercise is carried considering, $h = 1, h = 21$ (1 month), and $h = 63$ (3 months). We have a sample of $T = 1883$ yield curves, from which we set $T^* = 1400$. Table II shows the values $\hat{\lambda}$ and $(\hat{\lambda}_1, \hat{\lambda}_2)$ for the in-sample and for the out-of-sample exercises.

In-sample fitting

Table III shows the RMSE for all four term structure models. The first thing to observe is that the FF has the smallest RMSE in all maturities analyzed. In some cases it is only a little smaller than the results of the SV model, which also always delivers smaller RMSE than those reached by NS and BL. Incorporating the fifth factor resulted in a large improvement for the shortest (1 month) and longest (60 months) maturities—exactly those considered the most difficult to fit. For these two vertices, the SV delivered an RMSE that was, respectively, 75% and 170% higher than the FF model. In the same direction, the SV resulted in a much smaller RMSE, at these extreme vertices, than the BL and NS. It is strongly suggested that more flexibility improves the curve adjustment, especially at extreme maturities.

In order to compare the performance of each model over time, we computed the RMSE for the entire curve of yields at each period *t*. The FF delivered smaller RMSE than the SV in 84.65% of the time. The superiority against

Table III. RMSE of each term structure model in-sample adjustment

Model	Average	Maturities (months)									
		1	2	4	6	8	12	24	36	48	60
NS	0.1735	0.2755	0.1252	0.1464	0.1554	0.1201	0.1540	0.2903	0.2182	0.1034	0.2401
BL	0.1657	0.2513	0.0996	0.1430	0.1410	0.1134	0.1646	0.2485	0.2087	0.1058	0.2321
SV	0.1119	0.1236	0.0679	0.0977	0.0611	0.0705	0.1462	0.1284	0.1610	0.0988	0.1370
FF	0.0859	0.0706	0.0673	0.0680	0.0555	0.0641	0.1206	0.1264	0.0666	0.0829	0.0507

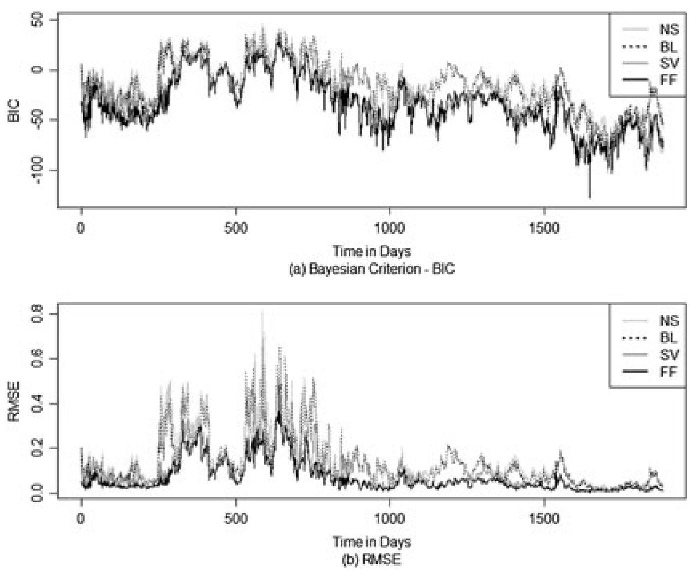

Figure 4. Fitting BIC and RMSE: NS class models

BL and NS occurred in 99.78% and 99.31% of the sample, respectively. Figure 4(a) presents the time series of these RMSE, from which we observe FF normally reaching the smallest value.

We also computed a time series of the Bayesian information criterion (BIC) to verify whether the greater flexibility that resulted in smaller RMSE compensates for losing degrees of freedom due to addition of extra parameters. Again, the result favors the FF 65.90% of the time compared to the SV model. The BIC of the FF was also smaller than those of BL and NS in 89.33% and 91.82% of our sample, respectively. Figure 4(b) presents the BIC time series, from which we can inspect the superiority of the FF model.

To gain better intuition into why more flexible models have delivered better results, we verify, in Figure 5, their fitting against different yield curve shapes. Figure 5(a), (b) and (c) shows that all models perform similarly well when facing less twisted curves. However, for curves presenting at least one inflection point, as indicated by Figure 5(d), (e) and (f), the most flexible models (SV and FF) do a better adjustment. In these last three figures we also observe that FF performs better than SV when exposed to more twisted curves.

Out-of-sample forecasting

Table IV reports the out-of-sample forecasting RMSE for each term structure model (equation (5)) and for each forecasting method (equation (8), for AR and QAR; equation (9) for VAR). Table V reports similar statistics for the random walk applied directly on yield levels. These tables need to be read together. Shaded boxes indicate the best forecasting model at each maturity; bold values, only present in Table IV, indicate outperformance over the RW forecasts.

As a pattern we observe the random walk reaching the lowest RMSE for 1-day horizon for all maturities but 24 and 60 months. For these two vertices, the best models were SV-AR and SV-QAR, respectively. The BL-QAR delivered the lowest RMSE for 1 month horizon for all maturities, except 1 month and 60 months, when the best models were NS-QAR and SV-QAR, respectively. Finally, the SV-QAR reached the lowest RMSE for the 3-month-ahead forecasting exercise for all maturities.

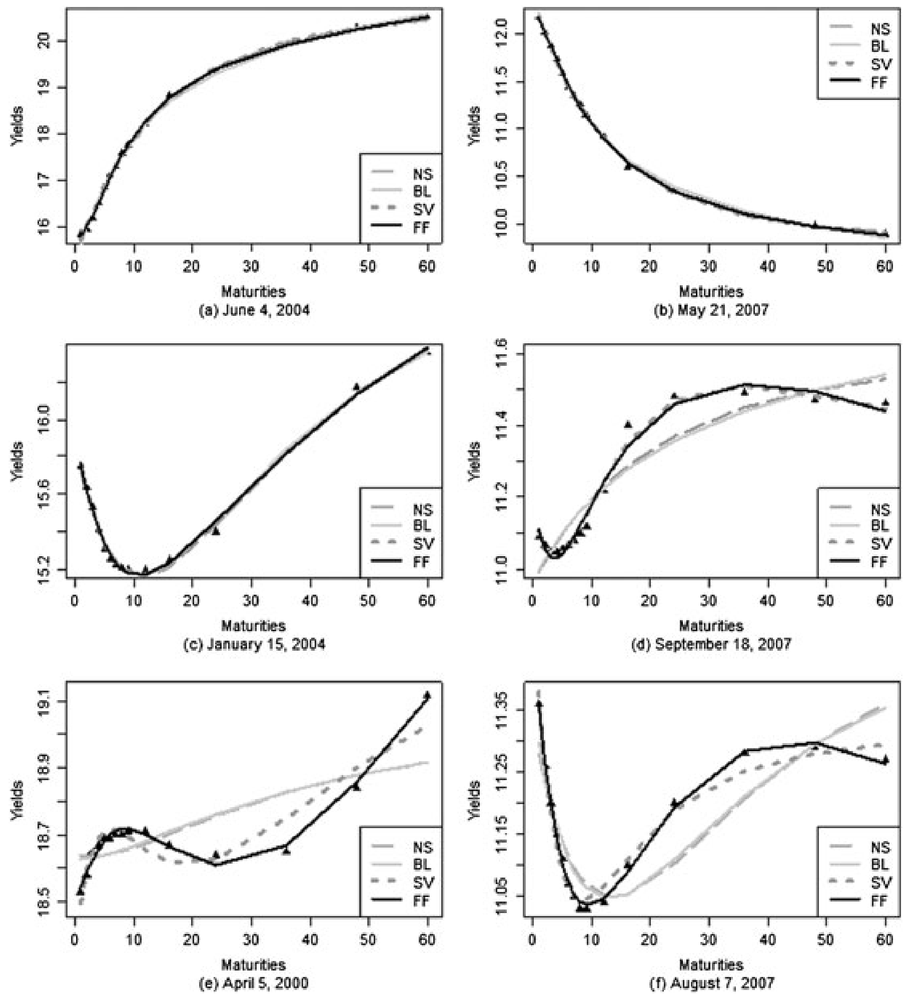

Figure 5. Fitted yield curves in specific days. Observed and fitted yield curves for the following days are plotted

Table IV. RMSE of out-of-sample forecasts: NS class models

Maturities (months)	AR(1)			VAR(1)			QAR(1)		
	1 day	1 month	3 months	1 day	1 month	3 months	1 day	1 month	3 months
					NS				
1	0.0976	0.4418	1.2760	0.1011	**0.1929**	**0.8368**	0.1015	**0.1914**	**0.8511**
3	0.0378	0.5216	1.3795	0.0411	0.4383	1.3625	0.0373	**0.2653**	**0.8623**
6	0.0955	0.6267	1.5268	0.1076	0.7214	1.9709	0.0940	**0.3616**	**0.9140**
12	0.0886	0.6696	1.6747	0.1001	0.9603	2.6254	0.0871	**0.4076**	**0.9905**
24	0.1570	0.7732	1.9533	0.1465	1.1207	3.0799	0.1590	**0.5457**	**1.2081**
60	0.1836	1.0848	2.5413	0.1901	1.2858	3.2215	0.1801	0.7297	1.5508
Average	0.1100	0.6863	1.7253	0.1144	0.7866	2.1828	0.1098	**0.4169**	**1.0628**
					BL				
1	0.0833	0.4906	1.3359	0.0859	**0.1957**	**0.7971**	0.0907	**0.1938**	**0.9175**
3	0.0398	0.5825	1.4968	0.0422	0.4404	1.3572	0.0376	**0.2521**	**0.8962**
6	0.0938	0.7069	1.7080	0.0998	0.7226	1.9853	0.0820	**0.3428**	**0.9188**
12	0.0889	0.7750	1.9182	0.0984	0.9592	2.6252	0.0872	**0.3971**	**0.9921**
24	0.1475	0.8804	2.1965	0.1439	1.1140	3.0292	0.1558	**0.5415**	**1.2039**
60	0.1780	1.1306	2.6899	0.1827	1.2943	3.2551	0.1700	0.7203	1.5272
Average	0.1052	0.7610	1.8909	0.1088	0.7877	2.1748	0.1039	**0.4079**	**1.0759**
					SV				
1	0.0634	0.6135	1.4585	0.0548	**0.2678**	**0.7627**	0.0589	0.4838	**0.6728**
3	0.0382	0.7009	1.6472	0.0386	0.4883	1.3676	0.0389	0.4897	**0.7116**
6	0.0604	0.8066	1.8198	0.0647	0.7691	2.0036	0.0560	0.5375	**0.7872**
12	0.0932	0.8218	1.8751	0.1087	1.0326	2.6143	0.0911	0.5208	**0.8615**
24	**0.1216**	0.8359	1.9631	0.1298	1.2261	3.0229	0.1223	0.5694	**1.0827**
60	0.1469	1.0090	2.3820	0.1627	1.4020	3.2344	**0.1451**	**0.6710**	**1.3443**
Average	0.0873	0.7979	1.8576	0.0932	0.8643	2.1676	0.0854	0.5454	**0.9100**
					FF				
1	0.1212	1.4974	1.3616	0.0456	**0.2512**	**0.7485**	0.0746	1.0858	1.6758
3	0.1084	1.2797	1.1305	0.0373	0.4887	1.3642	0.0842	1.0179	1.5496
6	0.1339	1.0427	**1.0080**	0.0630	0.8052	2.0445	0.0768	0.9188	1.4211
12	0.1501	0.8940	1.1139	0.1032	1.1138	2.7100	0.1175	0.9003	1.4218
24	0.1706	0.8361	1.3941	0.1280	1.3178	3.1171	0.1441	0.9300	1.5017
60	0.2059	0.8473	2.0040	0.1644	1.5334	3.4143	0.1605	0.8769	1.4778
Average	0.1484	1.0662	1.3354	0.0902	0.9184	2.2331	0.1096	0.9550	1.5080

Note: The table shows the out-of-sample forecasts RMSEs from each combination of term structure model (NS, BL, SV and FF) with a different forecasting method: AR(1), VAR(1) and QAR(1). Bold values indicate that the model beats the RW and shading indicates the best model for each maturity.

Table V. RMSE for the RW out-of-sample forecasts

Maturities (months)	RW		
	1 day	1 month	3 months
1	0.0386	0.3851	1.1251
3	0.0369	0.3832	1.1120
6	0.0522	0.4011	1.0812
12	0.0845	0.4650	1.1139
24	0.1234	0.5993	1.2872
60	0.1464	0.6821	1.3910
Average	0.0803	0.4860	1.1851

The results show that the various combinations of term structure model/econometric method rarely delivered RMSE lower than those of the random walk. For 1-day horizon forecasting, it occurred only on three occasions: for the 24 months maturity using SV-AR and SV-QAR and for 60 months maturity using SV-QAR. For 1 and 3 months forecasting horizon, the random walk RMSE was lower than almost all forecasting relying on AR and VAR.

Table VI. Ratios of forecast RMSEs and Diebold-Mariano tests

Maturities	SV-QAR(1)/RW			BL-QAR(1)RW			SV-QAR(1)/BL-QAR(1)		
	1 day	1 month	3 months	1 day	1 month	3 months	1 day	1 month	3 months
1	1.5257***	1.2564***	0.5979**	2.3491***	0.5032***	0.8155	0.6495***	2.4968***	0.7332**
3	1.0539	1.2779***	0.6399**	1.0167	0.6579***	0.8060	1.0366	1.9424***	0.794**
6	1.0714***	1.3398***	0.7281*	1.5698***	0.8545**	0.8498	0.6825***	1.5681***	0.8568
12	1.0781***	1.1201*	0.7734*	1.0319**	0.854**	0.8906	1.0448***	1.3116***	0.8684*
24	0.9907	0.95*	0.8411*	1.2625***	0.9035**	0.9353	0.7847***	1.0515	0.8994
60	0.9914	0.9838	0.9665	1.1614***	1.0560	1.0979	0.8536***	0.9316**	0.8803*

Note: The table presents prediction RMSEs ratios and the Diebold-Mariano forecast accuracy comparison tests for the best forecasting models under review. The null hypothesis is that the two forecasts have the same mean squared errors. Asterisks denote significance relative to the asymptotic null distribution: 10%*; 5%**; 1%***.

The exception occurred only for the 1-month maturity, in which case all four term structure models had smaller RMSE when relying on VAR.

Matters are quite different, however, when we use QAR to predict 1 and 3 months ahead. In these situations, the random walk always had a larger RMSE than those computed by at least one of the term structure models. NS-QAR and BL-QAR had smaller RMSE than RW for all maturities, expect for 60 months when the forecasting horizon was 1 month and 3 months, respectively. The use of SV-QAR resulted in smaller RMSE than that of the RW in absolutely all maturities when conducting a 3-month-ahead forecasting exercise. SV-QAR also had better RMSE in forecasting 24 and 60 months maturities for 1 day and 1 month forecasting horizon.

This advantage of the QAR has certainly to do with the robustness of the median, which is not affected by the presence of outliers and extreme values so commonly observed in financial data.

When we simply compare the models against themselves, we find that the BL-QAR generated, on average, a 40% more accurate 1-month-ahead forecast than the best non-QAR method, which is NS-AR. Similarly, comparison shows SV-QAR generating 32% more accurate 3-month-ahead prediction than the FF-AR, which was the best non-QAR method.

Interesting to note is that, while the inclusion of the second slope term, done by the FF model, improved the in-sample fitting, a similar gain was not translated into better forecasting performance. In general, the FF resulted in higher RMSE, which may have been caused by over-parameterization (Diebold and Li, 2006).

In Table VI we present the ratios of the forecasting RMSE for the RW, SV-QAR and BL-QAR. We only compare BL-QAR and SV-QAR against RW and against each other because these were the models delivering the smallest RMSE for 1 month and 3 months forecasting horizons, respectively. The SV-QAR also had the smallest average RMSE for 1 day forecasting horizon. Table VI also shows the results for the Diebold and Mariano (1995) (DM) test, which compares the forecasting accuracy based on the null hypothesis that both models generate equal forecasting mean squared errors. Asterisks indicate that the DM test rejects the null at 1% (***), 5% (**), and 10% (*) significance levels. Important to note is that when comparing SV-QAR/RW and BL-QAR/RW, a value larger (smaller) than one indicates the RW has the smallest (largest) RMSE. When contrasting SV-QAR against BL-QAR, a value larger (smaller) than one indicates BL-QAR has the smallest (largest) RMSE.

1-day-ahead forecasting

The RW was better than the SV-QAR in forecasting 1, 6 and 12 months maturities, if we consider a 1% confidence level, but the DM test does not allow concluding superiority of the RW forecasts for 3, 24 and 60 months maturities. Compared to BL-QAR, the random walk cannot be considered superior only when predicting the 3 months maturity yield.

When confronting SV-QAR against BL-QAR, the results are mixed but the DM test indicates superiority of the first, at 1% significance level, in forecasting 1, 6, 24 and 60 months maturities. BL-QAR seems better only in predicting the 12 months yield, and the test does not reject the null in the case of the 3 months maturity.

1-month-ahead forecasting

The superiority of BL-QAR over RW is confirmed, since the DM test rejected the null in favor of a smaller RMSE for the BL-QAR in all maturities but 60 months. In this last case, the null was not rejected. The opposite result was found when confronting SV-QAR against the RW, since this last method delivered better forecasting for all maturities, except for 24 and 60 months.

In the comparison between SV-QAR and BL-QAR, the test indicates the last model predicts more accurately, at 1% significance level, for all maturities, except for 24 and 60 months. For 24 months maturity the test did not reject the null, but for 60 months rejection occurred at 5%, indicating superiority of SV-QAR.

3-month-ahead forecasting

When comparing SV-QAR with RW, the test rejected the null for all maturities but 60 months, indicating a better forecasting performance of SV-QAR. For the longest maturity, the test did not reject the null. The DM also did not reject the null, at any maturity, when confronting BL-QAR with RW.

In the case of SV-QAR against BL-QAR, the test favored the first one for 1, 3, 12, and 60 months maturity. For 6 and 24 months there was no rejection.

CONCLUDING REMARKS

We compare, for the Brazilian yield curve data, the in-sample adjustment and the out-of-sample forecasting performance of some Nelson–Siegel class models, following Diebold and Li (2006). In order to improve the ability to fit more twisted curves, very commonly observed in term structures of emerging markets, we propose a new five-factor

(FF) model, which is a natural extension of Svensson (1994). We also innovate by using quantile autoregression (QAR), evaluated at the median, to forecast the yields.

For the in-sample fitting, the FF delivered the best results as indicated by RMSE and BIC statistics. Visual inspection of a few very twisted curves suggests this superiority arriving from better adjustment at the most extreme yields.

The FF performed poorly in out-of-sample forecasting, which may be due to over-parameterization (Diebold and Li, 2006).

For 1-day-ahead forecasting, no combination of term structure model/forecasting method beat the random walk applied directly to the yields. However, the Diebold and Mariano test indicated no superiority of the random walk forecasts against the Svensson‘s model (SV) combined with a QAR, for the 3, 24 and 60 months maturity.

For 1-month-ahead forecasting, we found the Bliss model (BL)-QAR delivering superior results to the random walk for all maturities but 60 months, in which case we found no difference between both methods.

For 3-month-ahead forecasting, the best model was SV-QAR, which beat the random walk in all maturities but 60 months.

While the best term structure model to beat the random walk forecasts depends on the prediction horizon, the forecasts based on QAR evaluated at the median were consistently better. The robustness of this estimator against outliers probably justifies our findings, especially because we have dealt with very volatile financial data, characterized by the presence of extreme values that tend to bias mean estimators.

The analysis of this paper can be extended in at least three directions. First, no-arbitrage restrictions (Christensen *et al.*, 2007) could be applied to more flexible models such as FF, SV and BL. Second, macroeconomic variables should be considered as natural covariates to improve the in-sample adjustment and the out-of-sample forecasts, as suggested by Diebold *et al.* (2006). Finally, quantile regression, or any other robust estimation procedure, should be included among our standard forecasting methods menu.

REFERENCES

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

Bank for International Settlements. 2005. Zero-coupon yield curves: technical documentation. *BIS papers No. 25* Available: http://www.bis.org/publ/bppdf/bispap25.pdf [11 September 2011].

Bliss RR. 1997. Testing term structure estimation methods. *Advances in Futures and Options Research* **9**: 197–231.

Bolder DJ. 2006. *Modelling term-structure dynamics for risk management: a practitioner's perspective*: Bank of Canada Working Paper No. 2006–48.

Christensen J, Diebold F, Rudebusch G. 2007. *The affine arbitrage-free class of Nelson–Siegel term structure models*. NBER Working Papers No. 13611.

Cox J, Ingersoll JE, Ross SA. 1985. A theory of the term structure of interest rates. *Econometrica* **53**: 385–407.

Diebold FX, Li C. 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* **130**: 337–364.

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–263.

Diebold FX, Rudebusch GD, Aruoba B. 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* **131**: 309–338.

Duffee G. 2002. Term Premia and interest rate forecasts in affine models. *Journal of Finance* **57**: 405–443.

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

Estrella A, Hardouvelis G. 1991. The yield curve as a predictor of real economic activity. *Journal of Finance* **46**: 555–576.

Estrella A, Mishkin FS. 1996. The yield curve as a predictor of U.S. recessions. *Current Issues in Economics and Finance – Federal Reserve Bank of New York* **2**: 1–6.

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics* **1**: 45–61.

Fisher M, Nychka D, Zervos D. 1995. *Fitting the term structure of interest rates with smoothing splines*. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series No. 1995–1.

Gimeno R, Nave J. 2006. *Genetic algorithm estimation of interest rate term structure*. Documentos de Trabajo del Banco de España No. 0634.

Gürkaynak R, Sack B, Wright J. 2007. The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics* **54**: 2291–2304.

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**: 77–105.

Hördahl P, Tristani O, Vestin D. 2005. A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics* **131**: 405–444.

Hotta LK. 1993. The effect of additive outliers on the estimates of aggregated and disaggregated ARIMA models. *International Journal of Forecasting* **9**: 85–93.

Hull J, White A. 1990. Pricing interest-rate-derivative securities. *Review of Financial Studies* **3**: 573–592.

Koenker R, Xiao Z. 2002. Inference on the quantile regression process. *Econometrica* **70**: 1583–1612.

Koenker R, Xiao Z. 2004. Unit root quantile autoregression inference. *Journal of the American Statistical Association* **99**: 775–787.

Koenker R, Xiao Z. 2006. Quantile autoregression. *Journal of the American Statistical Association* **101**: 980–990.

Laurini M, Hotta L. 2007. *Bayesian extensions to the Diebold and Li term structure model*. IBMEC Working Paper No. 40.

Ledolter J. 1989. The effect of additive outliers on the forecasts from ARIMA models. *International Journal of Forecasting* **5**: 231–240.
Litterman R, Scheinkman J. 1991. Common factors affecting bond returns. *Journal of Fixed Income* **1**: 54–61.
McCulloch JH. 1971. Measuring the term structure of interest rates. *Journal of Business* **44**: 19–31.
McCulloch JH. 1975. The tax adjusted yield curve. *Journal of Finance* **30**: 811–830.
Moench E. 2008. Forecasting the yield curve in a data-rich environment: a no-arbitrage factor augmented VAR approach. *Journal of Econometrics* **146**: 26–43.
Nelson CR, Siegel AF. 1987. Parsimonious modeling of yield curves. *Journal of Business* **60**: 473–489.
Rudebusch GD, Wu T. 2008. A macro-finance model of the term-structure, monetary policy, and the economy. *Economic Journal* **118**: 906–926.
Svensson LEO. 1994. *Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994*. NBER Working Paper Series No. 4871.
Uribe M, Yue V. 2006. Country spreads and emerging countries: who drives who?. *Journal of International Economics* **69**: 6–36.
Vasicek OA. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* **5**: 177–188.
Vasicek OA, Fong HG. 1982. Term structure modeling using exponential splines. *Journal of Finance* **37**: 339–348.
Wu T. 2006. Macro Factors and the Affine Term Structure of Interest Rates. *Journal of Money, Credit and Banking* **38**: 1847–1875.

Authors' biographies:

Rafael B. de Rezende is a PhD student at the Department of Finance, Stockholm School of Economics (Sweden). He holds a MA and BA in Economics from the Universidade Federal de Minas Gerais (Brazil). His research interests are in Monetary Economics, Macro-Finance, Time Series Econometrics and Forecasting.

Mauro S. Ferreira is an Associate Professor at the Universidade Federal de Minas Gerais (Brazil). He holds a PhD in Economics from the University of Illinois at Urbana-Champaign (2006), and a MA in Economics from the University of Sao Paulo (2000). Past positions: Chief-economist at the Federation of the Industries of Minas Gerais State (2007–2009), intern and consultant at the IMF (2004–2005), Visiting Assistant Professor at the Wesleyan University (Middletown, CT-USA) in the period of 2005–2006 and Visiting lecturer at the Trinity College (Hartford, CT-USA) in 2006.

Authors' addresses:

Rafael B. de Rezende, Department of Finance, Stockholm School of Economics, Sveavägen 65, SE-113 83 Stockholm, Sweden.

Mauro S. Ferreira, Department of Economics, Universidade Federal de Minas Gerais – CEDEPLAR, Av. Antonio Carlos 6627, Campus Pampulha, 31270-901, Belo Horizonte, Brazil.

