

Universidad del Rosario

Estimating and Forecasting the Term Structure of Interest Rates:US and Colombia Analysis

by

Cristhian Rodriguez

A thesis submitted in partial fulfillment for the
degree of Master in Quantitative Finance

in the

Faculty of Economics
Department of Finance

March 2016

Universidad del Rosario

Abstract

Faculty of Economics
Department of Finance

MSc in Quantitative Finance

by Cristhian Rodriguez

In this paper we use the most representative models that exist in the literature on term structure of interest rates. In particular, we explore affine one factor models and polynomial-type approximations such as Nelson and Siegel. Our empirical application considers monthly data of USA and Colombia for estimation and forecasting. We find that affine models do not provide adequate performance either in-sample or out-of-sample. On the contrary, parsimonious models such as Nelson and Siegel have adequate results in-sample, however out-of-sample they are not able to systematically improve upon random walk base forecast.

List of Figures

3.1	Colombian Volatility	11
3.2	USA Volatility	11
3.3	Colombian Average Fit:Nelson and Siegel	13
3.4	USA Average Fit:Nelson and Siegel	13
3.5	Colombian Average Fit:Affine Models	14
3.6	USA Average Fit:Affine Models	14
3.7	Colombian Shapes:Nelson and Siegel-KF	15
3.8	Colombian Shapes:Nelson and Siegel-OLS	15
3.9	USA Shapes:Nelson and Siegel-KF _a	16
3.10	USA Shapes:Nelson and Siegel-KF _b	16
3.11	USA Shapes:Nelson and Siegel-OLS _a	17
3.12	USA Shapes:Nelson and Siegel-OLS _b	17
3.13	Colombian Yield Curve-shapes-CIR	18
3.14	Colombian Yield Curve-shapes-VSK	18
3.15	USA Yield Curve-shapes-CIR(a)	19
3.16	USA Yield Curve-shapes-CIR(b)	19
3.17	USA Yield Curve-shapes-VSK(a)	20
3.18	USA Yield Curve-shapes-VSK(b)	20

List of Tables

4.1	Colombia RMSE	23
4.2	US RMSE	24

Contents

Abstract	i
List of Figures	ii
List of Tables	iii
1 Introduction	1
2 Term Structure Models	3
2.1 Affine Term Structure Models	3
2.1.1 Vasicek	4
2.1.2 Cox-Ingersoll-Ross (CIR)	5
2.2 Nelson and Siegel	5
2.3 Affine Models Estimation	7
Vasicek Model	7
CIR Model	7
2.4 Nelson and Siegel Estimation	7
Two Steps	8
One Step	8
3 Empirical application	10
3.1 Stylized Facts	10
The short end of the yield curve is more volatile than the long end.	10
Three main factors explain more than 95% of the changes in yield curve.	10
3.2 In Sample Results	12
The average yield curve is increasing and concave.	12
Yield curve assumes a variety of shapes in the sample.	12
4 Forecasting-Out of Sample	21
4.0.1 Affine Models	21
4.0.1.1 Nelson and Siegel	22
Two Steps	22
One Step	22
4.0.1.2 Benchmark models	22
VAR(1) On Yields Levels	22
Random Walk	22

4.1 Results	22
5 Conclusions	25

Chapter 1

Introduction

Modelling the term structure of interest rates has become a field of research on its own for different professionals, for example investors need to have an accurate analysis in pricing, investment and portfolio management decisions. In addition, central banks are very interested in the information that the yield curve conveys in terms of expectation of future rates. Different models has been developed over the last 30 years with the purpose of anticipate the dynamics of yield curve. In the literature, two types of models have surfaced affine models and Nelson and Siegel(1987)[9] type models.

When the objective is exclusively forecasting performance of the different models, the literature finds mixed results with respect to Nelson and Siegel type models and rather poor results for affine models. Nevertheless, the idea conceptual in-consistency of out-of-equilibrium models, creates some discomfort with Nelson and Siegel type models. However, empirical evidence show otherwise, Coroneo, Nyholm and Vidoka(2008)[1] showed that Nelson and Siegel is compatible with no-arbitrage constrains for US market; Dufee and Hopkins(2011)[7] inferred that omitting arbitrage free restrictions do not affect forecasting efficiency.

For affine models Duan and Simonato(1995)[5] used models of one factor as Vasicek(1977)[11] and CIR(1985)[2] for estimation analysis. On the other hand Dufee(2002)[6] finds poor forecasting performance out-of-sample. For Nelson and Siegel type model, Ullah, Matsuda and Tsukuda(2013) estimated the yield curve observing a good power of fit. Zivot and Choun-Yu(2011) conjectured good forecasting performance out-of-sample.

The purpose of this paper is to confirm some of these results with respect to US yield and evaluate the forecasting performance of these model for Colombian data. We analyze the stylized facts of the set of observed yields for both countries and use the models to determine forecasting performance in-sample, and more importantly out-of-sample. Our efforts are divided in two areas: first, the affine models represented by Vasicek(1977) and CIR(1985), and second, the polynomial form of Nelson and Siegel type models. The main tool for sequential estimation is the Kalman Filter and what we call two steps procedure that is a combination between OLS and VAR(1) of unobserved factors for one particular case of the exercise.

The results in the side of affine models confirm the hypothesis exposed in the literature. In sample we find a extreme weaknesses cause by the inability of replicating the stylized facts of term structure. In terms of out-of-sample more or less confirms the weakness of these models independent form the country of time frame. On the other side, Nelson and Siegel has an overall success in fitting the data in-sample. However, out-of-sample has poor performance with respect to simple benchmarks.

Chapter 2

Term Structure Models

Modelling term structure of interest rates is usually associated with deriving theoretically and/or empirically a functional relationship between yields and time to maturity, for example the yield curve of zero coupon bond. In order to achieve this goal we build two types of parametric models that will represent and capture the curve.

2.1 Affine Term Structure Models

Affine models are based on a particular form of the pricing equation that imposes an equilibrium where arbitrage opportunities are not possible. Two of the most well known models are the Vasicek(1987)[11] and CIR(1985)[2], used by practitioners and academics. In both cases, the price of the zero coupon bond follows a generalized geometric Brownian motion where $t < \tau$ and τ the maturity.

$$dP(t, \tau) = P(t, \tau) [\mu_\tau(t, r_t) dt + \sigma_\tau(t, r_t) dW_t] \quad (2.1)$$

The definition of a new *Itô* process W_t^Q with Q as a risk-neutral measure is required for solving the price equation.

$$dW_t^Q = \varphi_t dt + dW_t \quad (2.2)$$

When $\varphi_t = \lambda(t, r_t)$ and $\mu(t, r_t) - \sigma(t, r_t)\varphi = \mu - \lambda\sigma$, λ is the risk price market. Hence, the value of this zero coupon bond is represented as an expected value discounted value from a Q -martingale process.

$$P(t, \tau) = E[e^{-\int_t^\tau r_s ds} | r_t = r], t \in [0, \tau] \quad (2.3)$$

This model for the short rate provides an affine term structure model as long as the price has the following form,

$$P(t, \tau) = \exp(A[t, \tau] + B[t, \tau]r_t) \quad (2.4)$$

The bond price is an affine function of the short rate. The dynamic of short rate follows a diffusion process,

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t^Q \quad (2.5)$$

In most application estimation is performed on the implied yield rather than the observed prices, therefore we must relate the observed time-t compounded yield on a zero-coupon bond of maturity τ , $R(t, \tau)$, and the price equation.

$$R(t, \tau) = -\frac{1}{\tau} \ln(P(t, \tau)) \quad (2.6)$$

2.1.1 Vasicek

The diffusion process for this model allows the instantaneous spot rate to live in the support of $(-\infty, \infty)$. The Vasicek data generating process is also known as a continuous time Ornstein-Uhlenbeck process and is characterized by a mean reversion on the drift component of the diffusion equation. $\bar{\theta} = \theta - \frac{\lambda\sigma}{k}$

$$dr_t = k(\bar{\theta} - r_t)dt + \sigma dW_t^Q \quad (2.7)$$

Duffie and Kan(1996)[8] used one factor models for pricing under constrains of the neutral risk measures. They provide an analytical solution for the term structure equation (2.4) using expressions (2.8) and (2.9).

$$B[t, \tau] = \frac{1}{k}[1 - \exp(-k\tau)] \quad (2.8)$$

$$A[t, \tau] = \left[\theta + \frac{\lambda\sigma}{k} - \frac{\sigma^2}{2k^2} \right] [B[t, \tau] - \tau] - \frac{(\sigma B[t, \tau])^2}{4k} \quad (2.9)$$

2.1.2 Cox-Ingersoll-Ross (CIR)

The CIR model can be seen as a restricted version of the Vasicek(1977), because the instantaneous spot rate is forced to live in the positive support, $(0, \infty)$. The data generating process has a mean-reverting component in the drift equation, but in addition include a square root process in the diffusion. With $k > 0$, $2k\theta > \sigma^2$ guaranteeing the positiveness of short rate

$$dr_t = k(\bar{\theta} - r_t)dt + \sigma\sqrt{r_t}dW_t^Q \quad (2.10)$$

Duan and Simonato(1995) used an analytical solution for the term structure equation (2.4) using expressions (2.11) and (2.12), where the market price of risk is chosen as $\lambda\sqrt{r}$, and $[k + \sigma\lambda] \neq 0$.

$$B[t, \tau] = \frac{e^{\delta\tau} - 1}{\gamma(e^{\delta\tau} - 1) + \delta} \quad (2.11)$$

$$A[t, \tau] = \Gamma \ln \left[\frac{\delta e^{\delta\tau}}{\gamma(e^{\delta\tau} - 1) + \delta} \right] \quad (2.12)$$

$$\delta = \sqrt{(k + \lambda\sigma)^2 + 2\sigma^2}$$

$$\gamma = \frac{k + \lambda\sigma + \delta}{2}$$

$$\Gamma = \frac{2k\theta}{\sigma^2}$$

2.2 Nelson and Siegel

Nelson and Siegel(1987)[9] has been the preferred model by practitioner and macroeconomist, because is based on the objective of setting up all possible specifications that the curve might have in a parsimonious estimation. As in affine models, $P(t, \tau)$ is the price in time-t of a zero coupon bond of maturity τ . Here, the main difference is that

the dynamic for pricing does not guarantees an equilibrium without arbitrage opportunities¹.

$$P(t, \tau) = \exp(-\tau R(t, \tau)) \quad (2.13)$$

As the relation between yields to maturity and price is direct, we can obtain from the discount curve the instantaneous (nominal) forward rate curve represented as:

$$f(t, \tau) = -P'(t, \tau)/P(t, \tau) \quad (2.14)$$

The model propose a polynomial for the dynamic of forward rates with an exponential decay term.

$$f(t, \tau) = B1_t + B2_t e^{-\lambda_t \tau} + B3_t \lambda_t e^{-\lambda_t \tau} \quad (2.15)$$

The time-t compounded yield on a zero-coupon bond of maturity τ may be written as an equally-weighted average of forward rates.

$$R(t, \tau) = 1/\tau \int_0^\tau f(t, u) du \quad (2.16)$$

Using this representation it is straight forward to derive a functional representation for the yield curve. Equation (17) represents the term structure equation with the following features: as the curve begins in one at zero maturity and approaches zero at infinity maturity, being λ_t the exponential decay term that permits the factor loading $(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau})$ achieves its maximum.

$$R(t, \tau) = B1_t + B2_t \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + B3_t \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right) \quad (2.17)$$

The parameters $B1_t$, $B2_t$ and $B3_t$ are the level, slope and curvature of the yield curve respectively, together three components gives enough flexibility to the model for having an average upward and concave curve.

¹Although there is a possibility to re-write the Nelson-Siegel type model so as to find a no-arbitrage affine representation, see Dufee and Hopkins[7]

2.3 Affine Models Estimation

Affine models are generally considered as over-parameterized. Duan and Simonato(1995)[5] uses the Kalman filter as an optimal iterative process, based on the projection theorem, for estimation. In order to use the Kalman filter we must first write the model in state-space form given by the measurement and transition equations. The former is given by the yields for different maturities depending on unobserved errors and a functional relationship with the latter which is the short rate process.

$$R(t, \tau) = -\frac{1}{\tau}A[t, \tau] + \frac{1}{\tau}B[t, \tau]r_t + \epsilon_t \quad (2.18)$$

$$r_t = \alpha + \Upsilon r_{(t-1)} + \sqrt{\Phi} \eta_t \quad (2.19)$$

$R(t, \tau)$, $-\frac{1}{\tau}A[t, \tau]$, $\frac{1}{\tau}B[t, \tau]$ and ϵ_t are $(N \times 1)$ vectors according to the number of maturities. For this application we assume η_t and ϵ_t as iid $N(0,1)$ variables and not correlated between them. Below, for estimation the values of parameters are presented according to the type of affine model.

Vasicek Model

$$\begin{aligned} \alpha &= \theta(1 - e^{-kh}) \\ \Upsilon &= (e^{-kh}) \\ \Phi &= \frac{\sigma^2}{2k}(1 - e^{-2kh}) \end{aligned}$$

CIR Model

$$\begin{aligned} \alpha &= \theta(1 - e^{-kh}) \\ \Upsilon &= (e^{-kh}) \\ \Phi &= r_{(t-h)} \frac{\sigma^2}{k}(e^{-kh} - e^{-2kh}) + \theta \frac{\sigma^2}{2k}(1 - e^{-kh})^2 \end{aligned}$$

2.4 Nelson and Siegel Estimation

Nelson and Siegel fits the term structure using a smooth parametric function in a polynomial form that has three coefficients. Estimation of this parametric form is performed

using two methodologies.

Two Steps Diebold and Li(2006) use this methodology because it is easy to implement and also does not required sophisticated mathematical tools. First, they perform cross-sectional estimation by Non-Linear-Squared (NLS) at each time-t in the sample.

$$\min_{B1, B2, B3, \lambda} \sum_{i=1}^N (R(\hat{t}, \tau) - R(t, \tau))^2 \quad \text{for } t \text{ in } 1 : \tau \quad (2.20)$$

Second, they built a first order vector-autoregression (VAR(1)) for the series of estimated $\beta = \{B1, B2, B3\}$ represented in equation (2.22) with the aim of having parameters that relate the forecasting process between the factors. The result of the first step is (2.21) when the process find times series for each one of the parameters. Diebold and Li(2006)[3] find that the parameter λ might be fix through time without problems. Therefore, we fit as sample average of the series, $\lambda = \bar{\lambda}_t$.

$$R(\hat{t}, \tau) = B\hat{1}_t + B\hat{2}_t \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} \right) + B\hat{3}_t \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} - e^{-\hat{\lambda}\tau} \right) \quad (2.21)$$

$$\hat{\beta}_t = \hat{C} + \hat{\gamma}\beta_t \quad (2.22)$$

$R(\hat{t}, \tau)$ is a (T x N) matrix, $\{B1_t, B2_t, B3_t\}$ a (T x 1) vectors, \hat{C} a (3 x 1) vector and $\hat{\gamma}$ a (3 x 3) matrix of coefficients.

One Step As in affine models the use of a state space representation for the polynomial is correct because the jointly estimation reduces the possible bias of using two steps procedure. Diebold, Rudebusch and Aruoba(2006)[4] implement the filter where the unobserved state variables β are estimated with the use of the Kalman Filter in a dynamic system that simultaneously fits the yield curve.

First is the transition equation Where, η_t is iid N (0, ω) being ω a (3 x 3) covariance matrix, C a (3 x 1) vector and γ a (3 x 3) matrix of coefficients. Secondly is the measurement equation that keeps ϵ_t as iid N(0, ψ), being ψ a diagonal matrix of (N x N) variances, ζ a (N x 3) matrix of factor loadings, β_t a (3 x 1) vector and $R(t, \tau)$ a (N

x 1) vector. Besides, white noise and measurement disturbances have to be orthogonal between them and to the initial state.

$$\beta_t = C + \gamma\beta_{t-h} + \eta_t \quad (2.23)$$

$$R(t, \tau) = \zeta\beta_t + \epsilon_t \quad (2.24)$$

Chapter 3

Empirical application

The data consist of the zero coupon rates of Colombia and the United States available in their respectively central banks web page ¹, the frequency is monthly for both of them but the sample length and the number of maturities is different for each set of series. For the colombian data we have three maturities (one, five and ten years) and the sample is from January, 2003 until August, 2015. For the US we have ten maturities (one, three and six months also for one, two, three, five, seven, ten and twenty years) and the sample is from July, 2001 until August, 2015.

3.1 Stylized Facts

The short end of the yield curve is more volatile than the long end. For the exercise we capture data volatility as the conditional standard deviation represented in a GARCH(1,1) model, figures (3.1) and (3.2) exhibit that for Colombia the relationship is not clear in contrast with USA where the short yield maturity remarks more movements over the majority of sample against the long yield.

Three main factors explain more than 95% of the changes in yield curve. Despite available data for Colombia does not have enough maturities as USA only 3 against 10, in general terms the cumulative proportion of variance achieves the majority of explicative power in the second component.

¹<http://www.banrep.gov.co/es/tes> , <https://research.stlouisfed.org/fred2/categories/22>

FIGURE 3.1: Colombian Volatility

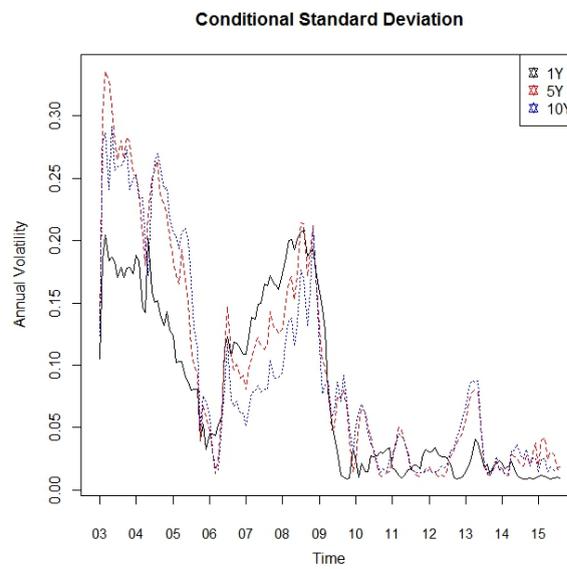
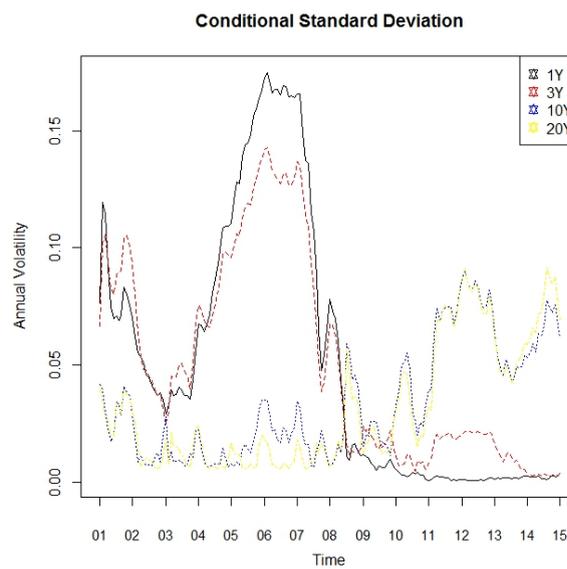


FIGURE 3.2: USA Volatility



Colombia

<i>Principal Components</i>	PC1	PC2	PC3
Standard deviation	1.46	0.846	0.39
Proportion of Variance	0.7109	0.2383	0.0508
Cumulative Proportion	0.7109	0.9492	1

USA

<i>Principal Components</i>	PC1	PC2	PC3	PC4	PC10
Standard deviation	2.6398	1.5141	0.6539	0.4398	0.0558
Proportion of Variance	0.6969	0.2293	0.0428	0.0193	0.0003
Cumulative Proportion	0.6969	0.9261	0.9689	0.9882	1

3.2 In Sample Results

The data from the term structure of interest rates, is very important for investors and policymaker. From the point of view of professional forecasters the objective is that a good model should be able to replicate the historical regularities.

The average yield curve is increasing an concave. Figures (3.3) and (3.4) represents Nelson and Siegel for one and two steps. The parameters $\hat{\lambda}$ for Colombia and USA are 0.999934 and 0.5247941 respectively, the numbers are obtained after Non-Linear-Squared estimation and bringing out the mean of λ_t series. Figures (3.5) and (3.6) reproduce affine models conduct more specifically CIR and Vasicek. The evidence of graphics is clear because exposes limitations of affine models against Nelson and Siegel either in Colombia and USA scenario.

Yield curve assumes a variety of shapes in the sample. Figures (3.7) to (3.12) personify for Nelson-Siegel the movements in each one of the curves through time , either using OLS kalman filter or OLS two steps estimation. Figures (3.13) to (3.18) typify the comportment for affine models. Results are definitive against Vasicek and CIR considering the great numbers of disparities throughout the sample.

FIGURE 3.3: Average Yield curve of Colombia-Nelson and Siegel

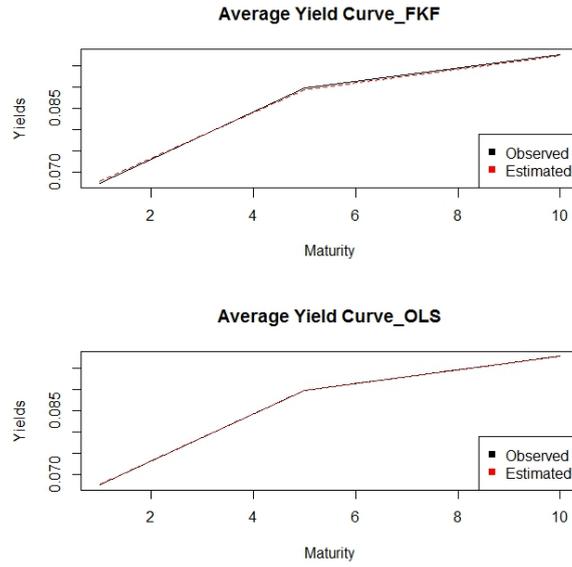


FIGURE 3.4: Average Yield curve of USA-Nelson and Siegel

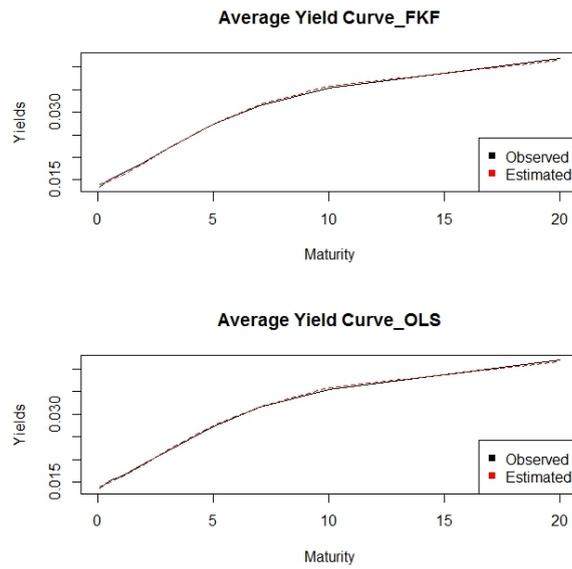


FIGURE 3.5: Average Yield curve of Colombia-Affine Models

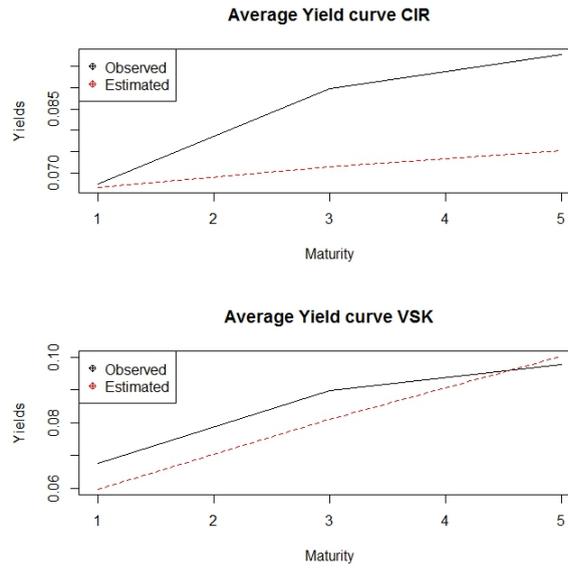


FIGURE 3.6: Average Yield curve of USA-Affine Models

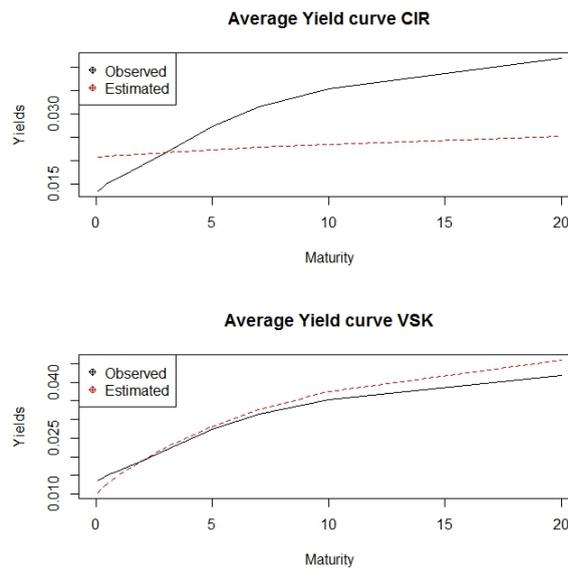


FIGURE 3.7: Colombian Yield Curve-Nelson and Siegel-Kalman Filter

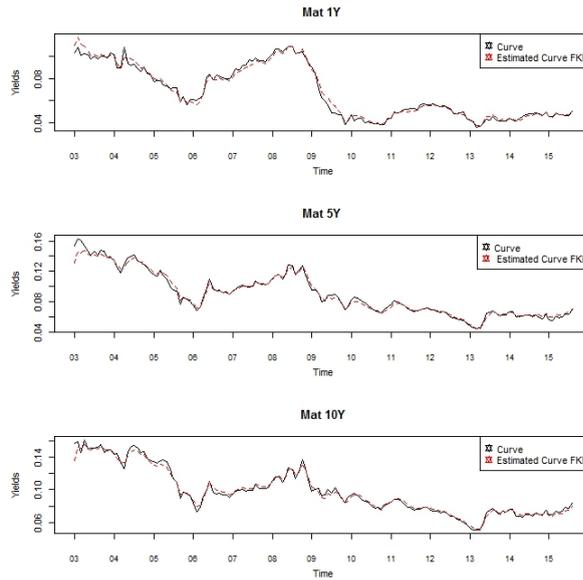


FIGURE 3.8: Colombian Yield Curve-Nelson and Siegel-OLS

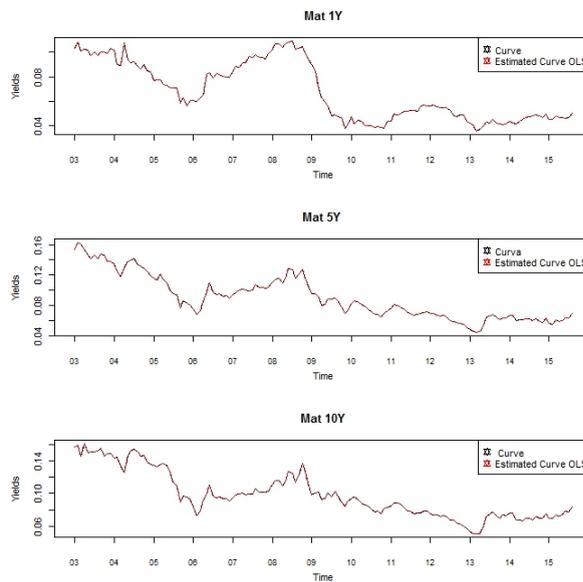


FIGURE 3.9: USA Yield Curve-Nelson and Siegel-kalman filter(a)

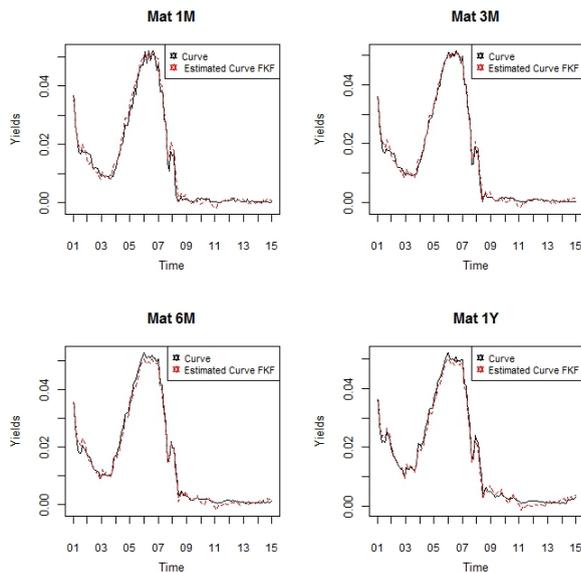


FIGURE 3.10: USA Yield Curve-Nelson and Siegel-kalman filter(b)

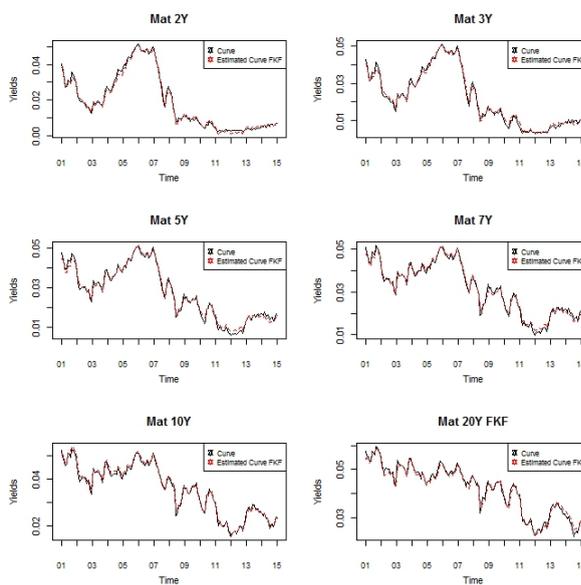


FIGURE 3.11: USA Yield Curve-Nelson and Siegel- OLS(a)

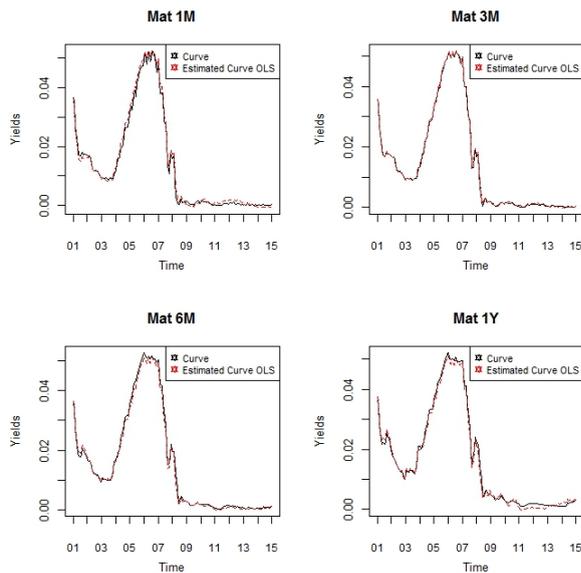


FIGURE 3.12: USA Yield Curve-Nelson and Siegel- OLS(b)

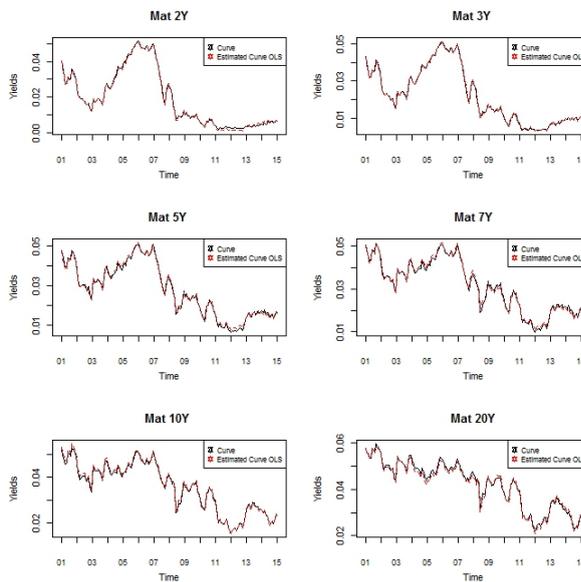


FIGURE 3.13: Colombian Yield Curve-shapes-CIR

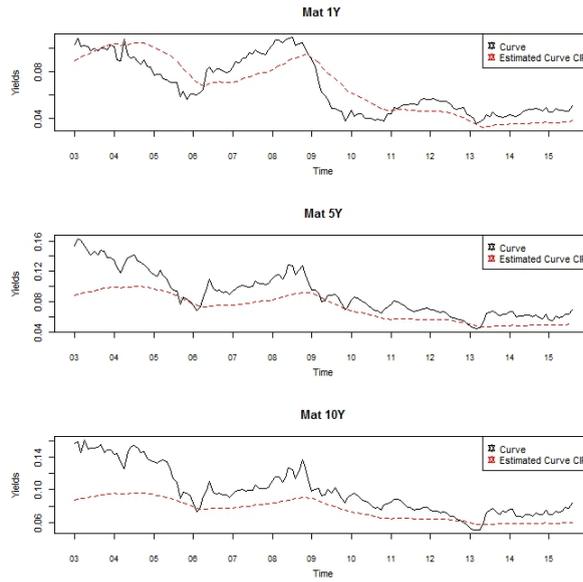


FIGURE 3.14: Colombian Yield Curve-shapes-VSK

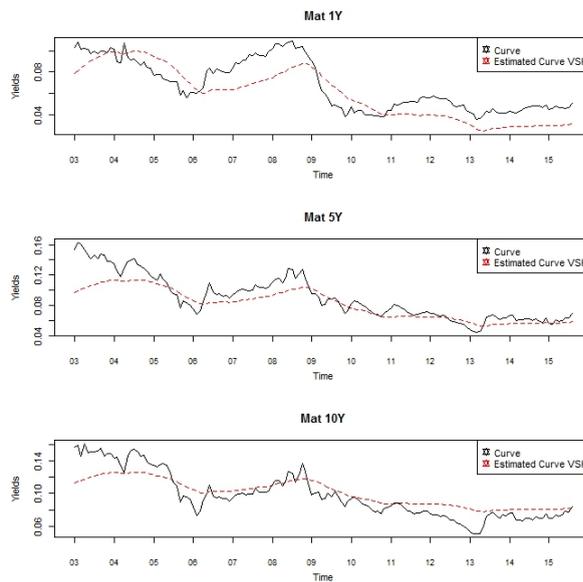


FIGURE 3.15: USA Yield Curve-shapes-CIR(a)

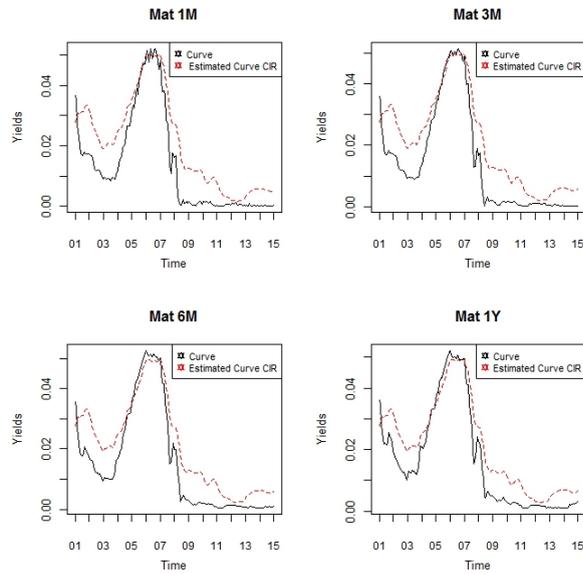


FIGURE 3.16: USA Yield Curve-shapes-CIR(b)

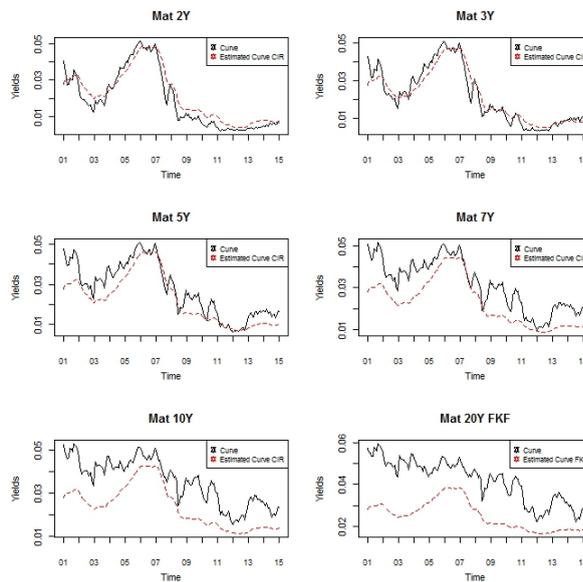


FIGURE 3.17: USA Yield Curve-shapes-VSK(a)

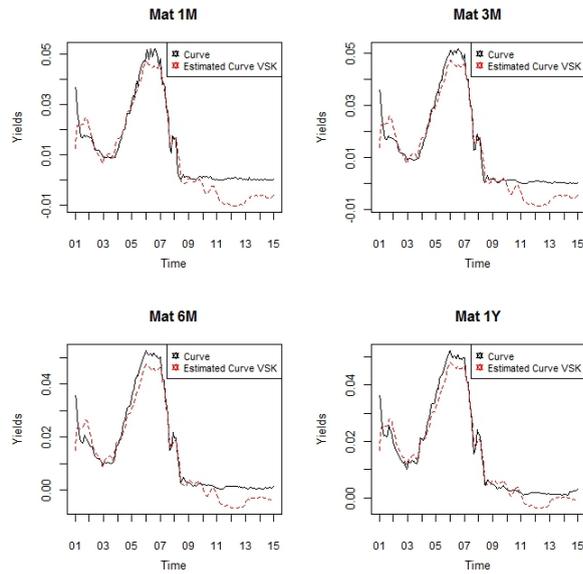
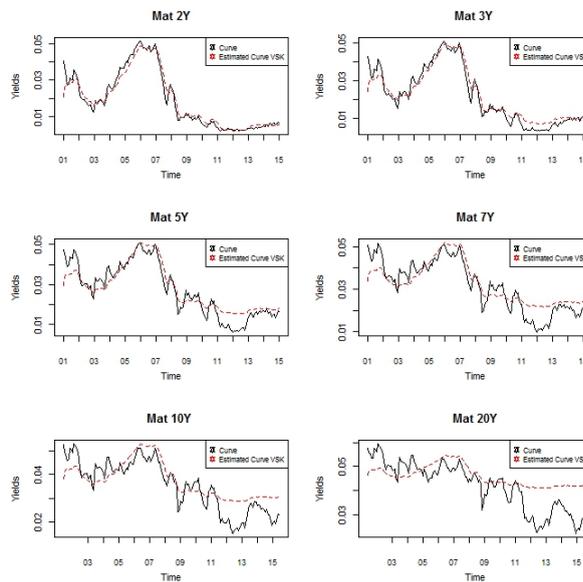


FIGURE 3.18: USA Yield Curve-shapes-VSK(b)



Chapter 4

Forecasting-Out of Sample

In this section we evaluate the performance of affine and Nelson Siegel type models and compare them against different benchmarks; such as a random walk and vector-autoregression of the level yields. The loss function to evaluate the performance in all models is the root mean squared error represented as:

$$RMSE = \sqrt{[R(t + \hat{h}, \tau) - R(t + h, \tau)]^2} \quad (4.1)$$

Where, h is the length of steps ahead that we take for forecast evaluation, one, six and twelve respectively in all available maturities. We use an expanding data window beginning in January of 2010 (2010:01) until July and August of 2015(2015:07;08) for USA and Colombia, respectively.

4.0.1 Affine Models

Forecasting affine models is made easier thanks to the iterative process of the kalman filter, where the state space representation updates and evaluates the likelihood function through the use of state variables distributions conditional on previous estimates values.

$$r_{t+\hat{h}} = \hat{\alpha} + \hat{\Upsilon} r_t + \sqrt{\hat{\Phi}} \eta_{t+\hat{h}} \quad (4.2)$$

$$R(t + \hat{h}, \tau) = -\frac{1}{\tau} A[\hat{t}, \tau] + \frac{1}{\tau} B[\hat{t}, \tau] r_{t+\hat{h}} + \epsilon_{t+\hat{h}} \quad (4.3)$$

4.0.1.1 Nelson and Siegel

Forecasting the yield curve requires that the unobserved level, slope and curvature variables have been predicted previously. As the polynomial system may be estimated by two different methodologies we can either forecast a VAR of unobserved vectors or in the case of kalman filter the transition vector.

Two Steps

$$\hat{\beta}_{t+h} = \hat{C} + \hat{\gamma}\beta_t \quad (4.4)$$

$$R(t + \hat{h}, \tau) = B1_{t+h} + B2_{t+h} \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} \right) + B3_{t+h} \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} - e^{-\hat{\lambda}\tau} \right) \quad (4.5)$$

One Step

$$\hat{\beta}_{t+h} = \hat{C} + \hat{\gamma}\beta_t + \eta_{t+h} \quad (4.6)$$

$$R(t + \hat{h}, \tau) = \hat{\zeta}\hat{\beta}_{t+h} + \epsilon_{t+h} \quad (4.7)$$

4.0.1.2 Benchmark models

Other models are taken as reference for their easiness in the estimation and also because not required strong fundamental theory. The chosen models are the VAR(1) on yields levels and the famous random walk.

VAR(1) On Yields Levels

$$R(t + \hat{h}, \tau) = \hat{A} + \hat{\omega} R(t, \tau) \quad (4.8)$$

Random Walk

$$R(t + \hat{h}, \tau) = R(t, \tau) \quad (4.9)$$

4.1 Results

Table (4.1) stacks the results for Colombia and allows to see that in average a random walk is hard to be defeated for any kind of specification model. Nevertheless, Nelson and

Siegel-One Step model is able to do it for long horizons of forecasting, other conclusion is that affine models have a poor performance out of sample.

TABLE 4.1: Colombia RMSE

	Maturities	One Month ahead	Six Months ahead	Twelve Months ahead
VAR(1)	1Y	0.1810	0.5093	0.8323
Yields	5Y	0.2952	0.8744	0.9132
On Levels	10Y	0.3120	0.9212	0.9457
	Average	0.2627	0.7683	0.8971
Random Walk	1Y	0.1750	0.4832	0.7372
RW	5Y	0.2772	0.7557	0.8086
	10Y	0.2741	0.8037	0.9654
	Average	0.2421	0.6808	0.8371
Nelson-Siegel	1Y	0.1810	0.5093	0.8323
Two Steps	5Y	0.2952	0.8744	0.9132
	10Y	0.3120	0.9212	0.9457
	Average	0.2627	0.7683	0.8971
Nelson-Siegel	1Y	0.2562	0.5118	0.7378
One Step	5Y	0.4731	0.8072	0.8177
	10Y	0.5034	0.8706	0.8929
	Average	0.4109	0.7299	0.8161
Vasicek	1Y	1.5302	1.9164	2.3739
VSK	5Y	0.7677	0.9471	1.1301
	10Y	1.3795	1.3124	1.1722
	Average	1.2258	1.3920	1.5588
CIR	1Y	1.222	1.490	1.800
	5Y	1.267	1.619	1.935
	10Y	1.318	1.543	1.797
	Average	1.2716	1.5505	1.8440

Table (4.2) on the other side stacks the results for USA, the conclusions are almost the same. first, the confirmation about the weaknesses of affine models for forecasting. Second, the impossibility of any model against the random walk even though for Colombia Nelson and Siegel-One Step do it, within USA is unable.

TABLE 4.2: US RMSE

	Maturities	One Month ahead	Six Months ahead	Twelve Months ahead
VAR(1)	1M	0.0415	0.35003	0.5585
Yields	3M	0.0579	0.3730	0.5763
On Levels	6M	0.0659	0.3796	0.5850
	1Y	0.0651	0.3380	0.5464
	2Y	0.0987	0.3282	0.5327
	3Y	0.1326	0.3827	0.5930
	5Y	0.1670	0.4877	0.6677
	7Y	0.1802	0.5278	0.6739
	10Y	0.1734	0.5026	0.5956
	20Y	0.1695	0.5156	0.6539
	Average	0.1152	0.4185	0.5983
Random Walk	1M	0.0212	0.0361	0.0498
RW	3M	0.0131	0.0356	0.0525
	6M	0.0154	0.0421	0.0579
	1Y	0.0187	0.0576	0.0768
	2Y	0.0603	0.1303	0.1718
	3Y	0.0972	0.2397	0.2977
	5Y	0.1399	0.3945	0.4770
	7Y	0.1536	0.4666	0.5843
	10Y	0.1506	0.5031	0.6300
	20Y	0.1497	0.5361	0.6852
	Average	0.0820	0.2452	0.3083
Nelson-Siegel	1M	0.0944	0.3413	0.4280
Two Steps	3M	0.0824	0.3109	0.3899
	6M	0.0740	0.2607	0.3435
	1Y	0.0942	0.2033	0.2819
	2Y	0.0932	0.2222	0.2003
	3Y	0.1088	0.3238	0.4171
	5Y	0.1720	0.4468	0.5662
	7Y	0.1742	0.4798	0.6432
	10Y	0.1533	0.5160	0.7210
	20Y	0.1854	0.6026	0.8352
	Average	0.1232	0.3707	0.4916
Nelson-Siegel	1M	0.1196	0.1657	0.1939
One Step	3M	0.0956	0.1363	0.1625
	6M	0.0846	0.1066	0.1334
	1Y	0.1018	0.1037	0.1193
	2Y	0.1214	0.1613	0.2087
	3Y	0.1511	0.2734	0.3759
	5Y	0.2324	0.4140	0.5608
	7Y	0.2433	0.4577	0.6432
	10Y	0.2407	0.5113	0.7120
	20Y	0.2975	0.5966	0.8054
	Average	0.1688	0.2927	0.3915
Vasicek	1M	0.8490	0.8910	0.9424
VSK	3M	0.6912	0.7284	0.7734
	6M	0.5712	0.6071	0.6435
	1Y	0.3419	0.3717	0.3926
	2Y	0.1344	0.2161	0.1876
	3Y	0.3132	0.4030	0.3885
	5Y	0.5487	0.6466	0.6615
	7Y	0.6700	0.7675	0.8017
	10Y	0.8280	0.9248	0.9694
	20Y	1.3143	1.4120	1.4831
	Average	0.7297	0.7752	0.8026
CIR	1M	0.7767	0.7079	0.6449
	3M	0.7716	0.7234	0.6493
	6M	0.7370	0.6861	0.6092
	1Y	0.6738	0.6144	0.5639
	2Y	0.4502	0.4598	0.4247
	3Y	0.2767	0.3409	0.3826
	5Y	0.5561	0.6154	0.7151
	7Y	1.0140	1.0081	1.1058
	10Y	1.4554	1.4471	1.5088
	20Y	2.002	2.0096	2.0193
	Average	0.8722	0.8613	0.8624

Chapter 5

Conclusions

In this article we worked with term structure of interest rates for different countries, the analysis was focused in the use of different models well-known by practitioners and academics. We estimated and made forecast of the curves through methodologies as Kalman Filter or two steps representation in the case of Nelson and Siegel model.

We found that Nelson-Siegel is able to have a good performance in sample scenario. Nonetheless, out of sample the results became worse, being almost impossible to defeat a random walk. In the case of affine models neither in sample or out of it the performance is even acceptable.

In future research we plan to incorporate new methodologies that allows us to find optimal results out sample and later find a loss function beyond the root mean squared error.

Bibliography

- [1] CORONEO, L., NYHOLM, K., VIDOVA-KOLEVA R., 2008. How arbitrage-free is the Nelson-Siegel model?, *European Central Bank Working Paper Series N 874*.
- [2] COX J.C., INGERSOLL J.E., ROSS J.S., 1985. A theory of the term structure of interest rates, *Econometrica*, 53(2), pp. 385-408
- [3] DIEBOLD, F.X., LI, C., 2006. Forecasting the term structure of government bond yields, *Econometrics*, 130, pp. 337-364.
- [4] DIEBOLD, F.X., RUDEBUSCH G.D., ARUOBA S.B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach, *Journal of Econometrics*, 131, pp. 309-338.
- [5] DUAN J.C., SIMONATO J.G., 1995. Estimating and testing exponential-affine term structure model by kalman filter, *CIRANO*, ISSN 1198-8177
- [6] DUFFEE, G.R., 2002. Term Premia and Interest Rate Forecasts in Affine Models, *The Journal of Finance*, pp. 405-443.
- [7] DUFFEE, G.R., HOPKINS J., 2011. Forecasting with the term structure: The role of no-arbitrage restrictions.
- [8] DUFFIE, D., KAN R., 1996. A yield factor model of interest rates, *Mathematical Finance*, 6(4), pp. 379-406.
- [9] NELSON, C. R., SIEGEL, A. F., 1987. Parsimonious Modeling of Yield Curves, *Journal of Business*, 60, pp. 473-489.
- [10] ULLAH, W., MATSUDA, Y., TSUKUDA, Y., 2013. Term Structure Modelling and Forecasting of government bond yields: Does a good in sample fit imply reasonable out-of sample forecasts, *Economic Papers*, 32(4), pp. 535-560.
- [11] VASICEK O., 1977. An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5, pp. 177-188.

-
- [12] ZIVOT E. CHOUN-YU W., 2011. Forecasting the term structures of Treasury and corporate yields using dynamic Nelson-Siegel models, *International Journal of Forecasting*, 27, pp. 579-511.