IS IT POSSIBLE TO MAKE A PROFIT WITH NELSON-SIEGEL TERM STRUCTURE MODEL IN THE FIXED INCOME MARKETS?

HAFİZE KURT

110664033

İSTANBUL BİLGİ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ
ULUSLARARASI FİNANS YÜKSEK LİSANS PROGRAMI

Okan AYBAR

2013

IS IT POSSIBLE TO MAKE A PROFIT WITH NELSON-SIEGEL TERM STRUCTURE MODEL IN THE FIXED INCOME MARKETS?

NELSON - SIEGEL TERM STRUCTURE MODELİ İLE SABİT GETİRİLİ PİYASALARDA KAR ELDE ETMEK MÜMKÜN MÜDÜR?

HAFİZE KURT

110664033

Okan Aybar :..............................

Prof. Dr. Oral Erdoğan :..............................

Kenan Tata :..............................

Tezin Onaylandığı Tarih :..............................

Toplam Sayfa Sayısı :..............................

Anahtar Kelimer (Türkçe)

1) Nelson Siegel
2) Vade Yapısı Modelleri
3) Ticaret Stratejisi
4) Diebold&Li
5) Kuponsuz Tahviller

Anahtar Kelimer (İngilizce)

1) Nelson Siegel
2) Term Structure Models
3)Trading Strategy
4) Diebold&Li
5) Zero Coupon Bond

ÖZET

Bu tezde Nelson - Siegel vade yapısı modeli gelecek getirileri tahmin etmek amacıyla tanıtılmış ve AR (1) yapısı da gelecekteki betaları tahmin etmek için uygulanmıştır. Daha sonra Nelson - Siegel - AR (1) modeliyle aylık getiri sağlayan Amerikan devlet iskontolu tahvilleri üzerinden bir ticaret stratejisi geliştirilmiştir. Bu önerilen ticaret stratejisi, bu tezde basit bir şekilde örnek veriler üzerinde denenmiştir. Karşılaştırmalı ve fiyatlandırmalı analizler aracılığıyla, elde edilen sonucun doğruluğu ve stratejinin iyi bir performans sunduğu teyit edilmeye çalışılmıştır. İlk analizde, bu ticaret stratejisinin etkileri iki farklı portföyde kıyaslanmıştır. Portföylerden ilki S&P500 endeksinden oluşturulmuşken, ikinci portföy "Fama - Bliss" in iskontolu tahvil fiyatlarıyla oluşturulmuştur. Fiyatlama analizinde, "Fama - French"in tahvil fiyatları için ortaya koyduğu iki faktör modeli kullanılmıştır ve bu analizin sonucunda portföyün olağandışı getiriler sağladığı ortaya çıkmıştır. Çünkü görülmüştür ki, kesişim noktası 0'dan belirgin bir şekilde farklıdır. Bu iki analiz ticaret stratejisinin ispat edilebilir bir performansa sahip olduğunu teyit etmektedir. Ayrıca, işlem maliyeti olmadığı farzedilmiştir.

ABSTRACT

In this thesis the Nelson-Siegel term structure model is introduced to forecast the future yields and then AR(1) structure is imposed to forecast the future betas. Then a trading strategy on monthly yield U.S. government zero coupon bonds is built by Nelson-Siegel – AR(1) model. The trading strategy that is proposed in this thesis performed simply over sample data. Comperative and pricing analyses are used to confirm that result by comparing strategy returns and the performance of the strategy. In the first analysis, the trading strategy is compared with two protfolios. The first one is invested on the index S&P500 and the second one comes from the Fama-Bliss discount bond prices. In pricing analysis, Fama-French's two factor model for bond prices is used and as a result of this analysis the protfolio is found to earn abnormal returns as the intercept is significantly different from zero. These two analyses confirm that the trading strategy has a poor performance. It is also assumed that there are no transaction costs.

Key words: Nelson Siegel, terms structure models, trading strategy, Diebold&Li, zero coupon bond

Table of Contents

List of Abbreviations

S&P 500	Standard & Poors 500
CIR	Cox, Ingersoll and Ross Model
AR	Autoregressive
VAR	Vector Autoregressive
ZCB	Zero Coupon Bond
DNS	Dynamic Nelson Siegel
OLS	Ordinary Least Square
WRDS	Wharton Research Data Services
RMSE	Root Mean Square Error
FRMSE	Forecasted Root Mean Square Error

1. INTRODUCTION

The idea of this paper is to build a trading strategy and test the profitability of this strategy based on forecasts of the yield curve for zero coupon government bonds. The goal of the trading strategy is to exploit the differences between the zero coupon bonds (ZCB) prices predicted by the forecasts and the ZCB prices realized on the market. In order to achieve this goal, first it is needed to define a way to forecast the future yield curve, build a trading rule and finally analyze the payoff of this strategy.

Until now, several ways have been proposed to forecast future yields. One of them is the power of forward interest rates in forecasting future spot rates which has been confirmed by several papers such as Fama, E. F. (1984) "The Information in the Term Structure". Another way is to consider a term structure model and use it to forecast future yield curve. This second way was choosen to proceed. The model should be able to nicely fit the present yield curve and explain its movements over time. The idea behind using a model to forecast the yield curve is that few factors can explain the greatest part of its variation. There are several approaches to model the term structure, the three principals are: no-arbitrage models, that ensures a certain consistency between the parameters that describe the dynamic evolution of the yield curve factors under the risk-neutral measure, and the translation of yield curve factors into yields under the physical measure; there are equilibrium models, they explain the dynamic of the short time interest rate and from it all the yield curve is derived, part of this group are the Vasicek model and CIR (Cox, Ingersoll, Ross) model. A third way can be the one proposed by Charles Nelson and Andrew Siegel (1987) that model the yield curve with three dynamic parameters. The advantage of this model is that it is a parsimonious one and thus is a natural candidate to be used for out-of-sample forecasting purpose. Therefore, Nelson-Siegel model is one of the most used by practitioners.

The most extensive documentation of the strong predictive content of the spread for output, containing its ability to forcaste a dual recession indicator in profit regressions are provided by Estrella and Hardouvelis (1991) and Estrella and Mishkin (1995). Arturo Estrella and Frederic S. Mishkin also analyse the out-of-sample performance of various financial variables as predictors of U.S. recessions (1996). They concluded that the slope of the yield curve appears as the clear individual option and generally performs better by itself out of sample than in conjunction with other variables. Oral Erdoğan, Paul Benett and Cenktan Özyıldırım extended the benchmark Estrella and Hardouvelis (1991) term spread approach to recession forecasting by including the stock market macro liquidity deviation factor (2010). They used a dependent dummy variable for crisis based on NBER business cycle dates between 1959Q1 and 2011Q4. They found that assembly the yield curve and the market liquidity deviation measures developed their in-sample and out-of-sample ability to predict the onset of a US recession. Therefore, Nelson-Siegel model will be examined on yield curves in this thesis substantially.

The reason of chosing the Nelson-Siegel model is the ease to be “converted” from a fitting model to a forecasting tool. The Nelson-Siegel model is a parsimonious model, just three dynamic factors to be estimated, this is a good hope for powerful forecast results.

In their highly cited article, Litterman and Scheinkman (1991) point out that most of the variation of the term structure of interest rates should be attributed to the change in three factors and these three factors can be interpreted as a level, a steepness (slope) and a curvature factor; the so called "shape factors". The level factor catches the parallel shift of the curve, the steepness shows the change in slope of the curve and the curvature represents the change in the curvature at medium term maturities. Jointly, these three factors explain more than 95% of the entire movements of the yield curve. The model which will be used in this work should be able to capture these dynamics of the change in the shape of the yield curve. As further explaination, the parameters in Nelson-Siegel model can be interpreted as "shape factors".

The Nelson-Siegel model is able to fit different shapes of the term structure: downward sloping, upward sloping, humped and S shaped. It is in this sense that the model is a good tool to catch all the possible movements of the yield curve and a good candidate for good forecast for the yield curve. The forecasts from the Nelson-Siegel model will be the building block for the trading strategy.

Most of the authors use the arbitrage-free Nelsol-Siegel model in their papers. For example, Christensen, Diebold, and Rudebusch (2007) have modified the Nelson-Siegel framework to impose the abritrage-free condition. They concluded that a version of dynamic Nelson-Siegel model can be made arbitrage free by adding an additional factor. On the other hand, some others take the Nelson-Siegel model which is not arbitrage-free as a basis by constructing the yield curve. Theoretically, the Nelson-Siegel model does not ensure the absence of arbitrage opportunities, as shown by Bjork and Christensen (1999) and Filipovic(1999). Laura Coroneo, Ken Nyholm and Rositsa Vidova-Koleva (2008) test whether the Nelson and Siegel (1987) yield curve model is arbitrage-free. Using a non-parametric resampling technique and zero-coupon yield curve data from the U.S. market, they find that the no-arbitrage parameters (level, slope, curvature) are not statistically different from the Nelson-Siegel model, at a 95 percent confidence level. Moreover, in an out-of-sample forecasting experiment, they show that the performance of the Nelson-Siegel model is as good as the no-arbitrage counterpart. Therefore, they concluded that the Nelson and Siegel yield curve model is compatible with arbitrage-freeness on the U.S. market. Also, Diebold and Li (2006) defend that the arbitrage-free term structure litrature has little to say about forecasting. Moreover, they showed that their model, which is not arbitrage-free, can produce good forecasts. The model is also not arbitrage-free because Diebold and Li's paper will be followed.

This thesis will be organized as follows: after a brief review of literature in Section 2, the Nelson-Siegel – AR(1) model will be introduced in Section 3. Section 4 presents and analyzes the data. Then, the estimation, test and analysis of the forecasting characteristics of the model by using Nelson-Siegel term structure model in the fixed income markets will be made to fit the yield curve and to forecast the future level of interest rates adhering to the paper of Diebold and Li (2006) in Section 5. Section 6 will be generated by introducing the trading strategy and adding two comparisons in order to be certain about the results. The last part, Section 7, includes conclusion and final comments.

2. LITERATURE REVIEW

The fixed income market (also known as the bond market) is a financial market where all interest rate financial instruments, like a bond, swaps or swaptions. Participants buy and sell debt securities usually in the form of bonds. The bond market is very large and still growing. As of 2009, the size of the worldwide bond market (total debt outstanding) is estimated to be $82.2 trillion. The main part of the bond market is government bond market because of its size, liquidity and lack of credit risk. However, the most important reason is the sensitivity to interest rates. There is an inverse relationship between bond valuation and interest rates, therefore the bond market indicates changes in interest rates or the shape of the yield curve.

The objective of term structure modeling from a practitioner's viewpoint is to develop a parsimonious representation of the yield curve matching the time series and the cross Sectional variation of bond yields. Academic literature and practitioner oriented publications show that the parametric yield curve model intoduced by Nelson and Siegel (1987) and re-interpreted by Diebold and Li (2003) as a modern three-factor model of level, slope (the difference between the long and the short end of the yield curve) and curvature. Due to its intuitive appeal and implementational easiness the Nelson- Siegel model provides a remarkably good fit to the cross Section of yields in many countries and has become a widely used specification among financial market practitioners and central banks. For example, Diebold, Rudebusch, and Aruoba (2005) use the model to study the interactions between the macro economy and the yield curve (Diebold, Piazzesi, and Rudebusch (2005)). They find that the level factor is highly correlated with inflation, and the slope factor is highly correlated with real activity while the curvature factor does not related to any of the main macroeconomic variables. Whereas Diebold, Ji, and Li (2006) apply it to identify systematic risk sources and to construct a generalized duration measure. Moreover, Fabozzi, Martellini, and Priaulet (2005) show that the Nelson-Siegel model produces forecasts that are not only statistically accurate but also economically meaningful. Besides, Pooter (2007) worked with this model for a different purpose. He examines several models within the Nelson-Siegel class for their in-sample and out-of-sample performance to try and evaluate the trade-off between in-sample fit and out-of-sample forecasting performance by using a sample of U.S. Treasury zero-coupon bond yields.

Diebold and Li (2006) develop a dynamic version of this model and show that the yields have a standard interpretation of level, slope, and curvature. Such a dynamic Nelson-Siegel (DNS) model is easy to estimate and forecasts the yield curve quite well. Despite its good empirical performance, however, the DNS model does not impose the presumably desirable theoretical restriction of absence of arbitrage. Federal Reserve Board, the European Central Bank and many other central banks use the Svensson (1995) extension to the Nelson-Siegel curve that adds a second curvature term, which allows for a better fit at long maturities. Following the model introduced by Diebold and Li (2006), Diebold, Christensen and Rudebusch (2008) developed a dynamic version of this model, the DNSS model, which corresponds to a modern four-factor term structure model. Its dynamic form does not enforce arbitrage-free consistency over time. Furthermore, they showed that the factor loadings of the Svensson generalization cannot be obtained in a standard finance arbitrage-free affine term structure representation. Therefore, they introduced a related generalized Nelson-Siegel model on which the no-arbitrage condition can be imposed.

The three factors in the model are the short term rate of interest (Vasicek (1977), Cox-Ingersoll-Ross (1985), Hull and White (1990)), the long run mean of the short term rate (Balduzzi, Bertola, and Foresi (1993), Naik and Lee (1993), Balduzzi, Foresi, Das and Sundaram (1996)), and the volatility of the short term rate (Longstaff and Schwartz (1992), Fong, Vasicek and Yoo (1992)). The more factors, the richer the time series and cross Sectional properties of bond returns the model can accomodate. So far, practicality has kept researchers in the field from going beyond three factor models. The simple estimation model should make the implementation of three factor models easier. That no factor other than the short rate is clearly dominant across yields of all maturities is an indication that all three factors play an important role in modeling the yield curve.

3. NELSON-SIEGEL MODEL

Nelson and Siegel introduce a specification for the forward interest rates curve; the forward rate curve is given by:

$$f_t(\tau) = \beta_{0t} + \beta_{1t} e^{-\lambda_t \tau} + \beta_{2t}(\lambda_t e^{-\lambda_t \tau})$$

where τ is a time constant, λ_t is a decay parameter and β_{0t} , β_{1t} and β_{2t} are the three dynamic parameters.

The relation between the forward interest rate and the yield to maturity for a bond is provided by classical bond pricing literature, the relation goes as follow:

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

This formula implies that the zero coupon yield is an equally weighted average of forward rates. The yield curve can be calculated by integrating the forward curve from 0 to τ and dividing by τ, the yield is simply the average of forward rates, the resulting formulation for the yield curve that will bealways refered to this variation of the Nelson-Siegel model which is:

$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$$

This is the Nelson-Siegel formulation for the yield curve, as maturity became larger, $y_t(\tau)$ goes to β_{0t}; if maturity gets smaller, $y_t(\tau)$ tends to $\beta_{0t} + \beta_{1t}$. It is important to highlight how this specification is able to explain several important characteristics found empirically in yield curves.

The interpretation for the three coefficients β_{0t} , β_{1t} and β_{2t} as the "shape factors" that Litterman and Scheinkman described in their paper. β_{0t} is the coefficient of a constant variable of value 1, that does not decay to zero as maturity gets larger and larger. The interpretion is for it as a level factor since this coefficient affects the yield curve at every maturity in the same way, with the same magnitude. β_{1t} is the loading on the factor $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$ that decay monotonically to 0 as maturity gets larger and larger, hence it affects the yield curve mostly at short maturities, and for this reason can be seen as short-term factor. This coefficient, affecting the short-term maturities in a heavier way than the longer ones, change the slope of the yield curve and can be interpreted as the steepness factor. β_{2t} is the parameter on $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}$, this factor increase from the low maturities until medium maturities and then decay to 0 as maturity increases, for this behavior the factor can be seen as a medium-term factor; by affecting the intermediate maturities it affects the curvature of the yield curve and can be interpreted as the curvature factor.

Diebold and Li (2006) suggest to fix λ_t in the three-factor model to a pre-specified value, which is the same for every t, instead of treating it as an unknown parameter. By doing so, the nonlinear measurement equations become linear in the state vector which can then be estimated using straightforward cross-Sectional OLS. The decay parameter λ_t determines the (medium-term) maturity at which the factor loading on the curvature factor β_{2t} is at its maximum. The value of 0.0609 that Diebold and Li (2006) use for λ_t is such that this maximum is reached at a 30-month maturity and since the model I have used in this paper is closest to the one in Diebold and Li (2006) I have found it appropriate to set the same value of λ_t. Larger values for λ_t produce slower decaying factor loadings with the curvature factor achieving its maximum at a longer maturity and vice versa. Then, the fitted yield curve just derives from the combination of the estimated betas and the factors through the Nelson-Siegel model.

The final target that my model should achieve is to forecast future interest rates using the Nelson-Siegel model in this scope. With the λ_t fixed and time invariant, the yield is determined exclusively by the interaction between deterministic factors and the three dynamic betas, hence to forecast a yield for h periods ahead I just need to forecast the future betas and then through the Nelson-Siegel model obtain the forecasted yield curve. To forecast the betas a parametric structure should be imposed on their evolution over time. As Diebold and Li proposed in their paper, an AR(1) structure is imposed, and the future beta depends on its previous value.

The formula for the AR(1) model applied to β s is:

$$\beta_{i,t+h/t} = c_i + \gamma_i \beta_{i,t} + \varepsilon_{i,t}$$

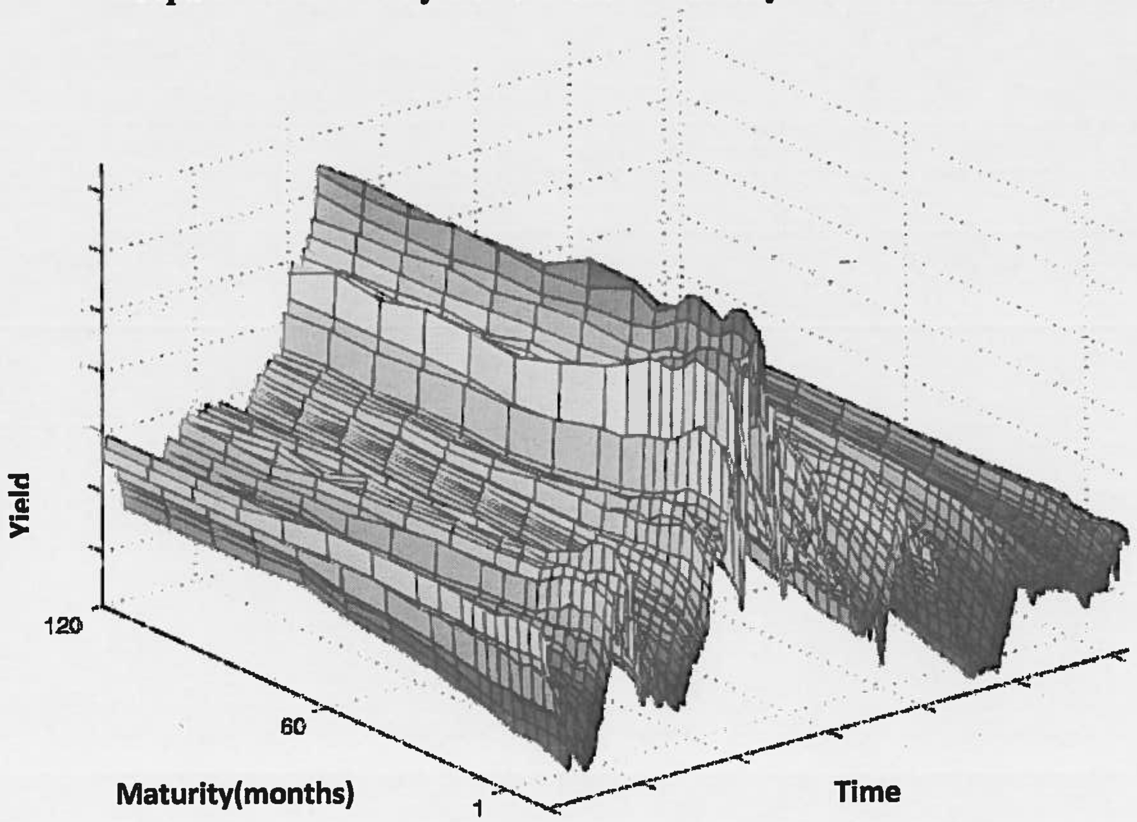
where i=0,1,2 and $\varepsilon_{i,t}$ is normally distributed error ($\varepsilon_{i,t} \sim N(0,{\sigma_i}^2)$). $\varepsilon_{i,t}$ is identically and independently distributed; $E\left[\beta_{i,t}\varepsilon_{i,t}\right] = 0$. c_i *and* γ_i have to be estimated by an OLS regression of $\beta_{i,t}$ s on an intercept and beta lagged value $\beta_{i,t-h}$.

Is the AR(1) structure realistic and a good structure to forecast the betas? Diebold and Li tried several structures to forecast the yield curve such as random walk, slope regression, AR(1) on yield levels, and VAR on factors. They found that AR(1) for the betas is the best structure to forecast future yields. They also provided some evidence on the goodness of fit of the AR(1) models fit to the estimated level, slope and curvature factors, showing residual autocorrelation functions. The autocorrelations are very small, indicating that the models accurately describe the conditional means of level, slope and curvature. As simple AR(1) factor structure satisfies the "parsimony principle", it indicates that imposing restrictions, both restrictive and unrealistic, can help to produce good forecasting models. Moreover, they use AR(1) structure instead of VAR since unrestricted VARs tend to produce poor forecasts of economic variables even when there is important cross-variable interaction, due to the large number of included parameters and the resulting potential for in-sample overfitting.

4. DATA

Data used in this thesis is monthly yield U.S. Treasuries that cover a period from January 1996 to December 2009 with a total of 168 time observations. Each of them is composed by 9 yields on maturities of 3, 6, 12, 24, 36, 60, 84, 120, 240 months. The data comes from the Federal Reserve Bank Reports in Wharton Research Data Services (WRDS) which is called as "Treasury constant maturity" dataset. The WRDS RATES database contains selected interest rates for U.S. Treasuries and private money market and capital market instruments. All rates are reported in annual terms.

Graph 1: Entire set of yield curves from January 1996 to December 2009



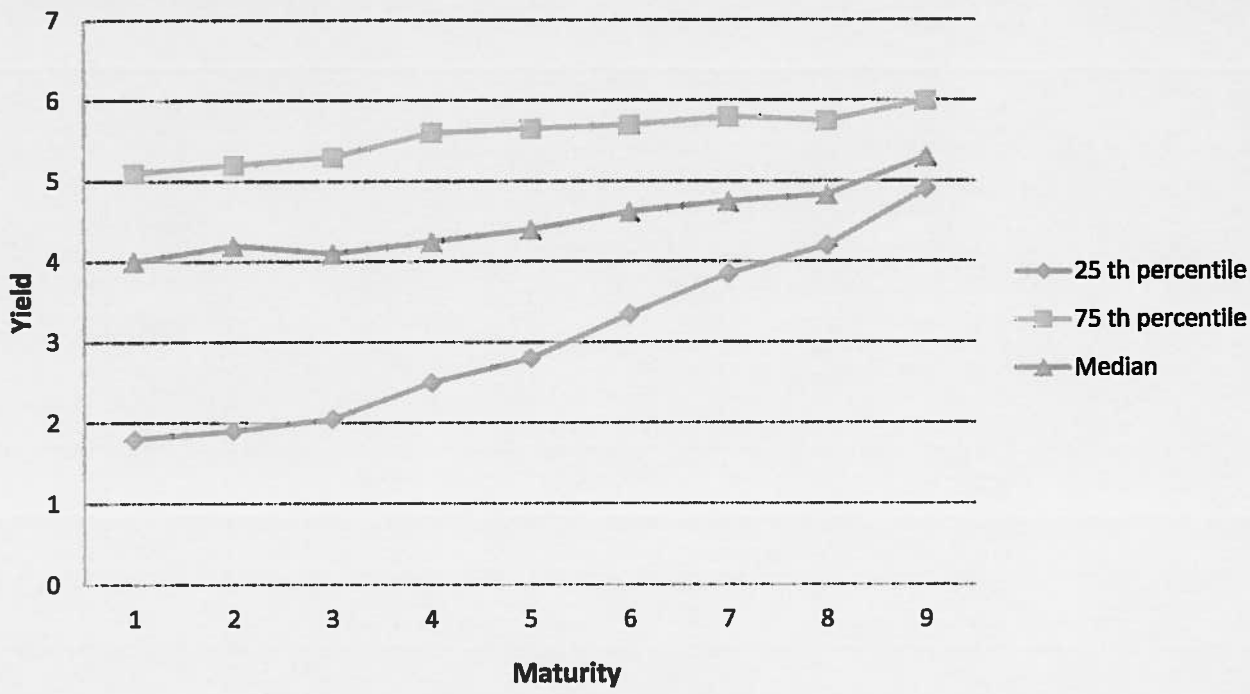
Graph 1 shows the complete series of yield curves for each monthly observation over the full length of the sample. It is important to note the great volatility that affected interest rates over the sample period.

In the Table 1 a summary of the principal statistics from the yield curves is putted, as it can be expected that the average short-term yield is lower than the average long-term yield and the volatility of short-term interest rate is much higher than the volatility of longer maturity interest rates. Additionally, the 3-months interest rate fluctuated from a maximum value of 6.36% to a minimum value of just 0.03%. It has huge volatility but it is known that the short term interest rates fluctuate more than long term interest rates. Another important characteristic of the data is the various forms that the yield curves assumed over the months. It can be said that increasing and decreasing yield curves, both concave and convex, would be interesting to see if the model is able to capture so different shapes. Nelson and Siegel had already wanted to find a yield curve model that would be simple but flexible enough to generate the range of shapes. An important aspect to note is that the maturities of the yields from the dataset are not equally spaced in time since there are closer maturities at the short end of the yield curve. Therefore, in the estimation the short-term interest rates will have more weight.

Table 1: Yields statistics (all the data are percentage values)

Maturities(months)	Mean	St. Dev.	Min	Max
3	3.40	1.91	0.03	6.36
6	3.55	1.91	0.15	6.39
12	3.67	1.85	0.31	6.33
24	3.94	1.75	0.80	6.81
36	4.12	1.60	1.07	6.77
60	4.44	1.34	1.52	6.76
84	4.71	1.19	1.89	6.86
120	4.88	1.01	2.42	6.91
240	5.41	0.86	3.18	7.22

Graph 2: Yields distribution



In the last graph, it can be clearly seen that how the area where half of the observations is much wider for the short-term rates with respect to the 120 or 240 months yields, confirming the greater volatility for short-term interest rates.

The ZCBs

To implement the trading strategy, some traded assets are needed to buy and sell, to make the analysis more consistent possible to be constructed artificial zero coupon bonds (ZCB):

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

where $P_t(\tau)$ denotes the price of a τ-period ZCB. And $y_t(\tau)$ denotes its continuosly compounded zero coupon nominal yield to maturity.

5. MODEL ESTIMATION AND FORECAST

As it has already mentioned that fitting the yield curve and trying to forecast the future level of interest rates will be made by using the Nelson-Siegel model. Following the paper of Diebold and Li (2006), the same formulation of the Nelson and Siegel model will be implemented:

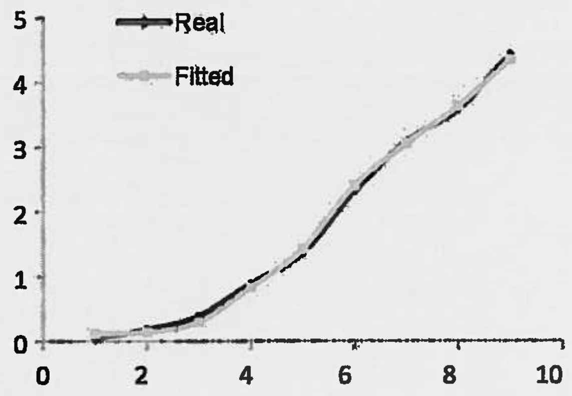
$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right)$$

5.1) IN-SAMPLE FIT

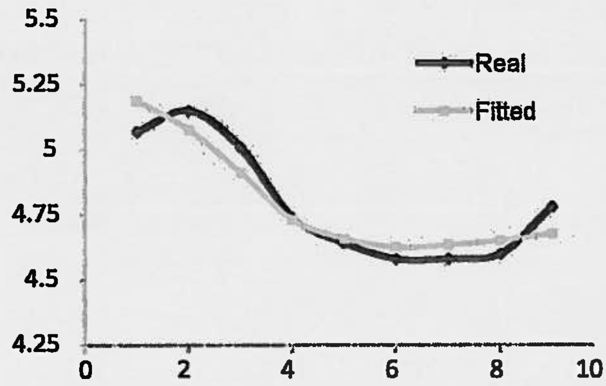
Nelson-Siegel – AR(1) model will be used to forecast future yields; but, first of all, it is reasonable to analyze the in-sample fit of the model because a necessary condition for a good forecasting model is its capacity to fit the data at a present date.

In the following examples (Graph 3) of fitted yield curves, it can be seen that the model is able to fit the data over different shapes. Furthermore, it is flexible enough to well fit increasing, decreasing, concave and convex yield curves. Besides, there is an irregular curve in Graph 3/c: May 2000 ; however, it is a very rare case.

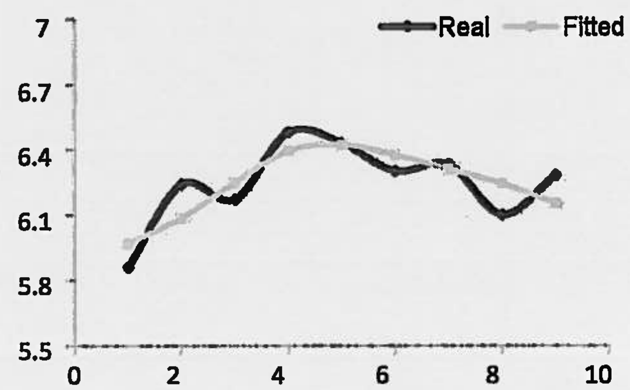
Graph 3/a: December 2009



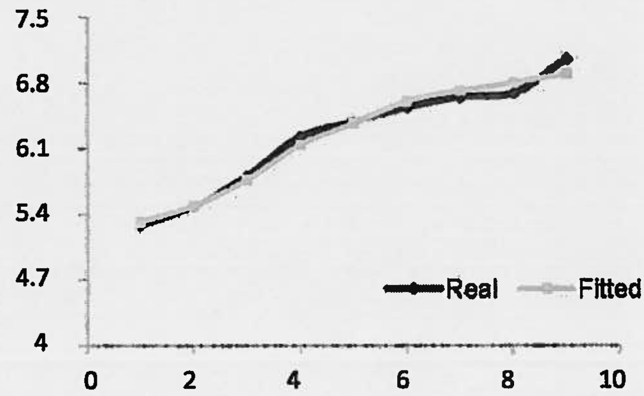
Graph 3/b: October 2006



Graph 3/c: May 2000

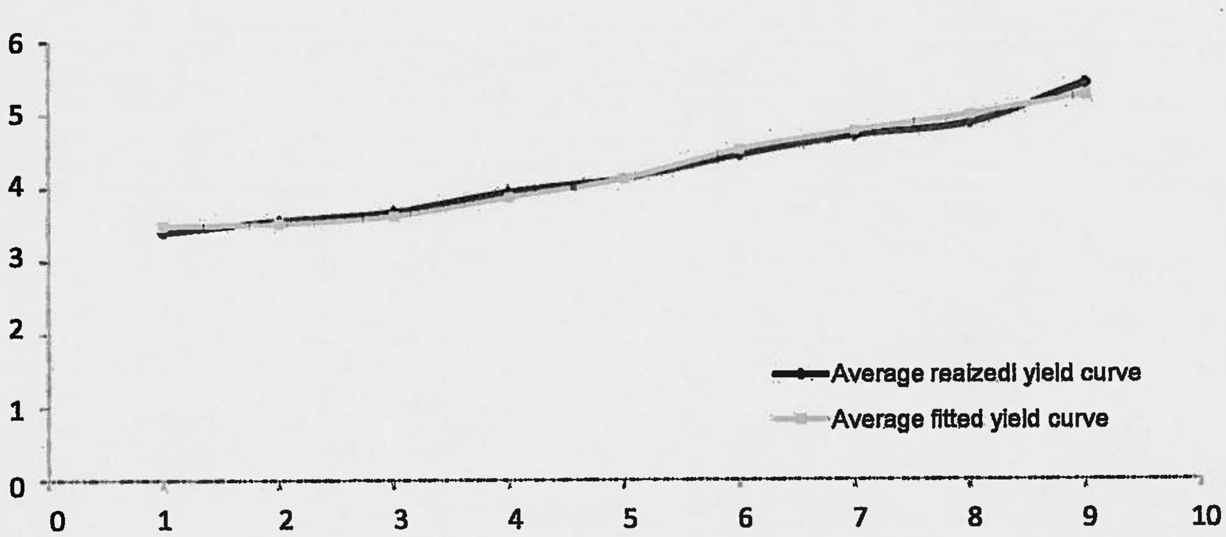


Graph 3/d: February 1997



In Graph 4, the average fitted yield curve and the average realized yield curve can be seen. As one can easily notice, the model on average has a very good performance to fit the realized yield curves because the two curves are very close each other.

Graph 4: Average realized and fitted yield curves



Three statistics are used in Ordinary Least Squares (OLS) regression to evaluate model fit: R-squared, the F-test, and the Root Mean Square Error (RMSE). In one hand, R-squared indicates the goodness of fit of the model while the F-test determines whether the proposed relationship between the response variable and the set of predictors is statistically reliable, and can be useful when the research objective is either prediction or explanation. On the other hand, the Root Mean Square Error, which measures the quality of the fit between the actual data and the predicted model, is the square root of the variance of the residuals. Therefore, the RMSE can be interpreted as the standard deviation of the unexplained variance, and has the useful property of being in the same unit as the response variable. It also indicates how close the observed data points are to the model's predicted values. Lower values of the RMSE imply better fit. Moreover, the RMSE is a good measure of how accurately the model predicts the response, and is the most important criterion for fit if the main purpose of the model is prediction. Considering these properties, the RMSE is calculated to show the accuracy of the model used in this paper.

$$\text{RMSE} = \sqrt{MSE} = \sqrt{1/_T \sum_{t=1}^{T} (fittedyield - realizedyield)^2}$$

The Table 2 summarizes the statistics of residuals from the regressions.

Table 2: Regression residuals statistics

Maturity (months)	Mean	St. Dev.	Min	Max	RMSE
3	0.076	0.067	-0.076	0.349	0.100
6	-0.036	0.049	-0.207	0.064	0.063
12	-0.053	0.075	-0.368	0.147	0.089
24	-0.068	0.035	-0.161	0.013	0.077
36	-0.002	0.033	-0.077	0.176	0.032
60	0.067	0.047	-0.061	0.201	0.077
84	0.057	0.059	-0.077	0.244	0.084
120	0.108	0.056	-0.049	0.270	0.122
240	-0.151	0.056	-0.310	-0.027	0.161

By looking at Table 2, it can be inferred that the model fits data well: the means of the residuals for all maturities are close to zero and the standard deviations are low. More specifically, Table 2 shows that the averages of the residuals from the fitted Nelson-Siegel model, for the included maturities, are all lower than 16 basis points (bp), in absolute value. For the Nelson-Siegel model the worst fitted maturity is the 240 month segment with a mean of the residuals close to -16 bp. A similar observation is made for the Nelson-Siegel model by Diebold and Li (2006). Substantially, the RMSEs, which are close to 0, give the conclusion that the model fits data well.

Last but not least, the statistics of the estimated 168 betas, three for each period, are summarized in the Table 3 below. In this table, the three-factor Nelson-Siegel model was fitted by using monthly yield data with λ_t fixed at 0.0609 and descriptive statistics for the estimated factors $\hat{\beta}_{0,t}$, $\hat{\beta}_{1,t}$ and $\hat{\beta}_{2,t}$.

Table 3: Statistics estimated betas

	Mean	St. Dev	Min	Max
Beta 0	5.534	0.776	3.383	7.213
Beta 1	-2.076	1.704	-5.004	0.683
Beta 2	-1.953	2.129	-6.125	2.136

5.2) OUT-OF SAMPLE FORECASTS

After the analysis of in-sample fit, now begining to forecast the future value of the yield curves in the Nelson-Siegel framework with fixed lambda can be started. In the light of information provided in Section 3, forecasting the yield curve is equivalent to forecasting the three betas. Therefore, it is suitable to annex an AR(1) structure to the model in order to forecast the future value of the betas.

The yield forecasts based on underlying AR(1) factor specifications are:

$$\hat{y}_{t+h/t}(\tau) = \hat{\beta}_{0,t+h/t} + \hat{\beta}_{1,t+h/t}(\frac{1-e^{-\lambda\tau}}{\lambda\tau}) + \hat{\beta}_{2,t+h/t}(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau})$$

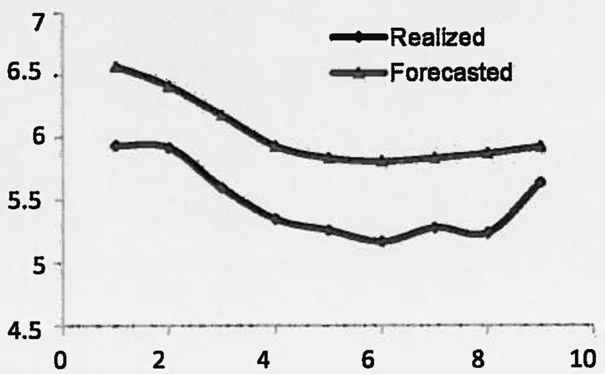
where $\hat{\beta}_{i,t+h/t} = \hat{c}_i + \hat{\gamma}_i\hat{\beta}_{i,t}$ i = 0,1,2

and $\hat{c}_i$ and $\hat{\gamma}_i$ are obtained by regressing $\hat{\beta}_{i,t}$ on an intercept and $\hat{\beta}_{i,t-h}$.

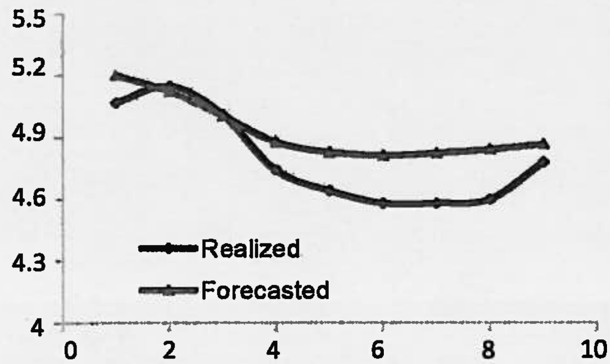
For ordinary least square regression, it has been made use of a rolling window of twelve observations. After obtaining the parameters for time t, they will be used to forecast the beta at time t+h. For instance, I used the observations from 1 to 12 in order to estimate c and γ . Later on, these coefficients are utilized to get the forecast for betas for period 13; to obtain the h periods ahead forecast of the yields, it is just needed to plug the forecasted betas in the Nelson -Siegel model.

The following four graphs are examples of the forecasted yield curve 1-month ahead:

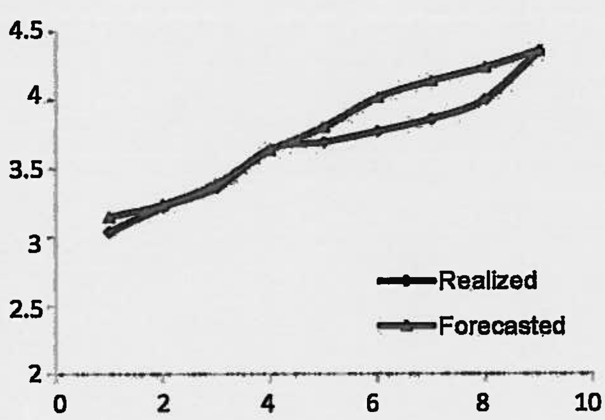
Graph 5/a: December 2000



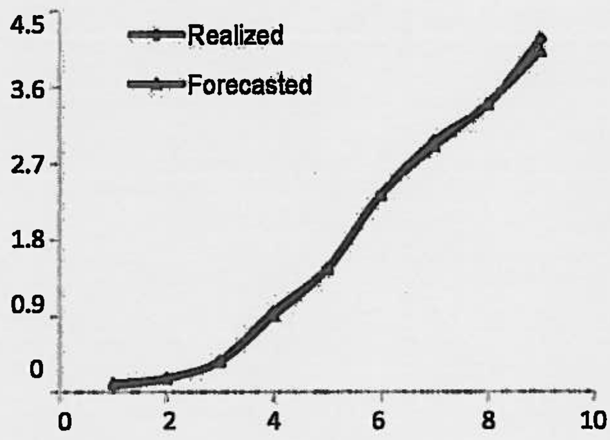
Graph 5/b: October 2006



Graph 5/c: May 2005



Graph 5/d: September 2009



In the above first two graphs (*Graph 5/a:* December 2000 and *Graph 5/b:* October 2006), it can be observed two bad results with forecasts far away from the realized yield curves while the latter two graphs (*Graph 5/c:* May 2005 and *Graph 5/d:* September 2009) indicate very good results with forecasts very close to the realized yield curves (actually Graph 5/d shows a sort of "miraculous" forecast).

Table 4, Table 5 and Table 6 show the statistics of the residuals for 1-month ahead forecast, 6-months ahead forecast and 1-year ahead forecast, respectively. In this case, the Forecasted Root Mean Square Error (FRMSE) will be used:

$$\text{FRMSE} = \sqrt{1/_T \sum_{t=1}^{T} (forecastedyield - realizedyield)^2}$$

Table 4: 1-month ahead forecast residuals statistics

Maturity (months)	Mean	St. Dev	Min	Max	FRMSE
3	0.110	0.281	-0.688	0.848	0.301
6	-0.007	0.241	-0.715	0.764	0.240
12	-0.022	0.252	-0.900	0.790	0.250
24	-0.035	0.287	-0.921	0.676	0.288
36	0.029	0.289	-0.724	0.754	0.290
60	0.093	0.279	-0.564	1.083	0.293
84	0.081	0.270	-0.594	1.247	0.281
120	0.131	0.245	-0.462	1.189	0.277
240	-0.133	0.232	-0.749	1.016	0.266

Table 5: 6-months ahead forecast residual statistics

Maturity (months)	Mean	St. Dev	Min	Max	FRMSE
3	0.250	1.417	-4.515	5.744	1.435
6	0.164	1.386	-4.396	5.316	1.390
12	0.194	1.334	-4.096	4.493	1.343
24	0.207	1.256	-3.500	3.618	1.269
36	0.261	1.152	-2.794	3.408	1.177
60	0.279	0.969	-1.915	2.937	1.005
84	0.227	0.853	-2.008	2.531	0.880
120	0.239	0.731	-1.641	2.241	0.767
240	-0.074	0.600	-2.013	1.381	0.603

Table 6: 1-year ahead forecast residual statistics

Maturity (months)	Mean	St. Dev	Min	Max	FRMSE
3	0.998	3.225	-4.842	10.351	3.364
6	0.811	3.032	-4.429	9.404	3.128
12	0.686	2.684	-3.565	8.133	2.760
24	0.508	2.250	-3.388	6.330	2.298
36	0.455	1.923	-2.893	5.172	1.969
60	0.372	1.453	-2.475	3.735	1.495
84	0.283	1.206	-2.417	2.858	1.235
120	0.266	0.973	-2.119	2.265	1.005
240	-0.071	0.767	-2.339	1.185	0.767

As it can be seen from the above tables (Table 4, Table 5 and Table 6) , the general forecast accuracy of the Nelson-Siegel – AR(1) model is better for one month ahead forecasts with the FRMSE sensibly lower with respect to the 6-months ahead forecasts and 1-year ahead forecasts. In their paper, Diebold and Li (2006) also show that the results of out-of-sample 1-month ahead forecast for Nelson-Siegel with AR(1) are much lower than 6-month and 12-month.

Having analyzed the statistical property of forecasts with different horizon periods, in the next part the most accurate one, the one month ahead forecast, will be used to build a new trading strategy.

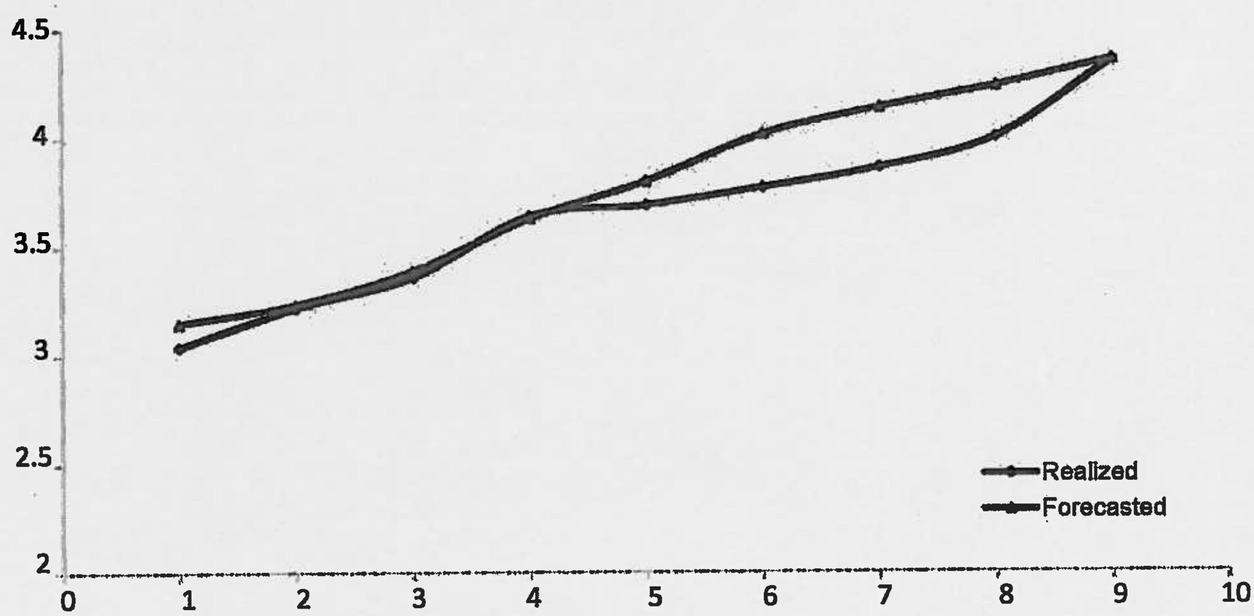
6. IMPLEMENTING THE TRADING STRATEGY

Until that section, the forecasts that the model created and the data from the real yield curves has been obtained. Based on the discrepancy between what the model predict and what are the yields on the market, it is willing to create a trading rule to exploit those mispricing opportunities. The idea is to buy what, based on my forecasts, is cheap and to sell what is expensive.

The trading strategy works as follows: At date t, the strategy forecasts the yield curve for the date t+h through the Nelson-Siegel – AR(1) model, compares the value at the maturities 3, 6, 12, 24, 36, 60, 84, 120 and 240 months of the forecasted yield curve with the present yield curve. If the yield forecasted is higher than the present real yield at a given maturity, then the strategy suggests to sell the ZCB with that maturity or if the yield forecasted is lower than the present yield for a given maturity, then it offers to buy the correspondent ZCB (lower interest rates imply higher ZCB prices and vice versa).

Graph 6 (yield curve May 2005) can be seen as an example of the trading strategy explained above. One can realize that the yield at maturity 5 years is lower than the one forecasted for the following period. Hence, the model expects an increase in the 5 years yield. In order to exploit this mispricing opportunity, the strategy tells us to sell the ZCB at 5 years to profit from the expected decrease in the ZCB price. After each period, one must close all the positions and re-run the same strategy. For instance, if at time t, given the signals of the model, the ZCB with maturity 3 months was bought; at t+1 the position selling the ZCB will be closed and re-build the portfolio following the new signals of my strategy for the period ending at t+2.

Graph 6: Yield curve May 2005



When the portfolio was built, all the monetary value of it will be invested that obtaining from the previous period. The same weight is given to each ZCB and this means, for instance, if the portfolio is invested in 4 ZCBs on different maturities, then each ZCB will have a share of the portfolio value equal to 25%.

6.1) ONE MONTH HOLDING PERIOD STARTEGY

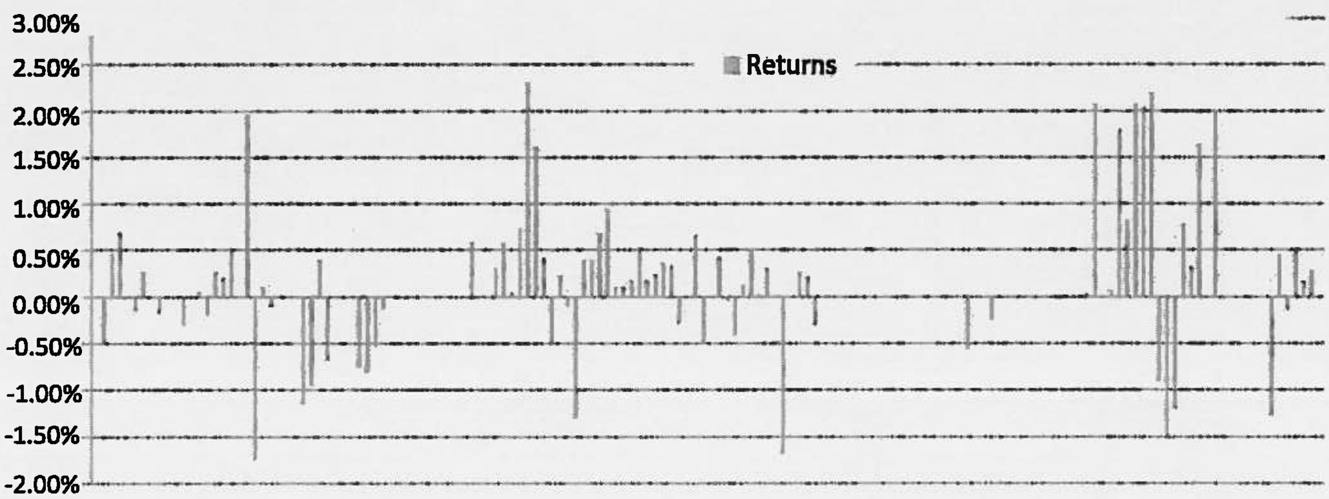
Implementing this strategy over a holding period of 1 month means that the forecast is for 1-month ahead and the portfolio is rebalanced every month. In this strategy, only for long positions are allowed, so every month ZCBs can be only bought and not short them. The initial capital invested has normalized to 1, and the portfolio is equally weighted among all the ZCBs in which the strategy tells us to invest.

Table 7: Statistic portfolio invested with the strategy

Average return per month	0.154%
Average return per year	1.851%
Overall return over the 154 months	23.749%
Volatility over the 154 months	0.824%
Average year volatility	0.678%

Over the entire period which is analyzed, the strategy earned a return of approximately 23%. If it is considered that the strategy runs for 154 months, the results seem not to be satisfying, and do not cover the cost of inflation. The Graph 7 shows the returns of the strategy for each single month over the entire period.

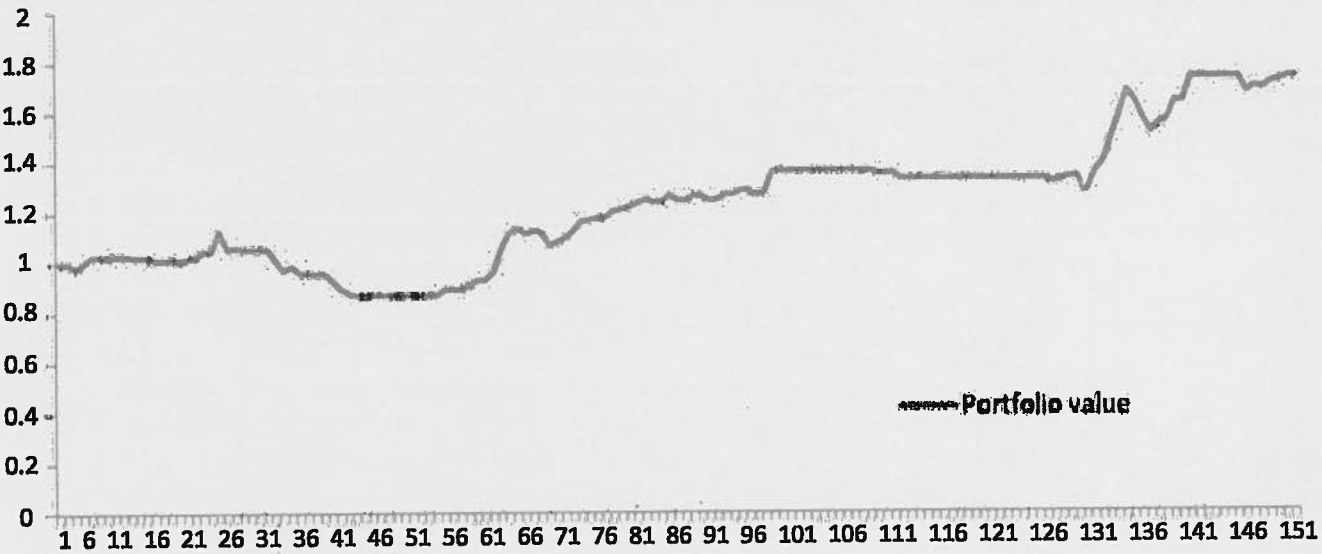
Graph 7: Strategy returns



In Graph 7, the highest monthly return is 2.3% as the lowest monthly return is -1.7%. In this part, the aim is to analyze how the results of the strategy evolved during all the investing periods. As it can be understood from the graph above, there are periods where the returns are positive and other where the returns are negative. It is important to note that when the strategy reached negative returns this means not only that the model bad forecasted the future yield curve but also wrongly predicted the direction of the change. In contrast to the statistical analysis of the forecasting properties of the model in previous section, here it is not our business about being interested in the precision of the forecasts but, given the trading rule, just interested in if the model gets the right direction of the change. Several negative returns say that the model was not able to do that several times along the sample period. At the end of this section will be explained why this happened.

In the Graph 8, the cumulative value of the portfolio (with initial value normalized to 1) over 154 months is plotted. This graph is even more clear to see that how the profitability of the strategy depends on the period that should be considered. In particular, the observations from 26 to 51 form the period where the strategy earned negative returns while after the period 51 there is a strong trend with positive returns. It is noteworthy to underline that there are several periods in which the evolution of portfolio value is almost flat. Those periods correspond to 0 investments in ZCBs. In other words, the strategy tells us to stay in cash and not to invest at those periods, which can be seen as a limitation as well.

Graph 8: Value **evolution portfolio invested with the strategy**



6.2) ANALYSIS ON THE PROFITABILITY OF THE STRATEGY

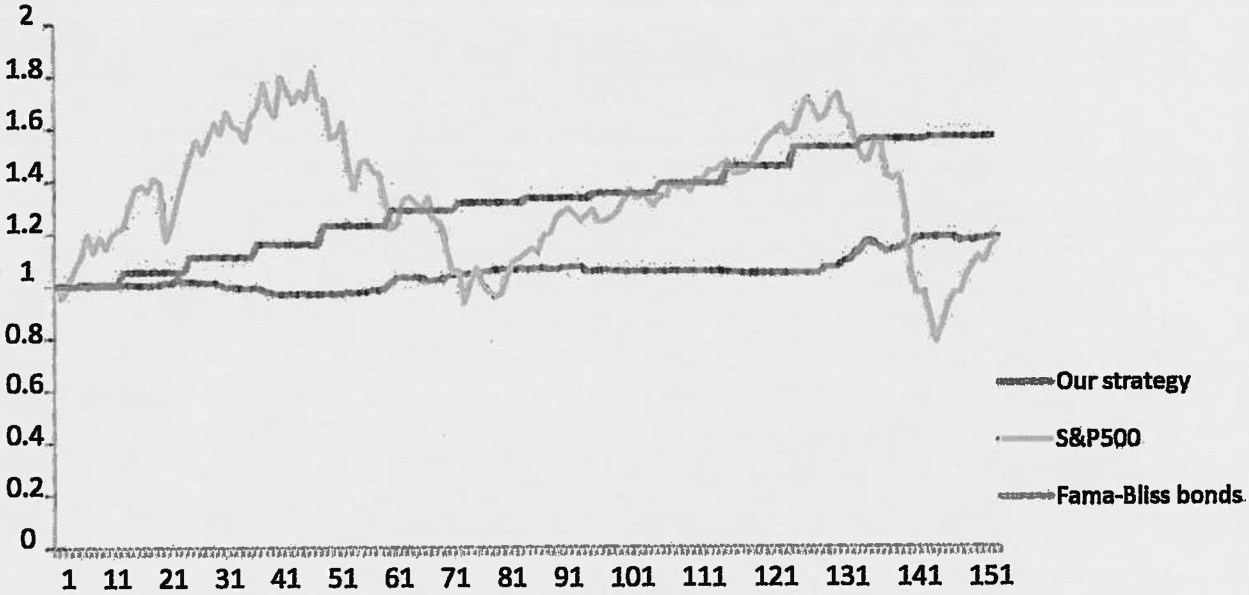
To confirm the result of one-month holding period strategy, two different analyses will be implemented: the first one compares the strategy returns with some benchmark investments and also the returns of the strategy with the benchmark returns. In the second analysis a multifactor model is selected. The price the return of portfolio and asset correctly was assumed, and it isimportant to note if the portfolio generated some abnormal returns.

Analysis I

In this part of the analysis, two benchmark portfolios was selected. The first one is a portfolio which is ideally invested on the index S&P500 and rebalanced each month with initial value normalized to 1. The second portfolio is an investment on the artificial 1-year maturity bond that comes from the Fama-Bliss (which consist of sequentially extracting forward rates from bond prices with successively longer maturities) discounted bond prices dataset in CRSP (The Center of Research in Security Prices) monthly treasury. The 1-year bond will be bought and holding it until maturity and then reinvest the whole portfolio value with a new 1-year bond that holding until maturity, and so on. The portfolio on S&P500 as a benchmark for investment in risky assets will be considered, while the portfolio on the 1-year bond as a risk free benchmark.

Graph 9 is just a qualitative comparison of the strategy with two benchmarks, both of them have very different characteristics compared to the strategy. They are selected in order to compare the strategy with a "risky benchmark" and a "risk-free benchmark". As it can be seen from Graph 9, at the final date, the strategy slightly outperformed the S&P500 but underperformed a buy-and-hold strategy in the 1-year bond(Fama-Bliss). Looking at the patterns, it is obivous that the strategy underperformed the S&P500 for most of the time. In detail, there are just two periods where the strategy portfolio was above the S&P500 portfolio, from June-2001 to June-2002 and from September-2007 to December-2009. Those periods corresponded to two periods with very low interest rates after 2001 and 2008 crises. In these periods, the strategy told us to buy ZCB by expecting a decrease in interest rates and hence an increase in ZCB prices. With a strategy that does not allow for short selling it is clear that the only periods in which the strategy can deliver consistently positive returns is during a decrease of interest rates. This is obviously a limit of the strategy.

Graph 9: The strategy portfolio and two benchmark investments evolution



The key point to consider is the volatility of the strategy returns. In fact, it is much lower than the S&P500 index for our strategy and this is very intuitive. In the strategy, ZCB that have lower volatility than stock prices will be used for the investment. This is true especially for ZCB near the maturity, indeed the greatest part of the strategy gains comes from the investment in longer maturity ZCBs, the one with more volatile prices. This is another limitation to the strategy that have been used: the asset class in which only ZCBs are invested, which means no coupons, have no big increase or decrease in their prices and this characteristic of ZCB prevents the strategy from being able to achieve very high returns (and also very low returns).

Analysis II

The returns of the trading strategy will be regressed on a factor model that it is assumed that price bonds is correct. Following the paper by Fama and French (1993), their two factor model for bond returns is used:

$$R_p(t) - R_f(t) = mTERM(t) + dDEF(t)$$

R_f is the risk free return and the rate on the fed funds was taken as a proxy of that. TERM and DEF are the two factors that are calculated as the following:

TERM: is the difference between 20 years treasury bonds yields and 3 month treasury bills yields. TERM is a proxy of unexpected change in interest rates since the short term yield on the three month treasury bills is considered as a proxy for general level of expected return on bonds.

DEF: is the difference between an average of Aaa and Baa corporate bond yields and 20 years treasury bond yield[1]. DEF is a proxy of the change in the economic environment and the default probability.

The regression is:

$$R_p(t) - R_f(t) = \alpha + mTERM(t) + dDEF(t) + \varepsilon_t$$

[1] this is different from the factor calculated by Fama and French where they took a portfolio of corporate bonds and not just their yields, the difference is due to impossibility to access to data from portfolio of corporate bonds.

If the portfolio earns abnormal returns then the α should be significantly different from 0. The idea is that Fama and French two factor model correctly price bond returns if in the regression analysis the strategy shows an intercept significantly different from zero. It means that, due to the role of the trading strategy, significant intercept represents the excess adjusted return of the active strategy.

The Table 8 shows the principal statistics from the regression:

Table 8: Regression results of the coefficients

	Coefficients	t-stat
α	-0.62	-2.55
m	0.84	9.09
d	0.76	4.62

The average risk premiums for the term-structure factors are important while explaining the bond returns. Low average premiums will prevent TERM and DEF from explaining much cross-sectional variation in average returns; but, high volatility implies that the two factors can capture substantial common variation in returns. The low means and the high volatilities of TERM and DEF will be advantageous for explaining bond returns.

The slopes on TERM and DEF and adjusted R-square allow direct comparisons of the common variation in bond returns tracked by the term-structure variables. The model has an adjusted R-square of 81.3%. The high adjusted R-square and the highly significant coefficients m and d mean that the model actually fit quite well the data and these factors are important to explain bond returns. As one would expect, long-term bonds are more sensitive than short-term bonds to the shifts in interest rates measured by TERM. The slope of DEF (d) shows that a common 'default' risk in returns increases from government bond to corporate.

In fact, this analysis is not about the validity of the two-factor model which be taken as granted. What is important to look for is the intercept since it measures the abnormal return that the portfolio generated respect to "normal" bond dynamics. The intercept is close to 0 (-0.62% per month), so it is not surprising that the average excess return on the bond portfolio is close to 0. The α is significantly negative which means that the strategy earned negative abnormal returns, or put another way, the strategy excess risk-adjusted returns were negative.

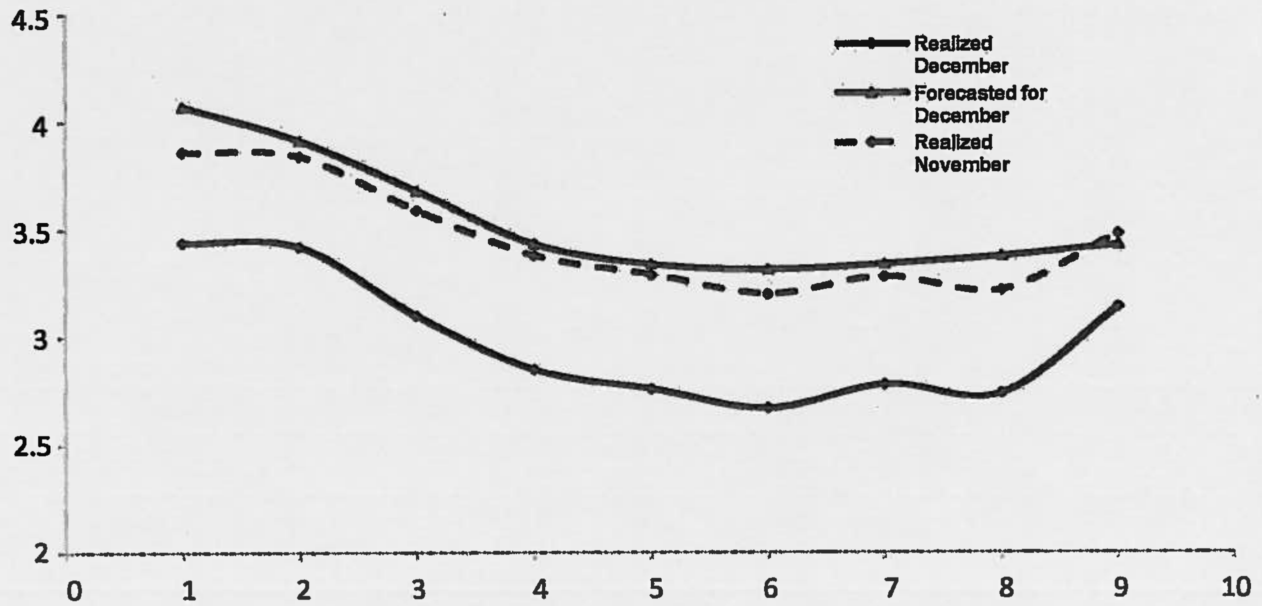
It can be concluded that this analysis confirms what was intuited in the previous comparative analysis. Both the comparative and pricing analysis confirmed a poor performance of the strategy compared with benchmark indicators in the first case and "normal" bond returns in the second case. It is going to be understood what went wrong in the trading strategy and therefore, the situation further in the next section will be diagnosed.

6.3) DIAGNOSIS

It is confirmed that the strategy has a poor performance with respect to the results of last two analyses and it will be looking for the reason why. In this section, it has been already mentioned the problem that sometimes the model misses the direction of the change of the yield curve. This may cause problems if one wants to apply the model as a forecast tool. Thus, the same graph in Section 5 (Graph 5/a) is reproposed, the yield curves realized and forecasted for December 2000 with an added yield curve at November 2000.

Given the yield curve at November 2000, ZCB for all maturities can be bought and the profit from the future decrease in all yields can be obtained. However, the model predicted an increase in the rates and hence, the created strategy was to keep our portfolio in cash.

Graph 10: Comparison among realized yield curves at November and December 2000 and the forecasted yield curve for December 2000



Why does this failure in forecast happen? The answer is in the particular structure that some forecasts for the betas was made. It is assumed that an AR(1) structure and what happened in this specific case is the following: as can be seen from Graph 11 in the two years before November 2000, the interest rates increased continuously and after that, they had started to decrease for the next three years. November 2000 is exactly at the peak of the curve. In an AR(1) structure, the variable forecasted depends only on its past values. If the rates increased continuously, an AR(1) structure will continue to predict an increase also in the following months. It will take a while to adapt to the change of economic environment. This is the biggest flow of this model; by imposing an AR(1) structure for forecasting future betas, it is just relying on past observations and the model can be completely missed other signals that come from other parts of the economy. Especially, the short term rates are highly influenced by central banks, the forecasts can be improved by including one or more macroeconomic variables that can predict the movement of central banks and in general the state of the economy.

Graph 11: 3-months interest rate evolution (January 1999 – January 2002)

Ideally, these variables should have very low importance during "normal" period since in normal condition my AR(1) structure performed well to predict future interest rates; but, be able to upgrade the model during a change in the economic cycle to make the adapted of the model quicker. The inclusion of some macroeconomic variables in the AR(1) structure to forecast future betas and future yield curves is my advice for future expansion of the Nelson-Siegel model as forecasting tool.

7. FINAL COMMENTS AND CONCLUSION

In this paper, all the steps to build and test a trading strategy was covered. The starting point is selecting a model for the dynamics of the interest rates. Later on, estimation and testing the model was made; it was found that the in-sample performance of the model is good, it is able to fit different shapes of the yield curve. Then, this model was expanded to create a way to forecast the future yields, and imposed an AR(1) structure on the betas. The finding is that the out-of-sample forecast performance is not as goog as than the in-sample fit; sometimes the model missed the direction of the change in the yield curve. This problem is due to my AR(1) structure on the betas. In this way, the model takes a while to switch from a regime of increasing interest rates to a decreasing interest rates environment (and vice versa). The problem was particularly serious in the trading strategy; sometimes buy opportunities are missed and other times buying zero coupon bond might lead a loss when their price dropped in the following periods. It can be concluded that the strategy based on the Nelson-Siegel – AR(1) model gives poor results, and the overall return on the entire periods was very low, although it is positive. The comparative analysis with benchmark indicators confirmed the poor performances of the strategy and the same from the pricing analysis of the alpha from the two factor model.

The problem with the strategy can have several sources: the first one is that the trading rule that was used in this work: allowing just for long positions can be actually main reason in missing profitable opportunities from the upward shift of the yield curve; hence the trading strategy can be improved by changing the trading rule, allowing for short selling or giving different portfolio weights to zero coupon bonds based on the strength of the buy or sell signals. The another problem concerns the asset class that was considered, zero coupon bonds have very low volatility, their prices do not move a lot up or down, this actually obstruct from reaching high returns, the strategy could be expanded by adding other assets like derivatives or coupon bonds. The third one is that as it was said a problem is the created model to forecast interest rates, in particular the AR(1) structure for the betas seem to keep the model slow to adapt in changing trends of interest rates; since trend in interest rates are driven by economic cycles I suggest to include some macroeconomic variables that can capture this aspect, this is the expansion that I suggest for the AR(1) structure for the betas.

As a conclusion, I want to highlight the take away from this paper. My trading rule was just a driving example on how to build one. The paper can be taken either as a guide for all the steps needed to rigorously test a trading strategy or as a mix of statistical and economical test of an interest rate model. In any case, the idea is that this paper can be replicated with a different trading rule, different asset classes and different models. What it is important to keep in mind is the importance of the assumptions behind the models; it is critical to understand them and recognize their implication. For the trading rules, instead, the only limitation is the investor's creativity.

References

- Erdogan, O., Bennett, P., Ozyildirim, C. (2013) Predicting U.S. Recessions: Yield Curve and Stock Market Liquidity Deviation as Leading Indicators, 20th Annual Conference of the Multinational Finance Society, Turkey 2013

- Balduzzi, P., Das, S. R., Foresi, S. & Sundaram, R. (1996). A Simple Approach to Three-factor Affine Term Structure Models. Journal of Fixed Income

- Balduzzi, P., Bertola, G. & Foresi, S. (1993). A Model of Target Changes and the Term Structure of Interest Rates. Working Paper

- Estrella, A., Hardouvelis, G. A., (1991). The Term Structure as a Predictor of Real Economic Activity. Journal of Finance

- Estrella, A., Mishkin, F. S., (1996). The Yield Curve as a Predictor of U.S. Recessions. Federal Reserve Bank of New York, Current Issues in Economic and Finance

- Björk, T. & Christensen, B. J. (1999) Interest Rate Dynamics and Consistent Forward Rate Curves. Mathematical Finance

- Coroneo, L., Nyholm, K. & Koleva, R. V. (2008). How Arbitrage-Free Is the Nelson-Siegel Model?. Working Paper Series

- Diebold, F. X., Christensen, J. H. E. & Rudebusch, G. D. (2008). An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model. Federal Reserve Bank of San Francisco, Working Paper Series

- Diebold, F. X. & Li, C. (2006). Forecasting the Term Structure of Government Bond Yields. Journal of Econometrics

- Diebold, F. X., Piazzesi, M. & Rudebusch, G. D. (2005). Modeling Bond Yields in Finance and Macroeconomics. The American Economic Review

- Diebold, F.X., Rudebusch, G.D. & Aruoba, S. B. (2005). The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. Journal of Econometrics

- Diebold, F.X., Ji, L. & Li, C. (2006). A Three-Factor Yield Curve Model: Non-Affine Structure, Systematic Risk Sources, and Generalized Duration. Macroeconomics, Finance and Econometrics

- Fabozzi, F. J., Martellini, L. & Priaulet, P. (2005). Predictability in the Shape of the Term Structure of Interest Rates. Journal of Fixed Income

- Fama, E. F. & French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics

- Fama, E. F. (1984). The Information in the Term Structure. Journal of Financial Economics

- Fama, E. F. & Bliss, R. R. (1987). The Information in Long-Maturity Forward Rates. American Economic Review

- Filipovic, D. (1999). A Note on the Nelson-Siegel Family. Mathematical Finance

- Fong, G., Vasicek, O. & Yoo, D. (1992). Omission Impossible

- Litterman, R. & Scheinkman, J. (1991). Common Factors Affecting Bond Returns. Journal of Fixed Income

- Longstaff, F. A. & Schwartz, E. S. (1993). Implementation of the Longstaff-Schwartz Interest Rate Model. Journal of Fixed Income

- Nelson, C. R. & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. The Journal of Business

- Nelson, C. R. & Siegel, A. F. (1988). Long-Term Behaviour of Yield Curves. The Journal of Financial and Quantitative Analysis

- Naik, V. & Lee, M. H. (1993). The Yield Curves and Bond Option Prices with Discrete Shifts in Economic Regimes. Working Paper

- Pooter, M., Ravazzolo, F. & Dijk, D. (2007). Predicting Term Structure of Interest Rates

- Pooter, M. (2007). Examining the Nelson-Siegel Class of Term Structure Models

- Swanson, N. R. & White, H. (1995). A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks. Journal of Business and Economic Statistics

- Veronesi, P. (2010). Fixed Income Markets. John Wiley & Sons Inc., Hoboken, New Jersey