A segmented Nelson and Siegel model for the term structure of interest rates

by:
Juan Felipe Peña Rojas

Advisors:
Carlos Castro
Cristhian Rodriguez.

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

Universidad Nuestra Señora Del Rosario
Faculty of Economics
Bogotá, Colombia
February 2019

Declaration of Authorship

I, JUAN FELIPE PEÑA ROJAS, declare that this thesis titled, 'A SEGMENTED NELSON AND SIEGEL MODEL FOR THE TERM STRUCTURE OF INTEREST RATES' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a master degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Singed:
Date: February 2019

Abstract

In the present academic work we implement the Nelson and Siegel Segmented Model (2017) in order to predict the structure of interest rates. On the other hand we compare the performance of the Segmented model with the Nelson and Siegel Classic model. The present work was done with daily data Colombian yields between the years 2013 and 2016, the data was obtained thanks to Precia who is the price provider for valuation in Colombia. We find that in the Segmented Model by locally adjusting the segments provides substantial improvements inside and outside the sample in comparison with the Classic model.

Keywords: Term structure, Nelson and Siegel model, Preferred habitat theory, Ordinary least squares, Root Mean Squared Error, Factor Loadings.

Contents

0.1 Introduction

The term structure of interest rates is the relationship between interest rates or bond yields and different terms or maturities. The term structure of interest rates is also known as a yield curve, and it plays a central role in an economy. The term structure reflects expectations of market participants about future changes in interest rates and their assessment of monetary policy conditions.

Most of the vast literature on the term structure of non defaultable securities is concerned with using observable security prices to estimate the fair market prices of other non-observable securities. This is extremely important because fixed-income securities and their derivatives trade only occasionally, and so must be priced based on other securities that do trade. A typical part of estimating the price at which a bond would trade involves decomposing its price into term and risk premiums. This analysis formally constrains the yield curve to be arbitrage-free.[ERS17]

The Nelson and Siegel model provides a parsimony specification to capture the differences in rates along the curve (for different maturities). Its implementation in one or two stages allows to recover the temporal variation of the factors maintaining the loads of the factors (loading factors) constant over time. The specification of the model and the estimation methods provide a strategy of simple implementation, which is why it also turns out to be a successful model outside the academia.

The preferred habitat theory of the term structure [MS66] advocates that local shocks may influence interest rates for each maturity. Empirical evidence related to this theory reveals that U.S. Treasury bonds' supply and demand shocks have nonnegligible effects on yield spreads, term structure movements, and bond risk premium. In an attempt to formalize the preferred habitat theory, [VV09] propose an equilibrium model in which demand directly influences and determines all yields in the term structure, in a dynamic way. According to this theory, the equilibrium yield rate for each term is determined by the demand and supply forces for that market, in other words, the preferences of investors on securities at that point in the curve. Investors can substitute preferences over terms that are not available in the market for a near term but available in the market.[DK13] note that investors act as arbitrators, guaranteeing the relationship between the demand for the securities and the returns along the curve; and on the other hand they guarantee that the curve is smooth, meaning that the yields for close periods are similar.

Inspired by the preferred habitat theory, and more specifically by its recent formalization by [VV09], [AAK+18] propose a class of models that separate the yield curve into segments which present their own local shocks, but which are simultaneously interconnected, composing the whole yield curve. The main objective of the family of segmented models is to address the following aspects:

1. The segmented model proposes to partition the segments of the curve in such a way that the dynamics of each segment can be determined by maturities that are represented in that segment.
2. The implementation of the segments along the curve must be globally consistent and smooth. This is achieved by ensuring that the rates of return are similar for the terms that connect the segments and therefore are close to each other. This is not equivalent to imposing non-arbitration restrictions.

0.2 Methodology

0.2.1 Nelson and Siegel Segmented Model

Let be $\tilde{\tau}$ a vector of m observable terms ($\{\tau_1 < \tau_2 < \tau_3 < \cdots < \tau_k\}$). The usual specification of a term structure model, for the set of yields $Y_t(\tilde{\tau})$, have a specification of a factorial model,

$$y_t(\tilde{\tau}) = W(\tilde{\tau})\beta_t + \varepsilon_t(\tilde{\tau}) \tag{1}$$

where $W(\tilde{\tau})$ is an mxK matrix of factor loadings and β_t is a K vector of no observable factors. For the three factors $(1, \frac{(1-\exp^{-\lambda\tau})}{\lambda\tau}, \frac{(1-\exp^{-\lambda\tau})}{\lambda\tau} - \exp^{-\lambda\tau})$ Nelson and Siegel model, the loadings matrix is determine with the triplet: the intercept, the slope and the curvature.

To define the segmented model, we consider three segments Short Term(ST), Middle Term(MT) and Long Term(LT) as show in Figure 1, then we define a set of nodes(border points of each segment) $\phi = \{\tau_0 < \tau_4 < \tau_7 < \tau_{11}\}$, with two external nodes(τ_0,τ_{11}) and two internal nodes(τ_4,τ_7). In this case the curve can be expressed as,

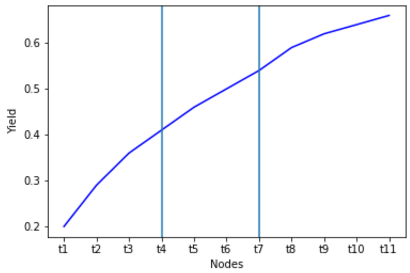

Figure 1: Representation of the 3 segments curve: Short Term, Middle Term and Long Term

$$y_t(\tilde{\tau}) = f_t^1 \mathbf{1}_{\tau_1 \le \tilde{\tau} \le \tau_4} + f_t^2 \mathbf{1}_{\tau_4 \le \tilde{\tau} \le \tau_7} + f_t^3 \mathbf{1}_{\tau_7 \le \tilde{\tau} \le \tau_{11}} + \varepsilon_t \tag{2}$$

where f_t^i function represent the curve of each of one of the segments, such that

$$f_t^i = a_t^i + b_t^i g(\tau) + c_t^i h(\tau) \tag{3}$$

where $g(\tau) = \frac{(1-\exp^{-\lambda\tau})}{\lambda\tau}$ and $h(\tau) = \frac{(1-\exp^{-\lambda\tau})}{\lambda\tau} - \exp^{-\lambda\tau}$.

0.2.2 Derivation

Now we are going to approach the methodology following an order according to the two aspects that are important in the segmented model which we highlighted in the in the Introduction Section.

How we build the Segmented model using Nelson and Siegel Factor Loadings?

The reason why we are going to estimate the Nelson and Siegel factors within each segment is to gather the notion that the behavior of that segment will be determined by the forces of supply and demand in that segment.

In our implementation we require 11 maturities in order to compose the yield curve ($\tau_i \;\; i \in [1, 11]$), where i's represent the position in the vector of available maturities. Then, we need to define exogenous k's elements of $\tau' s$ vector in which we are going to make the segmentation. In our case we define the set of partitions (Segments) with k=3 ($\phi = \{\tau_1, \tau_4, \tau_7, \tau_{11}\}$), which means we have (k+1=4) elements,ϕ are the nodes (border points) we choose for the model. On the other hand, the remaining elements of τ are treated as observed yields ($\tilde{\tau} = \{\tau_2, \tau_3, \tau_5, \tau_6, \tau_8, \tau_9 \tau_{10}\}$), which are assumed to be measured with error. Equation 4 is the matrix representation of equations 2 and 3.

$$y_t(\tilde{\tau}) = W(\tilde{\tau}) B_t + \epsilon_t(\tilde{\tau}) \tag{4}$$

$$y_t(\phi) = W(\phi) B_t \tag{5}$$

The matrix of loadings ($W(\tau_i) = \{1, g(\tau_i), h(\tau_i)\}$) is time invariant and only depends on maturities and segments, it contains the factor loadings consistent with the model of Nelson and Siegel. By contrary, the vector ($B_t = \{a_t^i, b_t^i, c_t^i\}$) is time variant and also depend on the segmentation i.

$$g(\tau_i) = \frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}$$
$$h(\tau_i) = \frac{1-e^{\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}$$

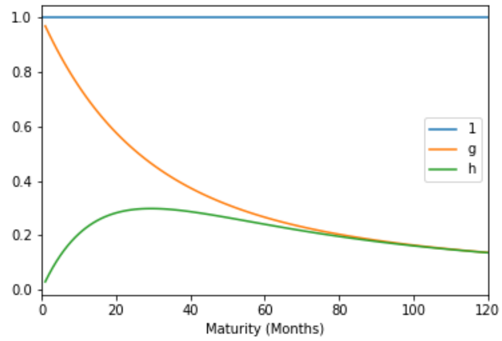

Figure 2: Factor Loadings Nelson and Siegel Model

How we guarantee smoothness across the term structure?

The implementation of the segments along the curve must be globally consistent and smooth. This is achieved by ensuring that the rates of return are similar for the terms that connect the segments and therefore are close to each other.

It is necessary to impose conditions to guarantee smoothness across the segments. We demand a matrix R of restrictions that are the constrains in the estimation problem.

$$y_t(\tau) = W(\tau)B_t, \;\; s.t \;\; R(\phi)B = 0$$

In order to create the matrix R of restrictions is necessary to know the first and second derivatives from the loadings that are placed inside the structure of R.

First Derivatives	***Second Derivatives***
$g'(\tau) = \frac{e^{-\lambda\tau}(1+\lambda\tau)}{\lambda\tau^2}$	$g''(\tau) = \frac{1}{\lambda\tau^3}[2 - e^{-\lambda\tau}((\lambda\tau)^2 + 2(\lambda\tau + 1))]$
$h'(\tau) = \frac{e^{-\lambda\tau}((\lambda\tau)^2+(\lambda\tau)+1)-1}{\lambda\tau^2}$	$h''(\tau) = \frac{1}{\lambda\tau^3}[2 - e^{-\lambda\tau}((\lambda\tau)^3 + (\lambda\tau)^2 + 2(\lambda\tau + 1))]$

R Matrix Construction

The first step is to guarantee a smoothness function through all the internal nodes in which the segmentation is done. In our case the internal nodes of ϕ are the set $\tilde{\phi} = \{\tau_4, \tau_7\}$ in which we need to impose three conditions that makes possible the continuity across segments. Besides, all the j's yields within the same segment need to be governed by the same function ($W(\tau_j) = f_t^i(\tau_j)$). With the purpose too be more clear with the notation we will refer to ST (Short Term), MT (Middle Term) and LT (Long Term) to the segments $\{\tau_1, \tau_4\}$, $\{\tau_4, \tau_7\}$ and $\{\tau_7, \tau_{11}\}$ respectively.

Smoothness Conditions

1. $f^{ST}(\tau_4) = f^{MT}(\tau_4), \quad f^{MT}(\tau_7) = f^{LT}(\tau_7)$
2. $f'^{ST}(\tau_4) = f'^{MT}(\tau_4), \quad f'^{MT}(\tau_7) = f'^{LT}(\tau_7)$
3. $f''^{ST}(\tau_4) = f''^{MT}(\tau_4), \quad f''^{MT}(\tau_7) = f''^{LT}(\tau_7)$

In total we have 3(k-1)=6 equations and 9 parameters ($B_t = \{a_t^1, b_t^1, c_t^1, a_t^2, b_t^2, c_t^2, a_t^3, b_t^3, c_t^3\}$) to be estimated, there are 9 parameters because we are going to estimate 3 factors for each of the 3 segments.

$$y_t(\phi) = W(\phi)B_t, \;\; s.t \;\; R(\tilde{\phi})B = 0$$

It is important to set the structural form of R and how to decomposed into a square invertible matrix R_1 and a complementary R_2 in order to reduce the dimensionality of parameters to be estimated.

$$X_i(\tau) = [1, g_i(\tau), h_i(\tau)] \quad X'_i(\tau) = [1, g'_i(\tau), h'_i(\tau)] \;\; ; X''_i(\tau) = [1, g''_i(\tau), h''_i(\tau)]$$

$$R = \begin{bmatrix} X_{ST}(\tau_4) & -X_{MT}(\tau_4) & 0_{1x3} \\ 0_{1x3} & X_{MT}(\tau_7) & -X_{LT}(\tau_7) \\ X'_{ST}(\tau_4) & -X'_{MT}(\tau_4) & 0_{1x3} \\ 0_{1x3} & X'_{MT}(\tau_7) & -X'_{LT}(\tau_7) \\ X''_{ST}(\tau_4) & -X''_{MT}(\tau_4) & 0_{1x3} \\ 0_{1x3} & X''_{MT}(\tau_7) & -X''_{LT}(\tau_7) \end{bmatrix}, \quad dim(R) = (6x9)$$

$$R_1 = \begin{bmatrix} X_{ST}(\tau_4) & -X_{MT}(\tau_4) \\ 0_{1x3} & X_{MT}(\tau_7) \\ X'_{ST}(\tau_4) & -X'_{MT}(\tau_4) \\ 0_{1x3} & X'_{MT}(\tau_7) \\ X''_{ST}(\tau_4) & -X''_{MT}(\tau_4) \\ 0_{1x3} & X''_{MT}(\tau_7) \end{bmatrix} \quad R_2 = \begin{bmatrix} 0_{1x3} \\ -X_{LT}(\tau_7) \\ 0_{1x3} \\ -X'_{LT}(\tau_7) \\ 0_{1x3} \\ -X''_{LT}(\tau_7) \end{bmatrix}, \quad \begin{pmatrix} dim(R_1) = (6x6) \\ dim(R_2) = (6x3) \end{pmatrix}$$

By construction all rows of R_1 are linearly independent, then, the matrix is invertible with rank equal to six. As we have divided the matrix R into two sub-matrix, we also require the same process for the vector B of factors. Hence, the original constrain is re-expressed from $(R(\tilde{\phi})B = 0)$ to $(R_1(\tilde{\phi})\theta_1 + R_2(\tilde{\phi})\theta_2 = 0)$, where the vectors $\{\theta_1, \theta_2\}$ are adjusted to match the dimensionality of sub-matrix $\{R_1, R_2\}$.

$$B = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \quad \theta_1 = \begin{pmatrix} F_t^{ST} \\ F_t^{MT} \end{pmatrix}, \quad \theta_2 = \left(F_t^{LT}\right), F_t^i = \begin{pmatrix} a_t^i \\ b_t^i \\ c_t^i \end{pmatrix}, \quad \begin{pmatrix} dim(\theta_1) = (6x1) \\ dim(\theta_2) = (3x1) \end{pmatrix}$$

We need to state that θ_1 is the vector that share neighborhood with the square invertible matrix R_1 and can be represented as an algebraic product of the form:

$$\theta_1 = -R_1^{-1} R_2 \theta_2, \quad dim(\theta_1) = (6x1)$$

Unconstrained Procedure

In the process we have some yields that are observed, these yields are located inside the segments of the term structure; In our sample the set is represented by $(\tilde{\tau} = \{\tau_2, \tau_3, \tau_5, \tau_6, \tau_8, \tau_9, \tau_{10}\})$ and we assume that are measured with error. Given the previous information, we implement the arrays division in B from R matrix to reconstruct the original representation known as measurement equation.

$$y_t(\tilde{\tau}) = W(\tilde{\tau})B_t + \epsilon_t(\tilde{\tau}), \quad B = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \quad dim(y_t(\tilde{\tau})) = (3x1)$$

$$y_t(\tilde{\tau}) = w_1(\tilde{\tau})\theta_1 + w_2(\tilde{\tau})\theta_2 + \epsilon_t(\tilde{\tau})$$

$$y_t(\tilde{\tau}) = [w_2(\tilde{\tau}) - w_1(\tilde{\tau})R_1^{-1}R_2]\theta_2 + \epsilon_t(\tilde{\tau})$$

$$y_t(\tilde{\tau}) = Z(\tilde{\tau})\theta_2 + \epsilon_t(\tilde{\tau}), \quad dim(Z(\tilde{\tau})) = m_x(k+1) = (3x3)$$

$$w_1(\tilde{\tau}) = \begin{bmatrix} X_3(\tau_3) & 0_{1x3} \\ 0_{1x3} & X_5(\tau_5) \\ 0_{1x3} & X_6(\tau_6) \end{bmatrix}, \quad dim(w_1(\tilde{\tau})) = (3x6)$$

$$w_2(\tilde{\tau}) = \begin{bmatrix} X_8(\tau_8) \\ X_9(\tau_9) \\ X_{10}(\tau_{10}) \end{bmatrix}, \quad dim(w_2(\tilde{\tau})) = (3x3)$$

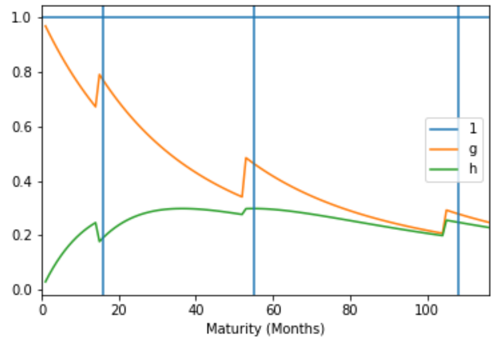

Figure 3: Factor Loadings Nelson and Siegel Strongly Segmented

0.2.3 Extension

Strongly Segmented Model

The strong segmentation factor loadings propose by [AAK+18] model as shown in Figure 3 are based on the Nelson and Siegel (1987) and Svensson (1994) models. However [AAK+18] allow different functional forms for the factor loadings to change within each segment. Despite possible discontinuities in the loading functions at nodes (border points), the smoothing restrictions guarantee that the yield curve remains continuous and smooth. In this model, each segment has its own dynamics and functional loadings, while smoothing constraints connect local to global dynamics across maturities, reinforcing the analogy with the preferred habitat theory. [AAK+18] propose the following form of the local loadings:

$$g_i(\tau) = \frac{(1-^{-\lambda\Lambda_i(\tau)})}{\lambda\Lambda_i\,(\tau)} \tag{6}$$

$$h_i(\tau) = \frac{(1-^{-\lambda\Lambda_i(\tau)})}{\lambda\Lambda_i\,(\tau)} - \lambda\Lambda_i\,(\tau) \tag{7}$$

where Λ_i are the functions that introduce the discontinuities in the loadings and the nodes, [AAK+18] adopt the following linear functional form:

$$\Lambda_i\,(\tau) = \quad \tau \quad - \quad \tau_i\,(1 \quad - \quad p)\,, \quad \tau \quad \in \quad A_i, \tau_i \quad \in \quad \phi \quad and \quad p \quad \in \quad [0,1] \tag{8}$$

The parameter p controls the degree of loading segmentation. To go from the strong segmentation model to the weak version we simply need to set $\Lambda_i\,(\tau) = \quad \tau$ for all i.

0.3 Data and Implementation

The present work was done with daily data Colombian yields between the years 2013 and 2016. The data was obtained thanks to the Colombian price provider PRECIA for yields with maturities of 1, 3 and 6 months and 1, 3, 5, 7, 9, 11, 13 and 15 years. The Colombian data compared to other countries does not have a high volume of transactions per year, therefore, to have enough data for the execution of the exercise, a pairing was performed with other PRECIA titles to obtained historical data.

The pairing work for the current exercise was as follows: titles were searched and paired against others with similar maturities of the following year and then year after year from 2013 to 2016. In other words, for 2013 a title was found with expiration to one year, for 2014 no title of these characteristics was found, but one with a maturity of 10 months, these two titles were paired, resulting in the historical for these two years. For the following years, starting in 2015, the match was made with maturities of one year, respectively, obtaining the historical base of this type of maturity (1 year maturity). This exercise was performed for all available maturities.

For empty periods (periods with out transactions), the information is completed by performing a spline interpolation. Given the nature of the Colombian market (small market), the short-term maturities(1, 3 and 6 months) are issued one time once a year, which is why Figure 5 has these strange behaviors, since a greater interpolation had to be done in the data for these cases.

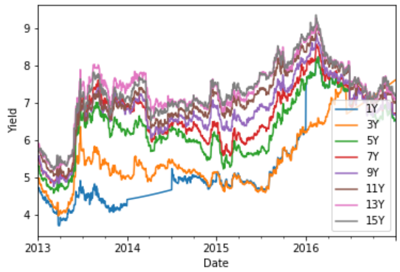

Figure 4: Representation of data provided by Precia for maturities of 1, 3, 5, 7, 9, 11, 13 and 15 years

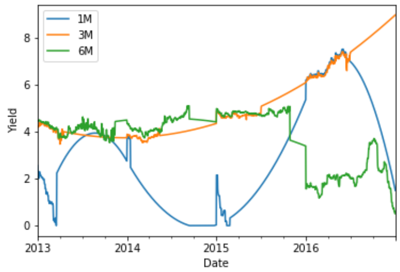

Figure 5: Representation of data provided by Precia for maturities of 1, 3, 6 months

0.4 Empirical Application

It is necessary to make an empirical application since the segmented model of Nelson and Siegel uses the long-term yields to find the adjusted values of the short and medium term yields and we believe that this methodology does not guarantee that the adjusted values for these yields be the right ones.

From [AAK+18] the restricted model of the paper and for three factor NS segmented model made of 3 segmented we have the following expression, for estimating the long term (L) factors, $\theta^L_{2,t}$ for each time period t

$$y_{t,L}(\tilde{\tau}) = [w_2(\tilde{\tau}) - w_1(\tilde{\tau})R_1^{-1}R_2]\theta^L_{2,t} + \epsilon_t(\tilde{\tau}) \tag{9}$$

$$y_{t,L}(\tilde{\tau}) = Z(\tilde{\tau})\theta^L_{2,t} + \epsilon_t(\tilde{\tau}), \;\; dim(Z(\tilde{\tau})) = m_x(k+1) = (3x3) \text{ invertible.}$$

$$w_1(\tilde{\tau}) = \begin{bmatrix} X_3(\tau_3) & 0_{1x3} \\ 0_{1x3} & X_5(\tau_5) \\ 0_{1x3} & X_6(\tau_6) \end{bmatrix}, \;\; dim(w_1(\tau)) = (3x6)$$

$$w_2(\tilde{\tau}) = \begin{bmatrix} X_9(\tau_9) \\ X_{10}(\tau_{10}) \\ X_{11}(\tau_{11}) \end{bmatrix}, \;\; dim(w_2(\tau)) = (3x3)$$

where $y_{t,L}(\tau) = (y_{t,L}(\tau_9), y_{t,L}(\tau_{10}), y_{t,L}(\tau_{11}))$ is the vector of long term yields. Note that for the estimation it is required that $Z(\tilde{\tau})$ is invertible. We use OLS to estimate the long term (L) factors $\theta^L_{2,t}$. We can obtain the short and medium term (S/M) factors from the restriction,

$$\theta^{S/M}_{1,t} = -R_1^{-1} R_2 \theta^L_{2,t}, \;\; dim(\theta_1) = [(6x6)_x (6x3)_x (3x1)] = (6x1)$$

Note that in the restricted version of the model we are not using the information of the other yields (for example in this case $y_t(\tau) = (y_t(\tau_1), \ldots, y_t(\tau_8)))$ to fit the parameters, therefore there is no reason to get proper fitted values for these yield, even though we can recover them.

For the fitted yields, you can follow a similar procedure for the rest of the long term yields $(y_{t,L}(\tau_7), y_{t,L}(\tau_8))$, you just need to adjust the variables in $Z(\tilde{\tau})$. From [AAK+18] it is mentioned that the equation 9 is valid for any pair of maturities and yields, that is redefining w_2 and w_1,

$$\begin{aligned} y_{t,L}(\tilde{\tau}) &= [w_2(\tilde{\tau}) - w_1(\tilde{\tau}) R_1^{-1} R_2] \theta^L_{2,t} + \epsilon_t(\tilde{\tau}) \\ y_{t,L}(\tilde{\tau}) &= Z(\tilde{\tau}) \theta^L_{2,t} + \epsilon_t(\tilde{\tau}), \;\; dim(Z(\tilde{\tau})) = m_x(k+1) = (2x3). \\ w_1(\tilde{\tau}) &= \begin{bmatrix} X_3(\tau_3) & 0_{1x3} \\ 0_{1x3} & X_6(\tau_6) \end{bmatrix}, \;\; dim(w_1(\tilde{\tau})) = (2x6) \\ w_2(\tilde{\tau}) &= \begin{bmatrix} X_7(\tau_7) \\ X_8(\tau_8) \end{bmatrix}, \;\; dim(w_2(\tilde{\tau})) = (2x3) \end{aligned}$$

the fitted short and medium term yield we can derive a similar setup $(y_{t,S}(\tau_1), \ldots, y_{t,M}(\tau_6))$ that is redefining w_2 and w_1, and split $\theta^{S/M}_{1,t}$ into the short and medium term yields, $\theta^S_{1,t}$, $\theta^M_{1,t}$ to get three dimensional parameter spaces.

For the medium term (M) we have w_2 and w_1

$$\begin{aligned} y_{t,M}(\tilde{\tau}) &= [w_2(\tilde{\tau}) - w_1(\tilde{\tau}) R_1^{-1} R_2] \theta^M_{1,t} + \epsilon_t(\tilde{\tau}) \\ y_{t,M}(\tilde{\tau}) &= Z(\tilde{\tau}) \theta^M_{1,t} + \epsilon_t(\tilde{\tau}), \;\; dim(Z(\tilde{\tau})) = m_x(k+1) = (3x3). \\ w_1(\tilde{\tau}) &= \begin{bmatrix} X_3(\tau_3) & 0_{1x3} \\ 0_{1x3} & X_5(\tau_5) \\ 0_{1x3} & X_6(\tau_6) \end{bmatrix}, \;\; dim(w_1(\tilde{\tau})) = (3x6) \\ w_2(\tilde{\tau}) &= \begin{bmatrix} X_7(\tau_7) \\ X_8(\tau_8) \\ X_9(\tau_9) \end{bmatrix}, \;\; dim(w_2(\tilde{\tau})) = (3x3) \end{aligned}$$

For the short term (S) we have w_2 and w_1

$$\begin{aligned} y_{t,S}(\tilde{\tau}) &= [w_2(\tilde{\tau}) - w_1(\tilde{\tau}) R_1^{-1} R_2] \theta^S_{1,t} + \epsilon_t(\tilde{\tau}) \\ y_{t,S}(\tilde{\tau}) &= Z(\tilde{\tau}) \theta^S_{1,t} + \epsilon_t(\tilde{\tau}), \;\; dim(Z(\tilde{\tau})) = m_x(k+1) = (3x3). \\ w_1(\tilde{\tau}) &= \begin{bmatrix} X_1(\tau_1) & 0_{1x3} \\ X_2(\tau_2) & 0_{1x3} \\ 0_{1x3} & X_4(\tau_4) \end{bmatrix}, \;\; dim(w_1(\tilde{\tau})) = (3x6) \\ w_2(\tilde{\tau}) &= \begin{bmatrix} X_7(\tau_7) \\ X_8(\tau_8) \\ X_9(\tau_9) \end{bmatrix}, \;\; dim(w_2(\tilde{\tau})) = (3x3) \end{aligned}$$

Note that the difficulty with this methodology is that we have arbitrary ways to define w_2 and w_1 and therefore the results may be sensitive to the way we define these two matrices. Also in this particular case since $Z(\tilde{\tau})$ is invertible you could use it to estimate $\theta^S_{1,t}$, $\theta^M_{1,t}$, but then these would be free parameters and may violate the restriction although they will provide a better fit for the yields.
[1]

[1] For practical purposes we will define the Nelson and Siegel Three Factor models as NS3

0.5 In Sample Performance

To measure the performance of the segmented models and compare them with the classic model of Nelson and Siegel within the sample for the years between 2013 and 2016 (i.e. 1461 days) we use the RMSE(Root Mean Squared Error) in all available dates for all the available maturities.

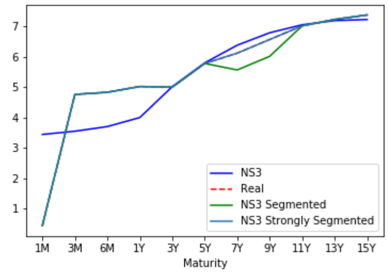

Figure 6: Comparison between real yield and fitted yield of NS3, NS3 Segmented and NS3 Strongly Segmented models in period 2015-03-12

Table 1 compares the RMSE(Root Mean Square Error) of NS3 ,NS3 Segmented and NS3 strongly segmented models. Although the NS3 Segmented and NS3 strongly models errors are almost equal to 0, for the 7 and 9 year maturities the RMSE for the NS3 Segmented model are greater than 0 (This may be due to the way in which matrices w_2 and w_1 were constructed as explained in the section Empirical Application), therefore the inside sample performance of NS3 strongly Segmented is better than the NS3 Segmented model.

The segmented models in comparison with the classic show a better performance in all available maturities. The results inside the sample of 1, 3 and 6 month maturities of NS3 show a higher than normal RMSE, possibly due to the way in which the data were constructed for these periods(See Figure 5 and Section Data and Implementation).

Maturity	NS3	NS3 Segmented	NS3 Strongly Segmented
1M	151.773420	2.449974e-11	6.879434e-13
3M	97.952827	2.412067e-11	7.182330e-13
6M	145.172881	2.173400e-12	6.270366e-13
1Y	78.856673	2.379020e-11	6.858002e-13
3Y	25.724497	1.964510e-12	5.746715e-13
5Y	13.969507	1.892531e-12	5.766169e-13
7Y	15.462041	23.35683	4.839083e-14
9Y	22.551928	23.35683	4.699260e-14
11Y	16.919568	3.323526e-13	1.851132e-13
13Y	9.448176	2.945397e-13	1.755687e-13
15Y	12.553559	3.369728e-13	1.713741e-13

Table 1: RMSE on Insample Data between 2013-01-01 and 2016-12-31

0.6 Forecasting Out of Sample

Prediction is one of the main objectives of multivariate analysis of time series. To predict the yield curve it is necessary to first calculate the intercept, the slope and the curvature variables by using ordinary least squares. We are going to use a VAR(1) model ($B_{t+1} = \Phi B_t + \varepsilon_t$) to predict the unobserved values. The loos function to evaluate the model will be the root mean squared error

$$RMSE = \sqrt{\left(\hat{y}_{t+h}\left(\tilde{\tau}\right) - y_{t+h}\left(\tilde{\tau}\right)\right)^2}$$

where h is the length of steps ahead that have been taking for forecasting, h will be 1, 6 and 20 days for all available maturities. When working with time series forecasting we often have to choose between a few potential models and the method that we are going to use for the choice of these is going to be expanding window . For a better understanding of the expanding window method we will explain it in the following way: Suppose you have, for example, 100 observations of a time series. First you estimate the model with the first 90 observations to forecast the observation 91. Then you include the observation 91 in the estimation sample and estimate the model again to forecast the observation 92. The process is repeated until you have a forecast for all 10 out of sample observations. Having already explained the methodology, for the present work we will predict the last 64 days of our data; that is to predict the days between 2016-10-28 and 2016-12-31.

Maturity	NS3	NS3 Segmented	NS3 Strongly Segmented
1M	204.1277	3.4885	3.4885
3M	363.5557	0.1607	0.1607
6M	338.6277	4.7718	4.7718
1Y	129.5160	2.9995	2.9995
3Y	118.0157	0.2892	0.2892
5Y	12.8207	2.9654	2.9654
7Y	27.4832	9.7528	3.5973
9Y	46.2696	9.5275	3.7047
11Y	18.9471	3.4783	3.4783
13Y	6.0622	3.9374	3.9374
15Y	31.3396	3.6211	3.6211

Table 2: RMSE on OutSample Data 1 Day Ahead

Maturity	NS3	NS3 Segmented	NS3 Strongly Segmented
1M	217.8136	21.4938	21.4938
3M	362.5706	0.9725	0.9725
6M	346.0230	25.9774	25.9774
1Y	128.2185	10.7218	10.7218
3Y	118.8673	1.7191	1.7191
5Y	13.6771	10.9060	10.9060
7Y	28.2989	12.5120	12.2468
9Y	44.9838	12.8340	12.6727
11Y	20.2798	10.2613	10.2613
13Y	11.9426	11.9886	11.9886
15Y	30.2459	9.8814	9.8814

Table 3: RMSE on OutSample Data 6 Days Ahead

As shown in tables 2, 3 and 4 the results of the NS3 Segmented and NS3 strongly segmented models RMSE are almost the same, if not identical, both show a better performance than the NS3 model in all available maturities and in all forecasting windows, except in a 20 days ahead window the NS3 model has a better performance than the Segmented model in 5 and 13 year maturities. The results out of the sample of the 1, 3 and 6 month maturities of NS3 show a higher than normal RMSE, possibly due to the way in which the data were constructed for these periods(See Figure 5 and Section Data and Implementation).

Maturity	NS3	NS3 Segmented	NS3 Strongly Segmented
1M	234.6337	44.3104	44.3104
3M	361.2137	1.9699	1.9699
6M	355.6739	49.6339	49.6339
1Y	126.8352	18.9459	18.9459
3Y	119.4593	3.4911	3.4911
5Y	16.9207	19.0228	19.0228
7Y	31.2709	20.6599	20.5827
9Y	43.9985	21.2531	21.2240
11Y	23.1048	14.9574	14.9574
13Y	16.3404	16.9419	16.9419
15Y	29.5400	12.9844	12.9844

Table 4: RMSE on OutSample Data 20 Days Ahead

0.7 Conclusion

In this article we work with daily data of Colombian yields between the years 2013 and 2016, we compare the performance in sample and out the sample of the Nelson and Siegel Segmented model and the Nelson and Siegel Classic model. Inside the sample the Segmented and Strongly Segmented model had an outstanding performance predicting the fitted yield almost identical to the real yield in comparison with the Classic model, and out the sample the Segmented and the Strongly Segmented models overcome the classic model in almost all the maturities and forecasting windows.

The poor performance for cases less than 1 year (1, 3 and 6 months) in the Nelson and Siegel Classic model is due to the pairing of the data.

If the segmented models are analyzed against the classic model, locally adjusting the segments provides substantial improvements within and outside the sample.

We find that the model of Nelson and Siegel classic has a good behavior within the sample except when the data contains strange behaviors (Figure 5).

For the next stage of this exercise we will seek to work with a more complete data, that is, one that does not need to be paired, which will allow us to better see the performance of the model for the short term and confirm the results obtained here.

Bibliography

[AAK+18] C. Almeida, K. Ardison, D. Kubudi, A. Simonsen, and J. Vicente. Forecasting bond yields with segmented term structure models. *Journal of Financial Econometric*, 16(1):1–33, 2018.

[DK13] S. D'Amico and T.B. King. Flow and stock effects of large-scale treasury purchases: evidence on the importance of local supply. *Journal of Financial Economics*, 108:425–448, 2013.

[ERS17] R. Engle, G. Roussellet, and E. Siriwardane. Scenario generation for long run interest rate risk assessment. *Journal of Econometrics*, 201:333–347, 2017.

[MS66] F. Modigliani and R. Sutch. Innovations in interest-rate policy. *American Economic Review*, 56:178–197, 1966.

[VV09] D. Vayanos and J.L. Vila. A preferred-habitat model for the term structure of interest rates. 2009.