

LTH - FACULTY OF ENGINEERING

MASTER THESIS

FINANCIAL MODELLING

---

# Interest rate modelling

---

*Author:*

Khalil EL ADI

*Supervisor:*

Erik LINDSTRÖM

19 December 2019



# Contents

Abstract . . . . .	ii
Preface . . . . .	iv
<b>1 Introduction</b>	<b>1</b>
<b>2 Theory</b>	<b>3</b>
2.1 Short Rate Models . . . . .	3
2.1.1 The Vasicek Model . . . . .	4
2.1.2 The Cox–Ingersoll–Ross Model . . . . .	5
2.2 Anticipation Theory . . . . .	6
2.2.1 Stationary model . . . . .	7
2.2.2 Non Stationary Model . . . . .	8
2.3 Time Series-Regression Analysis . . . . .	9
2.3.1 Engle–Granger test . . . . .	10
2.3.2 Johansen test . . . . .	11
2.4 Term structure analysis: Nelson-Siegel Approach . . . . .	11
<b>3 Data</b>	<b>13</b>

<b>4</b>	<b>Calibration Method</b>	<b>17</b>
<b>5</b>	<b>Results and Analysis</b>	<b>19</b>
5.1	Anticipation Theory . . . . .	19
5.1.1	Stationary models . . . . .	19
5.1.2	Non-Stationary models . . . . .	21
5.2	Time Series Analysis . . . . .	23
5.2.1	United States . . . . .	24
5.2.2	Germany . . . . .	26
5.3	Term Structure Modeling . . . . .	29
5.3.1	US . . . . .	29
5.3.2	Germany . . . . .	31
5.3.3	Modeling of parameters . . . . .	32
<b>6</b>	<b>Conclusion</b>	<b>39</b>
	Bibliography . . . . .	41
	Annex - Computer Code . . . . .	43

-

## Abstract

Many models have been developed throughout the years to describe the evolution of short term rates. One of the famous models is the Vasicek model. It was first introduced in 1977 and describes interest rates as a mean reversion process which is a specific characteristic that sets it apart from other financial assets. This model has the ability to let interest rates to be negative which was perceived as a weakness of the model before the 2008 financial crisis. However, it has become one of its strengths in the current negative rates environment set by central banks in an effort to stimulate the economy. The Cox–Ingersoll–Ross model was introduced in 1985 as a way to limit interest rates from being negative. Many complex and dynamics models were developed in the years that followed to describe the movement of short term rates. However, fewer models were interested in the long term rates. This is due to the impact macroeconomic variables have on these rates.

In this thesis, we detail three approaches to the modeling of long term interest rates in the United States and Germany. The first one is based on anticipation of future short term rates, the second one looks at the effect on macroeconomic variables on these long term rates and the last one describes the evolution of the yield curve.

-

## Acknowledgement

First, I would like to thank Amundi Asset Management for giving me the opportunity to work on my master thesis and giving me all the tools and data needed for this project. I would like to thank my daily supervisor, Valentine Ainouz, and my colleague Delphine Georges, in particular, for their continuous support and encouragements that was very helpful. My sincere thanks also goes to my thesis professor Erik Lindström for his help, insight and feedback. Finally, I would like to thank both my family and friends for all the support in helping me achieve my dreams and become the person that I am today.

# Chapter 1

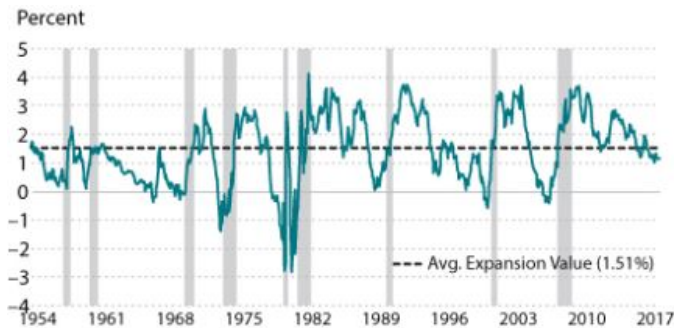
## Introduction

The main reason behind buying Bonds is to provide a predictable income stream for investors. Financial institutions, such as banks and asset management firms, use bonds (and sovereign bonds in particular), which are contracts that guarantee a payment and an interest rate at maturity, as a hedge and a safe asset in order to protect part of their investments or to cancel the discounting effect of inflation. However, it can be used for various other reasons such as speculation on the future trajectory of interest rates by traders in order to make a profit.

There are different types of bond. The safest ones are government bonds. These are debt instruments used by most administrations in order to raise money for various projects and close their budget deficit. Another type of bonds is corporate bonds. They are similar to government bonds and are used by companies as a way to raise capital instead of issuing equity. Bonds are subject to ratings by rating companies such as Moody's and S&P and could fit in two categories - Investment Grade or High Yield.

Modelling the fluctuations of interest rates has always been an important study subject to many research academics and financial institutions. The volatility of interest rates has an impact on the economy in general, and on companies' finances in particular, as many firms use bonds as a means to raise capital which makes them vulnerable to the changes in interest rates. For example, the US 10-year interest rate is considered to be a 'safe haven' for many investors. In addition, the spread between the 10-year and 2-year US bond yields is considered to be a good predictor of future recessions (*Fig.1*).





*Figure 1.1: US 2y10y spread (blue) and Recessions (Gray bars).  
Source: Recession Signals: The Yield Curve vs. Unemployment Rate Troughs, Economic Research Federal Reserve Bank Of St. Louis*

In this paper, we talk about about three different approaches to model and analyse the long term interest rates:

- Anticipation theory:

This approach is based on modelling the long term rates as an average of short term rates which are modelled using either a stationary model such as the Vasicek model or non stationary model.

- Time Series/Regression Analysis:

This approach is based on using macroeconomic variables that impact the long term rates.

- Term Structure Modelling:

This approach is based on using the level, slope and curvature of the yield curve in order to model the long term interest rates.

# Chapter 2

## Theory

### 2.1 Short Rate Models

Short rates models are mathematical processes used to describe the changes in the short rate  $r_t$ , which is the instantaneous spot rate of borrowing money for an infinitesimal short period of time  $dt$ . Although specifying the short rate impacts the yield curve, it mainly influences the short part of the curve while the long part changes are mostly impacted by macroeconomic events and investor appetite.

The price at time  $t$  of a zero-coupon bond that matures at time  $T$  - under no-arbitrage arguments and taking the evolution of  $r_t$  as a stochastic process under a risk-neutral measure  $Q$  - is given by:

$$P(t, T) = E^Q \left[ \exp \left( - \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right], \quad (2.1)$$

where  $P(T, T) = 1$ .

The yield curve is a line plot, at a specific date, of the interest rates of zero coupon bond with different maturities. From the previous equation, specifying a short rate model would give us the future prices of the zero coupon bonds.

We can get the forward rates based on bond prices using the formula below:

$$f(t, T) = -\frac{\partial}{\partial T} \ln(P(t, T)). \quad (2.2)$$

The short rate can then be derived as:

$$r_t = f(t, t) \quad (2.3)$$

There are many models used to analyze and evaluate the evolution of the short rate. In this thesis, we will look at both the Vasicek and CIR models. These models assume that the interest rates are Markovian processes.

The short rate models follow the following stochastic differential equation:

$$dr(t) = \mu(r, t)dt + \sigma(r, t)dW(t) \quad (2.4)$$

### 2.1.1 The Vasicek Model

The Vasicek model is a one-factor short rate model. It can be used for the pricing of interest rate derivatives. This model captures the mean reversion character of the short rates.

The stochastic process equation for this model is given by:

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \quad (2.5)$$

where  $\kappa$ ,  $\mu$  and  $\sigma$  are constants and  $W(t)$  is a Wiener process.

The  $\kappa(\mu - r_t)$  factor is the drift term and represent the expected change in interest rates.  $\mu$  is the long term equilibrium interest rate and  $\kappa$  is the speed at which the rate goes back to its long term equilibrium.

Using Ito's formula on  $X_t = r_t e^{\kappa t}$  :

$$dX_t = \mu \kappa e^{\kappa t} dt + \sigma e^{\kappa t} dW_t \quad (2.6)$$

Using the last two equations and a simple discretization, we get the following expression for  $r_t$ .

$$r_{t+\Delta t} = e^{-\kappa t} r_t + \mu(1 - e^{-\kappa \Delta t}) + \sigma e^{-\kappa(t+\Delta t)} \int_t^{t+\Delta t} e^{-\kappa t} dW_t \quad (2.7)$$

The mean and variance of  $r_t$  are:

$$E(r_{t+\Delta t} | r_t) = e^{-\kappa \Delta t} r_t + \mu(1 - e^{-\kappa \Delta t}) \quad (2.8)$$

$$V(r_{t+\Delta t} | r_t) = \left( \frac{\sigma^2}{2\kappa} \right) (1 - e^{-2\kappa \Delta t}) \quad (2.9)$$

When  $\kappa$  goes to zero, the expected value goes to  $r_t$  and the variance to zero. And when,  $\kappa$  goes to infinity, the expected value goes to  $\mu$  and the variance to zero  $\frac{\sigma^2}{2\kappa}$

### 2.1.2 The Cox–Ingersoll–Ross Model

Now, we consider the Cox–Ingersoll–Ross (CIR) model, the stochastic differential equation for this model is:

$$dr(t) = \kappa(\mu - r_t)dt + \sigma\sqrt{r_t}dW(t) \quad (2.10)$$

where  $\kappa$ ,  $\mu$  and  $\sigma$  are constants and  $W_t$  is a Wiener process.

One of the drawbacks of using the Vasicek model is that it allows the interest rate to be negative. The CIR model is an extension of the Vasicek model and solves the issue by adding the term  $\sqrt{r_t}$  to the standard deviation factor. The drift factor is the same as the Vasicek model.

Using Ito's formula on  $X_t = r_t e^{\kappa t}$  :

$$dX_t = \mu \kappa e^{\kappa t} dt + \sigma e^{\kappa t} \sqrt{X_t} dW_t \quad (2.11)$$

Using a simple discretization and the last two equations, we get a normally distributed  $r_t$ .

$$r_{t+\Delta t} = e^{-\kappa\Delta t}r_t + \mu(1 - e^{-\kappa\Delta t}) + \sigma e^{-\kappa(t+\Delta t)} \int_t^{t+\Delta t} e^{-\kappa t} \sqrt{r_t} dW_t \quad (2.12)$$

This means that the mean and variance of  $r_t$  are as follows:

$$E(r_{t+\Delta t} | r_t) = e^{-\kappa\Delta t}r_t + \mu(1 - e^{-\kappa\Delta t}) \quad (2.13)$$

$$V(r_{t+\Delta t} | r_t) = e^{-\kappa\Delta t}r_t \left( \frac{\sigma^2}{\kappa} \right) (1 - e^{-2\kappa\Delta t}) + \mu \left( \frac{\sigma^2}{2\kappa} \right) (1 - e^{-2\kappa\Delta t})^2 \quad (2.14)$$

When  $\kappa$  goes to zero, the expected value goes to  $r_t$  and the variance to zero. And when,  $\kappa$  goes to infinity, the expected value goes to  $\mu$  and the variance to zero  $\mu \frac{\sigma^2}{2\kappa}$

## 2.2 Anticipation Theory

The Anticipation Theory defines the long term interest rate as the average of future expected short term rates with an added risk premium that we assume to be constant. The risk premium is the compensation for investors to tolerate the extra risk holding an asset.

This theory has two main principles: no arbitrage and rational anticipation. We take  $R_t$ , the interest rate of a long term zero coupon bond at time  $t$ , as the average of expectations of future yields at times  $t, t + 1, \dots, t + n - 1$  plus a risk premium:

$$R_t = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i} + \lambda_T \quad (2.15)$$

The risk premium  $\lambda_T$  could change with different maturities  $T$ . However, we assume that it is a constant throughout time.

For coupon paying bonds, the long term interest rates are given by:

$$R_t = \frac{1 - \beta}{1 - \beta^n} \sum_{i=0}^{n-1} E_t \beta^i r_{t+i} + \lambda \quad (2.16)$$

Where  $\beta = \frac{1}{1+\bar{R}}$  and  $\bar{R}$  is the mean of the long term rates.

The main difficulty was choosing between either a stationary or non stationary approach for the short term interest rates models. Most short term interest rate models that have been developed are stationary models (Vasicek 1977, CIR 1985) even if most empirical tests fail to reject non-stationary models.

## 2.2.1 Stationary model

We use the stationary approach with the Vasicek model.

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \quad (2.17)$$

Using the discrete method, we have:

$$r_t = \kappa\mu dt + (1 - \kappa dt)r_{t-1} + \sigma\sqrt{t}e_t \quad (2.18)$$

We denote:  $a = (1 - \kappa dt)$ ,  $g = \kappa\mu dt$  and  $e_t$  is an i.i.d, standard normally distributed process. Thus, we have:

$$r_t = g + ar_{t-1} + \sigma\sqrt{t}e_t \quad (2.19)$$

Using this model, the expectation is given by:

$$E(r_{t+i} | r_{t-1}) = a^{i+1}r_{t-1} + \frac{(1 - a^{i+1})}{(1 - a)}g \quad (2.20)$$

We use this equation to get the formula for the long term rates using the anticipation theory.

$$\begin{aligned}
R_{t/t-1} &= E(R_t | r_{t-1}) = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i} + \lambda \\
&= \frac{g}{1-a} + \frac{a(1-a^n)}{n(1-a)} \left( r_{t-1} - \frac{g}{1-a} \right) + \lambda
\end{aligned} \tag{2.21}$$

Once the short term model has been estimated, we can use the previous equation to estimate the long term rate at time  $t+k$  where  $k \geq 0$  using the following formula:

$$\begin{aligned}
R_{t+k/t-1} &= E(R_{t+k} | r_{t-1}) \\
&= \frac{g}{1-a} + \frac{a^{k+1}(1-a^n)}{n(1-a)} \left( r_{t-1} - \frac{g}{1-a} \right) + \lambda
\end{aligned} \tag{2.22}$$

### 2.2.2 Non Stationary Model

We start by taking a process with a drift  $\alpha$  and a deterministic trend where  $\mu$ :

$$r_t = \alpha + \mu t + ar_{t-1} + \sigma e_t \tag{2.23}$$

We use the first differences approach:

$$dr_t = \mu + adr_{t-1} + \sigma e_t \tag{2.24}$$

Where  $e_t$  is *i.i.d.*, standard normally distributed process.

The expectation is given by:

$$\begin{aligned}
E(r_{t+i} | r_{t-1}) &= r_{t-1} + \sum_{k=0}^i E(\Delta r_{t+k}) \\
&= r_{t-1} + \sum_{k=0}^i (a^{k+1} \Delta r_{t-1} + \frac{(1-a^{k+1})}{(1-a)} \mu) \\
&= r_{t-1} + a \frac{(1-a^{i+1})}{(1-a)} (\Delta r_{t-1} - \frac{\mu}{(1-a)}) + (i+1) \frac{\mu}{(1-a)}
\end{aligned} \tag{2.25}$$

We use this equation to get the expression for the long term rates :

$$R_{t/t-1} = E(R_t | r_{t-1}) = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i} + \lambda \tag{2.26}$$

We take  $\mu = 0$  and get the following formula:

$$R_{t/t-1} = r_{t-1} + \frac{a}{(1-a)} \Delta r_{t-1} (1 - \frac{a(1-a^n)}{n(1-a)}) + \lambda \tag{2.27}$$

## 2.3 Time Series-Regression Analysis

Long term interest rates show a recent surprising dynamic in the United States and the euro-zone between 1986 and 2005. There had been a downward trend during the past 15 years which can mainly be explained by multiple fundamental factors such as the low macroeconomic volatility, low inflation anticipations and the change in strategy of communication of the various central banks regarding the monetary policies. However, these are not the only factors influencing the long term interest rates in the US and Europe. In the United States, the long term interest rates stayed constant around 4% even with the higher public deficit over the years. Same as in Europe where the interest rates stayed around 2 %.

This Time Series Analysis models the long term interest rates using certain macroeconomic factors that influence rates. Hence, the most important question is the choice of the different variables that could be used. The variables



believed to have an impact of the long rates can be classified into three categories:

- Fundamental variables: Inflation - Short rates
- Central bank variables: Fed Balance Sheet (for the US)
- Macroeconomic variables: GDP - Foreign Investments - Industrial Production
- Unemployment rate

The choice of the variables is based on the correlation with the long term rates but also on the independence between the different variables.

The long term interest rates, in the last decades, show a non-stationary dynamic that would dismiss using linear regression. However, we can use a cointegration analysis which is a method that estimates the long-run parameters in systems with non-stationary variables.

A set  $(X, Y, Z)$  are each integrated of order  $I(1)$  - this means that the first difference of a variable is a stationary series - if a linear combination  $aX + bY + cZ$  is stationary - of order  $I(0)$ . Formally, a set of variables  $(X_1, X_2, \dots, X_n)$ , integrated of at least order  $d$ , is said to be cointegrated if a linear combination of these variables is integrated of order less than  $d$ .

We study here two methods for testing cointegration: Engle–Granger test and Johansen test.

### 2.3.1 Engle–Granger test

We consider  $Y$  and  $X$  to be non-stationary variables where  $Y$  is the dependent variable and  $X$  is the explanatory variable.

The Engle–Granger method is a two-step method. The first step is to estimate  $\beta$  using an Ordinary Least Square in the following equation:

$$Y_t = \beta X_t + e_t \tag{2.28}$$

The next step is to test the stationarity of the residual  $e_t$ . It can be done using the Dickey–Fuller test or the Phillips–Perron test. Another regression is used

to test for a unit root for the process  $e_t$ .

$$e_t = \alpha e_{t-1} + u_t \quad (2.29)$$

A cointegration between the variables exists only if  $|\alpha| < 1$  .

### 2.3.2 Johansen test

The Johansen test is a method that can estimate more than one cointegration relationship between the variables. Let  $Y$  the  $(n \times 1)$  dependent variable and  $X$  is the  $(n \times p)$  matrix the  $p$  columns are explanatory variables. We consider:

$$X_t = AX_{t-1} + e_t \quad (2.30)$$

Where  $A$  is an  $(n \times n)$  matrix. We have thus:

$$\begin{aligned} \Delta X_t &= AX_{t-1} - X_{t-1} + e_t \\ &= \Gamma X_{t-1} + e_t \end{aligned} \quad (2.31)$$

Where  $\Gamma = A - I$  is an  $(n \times n)$  matrix.

The number of cointegrating relationships is given by the rank of the matrix  $\Gamma$ . If the rank of  $\Gamma$  is zero, then there is no cointegration. In case of a rank  $r=p$ , the cointegration is always true meaning that the variables must be stationary in the first place. Finally, the existence of stationary relationships is proven when the rank  $r < p$ .

## 2.4 Term structure analysis: Nelson-Siegel Approach

The Nelson and Siegel model is widely used to describe the interest rate yield curve. The Bank of International Settlements has reported that many Central

Banks use the Nelson-Siegel to model the interest rate term structure. The forward rate, based on this model, is given by the following formula:

$$f_t(\tau) = \beta_1(t) + \beta_2(t)e^{-\lambda\tau} + \beta_3(t)\lambda\tau e^{-\lambda\tau} \quad (2.32)$$

Where  $\tau$  is the maturity and  $\lambda$  is the decay factor.

Since the yield is given by the following fomula:

$$r_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u)du \quad (2.33)$$

We get:

$$r_t(\tau) = \beta_1(t) + \beta_2(t) \frac{1 - e^{-\lambda\tau}}{\lambda\tau} + \beta_3(t) \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \quad (2.34)$$

The Nelson and Siegen models the yield curve based on three main factors:

- The first factor is  $\beta_1(t)$ . It represents the level of the yield curve. It remains constant with regards to maturity. It is also viewed as the long term factor since the other factors tend to zero for very long maturities.
- The second factor is  $\beta_2(t)$  and represents the slope of the yield curve. It has an impact on the short term due to the fact that the term  $\frac{1-e^{-\lambda\tau}}{\lambda\tau}$ , next to  $\beta_2(t)$  in the formula, decays from 1 at time t=0 to 0 as time goes by.
- The third factor is  $\beta_3(t)$  and represents the curvature of the yield curve. It has an impact on the medium term as the term,  $(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau})$ , next to  $\beta_3(t)$ , increases to the maximum before decaying to zero in the long term. This factor captures what is called the 'butterfly effect' of the yield curve.

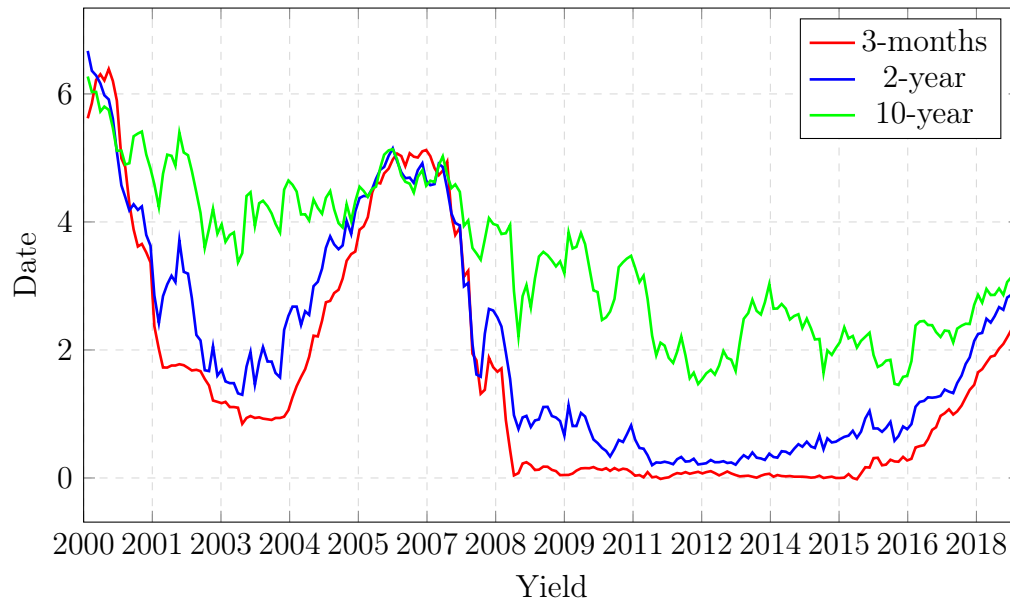
There is another factor that we need to talk about:  $\lambda$ . It is referred to as the decay factor. The estimation of this parameter is usually fixed first before the other factors  $(\beta_1(t), \beta_2(t), \beta_3(t))$ . By fixing  $\lambda$ , the estimation of the other factors becomes easier and could be done by an Ordinary Least Squares method.

# Chapter 3

## Data

This thesis presents multiple approaches to interest rate modelling and, as discussed before, we use data from both the United States and Europe. We decided to use and analyze the 3-months, 2-year and 10-year US treasury yields and German bond yields.

Figure 2 below presents the historical US treasury bond yields dating back to 2000:



*Figure 3.1: US 3-months, 2-year and 10-year treasury bond yields*

The 3-month and 2-year treasury yields are closely correlated to the Fed rates. Thus, we can clearly see that the low Fed rates since the financial crisis in 2008 have a clear impact on these short term yields. However, the 10-year yield is partially related to the Fed rates but is impacted by multiple other macroeconomic factors.

Figure 3 presents the historical German government bond yields dating back to 2000:

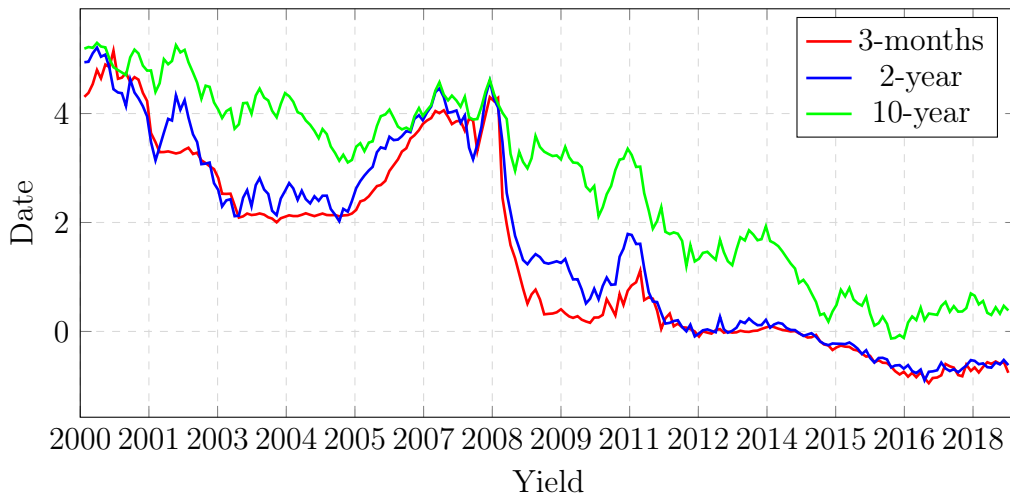


Figure 3.2: German 3-months, 2-year and 10-year Govt. bond yields

In The Nelson-Siegel model, we analyse the term structure of the US treasury bonds. Here is a representation of the US yield curve as of the 31/10/2018.

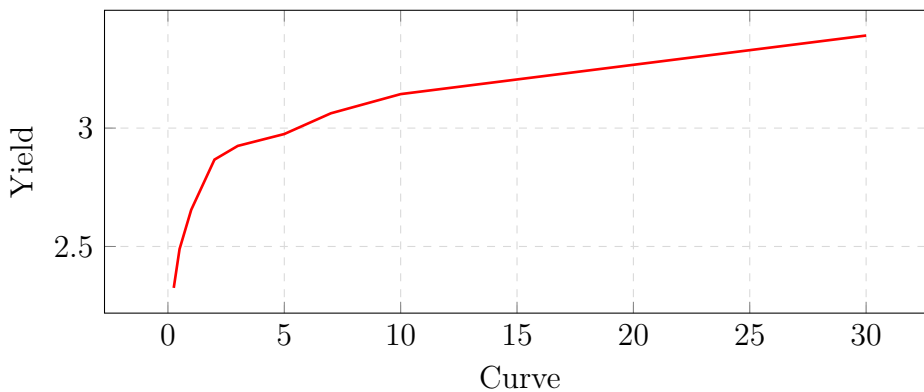
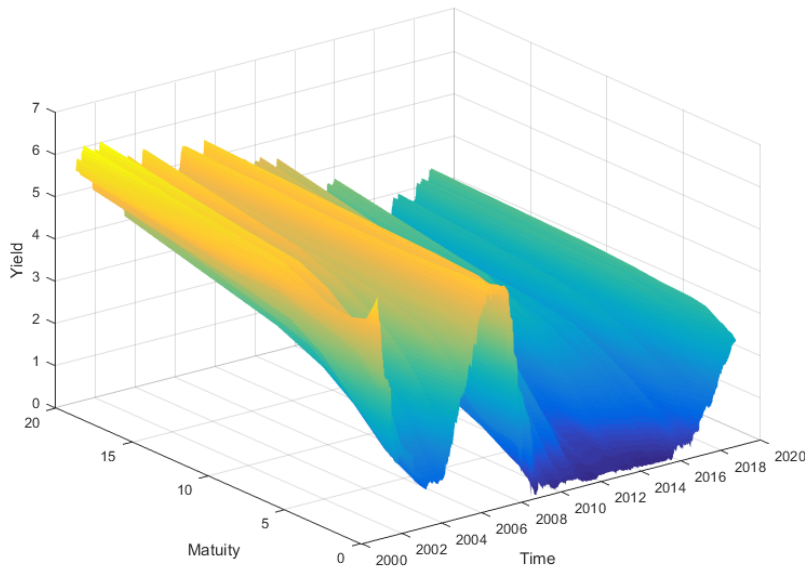


Figure 3.3: US yield curve as of the 31/10/2018

Since then, the short part of the yield curve has steepened more in 2019. The yield curve looked more like a 'butterfly' and was partially inverted in August 2019. As said before, this has been a good indicator of a recession but hasn't always been accurate.

Here is a representation of the interest rate surface which is a plot of the change of the yield curve over time.



*Figure 3.4: US yield surface*



# Chapter 4

## Calibration Method

Using the historical data from the previous chapter, we need to calibrate the different models that we relayed before.

There are two main methods to estimate the parameters of the different methods. The first one is the Ordinary Least Squares method. The principle of this method is to find the minimum of the sum of squared errors. We use this approach to find the  $\mathbb{P}$ -dynamics and the parameters  $\kappa$ ,  $\mu$  and  $\sigma$  for the Vasicek model. By using the Euler discretization, we get:

$$r_{t+\Delta t} = r_t + \kappa(\mu - r_t)\Delta t + \sigma(W_{t+\Delta t} - W_t) \quad (4.1)$$

The parameters in the drift factor,  $\kappa$  and  $\mu$  are estimated by solving the following equation:

$$(\hat{\kappa}, \hat{\mu}) = \arg \min_{(\hat{\kappa}, \hat{\mu})} \sum_{i=1}^{N-1} (r_{i+1} - r_i - \kappa(\mu - r_i)\Delta t)^2 \quad (4.2)$$

Once we have solved the equation above, we use the standard deviation of the residuals to estimate  $\sigma$ . The advantage of using this method is that there are only weak assumptions regarding the probability density of the errors.

The same approach is used to estimate the  $\mathbb{Q}$  implied dynamics in the Nelson-Siegel model. As said before, the parameter  $\lambda$  is fixed and the estimation



of  $(\beta_1(t), \beta_2(t), \beta_3(t))$  is made by an OLS estimator. The equation that we should minimize is thus:

$$\arg_{(\beta_1^{\hat{}}(t), \beta_2^{\hat{}}(t), \beta_3^{\hat{}}(t))} \min \sum_{i=1}^{N-1} \left( r_t(\tau_i) - \beta_1(t) - \beta_2(t) \frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - \beta_3(t) \left( \frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i} \right) \right)^2 \quad (4.3)$$

The second method to find parameters is the Maximum Likelihood (ML). This method was used in the time series approach. The principle of this method is to minimize the likelihood function (or log-likelihood) in order to estimate the parameters for the regressions in the Engle–Granger two-step method.

The likelihood function to be maximized is given by the following formula:

$$\hat{\beta} = \arg \max_{\beta} \sum_{i=1}^n \log(f(y_i | x_i; \beta)) \quad (4.4)$$

# Chapter 5

## Results and Analysis

### 5.1 Anticipation Theory

#### 5.1.1 Stationary models

We start our analysis by modelling the short term rates using the Vasicek model on the 3-month US treasury yields and German bonds:

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \quad (5.1)$$

Using the OLS method above, we find the following parameters for the Vasicek model:

Stationary Model		
	United States	Germany
$\kappa$	0.061	0.282
$\mu$	1.941	1.02
$\sigma$	0.745	1.18
$R^2$	0.01	0.01

Table 5.1: Short term model parameters for the 3-month interest rates

The  $R^2$ , which measures how close the data is to the fitted models, seems to be very low for both models. This could be related to the assumption of stationarity of the short term rates. We attempt to use a non-stationary model in the next section.

Using the equation (21), we have the following plots of the fitted US and German 10 year yields:

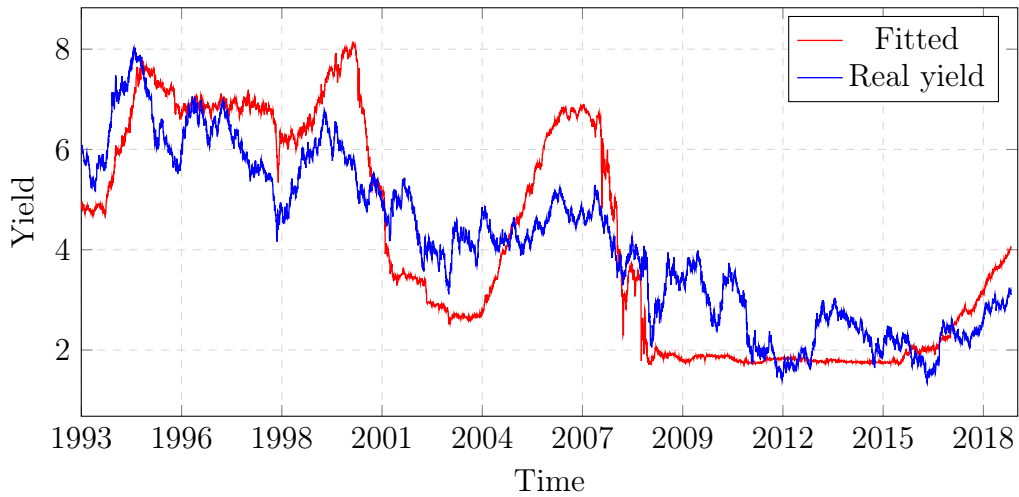


Figure 5.1: US fitted VS real yields

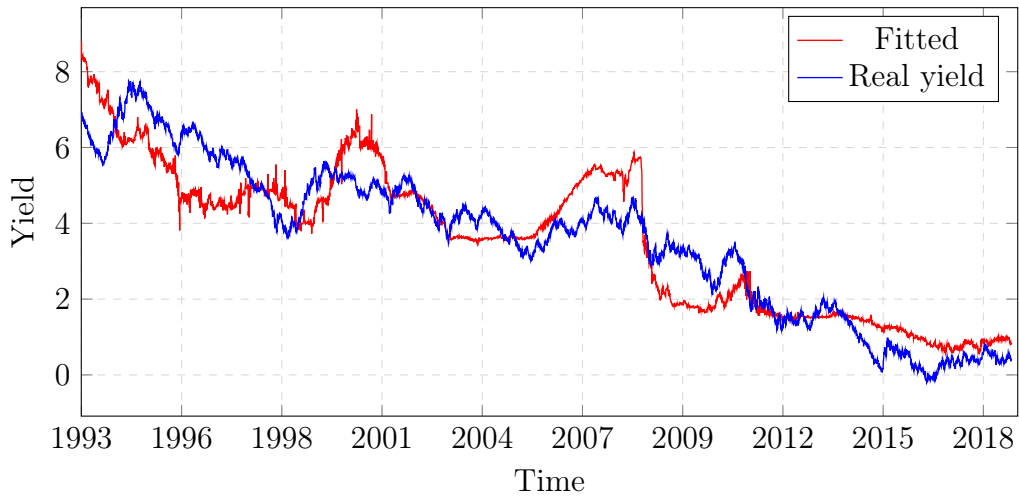


Figure 5.2: German Fitted VS real yields

### 5.1.2 Non-Stationary models

We use the equations (22) for the short rate process and get the following parameters.

Stationary Model		
	United States	Germany
$\mu$	-0.001	-0.002
$a$	0.09	-0.26
$R^2$	0.01	0.06

Table 5.2: Short term model parameters for the 3-month interest rates

The figure below shows the new fitted long term rates using the non-stationary short term rates and both equations (23),(24) and (25).

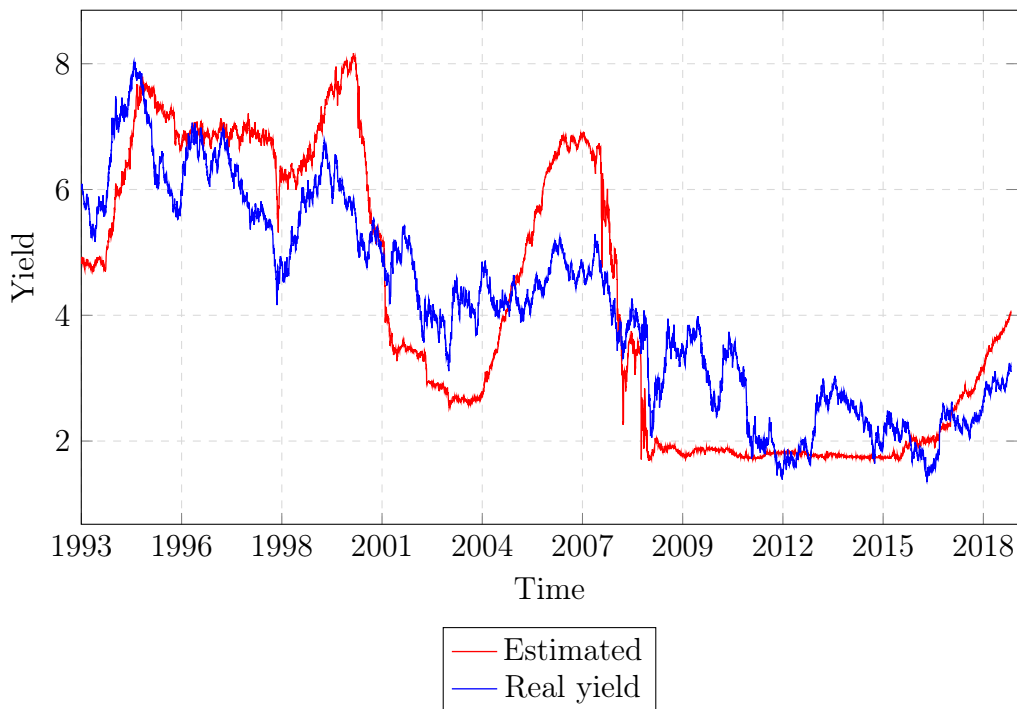


Figure 5.3: US estimated VS real yields

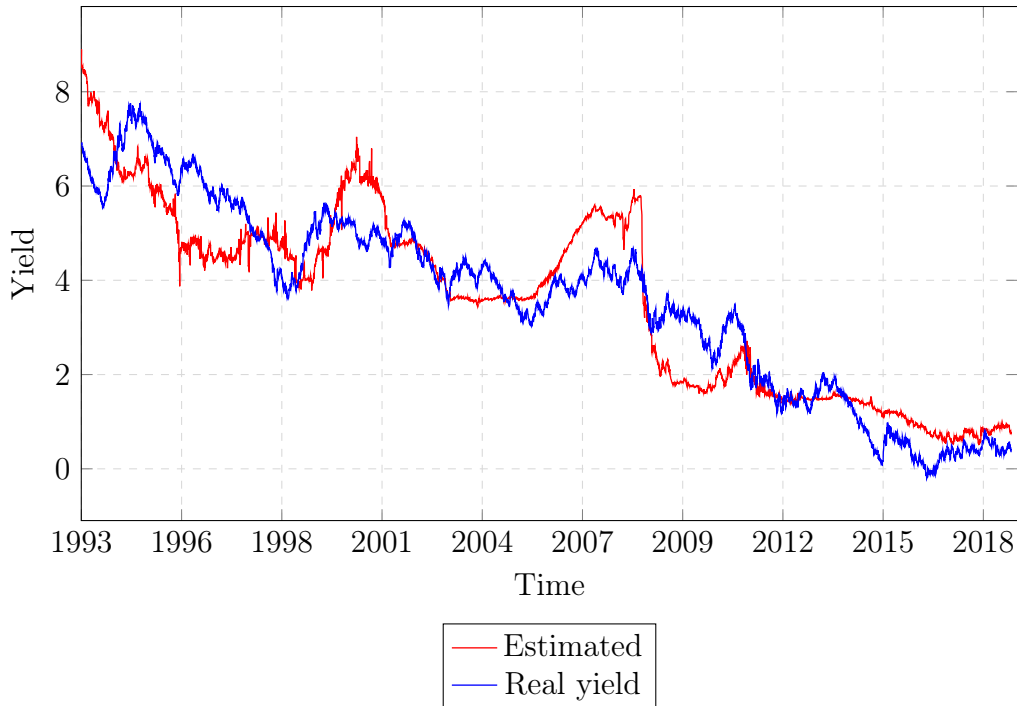


Figure 5.4: German estimated VS real yields

These plots are very similar the ones using the stationary models. This means that the models used for the modelling of the short term rates are not very accurate and thus do not give satisfactory results.

The estimated long term yields have similar dynamics of short term rates. Indeed, we can clearly see that the parameter ( $\alpha$ ) in both stationary and non stationary models is close to 0, which means that the modeling of the short term rates is not effective.

The two approaches give us clear over-estimation in certain years and under-estimation in some others. The results are a bit better for the German long term bonds than for the US bonds.

The Anticipation theory could be improved by taking more accurate short term models. The Vasicek model, for example, could be tweaked by allowing,  $\mu$ , which is the value that the short term rates revert back to, to be time dependent. This will allow rates to change over time taking into account the changes in Fed rates. We could also use two-factor models to ameliorate our short rate expectations. This theory could be further enhanced using a time

dependent risk premium, which could be modeled using macroeconomic or volatility factors.

The residual of these models are given in the figure below. As expected, the residuals are stationary but not white.

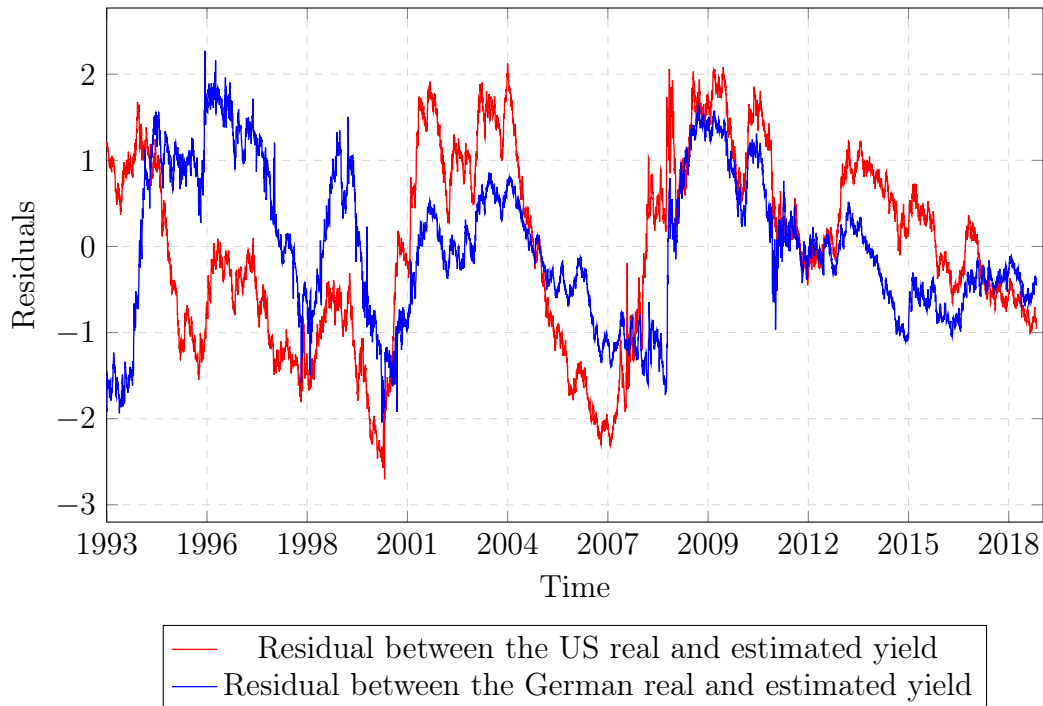


Figure 5.5: Residuals for the US and German sovereign bond yields fitted models

## 5.2 Time Series Analysis

In this approach, we use macroeconomic inputs to get a 'fair-value' of the long term interest rates. The fair value is the value where the bond yield should be based on the inputs that we have.

## 5.2.1 United States

We start by looking at the different inputs that we can use in our model. There are many variables that can be used here such as: Inflation, Short rates, FED balance sheet, GDP growth, Unemployment rate, Foreign Investments growth and Industrial Production (IP) index which is a measure of the production output of manufacturing, mining and utilities. *(All the data comes from Bloomberg)*

These variables can be divided into three categories: fundamental variables (inflation and Short term rates (3 month yields)), monetary variables (Fed rates and balance sheet) and other macroeconomic variables (GDP growth, Unemployment rate, Foreign Investments growth and Industrial Production index).

The first step is to look at the correlation between the long term yields and each of the variables. We also look at the correlation between the variables.

	yield10y	yield3m	Inf	Fed BS	F I	GDP	Unemploy
yield10y	1.0000	0.7951	0.4551	-0.8705	-0.7484	0.3153	-0.3447
yield3m	0.7951	1.0000	0.5285	-0.6759	-0.6085	0.5510	-0.6193
Inflation	0.4551	0.5285	1.0000	-0.4823	-0.2708	0.4107	-0.3308
Fed BS	-0.8705	-0.6759	-0.4823	1.0000	0.6551	-0.2317	0.1961
Foreign Inv	-0.7484	-0.6085	-0.2708	0.6551	1.0000	-0.3934	0.5175
GDP	0.3153	0.5510	0.4107	-0.2317	-0.3934	1.0000	-0.6782
Unemploy	-0.3447	-0.6193	-0.3308	0.1961	0.5175	-0.6782	1.0000

*Table 5.3: Correlation between the macroeconomic variables and the 3-month interest rates*

The first column represents the correlation between the 10 year US treasury yield and the rest of the variables. Based on the table above, we've decided to not take the Unemployment rate into account based on the possible dependence between this variable and the other ones.

The multiple unit root tests (Augmented Dickey–Fuller test and Phillips–Perron test) on the long term rates state that the rates not stationary and, thus, we use the Engle-Granger co-integration analysis.

We start by evaluating the integration orders of the variables. Using the

MATLAB function `jctest`, we conclude that the variables used here are all non-stationary and their first differences is stationary. This means that all the variables are integrated of order 1.

The `egcitest` function on MATLAB shows that there is a cointegration relationship in this case. We start by the first regression analysis:

```

yield10y ~ 1 + yield3m + Inflation_CPI_yoy + Foreign_Holdings + GDP_yoy + Fed_Balance_GDP
Estimated Coefficients:

```

	Estimate	SE	tStat	pValue
(Intercept)	5.9754	0.25428	23.499	1.4934e-61
yield3m	0.24863	0.027053	9.1905	3.2588e-17
Inflation_CPI_yoy	-0.00069619	0.028708	-0.024251	0.98068
Foreign_Holdings	-4.7232	0.69491	-6.7967	1.0181e-10
GDP_yoy	-0.074743	0.024064	-3.106	0.0021493
Fed_Balance_GDP	-0.073412	0.0062555	-11.736	6.006e-25

```

Number of observations: 223, Error degrees of freedom: 217
Root Mean Squared Error: 0.432
R-squared: 0.868, Adjusted R-Squared 0.865
F-statistic vs. constant model: 286, p-value = 2.28e-93

```

The plot of the estimated long term rates is given by:

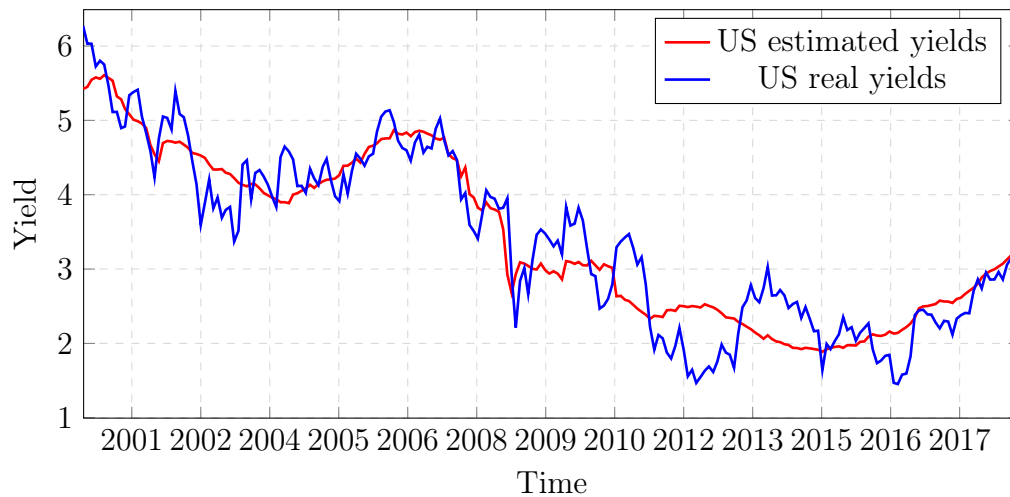
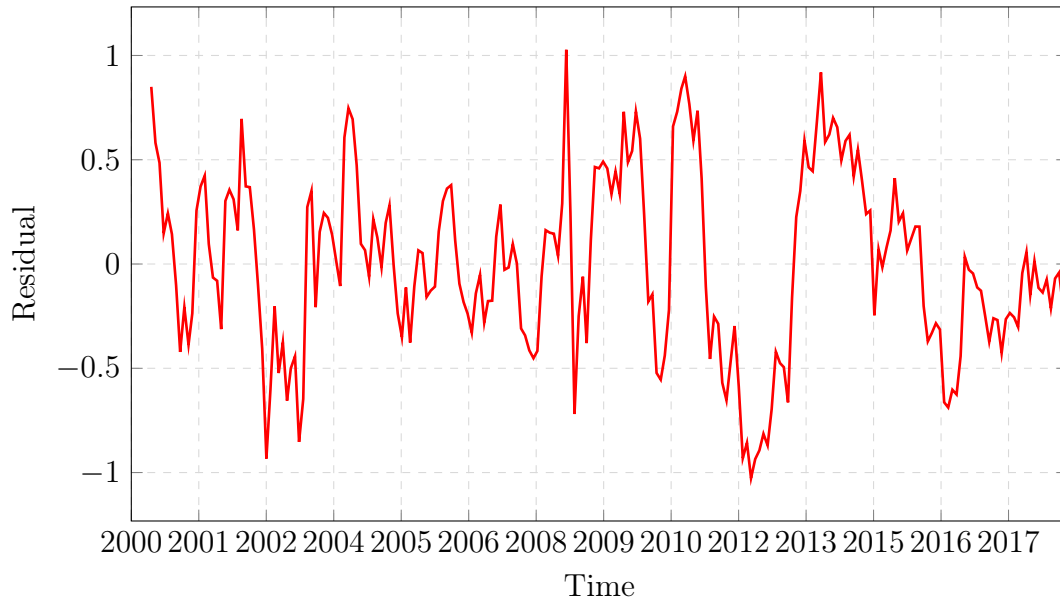


Figure 5.6: US estimated VS real 10 year yields



The residuals are shown in the plot below. We test the stationarity of this residual with both the Augmented Dickey–Fuller test and Phillips–Perron test. Both of these tests confirm that the residuals are stationary. However, there is clearly auto-regression features remaining.



*Figure 5.7: Residual of the model*

### 5.2.2 Germany

We use the same approach as before. The first step is to choose the inputs that we use in the model.

The first step is to look at the correlation between the long term yields and each of the variables. We take the one-year German bond yield, inflation rate, GDP and G4 balance sheet (Aggregate Central Bank balance sheets of the Fed, ECB, BOJ and BOE) as inputs. *(All the data comes from Bloomberg)* We also look at the correlations and the correlation between the variables.

	yield10y	yield1y	Inf	GDP	G4 BS
yield10y	1.0000	0.9252	0.3007	-0.0537	-0.9521
yield1y	0.9252	1.0000	0.3892	0.1574	-0.8912
Inflation	0.3007	0.3892	1.0000	0.3997	-0.2290
GDP	-0.0537	0.1574	0.3997	1.0000	0.0609
G4 Balance Sheet	-0.9521	-0.8912	-0.2290	0.0609	1.0000

Table 5.4: Short term model parameters for the 3-month interest rates

The next step is the cointegration analysis. For the German bonds yields, we use the Johansen Cointegration Test. (*The choice of this test is only based on a desire to try new methods*). By using the functions `jcontest` and `jcitest` in MATLAB, we know that the variables are all integrated of order 1 and that there is only one cointegration relationship between them.

The model in this case is given by:

$$\Delta y_t = A(By_{t-1} + c_0) + \sum_1^k C_k \Delta y_{t-k} + e_t \quad (5.2)$$

Where  $y_t$  is a (5 x 1) vector with all the variables,  $e_t$  is a white noise and  $By_{t-1} + c_0$  is the cointegration term.

The estimated parameters are:

$$B = [-1.3422 \ 0.5146 \ 1.2490 \ -0.4414 \ -0.1102]^T$$

$$A = [ \ 0.0144 \ -0.0467 \ -0.0696 \ -0.0115 \ 0.0620 ]$$

$$c_0 = 4.0833$$

The plot of the estimated long term rates and the residuals are in the next page. The residuals are stationary bu not completely white.

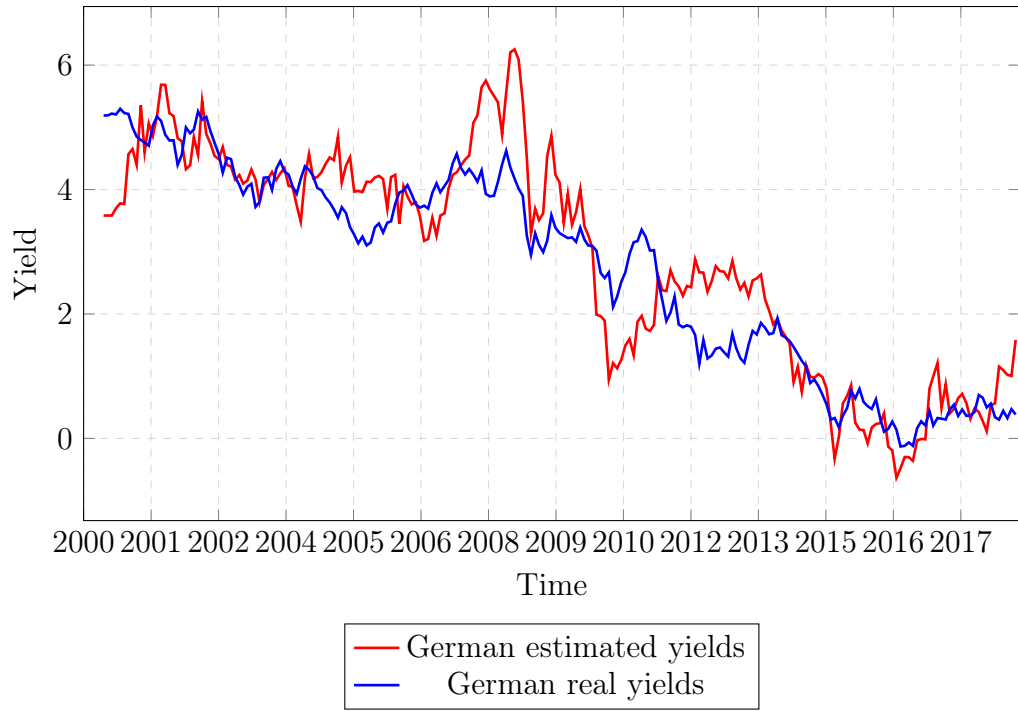


Figure 5.8: German estimated VS real 10 year yields

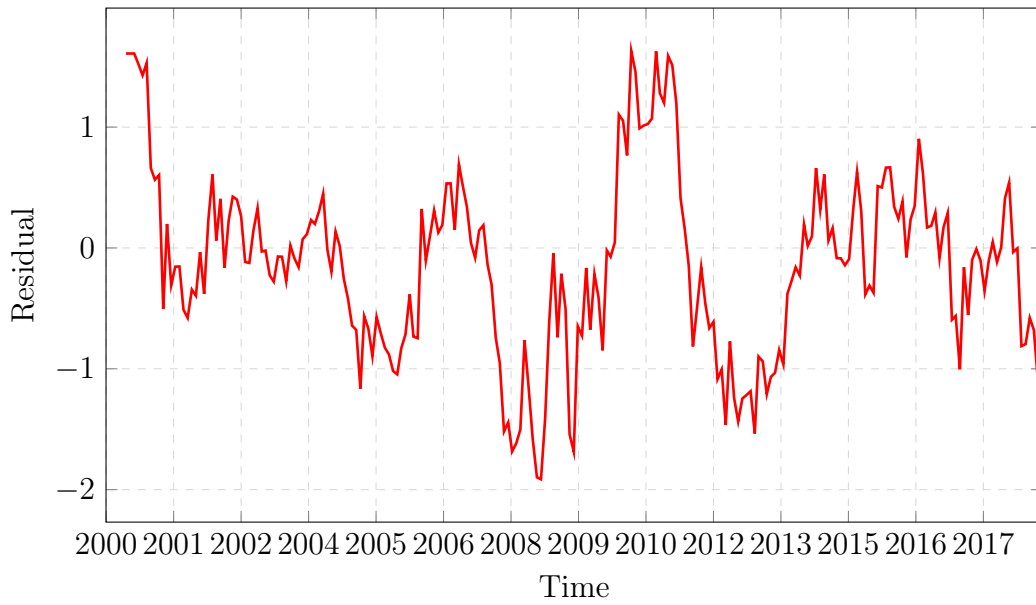


Figure 5.9: Residual fir the German sovereign bond yields fitted model

The main idea behind this approach is to find fair values of the yields based on the macroeconomic variables that we used. By comparing the results for both countries, we can clearly see that the fitted US yields fluctuate much less than for German yields. The low volatility of the US fitted yields is more in line with our goal to find accurate fair values. One possible explanation for this is the multiple variables used for the US relative to Germany.

We get better results using this method than the anticipation theory approach before. The strength of this approach is the accuracy. However, the main weakness is that the forecasts will be based on the anticipations of the different macroeconomic variables from economists. which could be inaccurate.

## 5.3 Term Structure Modeling

This last approach is widely used by central bankers to analyze the yield curve. The use of three factors - level, slope and curvature - helps catch all the effects of the term structure on the short, medium and long term. We fix  $\lambda = 0.2262$ . In this thesis, we will not get into detail on how this parameter has been estimated but it has been in discussed in the following paper. *Source: Calibrating the Dynamic Nelson-Siegel Model: A Practitioner Approach. Ibanez, Francisco Central Bank of Chile 14 December 2015*

### 5.3.1 US

As said before, we use the OLS method to estimate the parameters in the formula below.

$$r_t(\tau) = \beta_1(t) + \beta_2(t) \frac{1 - e^{-\lambda\tau}}{\lambda\tau} + \beta_3(t) \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \quad (5.3)$$

The figure below shows a chart of the three parameters discussed before:

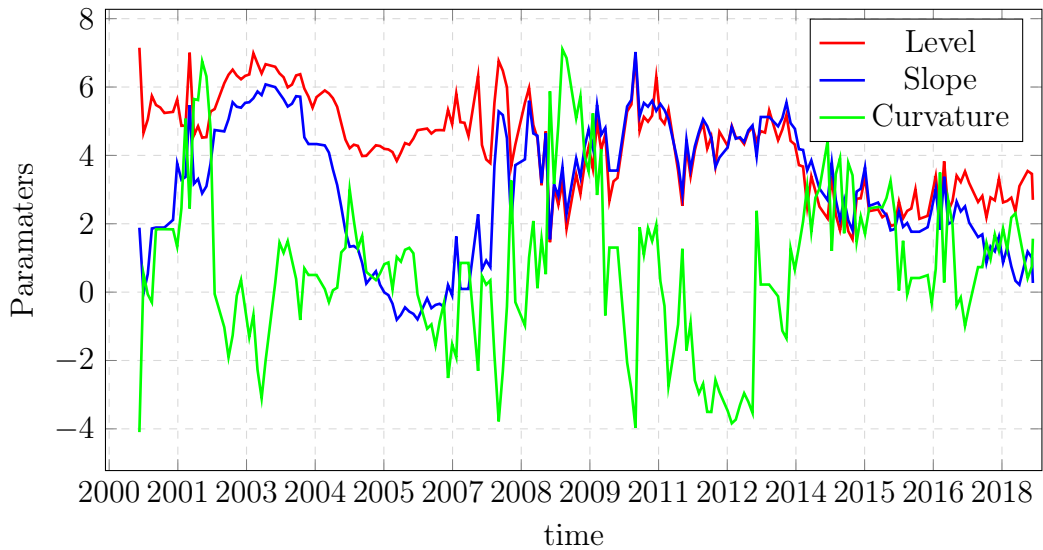


Figure 5.10: Estimated Nelson-Siegel parameters for the US treasury bond yields

The following figure presents a plot of the estimated interest rates:

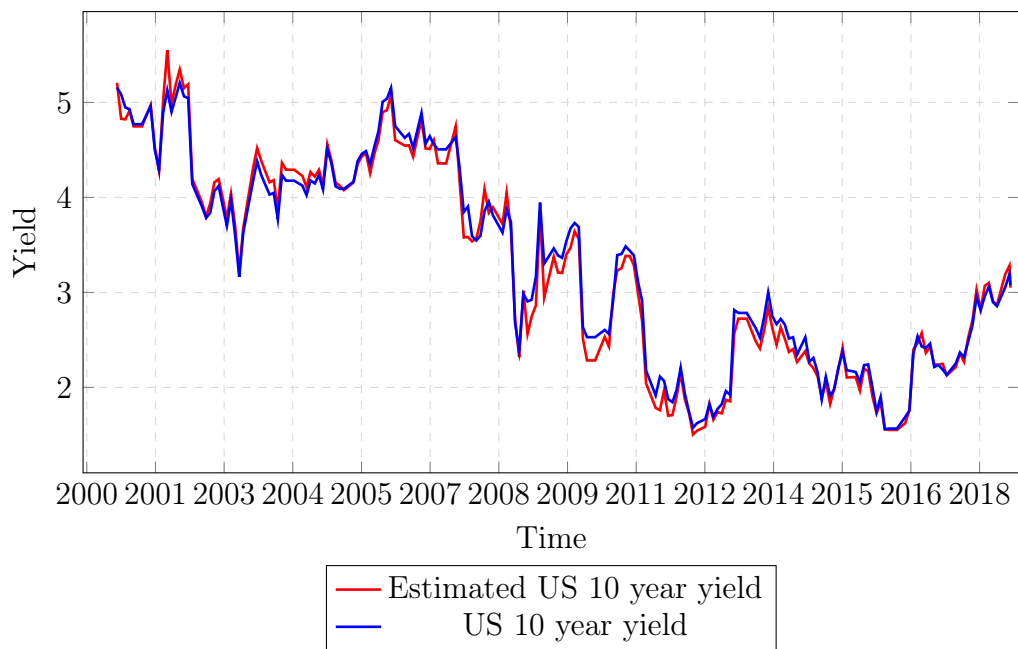


Figure 5.11: US fitted VS real yields

### 5.3.2 Germany

We use the same approach as in the previous case. We use the OLS method to estimate the parameters in the formula below.

$$r_t(\tau) = \beta_1(t) + \beta_2(t) \frac{1 - e^{-\lambda\tau}}{\lambda\tau} + \beta_3(t) \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \quad (5.4)$$

Here is a chart of the three parameters discussed before:

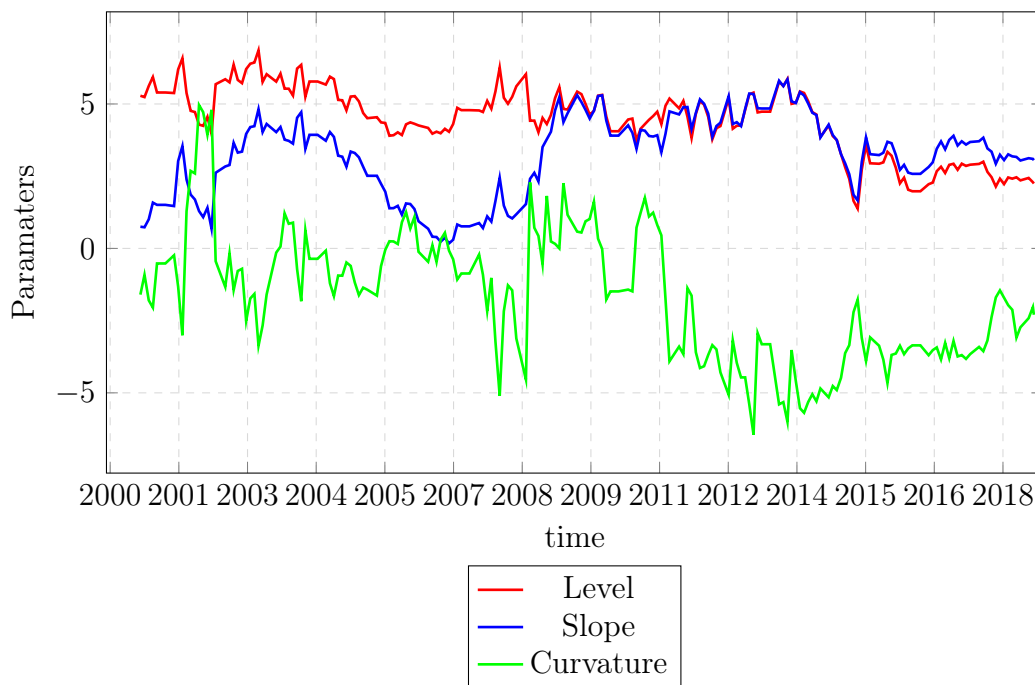


Figure 5.12: Estimated Nelson-Siegel parameters for the German treasury bond yields

The following plot is the representation of the estimated interest rates:



Figure 5.13: German fitted VS real yields

### 5.3.3 Modeling of parameters

In this section we attempt to model the different parameters based on the different macroeconomic inputs, that we discussed in the Time Series Analysis before. The analysis will be done to the US 2 year and 10 year treasury bond yields.

- The first factor: level. The level parameter is not stationary but the first differences is, which means that this factor is  $I(1)$ . The Engle–Granger test confirms the cointegration relationship with the following variables - Short term rates, inflation, GDP and Foreign investments referred to in the equation below as  $X_1, X_2$  and  $X_3$  respectively.

$$factor_1(t) = \alpha_1 + \alpha_2 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \quad (5.5)$$

We get the following parameters:

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	8.3068	0.67544	12.298	1.758e-26
yield3m	-0.17529	0.068251	-2.5683	0.010921
Inflation_yoy	0.54333	0.072297	7.5152	1.6608e-12
Foreign_Hold	-12.511	1.6178	-7.7335	4.4275e-13
GDP_yoy	-0.1566	0.05982	-2.6179	0.0094969

Number of observations: 213, Error degrees of freedom: 208  
 Root Mean Squared Error: 1.15  
 R-squared: 0.377, Adjusted R-Squared 0.365  
 F-statistic vs. constant model: 31.5, p-value = 1.65e-20

The following figure shows the estimated level.

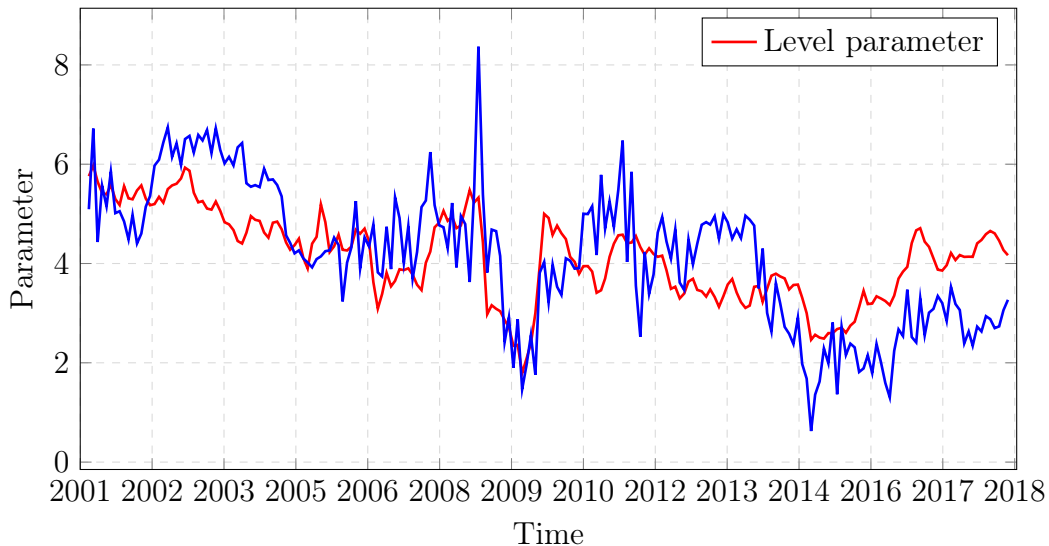


Figure 5.14: Level factor estimation

The model doesn't fit the data very well. One reason is the fact that the level parameter has a lot of noise. The residual analysis shows that the residual is stationary but not white.

- The second factor: slope. The modelling of the slope parameter is similar to the level factor since it is also  $I(1)$  and there is a cointegration rela-



relationship with the variables - Short term rates, inflation, GDP and Foreign investments. referred to in the equation below as  $X_1, X_2$  and  $X_3$  respectively.

$$factor_2(t) = \alpha_1 + \alpha_2 X_1 + \alpha_3 X_2 + \alpha_4 X_3 + \alpha_5 X_4 \quad (5.6)$$

We get the following parameters:

Estimated Coefficients:				
	Estimate	SE	tStat	pValue
(Intercept)	9.3698	0.6583	14.233	1.5078e-32
yield3m	-1.2236	0.066519	-18.394	1.6902e-45
Inflation_yoy	0.52118	0.070463	7.3966	3.3784e-12
Foreign_Hold	-14.259	1.5768	-9.0432	1.085e-16
GDP_yoy	-0.19992	0.058302	-3.429	0.00073036

Number of observations: 213, Error degrees of freedom: 208  
 Root Mean Squared Error: 1.12  
 R-squared: 0.675, Adjusted R-Squared 0.669  
 F-statistic vs. constant model: 108, p-value = 1.17e-49

The following graph shows the estimated level.

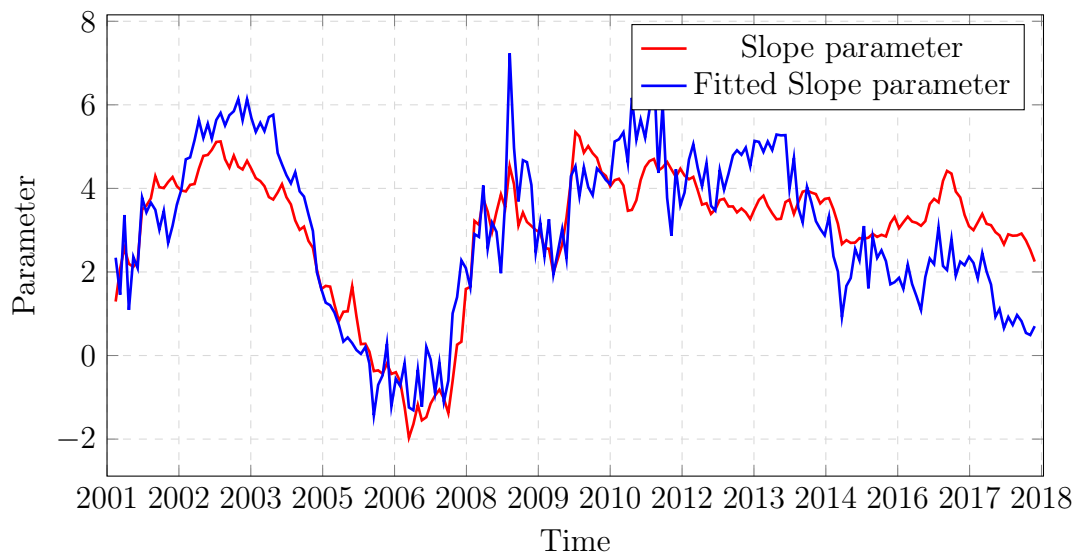


Figure 5.15: Slope factor estimation

The residual is stationary. However, it is not white. As said before, the slope of the yield curve is a good estimator of an upcoming recession. Thus, the modeling of this factor could help forecast a recession and thus be used in risk management. (For more: *Benzoni, Luca and Chyruk, Olena and Kelley, David, Why Does the Yield-Curve Slope Predict Recessions? (September 28, 2018)*)

- The last factor: curvature. Unlike the other factors, This factor is stationary and is the hardest one to model and predict. As said before, this parameter represents the medium term and captures the 'butterfly effect' of the yield curve. We decided to here an AR(2) model as it was more accurate than using macroeconomic variables in this case. The parameters are given by:

```

Discrete-time AR model: A(z)y(t) = e(t)
A(z) = 1 - 0.6412 z^-1 - 0.2162 z^-2

Sample time: 1 seconds

Parameterization:
Polynomial orders: na=2
Number of free coefficients: 2
Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:
Estimated using AR ('fb/now') on "parameter3".
Fit to estimation data: 39.88%
FPE: 2.315, MSE: 2.251

```

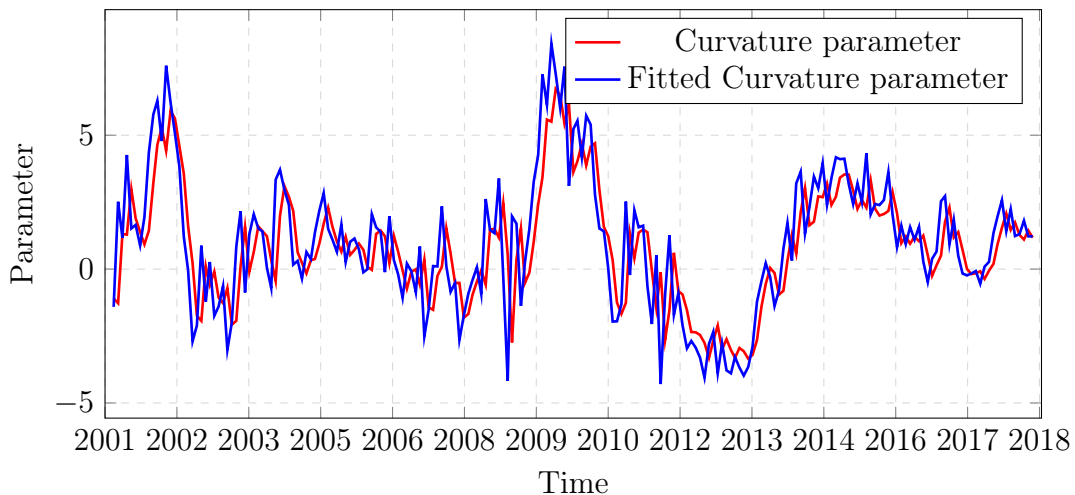
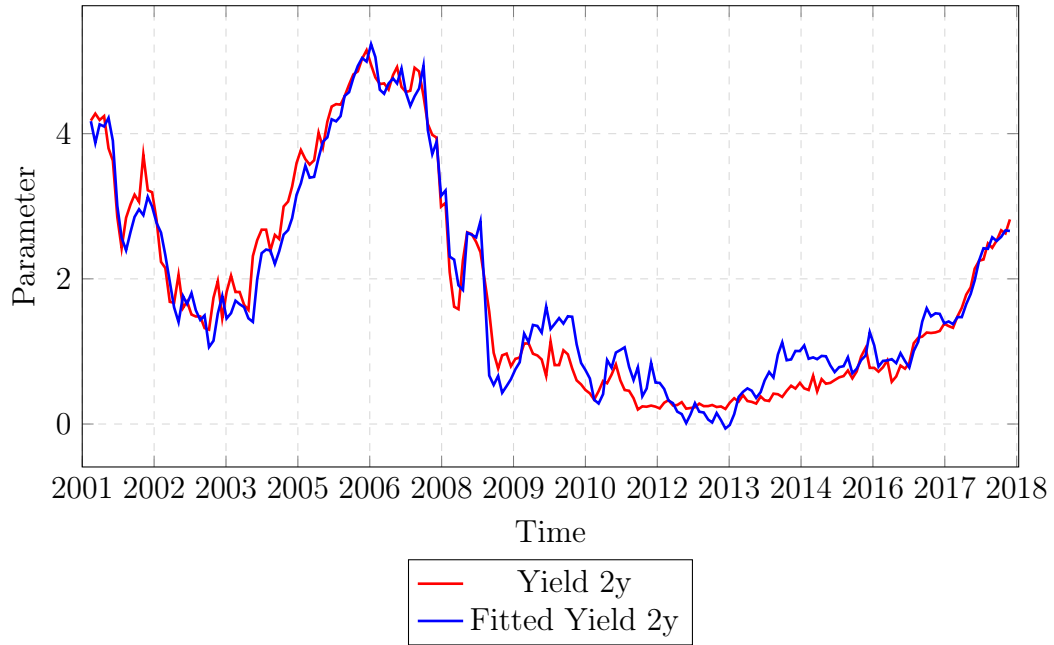
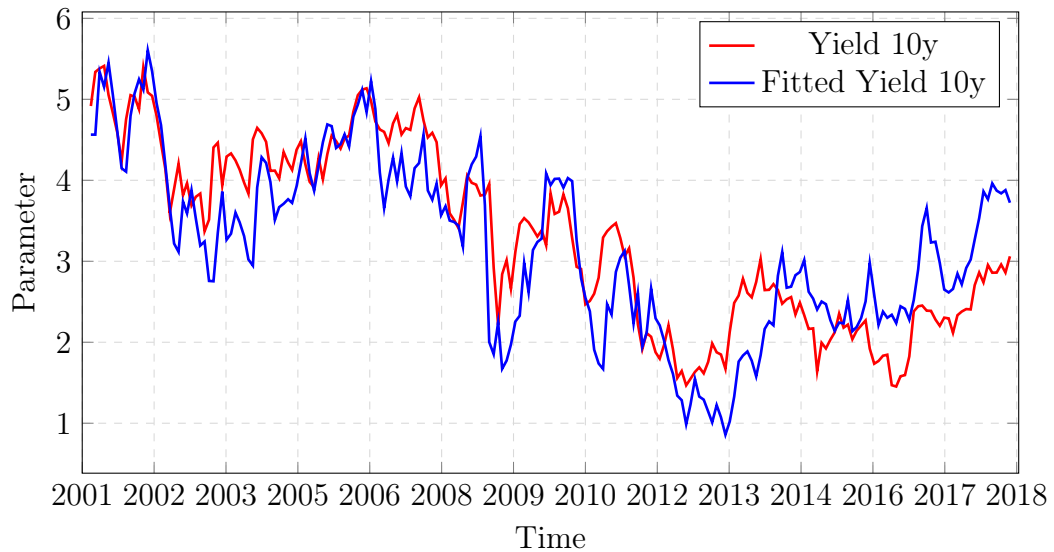


Figure 5.16: Curvature factor estimation

The 2 year and 10 year bond estimates based on the estimated factors is shown in the following graphs:



*Figure 5.17: US 2 year yield estimation*



*Figure 5.18: US 10 year yield estimation*

The results show that our estimates are overvalued in certain years and undervalued in other ones. The residuals from both these estimations are stationary but still not white.

This approach could be improved using better and more accurate macroeconomic inputs. However, I believe that this approach would not be well suited for forecasting purposes. The idea of using macroeconomic variables to forecast each factor could lead to major errors in estimation of the yields as each of the three factors would be inaccurate if one of the variables is badly forecasted.



# Chapter 6

## Conclusion

Financial companies have always been interested in modelling interest rates for various purposes and the choice of the interest rate models is not trivial. Financial intermediaries such as banks and brokers use interest rate models mainly for the pricing of financial derivatives such as caps, swaptions, bond options etc. Several models have been proposed such as Vasicek model (1977) and Cox, Ingersoll and Ross (1985) as short rate models for this purpose. Other financial companies, such as pension funds and asset management companies use them for predictions on the future fluctuations of bond yields using financial statistics and time series analysis models such as ARCH and GARCH models.

We used three approaches in this paper. The first one is the Anticipation theory, it uses the short term analysis using the Vasicek and CIR models and use the anticipations to estimate the long term interest rates. The second approach is using a Time Series/Regression Analysis using many macroeconomics inputs to get a estimate of the fair values of interest rates and forecasts. This approach is used a lot in Asset Management companies for speculative and hedging purposes and to get Quantitative Investment ideas about the future path of the interest rates. For example, BNP Paribas Quant team developed a indicator called MarFA that uses macroeconomic variable to estimate the Fair Value of an asset and determines if the current market price is over or undervalued. It has been a successful product that many clients of the bank are interested in. Our last approach is the Term Structure Modelling using the Nelson Siegel model. It is more broad approach as it uses the level, slope and curvature to model the interest rate surface.

To conclude, each of these models each have their strengths and weaknesses and their use is based on the purpose of the modelling.

# Bibliography

- [1] Annaert J, Claes AG, De Ceuster MJ, Zhang H. Estimating the Yield Curve Using the Nelson-Siegel Model. European Financial Management Association, Aarhus. 2010.
- [2] Koopman SJ, Mallee MI, Van der Wel M. Analyzing the term structure of interest rates using the dynamic Nelson–Siegel model with time-varying parameters. *Journal of Business Economic Statistics*. 2010 Jul 1;28(3):329-43.
- [3] BIS, MD. "Zero-coupon yield curves estimated by central banks." BIS papers 25 (2005).
- [4] Jondeau, E. and Ricart, R., 1999. Le contenu en information de la pente des taux: Application au cas des titres publics français. *Économie prévision*, 140(4), pp.1-20.
- [5] Hall, A.D., Anderson, H.M. and Granger, C.W., 1992. A cointegration analysis of treasury bill yields. *The review of Economics and Statistics*, pp.116-126.
- [6] Christensena J, Diebold D and Rudebusch G, The affine arbitrage-free class of Nelson–Siegel term structure models, *Journal of Econometrics* Volume 164, Issue 1, 1 September 2011, Pages 4-20
- [7] Diebold F.X. and Li C., 2006. Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, 130, 337-364.
- [8] Jondeau, E. and Sedillot, F., 1998. La prevision des taux longs français et allemands a partir d'un modele a anticipations rationnelles (No. 55).



- [9] KAMGNA, S.Y., NGUENANG, C., TALABONG, H. and OULD, I., 2009. Central Bank reaction function and monetary policy credibility: The case of BEAC.
- [10] Loubens A, Jardet C, Idier J. Les déterminants des taux d'intérêt à long terme aux États-Unis et dans la zone euro : une approche multivariée. In: *Économie prévision*, n185, 2008-4. pp. 13-32
- [11] Muni Toke I, 2011. *Modeles stochastiques de taux d'interets - Draft: 11 january 2011.*
- [12] Ricketts N , Real short-term interest rates and expected inflation: Measurement and interpretation, *Bank of Canada Review Summer 1996*
- [13] Sedillot F, La pente des taux contient-elle de l'information sur l'activité économique future ? - Direction des Études économiques et de la Recherche. *Bulletin De La Banque De France – N 63 – Mars 1999.*
- [14] Kliesen Kevin. *Recession Signals: The Yield Curve VS. Unemployment Rate Troughs*, Economic Research Federal Reserve Bank Of St. Louis, 2018.
- [15] Wheelock, David C. and Mark E. Wohar. "Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature." *Federal Reserve Bank of St. Louis Review*, September-October 2009, 91(5), pp. 419-40.
- [16] Prat, Georges. "Anticipations, prime de terme et maturité du titre long Que nous enseignent les données séculaires sur la structure des taux d'intérêt? États-Unis: 1873 à 1975." *Revue économique* (1992): 1037-1069.
- [17] Jondeau, E., 1998. Représentation VAR et test de la théorie des anticipations de la structure par terme. *Journal de la société française de statistique*, 139(1), pp.49-71.
- [18] Dewachter, H. and Lyrio, M., 2006. Macro factors and the term structure of interest rates. *Journal of Money, Credit, and Banking*, 38(1), pp.119-140.
- [19] Diebold, F.X., Piazzesi, M. and Rudebusch, G.D., 2005. Modeling bond yields in finance and macroeconomics. *American Economic Review*, 95(2), pp.415-420.

# Annex - Computer Code

```
%% Anticipation Theory: Stationary
%% US
%% Data
clear
clc
data = importdata('Anticipation_theory_data.mat');
N = length(data(1:end,1)) ;
data_US_3m = data(:,1);
data_US_10y = data(:,2);

% Vasicek model
yields = data_US_3m;
regressors = [ones(length(yields) - 1, 1) yields(1:end-1)];
[coefficients, interv, residuals] = regress(diff(yields), regressors); ...
% OLS regression
dt = 1/260;
tbl=table(diff(yields),ones(length(yields) - 1, 1),yields(1:end-1),...
'VariableNames',{'diff_yields','ones','yields'});
lm1 = fitlm(tbl,'diff_yields ~ yields','Intercept',true)

kappa = -coefficients(2)/dt;
mu = -coefficients(1)/coefficients(2);
sigma = std(residuals)/sqrt(dt);

kappah = -table2array(lm1.Coefficients(2,1))/dt;
muh=table2array(lm1.Coefficients(2,1))/table2array(lm1.Coefficients(1,1));
sigmah = std(table2array(lm1.Residuals(:,1)))/sqrt(dt);

Par = [kappa , mu, sigma];
Parh = [kappah , muh, sigmah];

plot(yields)

% Long term model
A = 1- kappa*dt;
G = kappa*mu*dt ;
n = 12/3*10;
R = G/(1-A) + A*(1-A^n)/(n*(1-A))*(yields(1:end-1)-G/(1-A));
lambda = mean(data_US_10y(2:end)-R)
plot(R+lambda)
hold on
plot(data_US_10y(2:end))

%% Germany
```

```

%% Data
clear
clc
data = importdata('Anticipation_theory_data.mat');
N = length(data(1:end,1)) ;
data_German_3m = data(:,3);
data_German_10y = data(:,4);

% Vasicek model
yields = data_German_3m;
regressors = [ones(length(yields) - 1, 1) yields(1:end-1)];

[coefficients, interv, residuals] = regress(diff(yields), regressors);...
% OLS regression
dt = 1/260; % daily data
tbl=table(diff(yields),ones(length(yields) - 1, 1),yields(1:end-1),...
'VariableNames',{ 'diff_yields', 'ones', 'yields' });
lm1 = fitlm(tbl, 'diff_yields ~ yields', 'Intercept', true)

kappa = -coefficients(2)/dt;
mu = -coefficients(1)/coefficients(2);
sigma = std(residuals)/sqrt(dt);

kappah = -table2array(lm1.Coefficients(2,1))/dt;
muh=table2array(lm1.Coefficients(2,1))/table2array(lm1.Coefficients(1,1));
sigmah = std(table2array(lm1.Residuals(:,1)))/sqrt(dt);

Par = [kappa , mu, sigma]
Parh = [kappah , muh, sigmah]

plot(yields)

% Long term model
A = 1- kappa*dt;
G = kappa*mu*dt ;
n = 12/3*10;
R = G/(1-A) + A*(1-A^n)/(n*(1-A))*(yields(1:end-1)-G/(1-A));
lambda = mean(data_German_10y(2:end)-R)
plot(R+lambda)
hold on
plot(data_German_10y(2:end))

%% Anticipation Theory: Non - Stationary
%% US
%% Data
clear
clc
data = importdata('Anticipation_theory_data.mat');
N = length(data(1:end,1)) ;
data_US_3m = data(:,1);
data_US_10y = data(:,2);

% short rate model
yields = data_US_3m;
regressors = [ones(length(yields) - 2, 1) yields(2:end-1)-yields(1:end-2)];
[coefficients, intervals, residuals, rint, stats] = ...
regress(yields(3:end)-yields(2:end-1), regressors); % OLS regression
dt = 1/260;
tbl=table(yields(3:end)-yields(2:end-1),yields(2:end-1)-yields(1:end-2),...
'VariableNames',{ 'diff_yields', 'yields' });

```

```

lm1 = fitlm(tbl, 'diff_yields ~ yields', 'Intercept', true)

kappah = -table2array(lm1.Coefficients(2,1))/dt;
muh=-table2array(lm1.Coefficients(2,1))/table2array(lm1.Coefficients(1,1));
sigmah = std(table2array(lm1.Residuals(:,1)))/sqrt(dt);

Par = coefficients
Parh = table2array(lm1.Coefficients)

plot(yields)

% Long term model
A = coefficients(2);
n = 12/3*10;
R = yields(2:end-1) + A/(1-A)*(yields(2:end-1)-yields(1:end-2))*...
    (1-A*(1-A^(n)))/n*(1-A));
lambda = mean(data_US_10y(3:end)-R)
plot(R+lambda)
hold on
plot(data_US_10y(2:end))

%% Germany
%% Data
clear
clc
data = importdata('Anticipation_theory_data.mat');
N = length(data(1:end,1)) ;
data_German_3m = data(:,3);
data_German_10y = data(:,4);

% short rate model
yields = data_German_3m;
regressors = [ones(length(yields) - 2, 1) yields(2:end-1)-yields(1:end-2)];
[coefficients, interv, residuals, rint, stats] = ...
    regress(yields(3:end)-yields(2:end-1), regressors); % OLS regression
dt = 1/260;
tbl=table(yields(3:end)-yields(2:end-1),yields(2:end-1)-yields(1:end-2),...
    'VariableNames',{'diff_yields','yields'});
lm1 = fitlm(tbl, 'diff_yields ~ yields', 'Intercept', true)

kappah = -table2array(lm1.Coefficients(2,1))/dt;
muh=-table2array(lm1.Coefficients(2,1))/table2array(lm1.Coefficients(1,1));
sigmah = std(table2array(lm1.Residuals(:,1)))/sqrt(dt);

Par = coefficients
Parh = table2array(lm1.Coefficients)

% Long term model
A = coefficients(2);
n = 12/3*10;
R = yields(2:end-1) + A/(1-A)*(yields(2:end-1)-yields(1:end-2))*...
    (1-A*(1-A^(n)))/n*(1-A));

lambda = mean(data_German_10y(3:end)-R)
plot(R+lambda)
hold on
plot(data_German_10y(3:end))

```

```

%%Time Series Analysis
%% US yields: LT
%% Read data
clear
clc
data = importdata('Time_Series_Analysis_data_US.mat');

% Data
yield3m = data(:,1); yield6m = data(:,2); yield1y = data(:,3); yield2y = ...
    data(:,4); yield3y = data(:,5); yield5y = data(:,6); yield7y = ...
    data(:,7); yield10y = data(:,8); yield30y = data(:,9); % US Yields
Inflation_CPI_yoy = data(:,15); % Inflation
Unemployment_rate = data(:,19); % Unemployment
US_Treasury_Major_Foreign_Holdings = data(:,24); Treasury_Outstanding ...
    = data(:,25); Foreign_Holdings = ...
    US_Treasury_Major_Foreign_Holdings./Treasury_Outstanding;% Foreign H
Eurodollar_Yields=(100-data(:,27)+100-data(:,28))/2; Eurodollar_1y = ...
    100-data(:,27); % Eurodollar Futures
FED_Balance_Sheet = data(:,42); Fed_Balance_GDP = data(:,44); GDP_yoy =...
    data(:,48); % GDP
IP_index_yoy=data(:,56);PMI_US_index=data(:,57);PMI_EU_index=data(:,58);
PME_Germany_index = data(:,59) ;% Manufacturing Indices

%% First Correlation between Variables

r = yield10y ;
N = length(r) ;
k = 0 ;

X=( [yield10y(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), ...
    Fed_Balance_GDP(1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N), ...
    Unemployment_rate(1+k:N)]);
corr(X);

%% Regression using Engle-Granger cointegration

Y=( [yield10y(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), ...
    Foreign_Holdings(1+k:N), GDP_yoy(1+k:N), Fed_Balance_GDP(1+k:N)]);
i = size(Y,2);
[h0,pValue0] = jctest(Y,1,'BVec',{[1 zeros(1,i-1)]',[0 1 ...
    zeros(1,i-2)]',[0 0 1 zeros(1,i-3)]',[0 0 0 1 zeros(1,i-4)]',...
    [0 0 0 0 1 zeros(1,i-5)]',[0 0 0 0 0 1 zeros(1,i-6)]'],'Alpha',0.075);
[h,pValue,stat,cValue,reg1,reg2] = egcitest(Y);

tbl=table(r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), ...
    Foreign_Holdings(1+k:N), GDP_yoy(1+k:N), Fed_Balance_GDP(1+k:N), ...
    'VariableNames',{'yield10y','yield3m','Inflation_yoy',...
    'Foreign_Hold','GDP_yoy','Fed_BS'});
lm1 = fitlm(tbl,...
    'yield10y ~ yield3m+Inflation_yoy+Foreign_Hold +GDP_yoy + Fed_BS'...
    ,'Intercept', true);
lm1

figure
ts1 = timeseries(lm1.Fitted);
plot(ts1)
hold on

```

```

ts2 = timeseries(r(1+k:N));
plot(ts2)

legend('Yield 10y fair value','Yield 10y')

residual = r(1+k:N)-lm1.Fitted ;
figure
plot(residual)
[adftest(residual),pptest(residual)];
archtest(residual)

writetable(table(flipud(ts2.Data)), 'Data_Time_Series.xlsx', 'Range', 'B2'...
, 'WriteVariableNames', false);
writetable(table(flipud(ts1.Data)), 'Data_Time_Series.xlsx', 'Range', 'C2'...
, 'WriteVariableNames', false);
writetable(table(flipud(residual)), 'Data_Time_Series.xlsx', 'Range', 'D2'...
, 'WriteVariableNames', false);

%% Germany yields: LT
%% Read data
clear all
clc
data = importdata('Time_Series_Analysis_data_G.mat');

%Data
yield3m = data(:,1); yield2y = data(:,2); yield3y = data(:,3); yield4y = ...
data(:,4); yield5y = data(:,5); yield7y = data(:,6); yield8y = ...
data(:,7); yield10y = data(:,8); yield30y = data(:,9); % US Yields
Euribor_rates = data(:,30); ECB_rate = data(:,12) ; Euribor_rates_1y = ...
data(:,30); % EU rates
Inflation_CPI_EU_yoy = data(:,14); Inflation_CPI_Germany_yoy = ...
data(:,15); % Inflation
Unemployment_EU = data(:,19); Unemployment_Germany=data(:,20); % Unemploy
G4_Balance_Sheet_percentage_of_GDP = data(:,22); GDP_yoy = data(:,34)*100 ;
% GDP

%% First Correlation between Variables

r = yield10y ;
N = length(r);
k = 0 ;

X=( [yield10y(1+k:N), yield3m(1+k:N), Inflation_CPI_Germany_yoy(1+k:N) , ...
GDP_yoy(1+k:N), G4_Balance_Sheet_percentage_of_GDP(1+k:N)] );
corr(X);

Y=( [yield10y(1+k:N), yield3m(1+k:N), Inflation_CPI_Germany_yoy(1+k:N) , ...
GDP_yoy(1+k:N)] );
i = size(Y,2);
[h0,pValue0] = jcontest(Y,1,'BVec', {[1 zeros(1,i-1)]}', ...
[0 1 zeros(1,i-2)]', [0 0 1 zeros(1,i-3)]'}, 'Alpha', 0.05);

%% Cointegration using Johansen test
Y=X;
[h,pValue,stat,cValue,mles] = jcitest(Y,'model','H1*', 'display','full');h
mles.r1.paramNames
YLag = Y(1:end-1,:);
T = size(YLag,1);

```

```

A = mles.r1.paramVals.A;
B = mles.r1.paramVals.B; B=B(:,end);
c0 = mles.r1.paramVals.c0';
param = [-c0/B(1);-B(2:end)/B(1)]';
residual = YLag*B/B(1)+repmat(c0',T,1)/B(1);
Yhat = [ones(size(YLag,1),1) YLag(:,2:end)]*param';
plot(Yhat)
hold on
plot(yield10y)
figure
plot(residual)
[adftest(residual),pptest(residual)]
archtest(residual)

writetable(table(flipud(yield10y)), 'Data_Time_Series.xlsx', 'Range', 'E2', ...
'WriteVariableNames', false);
writetable(table(flipud(Yhat)), 'Data_Time_Series.xlsx', 'Range', 'F2', ...
'WriteVariableNames', false);
writetable(table(flipud(residual)), 'Data_Time_Series.xlsx', 'Range', 'G2', ...
'WriteVariableNames', false);

```

```

%% Nelson Siegel Approach
%% Parameters
clc, clear
terms = [3/12 6/12 1 2 3 5 7 10 20] ;
spot = [3/12 6/12 9/12 12/12 15/12 18/12 21/12 24/12 30/12 36/12 4:1:20] ;
lambda = 0.2262 ;

%% Yields base construction
clc
nseries=length(terms); % We count the number of series.

data = importdata('Nelson_Siegel_data.mat');
times = importdata('Nelson_Siegel_time.mat');
N = length(data(1:end,2)) ;
data1 = data(1:N,1:9); % US
%cmt = data(1:N,10:18); % Germany
yields = data1(:,1:end);
dates = datenum(times);

%% Bootstrapping the yields.
time=datetime(dates, 'ConvertFrom', 'datenum');
spot_surface = zeros(length(dates),length(spot));
for i=1:length(dates)
    yield_matrix=[yields(i,:)'/100 zeros(length(terms),1)];
    for j=1:length(terms)
        yield_matrix(j,2)=addtodate(dates(i,1), terms(j)*12, 'month');
    end
    yield_matrix=yield_matrix(~any(isnan(yield_matrix),2),:);
    [z,t]=pyld2zero(yield_matrix(:,1), yield_matrix(:,2), dates(i));
    t=yearfrac(dates(i),t);
    spot_surface(i,:)=spline(t,z,spot)*100;
end
clc, clear i j z t yield_matrix

%% Curve Fitting.
term_factor=zeros(length(spot),3);
term_factor(:,1)=1;
term_factor(:,2)=-(1-exp(-lambda.*spot'))./(lambda.*spot');

```

```

term_factor(:,3)=(1-exp(-lambda.*spot'))./(lambda.*spot')...
-exp(-lambda.*spot');
lsc=zeros(length(dates),3); % Time series vector

%% Estimation
bounds=[0 -15 -15 ; 15 15 15];% Parametersboundries.
rest=[-1,-1,0];%(b1+b2)>0
opt=optimoptions(@lsqlin,'Display','off','Algorithm','interior-point');
for i=1:length(dates)
    lsc(i,:)=lsqlin(term_factor,spot_surface(i,:),rest,0,[],[],bounds...
        (1,:),bounds(2,:),[],opt);
end
plot(time(1:15:end),lsc(1:15:end,:))
legend('Level','Slope','Curvature')
Yhat = lsc * term_factor'; Yhat = Yhat(:,17);

plot(Yhat)
hold on
plot(yields(:,8))

%% Term structure plot
clc
clear X Y
X= repmat(time,1,size(yields,2))
Y= repmat(terms,size(yields,1),1);
surface(X,Y,yields)

%% Modeling of the parameters
%% Take monthly values
range = datetime('01/13/2001'):calmonths(1):datetime('10/30/2018');
par1=table(time,lsc(:,1));
par2=table(time,lsc(:,2));
par3=table(time,lsc(:,3));

for i=2:length(range)
    if isempty(lsc(find(par1.time==range(i),1)))==0
        parameter1(i)=lsc((find(par1.time==range(i))),1);
        parameter2(i)=lsc((find(par2.time==range(i))),2);
        parameter3(i)=lsc((find(par3.time==range(i))),3);
    else
        if isempty(lsc(find(par1.time==range(i)-1,1)))==0
            parameter1(i)=lsc((find(par1.time==range(i)-1)),1);
            parameter2(i)=lsc((find(par2.time==range(i)-1)),2);
            parameter3(i)=lsc((find(par3.time==range(i)-1)),3);
        else
            if isempty(lsc(find(par1.time==range(i)-2,1)))==0
                parameter1(i)=lsc((find(par1.time==range(i)-2)),1);
                parameter2(i)=lsc((find(par2.time==range(i)-2)),2);
                parameter3(i)=lsc((find(par3.time==range(i)-2)),3);
            else
                parameter1(i)=lsc((find(par1.time==range(i)-3)),1);
                parameter2(i)=lsc((find(par2.time==range(i)-3)),2);
                parameter3(i)=lsc((find(par3.time==range(i)-3)),3);
            end
        end
    end
end
end
end
end

```



```

%% Read data US

clear
clc
data = importdata('Time_Series_Analysis_data_US.mat');

% Data
yield3m = data(:,1); yield6m = data(:,2); yield1y = data(:,3); yield2y = ...
    data(:,4); yield3y = data(:,5); yield5y = data(:,6); yield7y = ...
    data(:,7); yield10y = data(:,8); yield30y = data(:,9); % US Yields
Inflation_CPI_yoy = data(:,15); % Inflation
Unemployment_rate = data(:,19); % Unemployment
US_Treasury_Major_Foreign_Holdings = data(:,24); Treasury_Outstanding ...
    = data(:,25); Foreign_Holdings = ...
    US_Treasury_Major_Foreign_Holdings./Treasury_Outstanding;% Foreign H
Eurodollar_Yields=(100-data(:,27)+100-data(:,28))/2; Eurodollar_1y = ...
    100-data(:,27); % Eurodollar Futures
FED_Balance_Sheet = data(:,42); Fed_Balance_GDP = data(:,44); GDP_yoy =...
    data(:,48); % GDP
IP_index_yoy=data(:,56);PMI_US_index=data(:,57);PMI_EU_index=data(:,58);
PME_Germany_index = data(:,59) ;% Manufacturing Indices

%% Modelling of the level, slope and curvature – AR(1)+endogenous variables
%% Level
r = parameter1';
N = length(r);
k=0;
Y = [r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), Fed_Balance_GDP...
    (1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N)];
tbl=table(r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), ...
    Fed_Balance_GDP(1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N), ...
    Unemployment_rate(1+k:N), 'VariableNames', {'parameter', 'yield3m', ...
    'Inflation_yoy', 'Fed_Balance_GDP', 'Foreign_Hold', 'GDP_yoy', ...
    'Unemployment_rate'});
lm1 = fitlm(tbl, 'parameter~Inflation_yoy+GDP_yoy +yield3m+Foreign_Hold'...
    , 'Intercept', true);
lm1

plot(lm1.Fitted)
hold on
plot(r(1+k:N))

%% Slope
r = parameter2';
N = length(r);
k=0;
Y = [r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), Fed_Balance_GDP...
    (1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N)];
tb2=table(r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), ...
    Fed_Balance_GDP(1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N), ...
    Unemployment_rate(1+k:N), 'VariableNames', {'parameter', 'yield3m', ...
    'Inflation_yoy', 'Fed_Balance_GDP', 'Foreign_Hold', 'GDP_yoy', ...
    'Unemployment_rate'});
lm2 = fitlm(tb2, 'parameter~Inflation_yoy+GDP_yoy +yield3m+Foreign_Hold'...
    , 'Intercept', true);
lm2

```

```

plot(lm2.Fitted)
hold on
plot(r(1+k:N))

%% Curvature
r = parameter3';
N = length(r);
k=0;
Y = [r(1+k:N), yield3m(1+k:N), Inflation_CPI_yoy(1+k:N), Fed_Balance_GDP...
      (1+k:N), Foreign_Holdings(1+k:N), GDP_yoy(1+k:N)];
th = ar(parameter3,2);
par = th.a;
p_pred = [-par(2) -par(3)]* [r(2:N-1) r(1:N-2)]';
plot(p_pred)
hold on
plot(r(3:N))

%% Plots
yields = lsc*term_factor';
lsc_pred = [lm1.Fitted lm2.Fitted lm3.Fitted];
yields_pred = lsc_pred*term_factor';

ts1 = timeseries(yields_pred(:,8));
plot(ts1)
hold on
ts2 = timeseries(yield2y);
plot(ts2)

legend('Yield 2y fair value', 'Yield 2y')

figure

ts1 = timeseries(yields_pred(:,17));
plot(ts1)
hold on
ts2 = timeseries(yield10y);
plot(ts2)

legend('Yield 10y fair value', 'Yield 10y')

```