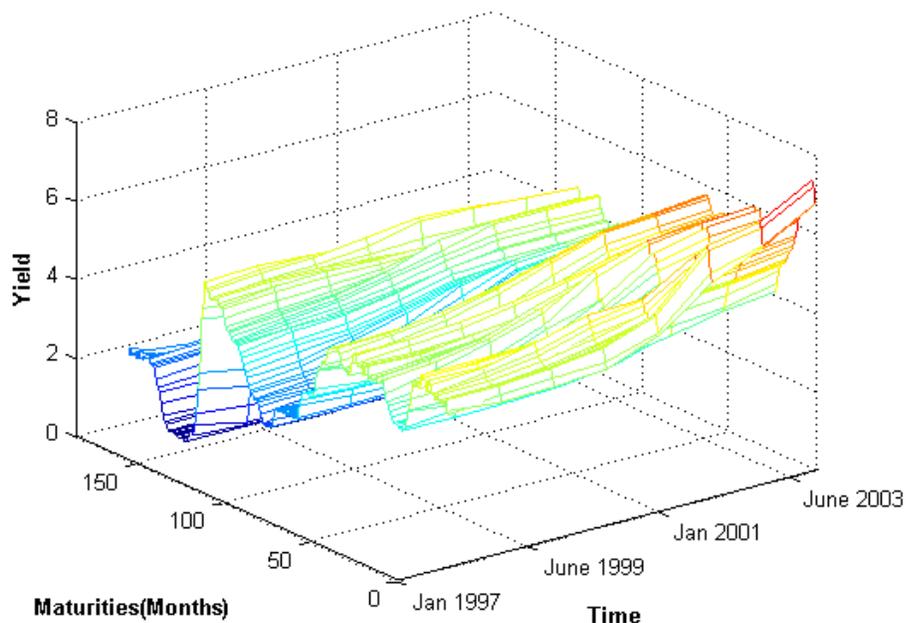

Swedish Bonds Term Structure Modeling with The Nelson Siegel Model

THE Monthly Swedish Short and Long Yield from year 1997 to 2011



Malick Senghore

Bachelors Thesis (2013). Lund University, Sweden.

CONTENTS

ACKNOWLEDGEMENT	1
1 BACKGROUND AND INTRODUCTION	2
1.1 Background	2
1.2 Introduction	3
2 INTEREST RATES THEORY	4
2.1 Discount factor	4
2.2 The Stochastic discount factor	5
2.3 Zero Coupon bonds	6
2.3.1 The Yield Curve	7
2.4 Coupon Bonds	9
2.4.1 Fixed Coupon Bonds	10
2.4.2 The Yield To Maturity	11
2.4.3 Floating rate notes and Variable rate notes	12
2.5 The Day-Count Convention and The Compounding Types	12
2.5.1 Day-Count Convention (year fraction)	13
2.5.2 Compounding Types	13
2.5.3 Continuously compounded interest rates	13
2.5.4 Simply compounded rate	14
2.5.5 Annually-compounded spot interest rate	15
2.5.6 k-times-per-year compounded spot interest rate	15
2.6 Interest rates assets	16
2.6.1 Forward Rate	16
2.6.2 Forward Rate Agreement (FRA)	16
2.6.3 Interest rate Swaps	18

2.6.4	Future Contract	20
2.6.5	Caps and Caplets	20
2.6.6	Floor and Floorlets	21
2.7	Swaptions	21
3	LINEAR SYSTEMS	23
3.1	Linear processes	23
3.2	White Noise	23
3.2.1	Moving Averages (MA)	24
3.2.2	Autoregressive processes AR	24
3.2.3	ARMA processes	25
4	DYNAMIC MODELS	26
4.1	Vector Autoregressive Process (VAR)	26
4.2	State Space Modeling	26
4.2.1	The Kalman Filter	27
5	THE NELSON-SIEGEL MODELS	30
5.1	Nelson-Siegel Model Classes	31
5.1.1	Three-factor base model	31
5.1.2	Two-Factor Model	33
5.1.3	Four-Factor Model	33
6	APPLICATION OF THE NELSON SIEGEL MODEL	36
6.1	Intuition Behind The Three Factor Nelson Siegel Model	36
6.2	Data	37
6.3	Descriptive Statistics, data	39
6.4	Estimation of model parameters	40
6.5	Analysis	41
6.5.1	Modeling the Estimated Factors	44
6.6	Conclusion	49
	BIBLIOGRAPHY	50

ACKNOWLEDGEMENT

My sincere thanks and gratitude to lecturer and supervisor Magnus Wiktorsson, without whom this thesis would not have materialized. Thank you very much for your help, support and patient.

BACKGROUND AND INTRODUCTION

1.1 Background

Fixed income securities, also known as *bonds*, are one of the most commonly used loan instruments by governments, private companies and financial intermediaries to raise funds. For instance, they are used by governments to finance budget deficits and public expenditures, by private companies to finance current and future investments and by banks and other financial intermediaries to finance loans to household and companies.

The price paid or required by the participants in the credit market, where fixed income securities are traded, is referred to as *interest*. In other words, interest is the compensation required to loan out money, or the cost of taking a loan.

Interest rates affects many aspects of our lives in many different ways. From decisions on mortgages, buying and selling of assets (both real and financial), private and government investments, monetary policies to stimulate economic growth both in the short- and long run, etc. It is therefore imperative to understand interest rate and its role in the financial system and the society as a whole. The knowledge on interest rates will help us to take rational decisions on private economic issues, as well as in understanding and taking part in political debates on macroeconomic issues. Moreover, the very idea that money can be traded at a price is fascinating in itself.

There exist a variety of fixed income securities in the market for both short- and long maturities as well as fixed or variable interest rates and also a variety of models to model, estimate and forecast their term structures.

1.2 Introduction

Swedish government bonds will be the main focus of this thesis. The money market in Sweden, (i.e the market in which Swedish government bonds with maturities less a year are traded), develops from a regulated system in the seventies to a well developed and functioning international market as we know it today.

Under the early eighties, the fixed security market starts to take shape in Sweden, with the introduction of *certificates of deposits* issued by banks. The banks were able to loan money, from the general public, for a fixed interest and with minimal risk. In 1982, the first *zero-coupon bonds* were issued by the Swedish government, thus enabling them to loan money directly from companies and private investors instead of the banks. In 1983, the concept of zero-coupon bonds was then extended to other sectors of the society, municipalities started issuing *municipal bonds*, companies started issuing *commercial papers*, financial companies started issuing *market evidence* in other to finance their lending to households and companies, Swedish Housing Finance Institutions started issuing *mortgage bonds*. The first Swedish government *coupon bonds*, called *National Bond 1001* with maturity date 1987, were also issued by the Swedish government through the Swedish National Debt Office in 1983.

The development of the money market in Sweden can be attributed to three main factors, namely, the emergence of a securities market, deregulation of the financial system and the arrival of a derivatives market in Sweden. Most of the development occurs in the early eighties.

In this thesis, we fit Swedish government bonds with a well known parametric term structure model namely, the *Nelson Siegel model*. The Nelson-Siegel Model classes are very popular and widely used by practitioners and Central Banks. Their popularity is buried in the model's simplicity and ability to fit the various shapes and forms the term structures of interest rates can exhibit and providing economically and statistically meaningful and correct results.

The *Diebold and Li (2006)* two-step approach of the three-factor Nelson Siegel model will be applied to Swedish government bonds and we will show that this approach models well the yields on Swedish government bonds.

We now proceed by going through the theory on interest rates, followed by a brief description of linear and dynamic stationary processes. In the dynamic system, *state space* representation will be introduced and the *state equation* used to represent the dynamics of the model coefficients. The Nelson-Siegel model classes will be introduced and presented as factor models. Finally, the three-factor Nelson-Siegel model will be applied to Swedish government bonds. Using ordinary least squares, factor dynamics are obtained for each cross-section of the data. A time series model, namely *independent stable AR(1) processes*, is then fitted to the factors. The results obtained will be contrasted with the data and an *in-sample-fit* reported.

INTEREST RATES THEORY

In every well developed market economy, there is the need for households, firms and governments to finance loans, investments and budget deficits by borrowing money at a price. There is also the need to save excess income for future consumption, profits for future investments and budget surpluses for future government expenditures. The price demanded to lend out money or the cost incurred for taking a loan is referred to as *Interest rate*.

The need to borrow and lend money is satisfied in the *credit market*, where the actors acquire financial assets such as bank accounts, stocks, bonds, etc through financial intermediaries, by entering in a contract or an agreement. These financial assets gives the holder the *right/option* and the issuer the *obligation* to fulfill the contract/agreement at a pre-specified date stipulated in the contract.

We will start our journey to understanding interest rates by going through some basic interest rate theory.

We are all familiar with the saying that a dollar today is worth more than a dollar tomorrow. This is normally shown to be true by using a simple but powerful formula called *the present value formula*. This leads us to the discounting factor that depends on the *simple- or effective interest rate*, the *day-count convention* and the *time to maturity* of the future cash flow.

2.1 Discount factor

To enable us to make a decision on whether to invest on a financial/real asset or not, we need to know how much consumption we are forgoing today in order to increase our consumption in the future. This can be computed by discounting all the future cash flows of our investments with a *discounting factor*.

The discount factor helps to relate the present value of a future cash flow. It is a function of the future cash expected, time to maturity and the interest rate on the instrument in question. The bank account, also known as the *money market account* will be used to describe the discount factor.

The bank account describes how an amount of money deposited in a bank, with a fixed positive instantaneous rate r , grows over time. If we let B_t denote the value of a bank account at time ($t \geq 0$), then the bank account grows according to the

differential equation:

$$dB_t = rB_t dt \quad (2.1.1)$$

$$B_0 = 1 \quad (2.1.2)$$

where r is a positive function of time.

This implies that, the amount of money in the bank account at time t will grow to:

$$B_t = \exp\left(\int_0^t r ds\right) \quad (2.1.3)$$

The instantaneous interest rate r_t used in the discounting factor to obtain the present value of the money market account is composed of, amongst other things, the compensation for the lender for forgoing consumption today for future date, the compensation for increases in future price levels (i.e. inflation) as well as a risk premium for the lenders exposure to default risk. All these three components together is what is collectively referred to the *interest rate* on a loan or an *I.O.U* instrument.

Note: *A promise to pay a debt, especially a signed paper stating the specific amount owed and often bearing the letters I.O.U*

2.2 The Stochastic discount factor

Our goal in interest rate theory is to study and understand the variability of interest rates. We therefore allow the interest rates to be stochastic, as indeed, they are the underlying assets of interest.

When the spot rate in the discount factor is allowed to be stochastic, we thus obtain a *stochastic discount factor*.

The Stochastic discount factor $D(t, T)$, is basically used to relate amounts of money at two different time points t and T . It is the amount at time t that is equivalent to one unit of currency (*the nominal amount*) payable at time T , and is given by:

$$D(t, T) = \frac{B_t}{B_T} = \exp\left(\int_t^T r_s ds\right)$$

The stochastic discount factor leads us to the simplest form of loans in the money market, namely the *Zero-Coupon Bonds*. Zero-coupon bonds provides their holder with a deterministic amount, that is known when the bond is issued. Zero-coupon bonds are also referred to as *discount bonds* as they are traded at a discount to their face value. They are one of the main building blocks of interest rate theory.

2.3 Zero Coupon bonds

Zero-coupon bonds are the simplest form of loan in the credit market. They have only one payment stream under their whole lifetime, that is, the face value of the bond payable to the bond holder at maturity. However, zero-coupon bonds are in practice not directly observable in the market and long maturities zero-coupon bonds are not traded at all, due to for example *credit-risk*.

The price of a zero-coupon bond is obtained by discounting its nominal values with a stochastic discounting factor. The price is therefore a function of the *zero-coupon interest rate*, time to maturity and the face value of the bond.

The short rates are the rates that prevails in the money market and can directly be control by Central Banks policies. This can be done through *monetary policies*, given a variable foreign exchange rate policy, by *open market interventions* or through *repurchase agreement rates*. For example, in Sweden, increasing/decreasing the money supply lowers/increases the short rate.

The long rates are governed by the expected short-rates, inflation and the risk premium on the bond. The risk premium, which can be either positive or negative, is the expected returns on a bond with a fixed rate, in relation to a series of short-term investments.

Formally, a zero coupon bond that matures at time T (also called a T -bond) is defined as follows:

Definition 2.1. *A T -maturity zero-coupon bond is a contract that guarantees its holder the payment of one unit of currency (the nominal amount) at time T , with no intermediate payments. The contract value, i.e the price of a zero coupon bond at time $t < T$ is denoted by $P(t,T)$ and $P(T,T) = 1$ for all maturity times T is equal to the present value of the nominal amount which can be written as:*

$$P(t,T) = \frac{P(T,T)}{\left(1 + r \times \frac{d}{360}\right)^{T-t}} \quad (2.3.1)$$

where r is a deterministic interest rate and d refers to number of days remaining for the bond to mature. Note that r is also the internal rate of return on the bond, i.e. the bond yield.

The discount factor and the zero-coupon bond price are very much related, in the sense that they both give the present value of a future cash flow, but with a significant difference been, the price of a zero-coupon bond is a *value of a contract* and the discount factor is an *amount of money*.

Normally zero-coupon bonds are treated as risk free assets, due to the deterministic cash flow they provide its holder. They are therefore widely used in portfolio diversification and hedging. However, the cash flow from bonds in general is only deterministic if it is exercise at maturity, otherwise there are basically three types of risk associated with bonds namely, *interest rates risk*, *credit risk* and *inflation risk*.

As mentioned earlier, the price of a zero-coupon bond is a stochastic process with two random variables, t and T . If T is fixed, then the contract price, denoted by $P(t, T)$, will be a scalar stochastic process. It gives the prices at different times for a bond with fixed maturity T and its trajectory is very irregular.

On the other hand, if t is fixed, then the contract price is a smooth function of maturities, for $T > t$. Given a fixed t , a fundamental curve which can be obtained from observed interest rates is referred to as the zero-coupon curve or *the yield curve*.

2.3.1 The Yield Curve

The yield curve describes the relationship between the returns on bonds (i.e the yield) with the same credit-risk but with different maturities. It is widely used to compare yields offered for different maturities. The yield curve is also known as the *term structure of the discount factors*. Formally we can define the yield curve at time t as:

Definition 2.2. *The yield curve at time t , is the graph of the function:*

$$T \mapsto P(t, T), T > t \quad (2.3.2)$$

which decreases with maturity T .

In the Swedish bond market, bonds with maturities less than a year are quoted as *simple interest rates* and those with maturities more than a year are reported as *effective rates*. In constructing a yield curve for the whole tenor structure, we therefore need to convert the simple rates into effective rates in order to make sure that we are comparing rates with the same credit-risk. Theoretically, only effective rates for zero-coupon bonds guaranteed risk-free return if held until maturity.

Note, the term *yield curve* is often used to denote several different curves deduced from quotes from the interest-rate-market. That is to say, there exist other representations of the yield curve different from the one defined in definition 2.2. Thus, a yield curve that is a function both the short and long rate can be described as follows:

Definition 2.3. *The yield curve at time t is the graph of simply compounded interest rates for maturities T less than one year and the annually compounded interest rates for maturities more than a year.*

$$T \mapsto \begin{cases} L(t, T) & t < T \leq t + 1 \text{ (years)}, \\ A(t, T) & T > t + 1 \text{ (years)}. \end{cases} \quad (2.3.3)$$

where $L(t, T)$ and $A(t, T)$ are as described in 2.5.4 and 2.5.8 respectively.

The slope of the yield-curve varies over time and assumes different shapes and forms, which is a reflexion of the values of the short and the long yields. The

yield-curve slope is sometimes used as a future gross domestic product and inflation development indicator. Depending on its slope, the yield-curve can be classified into three main categories namely, normal-, inverted- or flat yield.

Normal Yield Curve: A yield curve is called normal if it has a positive slope with short-rates being lower than long-rates. This can occur when an economy is moving from a period of recession into a period of growth. A normal yield curve can be indicating higher inflation rates are expected which increases the long-rates.

Inverted Yield Curve: The inverted yield curve has a negative slope and it occurs if short-term yields are higher than long-term yield. An inverted yield curve is usually observed when an economy is at the top of its business cycle or under great stress. This signals the coming of a weak business cycle and hence lower inflation rates thereby causing the long rates to fall.

Flat Yield Curve: If the yield curve has a zero slope, then we have a flat yield curve. This occurs if there is little or no significant difference between short-term yields and long-term yield.

In trying to understand the different slopes, shapes and forms assumed by the yield-curve, economists have come up with different explanations and hypotheses, amongst them include, the *expectation hypothesis* and the *preferred habitat hypothesis*.

The Expectation Hypothesis: This is the most well known hypothesis used in trying to explain the theory behind the shape of the yield curve. It is based on three main assumptions:

1. All investors are *risk neutral*, in the sense that they lack preferences for bonds with certain maturities and therefore do not demand risk compensations on their investments. It is also assumed that investors in the bond markets have *rational expectations*.
2. The yield curve conveys all the information in the market, i.e. the expectation hypothesis assumes that the bond market is *effective*.
3. If there exist arbitrage between bonds with different maturities, investors can establish a yield curve that is consistent with how rates are expected to develop over time.

The assumptions made in the expectation hypothesis implies that future interest rates can be read from the yield curve, and the expected returns are the same and independent of the investment strategy we choose. For example, a one month investment in a three-year bond gives the same return as a one month investment on a five-year bond.

The Preferred Habitat Hypothesis: This can be seen as an extension of the expectation hypothesis, where investors prefer certain maturities to others and therefore demand compensation, in the form of a risk premium, for investments in bonds with maturities that deviate from their preferences. This hypothesis can be categorized into two, depending on the preference of the investor.

1. **Liquidity preference theorem** implies that investors prefer bonds with short maturities to bonds with longer maturities (i.e. consumption today is preferred to consumption tomorrow). Risk-averse investors therefore demand a risk premium for investing in bonds with long maturities and the longer the maturity, the more risk premium is demanded. Here, the yield-curve tends to have a more positive slope than that of the expectation hypothesis.
2. **The market segmentation theorem** assumes that bond investors and issuer by law, preference or habit invest and issue bonds with certain types of maturities. For example, banks normally prefer short maturity bonds whereas insurance companies and pension funds prefer long maturity bonds.
Note that the yield-curve under the market segmentation theorem cannot be used to analyze market expectations of future interest rates, since it is not directly based on the outcome market forces, supply and demand.

It is important to note that Central Banks can also affect the shape of the yield curve through the *confidence and trustworthiness* the market has on their ability to stabilize inflation. For example, if the actors in the market believe that the Central Bank can stabilize the inflation rates to an acceptable level, then the yield curve reflects the expected future long rates which in turn becomes less volatile compared to the short rates.

On the other hand, if the market does not believe in the ability of the Central Bank to control the inflation rate, then both the long and short rates will be highly volatile.

It is appropriate to now mention another type of bond that is very much similar to a zero-coupon bond, and well traded in the markets, called the *coupon bonds*. The main difference between a coupon bond and a zero-coupon bond is that the former pays coupons to the bearer at periods stipulated in the contract, which can either be annually (as in Sweden) or every six months (as in the USA).

2.4 Coupon Bonds

Zero-coupon bonds are not very well traded in the markets and those markets that do trade in them, trade in zero-coupon bonds with very short maturity (i.e. less than a year). What is traded in the market are *coupon bonds* with maturities more than a year. In addition to the face value of the bond, the bearer of a coupon bond

receives periodic coupon payments called the *the coupon rates*.

It is worth mentioning that the coupon rate is different from the *market rate*. The market rate is the rate at which bonds are traded at the present moment, in the bond market. The market rates varies overtime and can therefore directly affects the bond price and hence it market value. The coupon rates on the other hand, are a fixed percentage of the nominal amount and is stipulated in the contract.

Coupon bonds can be classified into two main categories, depending on whether they pay a *fixed coupon* rate or a *variable coupon* rate. The two categories are describe below:

2.4.1 Fixed Coupon Bonds

This is the simplest coupon bond. It is a bond, which for some intermediate points in time, will provide the holder with predetermined payments.

Formally, a fixed coupon bond that matures at time T can be described as follows:

Definition 2.4. *Given a tenor structure, T_0, T_1, \dots, T_n , a coupon bond is a contract entered at time T_0 that guarantees its holder the payment of the nominal amount of the bond at time T_n , with intermediate determined payments (called coupons) C_t at times T_1, \dots, T_n .*

The contract value, i.e the price of a coupon bond at time T_0 is denoted by $P(t, T)$ and $P(T, T) = 1$ for all maturity times T , is equal to the sum of the discounted future cash flows which can be express as:

$$p(t) = \sum_{t=1}^n \frac{C_t}{(1+r_t)^t} + \frac{P(T, T)}{(1+r_n)^n} \quad (2.4.1)$$

where r_t for $t = 1, 2, \dots, n$, is the effective interest rate.

The fixed coupon bonds can be replicated by holding a portfolio of zero coupon bonds with maturities T_i , for $i = 1, 2, \dots, n$. That is, we pay C_i zero coupon bonds of maturities T_i for $i = 1, 2, \dots, n - 1$ and $A + C_n$ bonds with maturity T_n . With this portfolio, for a time $t < T_1$, we can price the coupon bonds as:

$$p(t) = A \times P(t, T_n) + \sum_{i=1}^n C_i \times P(t, T_i) \quad (2.4.2)$$

The coupon bonds are quoted mostly in term of returns on the face value A over a given period $[T_{i-1}, T_i]$ and not in monetary terms. For example, given that the i th coupon has a return equal to r_i , this implies

$$c_i = r_i \times (T_i - T_{i-1}) \times A \quad (2.4.3)$$

If the interval lengths are equal (i.e. $T_i = T_0 + i\delta$) and the coupon rates for each interval is equal to a common rate r , then we have a **standardize coupon bond**.

The price $p(t)$, for $t < T_1$, of a standardized coupon bond is given by:

$$p(t) = A \times \left(P(t, T_n) + r\delta \sum_{i=1}^n P(t, T_i) \right) \quad (2.4.4)$$

As observed from the pricing formula of the coupon bond above, different payments are discounted with different effective rates. This makes it difficult to relate the price of the bond to a single effective rates. This problem can be overcome by quoting the bond with a single interest rate that is derived from the bond market price. This single rate, with which all future cash flows of the bond, both coupons and face value, is discounted and re-invested is referred to as *the yield to maturity*.

2.4.2 The Yield To Maturity

The yield to maturity gives the bonds *internal rate of return*, i.e. the return on investments that describes the present values of all future cash flows. Thus, the yield to maturity is the single effective interest rate that makes the price of the bond today to be equal to its market price.

Given the price $P(t, T)$, of a bond and its coupon payments C_t , the yield to maturity can be obtained from the pricing formula of a coupon bond by solving for y in the formula below:

$$P(t, T) = \sum_{t=1}^n \frac{C_t}{(1+y)^i} + \frac{P(T, T)}{(1+y)^n} \quad (2.4.5)$$

One should be very careful in how the yield to maturity is interpreted. This is because the yield to maturity implicitly assumes that the yield curve is completely flat, i.e. the interest rates are the same for the whole investment period which is not the case. Future interest rates are generally unknown and varies depending mainly on what stage in the business cycle an economy is in.

The yield to maturity does not provide a deterministic return since there is always a *re-investment risk* for each and every coupon payment.

The yield to maturity tells us nothing about the interest rates on other bonds, as it only consider the bond for which its internal rate of return is been computed.

There is also the risk that different yield to maturity for different bonds can be interpreted as a possibility to reinvest coupons at different rates, whereas the coupons are re-invested at the market rates.

There is a yield to maturity that is of particular interest to investor, i.e. the **par yield**. The par yield is the coupon rate which makes the bond price equal to its nominal value.

There exist a variety of coupon bonds in the market, with some bonds having physical assets as their underlying. These types of bonds are called **mortgage backed/asset backed securities**. The advantages of investing in these types of bonds is vested on the credit worthiness of the issuer and the market value of the underlying. There also exist bonds whose value is tied to a certain price index.

These bonds guard the holders against inflation and are called **inflation-linked bonds**.

However, we are only interest in loan instruments issued by governments and will therefore not consider the above mentioned bonds. Government bonds do offer coupons with rates that varies depending on a benchmark rate. Bonds with varying rates are generally referred to as *floating rate notes*.

2.4.3 Floating rate notes and Variable rate notes

There are numerous coupon bonds for which the value of the coupon is not fixed at the time the bond is issued, but rather reset for every coupon period. Mostly, but not always, the resetting is determined by some financial benchmark, for example, in Sweden, the three months **STIBOR rates** is normally used. One of the simplest floating rate bonds quotes the **STIBOR** rate plus/minus a certain amount of basis points. The amount of basis points demanded by investors is solely determined by the credit worthiness of the bond issuer.

The floating rate bonds can be replicated by using a self-financial bond strategy, with an initial cost $P(t, T_{i-1})$ at time t and reinvesting the amount receive at time T_{i-1} in bonds that matures at time T_i . Thus the price $p(t)$ for the floating rate bonds, given that the coupon dates are equally spaced (i.e. $T_i = T_0 + i\delta$) and assuming that the face value equals to one, for time $t < T_1$ is given by:

$$p(t) = P(t, T_n) + \sum_{i=1}^n [P(t, T_{i-1}) - P(t, T_i)] = P(t, T_0). \quad (2.4.6)$$

In particular, if $t = T_0$, then $p(T_0) = 1$.

It is important to note that some assumptions must be made to guarantee the existence of a market that is sufficiently rich and regular where these bonds are traded. That is, we have to assumed that there exists a frictionless market for zero-coupon bonds for every maturity time T and that the price of a zero-coupon bond $P(t, t) = 1$ for all times t . The assumption that $P(t, t) = 1$ is necessary to ensure that we avoid arbitrage pricing. We also have to assumed that for each fixed time $t < T$, the price of a zero-coupon bond that matures at time T is differentiable with respect to the time of maturity T .

Now we discuss two fundamental features of the interest rates. These features are the *Day-Count convention* and the *Compounding-Type*. The role of these features were briefly observed in the present value formula. Below, we discuss them in more detail.

2.5 The Day-Count Convention and The Compounding Types

The compounding types and the day-count convention are the two fundamental properties of interest rates that are needed to enable us to used zero-coupon bonds to price interest rates.

2.5.1 Day-Count Convention (year fraction)

The day-count convention is a fraction of a year that helps us to compute the interest payable and the end of an interest- or loan period. It tells us how interest on an investment grows over time. The numerator represents the number of days in the interest- or loan period and the denominator represents the number of days in the reference period. We denote by $\tau(t, T)$, the chosen time measure between t and T , which is usually referred to as year fraction between the issuing date t and maturity date T . When t and T are less than one-day distant, then $\tau(t, T)$ is to be interpreted as the time difference, i.e. $(T-t)$ in years. It is important to mention that there are various type of day count conventions used, depending on the market type, country and currency been used. However, for our purpose, we will used the Swedish convention, where the reference period is *360 days*.

2.5.2 Compounding Types

The compounding type refers to how the interest rate is computed based on both the initial principal amount invested and the interest generated in earlier periods. Basically, the compounding types can be classified into four main categories, namely *continuously-compounded rates*, *simply-compounded rates*, *k-times-per-year compounded rates* and *annually-compounded rates*. Of these four compounding types, the simply-compounding type, also called the **LIBOR rates**, is the most commonly used both in theory and in practice. The compounding types can be express as *forward rates* or *spot rates*. Below, a description of the above mentioned compounding types are given.

2.5.3 Continuously compounded interest rates

Basically, continuously compounding rate is the constant rate prevailing on an investment on a zero-coupon bond at time t , for maturity at a future time interval $[S, T]$, that yields a unit of currency at time of maturity.

If the contracting date coincides with the start of the interval, then we have a continuously-compounded spot rate, otherwise a continuously-compounded forward rate. The continuously compounded spot and forward rates are describe below:

The continuously compounded forward rate contracted at time t for the period $[S, T]$ is defined as

$$R(t; S, T) = -\frac{\log P(t, T) - \log P(t, S)}{\tau(T, S)} \quad (2.5.1)$$

The continuously compounded spot rate contracted at time S for the period $[S, T]$ is defined as:

$$R(S; S, T) = -\frac{\log P(S, T)}{\tau(T, S)} \quad (2.5.2)$$

The continuously compounded spot rate is a constant rate, from which we can derive the price of a zero coupon bond as

$$R(S; S, T) \times \tau(S, T) = -\log P(S, T) \quad (2.5.3)$$

$$\Rightarrow P(S, T) = \exp(R(S, T) \times \tau(S, T)) \quad (2.5.4)$$

2.5.4 Simply compounded rate

When accruing occurs proportionally to the time of the investment then we have a simply compounded spot rate. The simply compounded rate is also referred to as the *LIBOR* rates, $L(t, T)$. This is the rate used most commonly in the market.

The LIBOR Interest Rates

LIBOR stands for *London Interbank Offer Rate*. It is the average rate with which Banks on the London money market are prepared to borrow and lend money to each other. The *LIBOR rates* are quoted in ten different currencies and comes in fifteen different maturities. Changes in the *LIBOR rates* are closely monitored by all actors in the financial market, because it is generally used as the base rate by banks and other financial institutions. Thus a change in the *LIBOR rates* will impact saving accounts, mortgages, loans, etc.

The *LIBOR* rate prevailing at time t for the maturity T , is the constant rate at which an investment has to be made to produce one unit of currency at maturity, starting from $P(t, T)$ units of currency at time t , when accruing occurs proportional to the investment time.

LIBOR are forward rates that can either be quoted as continuously compounded rates or simple rates. The simple rates notation is the one most commonly used in the markets, whereas the continuously compounded notation is used for theoretical purposes. The simple *LIBOR* rate can be quoted as forward rates or spot rates as follows:

The simple forward rate contracted at time t for the period $[S, T]$ is called the *LIBOR* forward rate and is defined as

$$L(t; S, T) = -\frac{P(t, T) - P(t, S)}{\tau(T, S)P(t, T)} \quad (2.5.5)$$

The simple spot rate contracted at time S for the period $[S, T]$ is called the *LIBOR* spot rate and it is defined as:

$$L(S; S, T) = -\frac{P(S, T) - P(S, S)}{\tau(T, S)P(S, T)} \quad (2.5.6)$$

2.5.5 Annually-compounded spot interest rate

The annually compounded spot rate prevailing at time t for maturity T is denoted by $A(t, T)$, and is the constant rate at which an investment has to be made to produce an amount of one unit of currency at maturity, starting from $P(t, T)$ units of currency at time t , when reinvesting the obtained amount once a year.

$$A(t, T) = \frac{1}{[P(t, T)]^{\frac{1}{\tau(t, T)}}} - 1 \quad (2.5.7)$$

$$\Rightarrow P(t, T) = \frac{1}{[1 + A(t, T)]^{\tau(t, T)}} \quad (2.5.8)$$

2.5.6 k-times-per-year compounded spot interest rate

The k-times-per-year compounded spot interest rate prevailing at time t for the maturity T is denoted by $A^k(t, T)$ and is the constant rate at which an investment has to be made to produce an amount of one unit of currency at maturity, starting from $P(t, T)$ unit of currency at time t , when reinvesting the obtained amount k times a year.

$$A^k(t, T) = \frac{k}{[P(t, T)]^{\frac{1}{k\tau(t, T)}}} - k \quad (2.5.9)$$

$$\Rightarrow P(t, T) = \frac{1}{\left[1 + \frac{A^k(t, T)}{k}\right]^{k\tau(t, T)}} \quad (2.5.10)$$

Observe that the continuously compounded spot interest rate can be obtained as the limit of the k-times-per-year compounded rates, where the number of compounding times k going to infinity. Indeed

$$\lim_{k \rightarrow +\infty} \frac{k}{[P(t, T)]^{\frac{1}{k\tau(t, T)}}} - k = R(t, T)$$

All the spot rates mentioned above are equivalent in infinitesimal time intervals. Indeed

$$r(t) = \lim_{T \rightarrow t^+} R(t, T) \quad (2.5.11)$$

$$= \lim_{T \rightarrow t^+} L(t, T) \quad (2.5.12)$$

$$= \lim_{T \rightarrow t^+} A(t, T) \quad (2.5.13)$$

$$= \lim_{T \rightarrow t^+} A^k(t, T) \text{ for each } k \quad (2.5.14)$$

2.6 Interest rates assets

Derivatives that have interest rates as underlying assets, (i.e. interest rate forwards/futures, caps/caplets, floors/floorlets, swaps, swap contracts, etc), are used by many actors in the financial market for varying purposes. For example, speculators in the interest rate market uses interest rate derivatives to beat the markets returns . Arbitragers used a combination of interest rate derivatives in other to make a risk-free profit and hedgers used them in their portfolios in other to reduced risk.

We will now introduce forward- and future contract that were first used to help farmers hedge their agricultural produces against price-risk . In our case, we will study these contracts with interest rates, rather than agricultural produces, as underlying assets.

2.6.1 Forward Rate

The basic construction for interest rates assets is the forward rate and it is used to adjust for interest rate risk. Forward rates are interest rates that can be locked in today for an investment in a future time period, and are set consistently with the current yield of discount factors. Generally speaking, the holder of a forward contract has the obligation to buy or sell a certain product at a future date for a given price. It is a bilateral agreement between two actors, (for example, two banks or a bank and a company). The forward contract can be constructed in any way that suits the two parties involved. Forwards are traded *Over The Counter* (O.T.C), i.e. no exchange market is involved. The forward contract has only one cash flow that is paid at the maturity date.

The forward rate is characterized by three time instances, namely the contract date t , the date the contract is effective S and the exercise date T , where $t < S < T$.

A forward rate can be defined as:

Definition 2.5. *Given three fixed time points $t < S < T$, a contract at time t which allows an investment of a unit amount of currency at time S , and gives a risk less deterministic rate of interest over the future interval $[S, T]$ is called **The forward rate**.*

Note that spot interest rates are forward rates where the time of contracting coincides with the start of the interval over which the interest rate is effective (i.e. $t = S$).

Another way to define the forward rates is through a *Forward Rate Agreement*. By demanding that the forward rate agreement be fair, we have the forward rate.

2.6.2 Forward Rate Agreement (FRA)

Basically the *FRA* allows one to lock-in the unknown interest rate in a future time interval with a desired value that is agreed upon today. A forward rate agreement is therefore a contract that involves three time instances:

1. t : the time at which the rate is considered
2. S : maturity time
3. T : expire time

such that $t < S < T$. The contract gives its holder an interest-rate payment for the period between T and S . At the maturity time S , a fixed payment based on a fixed rate K is exchanged against a floating payment based on the spot rate *LIBOR* that resets in S with maturity T , i.e. $L[S, T]$.

At time T one receives $(\tau(S, T) \times K \times A)$ units of currency and pays the amount $(\tau(S, T) \times L(S, T) \times A)$, where A is the contract nominal value. The value of the contract, which can be both positive (when $K > L(S, T)$) or negative (when $K < L(S, T)$), is given by :

$$A \times \tau(S, T) \times [K - L(S, T)] \quad (2.6.1)$$

Thus the total value of the forward rate agreement at time t is given by:

$$\mathbf{FRA}(t, S, T, \tau(S, T), A, K) = A \times [P(S, T) \times \tau(S, T) \times K - P(t, S) - P(t, T)]$$

The forward rates can be classified into two main groups namely, *simply-compounded forward interest rate* and *Spot forward interest rate*.

Simply-compounded forward interest rate

The rate K , that renders the FRA arbitrage free at the contract date is obtained by equating the value of the forward rate agreement to zero. This value thus obtained is called the *simply-compounded forward interest rate*. It is the value of the fixed rate in a prototypical FRA with expiry S and maturity T that renders the FRA a fair contract at time t .

The simply-compounded forward rate can be viewed as an estimate of the future spot rate *LIBOR/STIBOR*. It is a random quantity, based on market condition at that time. It can be shown that, under a suitable probability measure, the simply compounded forward interest rate at time t is the expectation of the of the spot *LIBOR* at time t .

The simply compounded forward interest rate prevailing at time t for expiry $S > t$ and maturity $T > S$ is denoted by $F(t; S, T)$ and is defined by :

$$F(t; S, T) := \frac{1}{\tau(S, T)} \left(\frac{P(t, S)}{P(t, T)} - 1 \right)$$

Using the simply compounded forward interest rate, the forward rate agreement can be written as:

$$\mathbf{FRA}(t, S, T, \tau(S, T), A, K) = A \times P(t, T) \times \tau(S, T) \times [K - F(t; S, T)]$$

Therefore to value the FRA, we just need to replace the LIBOR rate $L(S,T)$ in the payoff function with the corresponding forward rate $F(t;S,T)$ and then take the present value of the resulting deterministic quantity. When the maturity of the forward rate collapses towards its expiry, we have the instantaneous forward rate.

Instantaneous forward interest rate

The instantaneous forward interest rates are fundamental quantities in the theory of interest rates. It is mostly used to express absences of arbitrage opportunity of an interest rate model by relating certain quantities in the expression of the evolution of the instantaneous forward interest rate.

The instantaneous forward interest rate prevailing at time t for the maturity $T > t$ is denoted by $f(t, T)$ and is defined as

$$f(t, T) := \lim_{T \rightarrow S^+} F(t; S, T) = -\frac{\partial \log P(t, T)}{\partial T}$$

so that we also have

$$P(t, T) = \exp\left(-\int_t^T f(t, u) du\right)$$

$f(t, T)$ can be seen as forward rates at time t with maturity T very close to expiry S .

A generalization of the *FRA* is the **Interest rate swaps**.

2.6.3 Interest rate Swaps

A swap contract, in principle, is an agreement between two parties to exchange a series of cash flows. An interest rate swap, for a given time interval, is therefore a contract in which the parties exchange a series of interest rate flows with each other. In most cases, one party exchange a variable rate with a fixed rate. Observed that, under the swap agreement, no principle amount is paid out even though the swap rate is computed based this amount.

The swap is considered to be the simplest interest rate contract. A swap contract in which one is able to exchange a payment stream at a fixed rate, the **swap rate**, for a payment steam at a floating rate (typically the three months LIBOR/STIBOR rate) over the future time intervals $[T_0, T_1], [T_1, T_2], \dots, [T_{n-1}, T_n]$ is called *plain vanilla swap* or the *coupon swap*.

The swap rate is chosen such that the value of the swap equals zero at the time when the contract is made and its given by:

$$\text{swap rate} = \frac{1 - P(T_0, T_n)}{\delta \times \sum_{i=1}^n P(T_0, T_i)} \quad (2.6.2)$$

For face value equal one, the price of the swap contract $\Pi^{\text{swap}}(t)$, for $t < T_0$, of a swap derivative can be written in terms of the sum of prices of **FRA** weighted by the interval lengths $[T_i, T_{i-1}]$, for $i = 1, 2, \dots, n$. This sum can be separated into

two parts, one that identifies the **floating leg** and one that identifies the **fixed leg** as.

$$\Pi_t^{swap}(t) = \sum_{i=1}^n \Pi_t^{FRA}(T_{i-1}, T_i) \quad (2.6.3)$$

$$= \sum_{i=1}^n \left(P(t, T_{i-1}) - P(t, T_i) \right) - K \times \sum_{i=1}^n (T_i - T_{i-1}) \times P(t, T_i) \quad (2.6.4)$$

The first sum in the second equality is the floating leg and its a telescopic sum, thus only the first and last terms will remain. This gives the swap price, Π_t^{swap} as:

$$\Pi_t^{swap}(t, T_i, T_{i-1}) = \sum_{i=1}^n (T_i - T_{i-1}) \times P(t, T_i) \times \left(\frac{P(t, T_0) - P(t, T_n)}{\sum_{i=i}^n (T_i - T_{i-1}) \times P(t, T_i)} - K \right) \quad (2.6.5)$$

Note that $\sum_{i=1}^n (T_i - T_{i-1}) \times P(t, T_i)$ is known as the *present value of basic points* (**PVBP**). The swap rate which makes this contract free to enter at time $t < T_0$ is

$$\frac{P(t, T_0) - P(t, T_n)}{\sum_{i=1}^n (T_i - T_{i-1}) \times P(t, T_i)}$$

This leads to two main interest rate assets, i.e. a *prototypical payer interest rate swap* **PFS** and a *prototypical receiver interest rate swap* **RFS**.

Prototypical payer and receiver interest rate swap

This is a contract that exchanges payments between two different index legs, starting from a future time. At every time, T_i in a respecified sets of dates $T_{\alpha+1}, \dots, T_{\beta}$ the fixed leg, also known as the coupon bearing bond, pays out the amount

$$A \times \tau_i \times K$$

corresponding to a fixed interest rate K , a nominal value A and a year fraction τ_i between T_{i-1} and T_i , whereas the floating leg, also known as the floating rate note, pays the amount

$$A \times \tau_i \times L(T_{i-1} - T_i)$$

corresponding to the *LIBOR/STIBOR* rates $L(T_{i-1} - T_i)$, resetting at the previous instant T_{i-1} . Clearly, the floating leg rate resets at dates $T_{\alpha}, T_{\alpha+1}, \dots, T_{\beta-1}$ and pays at dates $T_{\alpha+1}, T_{\alpha+1}, \dots, T_{\beta}$

When the fixed leg is paid and the floating leg is received, we have a **payer interest rate swap**, whereas when the fixed leg is received and the floating leg paid we have a **receiver interest rate swap**.

Requiring that the interest rate swap be arbitrage free at time t leads us to the forward swap rate that is defined as follows:

The Forward Swap Rate $S_{\alpha,\beta}$ at time t , for the sets of times $T_{\alpha+1}, T_{\alpha+1}, \dots, T_{\beta}$ and year fractions τ , is the rate in the fixed leg of the interest rate swap that makes the interest rate swap a fair contract at the present time. This implies that

$$S_{\alpha,\beta} = \frac{P(t, T_{\alpha}) - P(t, T_{\beta})}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t, T_i)}$$

2.6.4 Future Contract

A futures contract supports the main characteristics of a forward contract, with some significant differences. Futures are standardized contracts, that are traded in well organized markets, in respect to a delivery date and a given amount. The price on a futures contract is usually smaller than that of a forward, since the *credit-risk* of future contracts are much lower than that of a forward contracts. In a futures contract, there is no delivery of products, instead a cash settlement of profits/loses takes place. There is also a daily contract settlement in the futures market in order to compensate for price changes. For example, if the price of a future contract increases from one day to another, then the issuer of the contract is obliged to compensate the holder for the price differences.

The most commonly traded derivative products in the interest rate market are *caps/floors* and *swaptions*. The underlying assets for the caps/floor are the *forward LIBOR rates* and the underlying for the swaptions are *swap rates*. As mentioned earlier, there are four significant dates in a FRA, namely, the start and the end date of the accrual period, the payoff date and the fixing date. The complexity of pricing derivative products increases depending on how these date are set. It is therefore important to know the length of time to maturity of the underlying interest rate swap, to determine the set of reset and payment dates. The set of all maturity dates in the contract is often referred to as a *tenor structure*.

Definition 2.6. A *tenor structure*, also known as an *accrual factor*, is a set of maturities $\bar{T} = [T_0, T_1, \dots, T_n]$, where the first element of the set T_0 is typically start date and the last element T_n is the end date of the contract. The difference between two elements in the tenor structure is called the *tenor* and is denoted by:

$$\tau_i = T_i - T_{i-1} \text{ for } i \in [1, \dots, N] \quad (2.6.6)$$

2.6.5 Caps and Caplets

The cap is one of the most traded of all interest rates derivatives. Generally speaking, an interest rate cap is an agreement between two parties (normally a bank and a company), that if the interest rate on the *LIBOR/STIBOR* with a certain maturity for a given period in the future, exceeds a certain level, then the issuer of the cap should compensate the buyer. Interest rate caps are *financial insurance contracts*

which protects the holder from having to pay more than a pre-specified rate, the **cap rate**, even though a floating rate loan or *I.O.U* was taken. The cap rate is determined from the difference between interest rate stipulated in the contract and actual interest. Cap contract stipulates, the level of interest rate from which the holder wants protection from, the principle amount and the duration. The cap is defined as follows:

Definition 2.7. For a given tenor structure $T_\alpha, T_{\alpha+1}, \dots, T_\beta$, the holder of a **cap** has the option to receive the forward LIBOR/STIBOR rate $\tau_i \times L[T_{i-1}; T_{i-1}, T_i]$ and pays a predefined rate $\tau_i \times K$ at the date T_i for $(i = \alpha + 1, \dots, \beta)$. In other words the cap is a sum of caplets with payoff function

$$\Pi^{\text{caplets}}(L[T_{i-1}; T_{i-1}, T_i]) = \tau_i(L[T_{i-1}; T_{i-1}, T_i] - K)^+$$

where the payoff date is T_{i-1} .

Note that the caps relate to the LIBOR/STIBOR as a *European call* option relates to the stock if the forward LIBOR/STIBOR rates increases significantly.

2.6.6 Floor and Floorlets

Interest rate floors are *financial insurance contract* which guarantee that the interest paid on a floating rate loan will never be below some predetermined level called the **floor rate**.

Definition 2.8. For a tenor structure $(T_\alpha, T_{\alpha+1}, \dots, T_\beta)$ the holder of a **floor** has the option to pay the forward LIBOR/STIBOR rate $\tau_i \times L[T_{i-1}; T_{i-1}, T_i]$ and receive the predefined rate $\tau_i \times K$ at the date T_i for $(i = \alpha + 1, \dots, \beta)$. In other words the floor is a sum of floorlets with payoff function

$$\Pi^{\text{floorlet}}(L[T_{i-1}; T_{i-1}, T_i]) = \tau_i(K - L[T_{i-1}; T_{i-1}, T_i])^+$$

where the payoff date is T_{i-1} .

Note also here that the floor relate to the LIBOR/STIBOR as a *European put* option relates to the stock.

2.7 Swaptions

The *swap options*, or as it is most commonly known the *swaptions*, are options with the swap rates as underlying instruments. Since the swap-rates can be characterized into two different legs, depending on which party receives or pays the fixed leg, we have two main types of swaptions, *the payer version* and *the receiver version*.

A European **payer swaptions** in an option that gives the holder the right, and not the obligation, to enter a payer interest rate swap at a given future time, the swaptions maturity, which usually coincides with the first reset date of the

underlying interest rate swap. The payer-swaption payoff, discounted from the maturity T_α to the current time is given by:

$$N \times D(t, T_\alpha) \times \left(\sum_{i=\alpha+1}^{\beta} P(T_\alpha, T_i) \times \tau_i \times (F(T_\alpha; T_{i-1}, T_i) - K) \right)^+ \quad (2.7.1)$$

It is important to note that the above payoff cannot be decomposed into its elementary products. This is the main difference between the caps/floors and the swaptions. This means that in order to value and manage swaptions contracts, we need to consider the joint action of the rates involved in the payoff, whereas in the cap/floor cases we can consider each caplet/floorlet on their own first and then put the results together to obtain the cap/floor. By considering the joint actions of the rates in the swaption, this may induce correlations between the rates. It is also important to note that a payer swaption also has a value less than the value of a corresponding cap contract, i.e.

$$\left(\sum_{i=\alpha+1}^{\beta} P(T_\alpha, T_i) \times \tau_i \times (F(T_\alpha; T_{i-1}, T_i) - K) \right)^+ \leq \sum_{i=1+\alpha}^{\beta} P(T_\alpha, T_i) \times \tau_i \times (F(T_\alpha; T_{i-1}, T_i) - K)^+$$

LINEAR SYSTEMS

3.1 Linear processes

Linear systems are used in modeling physical phenomenon as a realization of a *stochastic process*.

A stochastic process is a collection of random variables that is often used to represent the evolution of some random value, or system, over time.

Formally, a stochastic process is defined:

Definition 3.1. *Given a sample space Ω , a stochastic process defined on Ω with parameter space T is a family of random variables.*

$$\{X(t, \omega), t \in T, \omega \in \Omega\}$$

Two of the most important and commonly used linear processes that forms the basic elements in both linear and non-linear time series modeling are the *Moving Average* (MA) and the *Auto-Regressive*(AR) processes. They enable model identification and parameter estimations in time series data in a very simple and efficient way. For instance, the AR-process, which is a *Markov process*, is generated by passing a *white noise* through a recursive filter.

Before giving a brief description of these two building blocks, and their generalization, *i.e. the ARMA process*, we would first define *white noise*, which generate these processes, in both discrete and continuous time.

3.2 White Noise

Generally speaking, a white noise process is a random process which is generated by mutually uncorrelated zero mean random variables. In discrete time, white noise is defined as:

Definition 3.2. (*Discrete White Noise*)

If a sequence $\{e_t\}$ of uncorrelated random variables has zero mean and variance σ^2 ,

i.e.

$$E[e_t] = 0, \quad (3.2.1)$$

$$\text{Cov}[e_s, e_t] = \begin{cases} \sigma^2 & \text{if } s = t \\ 0 & \text{otherwise} \end{cases} \quad (3.2.2)$$

then the sequence $\{e_t\}$ is called a *white noise* or **innovation sequence**.

It is not as easy as in the discrete case, to define continuous white noise processes, however the formal definition given below is generally accepted.

Definition 3.3. (*Continuous White Noise*)

Continuous white noise is formally defined as a generalized process $\{e_t\}$ with autocovariance function

$$\gamma(\tau) = \sigma_e^2 \delta(\tau) \quad (3.2.3)$$

where $\delta(\tau)$ is the Dirac delta function defined as:

$$\delta(\tau) = \begin{cases} 1 & \text{for } \tau = 0 \\ 0 & \text{otherwise} \end{cases}$$

3.2.1 Moving Averages (MA)

A *MA-process* is characterized by the fact that its weights, which are rational functions of polynomials are zero from a given point. This given point gives the order of the process. Generally speaking, an MA-process is defined as

Definition 3.4. *The process X_t given by*

$$X_t = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}, \quad (3.2.4)$$

where e_t is white noise, is called a *Moving Average Process of order q* . An MA-process is **always stationary**.

3.2.2 Autoregressive processes AR

An *AR-process* is characterized by the fact that its weights, which are rational functions of polynomials are zero from a given point. The given point gives the order of the process. The process is called *autoregressive* because the value of the process at time t (i.e. X_t), can be seen as a regression on past values of the process. Formally, an AR-process is defined as

Definition 3.5. *The process X_t given by*

$$X_t + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} = e_t \quad (3.2.5)$$

where e_t is white noise, is called an *Autoregressive Process of order p* . An AR-process is **always invertible**.

A generalization of the *AR-process* can be made by letting the right hand side of equation (3.2.5) be an *MA-process*. The same procedure can be used to generalize the *MA-process* by letting the left hand side of equation (3.2.4) be an *AR-process*. In other words, a natural generalization of the two models is to combine them, which gives us an *ARMA-process*.

3.2.3 ARMA processes

An autoregressive moving average is obtained by passing a white noise through a recursive filter. Most non-linear models are a direct extension of an *ARMA-model*. This shows its strength in modeling both linear/non-linear and stationary/non-stationary time series data, thereby justifying its popularity.

Definition 3.6. *The process X_t given by*

$$X_t + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} = e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} \quad (3.2.6)$$

where e_t is white noise, is called an *Autoregressive Moving Average process of order p, q i.e. ARMA(p, q)*. An *ARMA-process* can be written in transfer function form from X_t to e_t as:

$$X_t = H(z^{-1})e_t \quad (3.2.7)$$

where $H(z^{-1}) = \frac{\theta(z^{-1})}{\phi(z^{-1})}$

An *ARMA(p, q) process is **stationary** if the roots of $\phi(Z^{-1})$ lies within the unit circle and **invertible** if the roots of $\theta(z^{-1})$ lies within the unit circle.*

The auto-covariance function for an *ARMA(p, q)-process* satisfies the linear difference equation

$$\gamma(k) + \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p) = \theta_k \gamma_{\epsilon X}(0) + \dots + \theta_q \gamma_{\epsilon X}(q-k) \text{ for } k = 0, 1, 2, \dots \quad (3.2.8)$$

If $p > q$, then we have

$$\gamma(k) + \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p) = 0 \text{ for } k = p, p+1, \dots \quad (3.2.9)$$

and if $p < q$ we obtain

$$\gamma(q+1) + \phi_1 \gamma(q) + \dots + \phi_p \gamma(q+1-p) = 0 \quad (3.2.10)$$

DYNAMIC MODELS

4.1 Vector Autoregressive Process (VAR)

We begin this chapter by discussing an extension of the univariate *AR-process* called the **Vector Autoregressive process (VAR)**. For modeling multivariate time series data, the *VAR* model is one of the most popular and easy to used model for describing dynamic behaviors, as well as, providing meaningful forecast of economic and financial time series data. Formally, a *VAR-process* is defined as:

Definition 4.1. Let $X_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ be a set of $(n \times 1)$ observations. Then a *VAR(p)* process is defined as:

$$X_t + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} \dots + \Phi_p X_{t-p} = e_t \quad (4.1.1)$$

where Φ_i are $(n \times n)$ coefficient matrices and e_t is an $(n \times 1)$ zero mean white noise vector process with constant covariance matrix Σ_e .

It is important to observed that the *VAR(·)* process can sometimes be too restrictive to properly represent the main characteristics of the data. Therefore, additional deterministic terms (such as linear trends) might be needed to represent the data. Moreover, external variables may also be added to the *VAR(·)* process for the data representation to be proper. Then equation(4.1.1) is therefore generalized as

Definition 4.2.

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \Pi D_t + G Y_t + e_t \quad (4.1.2)$$

where D_t is an $(l \times 1)$ deterministic matrix, Y_t represents the $(m \times 1)$ matrix of external variables and Π and G are parameter matrices.

4.2 State Space Modeling

As in most modeling procedure, the main aim is to find an appropriate way to model the relationship between input and output signals of a system. These types

of models focuses only on the *external description of the system*. To gain some insight on the *internal state* of the system under study, we resort to state space modeling. This is obtained by defining a state vector such that the dynamics of the system can be described by a *Markov process*.

A state space model is formulated in discrete time by using a (multivariate) difference equation or in continuous time by a (multivariate) differential equation. The state space representation describes the dynamics of the state vector X_t , and a static relation between the state vector and the (multivariate) observation Y_t . Thus a linear state space model consist of two sets of equations, that is *the system equation*

$$X_t = A_t X_{t-1} + B_t u_{t-1} + e_{1,t} \quad (4.2.1)$$

and the *observation equation*

$$Y_t = C X_t + e_{2,t} \quad (4.2.2)$$

where X_t is the N-dimensional random state vector that is not directly observable. u_t is a deterministic input vector, Y_t is a vector of observable stochastic output, and A_t , B_t , and C_t are deterministic matrices in which the parameters are embedded. Finally the processes $e_{1,t}$ and $e_{2,t}$ are uncorrelated white noise processes. Thus the system equation describes the evolution of the system states whereas the observation equation describes what can be directly measured.

For linear time systems, in which the system noise $e_{1,t}$ and the measurement noise $e_{2,t}$ are taken to be Gaussian with zero mean, the *Kalman filter* is used to estimate the hidden state vector and also for providing predictions. The Kalman filter is described below.

4.2.1 The Kalman Filter

For linear dynamic systems, the Kalman filter provides the optimal prediction and reconstruction of the latent state vector. The foundation of the Kalman filter is based of the *linear projection theorem* that is stated below:

Theorem 4.1. Let $\mathbf{Y} = (Y_1, \dots, Y_m)^T$ and $\mathbf{X} = (X_1, \dots, X_m)^T$ be random vectors, and let the $(m + n)$ -dimensional vector $(\mathbf{Y}, \mathbf{X})^T$ have the mean

$$\begin{pmatrix} \mu_Y \\ \mu_X \end{pmatrix} \text{ and covariance } \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix}$$

Define the linear projection of \mathbf{Y} on \mathbf{X}

$$E[\mathbf{Y}|\mathbf{X}] = \mathbf{a} + \mathbf{B}\mathbf{X} \quad (4.2.3)$$

Then the projection and the variance of the projection error is given by

$$E(\mathbf{Y}|\mathbf{X}) = \mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(\mathbf{X} - \mu_X) \quad (4.2.4)$$

$$\text{Var}(E(\mathbf{Y}|\mathbf{X})) = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{YX}^T \quad (4.2.5)$$

Finally, the projection error, $\mathbf{Y} - E(\mathbf{Y}|\mathbf{X})$, and \mathbf{X} are uncorrelated, i.e.

$$C(\mathbf{Y} - E(\mathbf{Y}|\mathbf{X}), \mathbf{X}) = 0 \quad (4.2.6)$$

From the above proposition, the Kalman filter equations for reconstructing, updating and predicting the latent states are generated. Since $e_{1,t}$ and $e_{2,t}$ in the state space model are assumed to be normally distributed, then $X_t|Y_t$ is also normally distributed and are thus completely characterized by its mean

$$\hat{X}_{t|t} = E(X_t|Y_t) \quad (4.2.7)$$

and variance

$$\Sigma_{t|t}^{xx} = Var(X_t|Y_t) \quad (4.2.8)$$

The optimal linear reconstruction of the states, which in linear time invariant systems is given by the Kalman filter, is obtained from

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Y_t - C\hat{X}_{t|t-1}) \quad (4.2.9)$$

and the variance of the reconstruction is given by

$$\Sigma_{t|t}^{xx} = \Sigma_{t|t-1}^{xx} - K_t \Sigma_{t|t-1}^{yy} K_t^T \quad (4.2.10)$$

$$= \Sigma_{t|t-1}^{xx} - K_t C \Sigma_{t|t-1}^{xx} \quad (4.2.11)$$

where the **Kalman gain** at time t K_t , is given by

$$K_t = \Sigma_{t|t-1}^{xx} C_t^T [C_t \Sigma_{t|t-1}^{xx} C_t^T + \Sigma_{2,t}]^{-1} \quad (4.2.12)$$

Now the one-step predictions is given by:

$$X_{t+1|t} = A_t \hat{X}_{t|t} + B_t u_t \quad (4.2.13)$$

$$\Sigma_{t+1|t}^{xx} = A_t \Sigma_{t|t}^{xx} A_t^T + \Sigma_{1,t} \quad (4.2.14)$$

$$Y_{t+1|t} = C_t \hat{X}_{t+1|t} \quad (4.2.15)$$

As can be observed, the Kalman filter is a recursive filter and thus initial conditions

$$\hat{X}_{1|0} = E(X) = \mu_0 \quad (4.2.16)$$

$$\Sigma_{1|0}^{xx} = Var(X_t) = \Sigma_0 \quad (4.2.17)$$

are needed. Finally, the innovation (i.e. the measurement error) is given by

$$\hat{Y}_{t+1|t} = Y_{t+1} - \hat{Y}_{t+1|t} \quad (4.2.18)$$

and its variance R_{t+1} is computed as

$$R_{t+1} = Var(\hat{Y}_{t+1|t}) \quad (4.2.19)$$

$$= \Sigma_{t+1|t}^{yy} \quad (4.2.20)$$

$$= C_t \Sigma_{t+1|t}^{xx} C_t^T + \Sigma_{2,t} \quad (4.2.21)$$

If the assumptions of normality and linearity are no longer valid, then the Kalman filter will no longer be valid for parameter estimations and state prediction. We therefore have to resort to non-linear state space modeling, which in a sense are approximative filters. The *Extended Kalman Filter* **EKF** and *Particle filter* are some of these non-linear filters that can be used for parameter estimations.

THE NELSON-SIEGEL MODELS

The need to accurately estimate and forecast the relationship that exist between interest rates and time to maturity (i.e. the term structure of interest rates) has lead to numerous research that produces a variety of model classes, of which the Nelson Siegel model classes are amongst the most popular ones.

The Nelson Siegel model classes are parametric models that smoothly fits the term structure of interest rates. They provides the desired flexibility that is needed to fit the varying shapes and forms the term structures of interest rates exhibits. The Nelson Siegel model, and its various extensions and restriction for the matter, has been shown to perform well in providing forecasts that are both statistically accurate and economically meaningful, especially for long forecast horizons.

The simplicity, flexibility, accuracy and economically meaningful result that is obtained from the Nelson Siegel model classes makes them one of the most widely used parametric model classes in estimating and forecasting the term structure of interest rates. Their popularity spreads to practitioners and central banks world-wide.

The above mentioned qualities of the model makes it easy to extend its applications to areas outside of term structure model estimation and forecasting. For example, the Nelson Siegel model was used by Diebold, Rudebusch, and Aruoba (2006 b) to study the interactions between the yield curve and the macro economy.

Nelson and Siegel (1987) suggested to fit the forward rate curve at a given point with approximating functions that consist of the product between a polynomial and an exponentially decaying term. By averaging over these forward rates, the Nelson Siegel Spot rate curve is obtained. Below we describe the Nelson Siegel model classes and its various extensions and restriction as factor models.

5.1 Nelson-Siegel Model Classes

It is important to note that all the classes of Nelson-Siegel models that are considered below are based on the assumption that all *bond prices are arbitrage free*. In other words, we assumed that, given the bond prices, there exist no opportunities to make a risk free profit.

The first model that will be discussed, namely *the three-factor model*, is considered to be the base model, from which all other variations are derived.

Dieblod and Li (2006) proposed modeling the dynamics of the model coefficients, in the Nelson Siegel model classes, and interpreted them as factors. They called the factors *level*, *slope* and *curvature* and implied that the term structure of interest rates, for a given maturity, can be seen as a sum of these different components. They named the components *the long-term*, *the short-term* and *medium-term* component.

5.1.1 Three-factor base model

In its original form, Nelson-Siegel model fits the yield-curve y at any point in time point in time τ with the simple functional form below, which is equation(2) in *Nelson and Siegel (1987)*

$$y(\tau) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} \right) + \beta_2 \left(\frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{-\frac{\tau}{\lambda}} \right)$$

where $y(\tau)$ is the spot-rate curve with τ denoting the time to maturity, and β_0 , β_1 and β_2 are model parameters, which in dynamic form are referred to as *level*, *slope* and *curvature* and λ is referred to as the *decay* or *shape* parameter. *Diebold and Li (2006)* provided a dynamic representation by replacing the parameters in the Nelson-Siegel model with time varying factors. This leads us to a *dynamic Nelson-Siegel model* given by:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} \right) + C_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} - e^{-\frac{\tau}{\lambda_t}} \right)$$

The Nelson Siegel model represents the spot-rate curve with only four parameters. However, it is still able to provide a good fit to the cross section of yields at a given point in time. The model factors are loaded by components called factor loadings, that have a clear interpretation base on their contribution to the term structure. The model's factor loadings are referred to *long-*, *short-* and *medium-term* component respectively.

The long term component, is the factor loading on the level factor. This component is a constant one and is the same for all maturities. The long term component governs the level of the term structure. Generally, it is defined as the interest rates with the longest maturities. Thus, an increase in the level factor increases all yield equally, as the factor loading on it is identical for all maturities.

The factor loading on the slope of the term structure is referred to as the short-term component. The short-term component starts at one and decays exponentially

to zero. The rate at which the short-term components decays from one is determined by the size of the *shape parameter* λ_t . Thus large values of λ_t induces a slower decay as small values induces faster decay.

The factor loading on the curvature factor is called the medium-term component. The medium-term component starts at zero, increases for medium maturities to a maximum that is determined by the shape parameter λ_t and then decays back to zero. The medium-term component gives the Nelson-Siegel model its flexibility by enabling the model to capture most of the varying shapes that the term structures of interest rates exhibits.

From the three factor model described above, we observed that both the slope and curvature factor loadings are governed by the same decaying parameter, i.e. λ_t . This is indeed a restriction to the model. It is worth mentioning that Nelson and Siegel (1987) tried lifting this restriction and concluded that the unrestricted model was over parameterized. This was shown by *Bliss (1997)* to be due to the fact that Nelson and Siegel used bonds with maturities less than one year. Bliss shows that, when bonds with longer maturities are used there is no over-parameterization issue. The model that follows from Bliss observation is discuss below.

The Three-Factor Model (Bliss (1997))

Bliss estimates the term structures of interest rates by lifting the restriction, in the three-factor model, namely, the slope and the curvature factor loadings been governed by the same decay parameter λ_t . Bliss therefore allows for the slope and curvature to decay at a different rates.

Bliss three-factor model in its dynamic form can be represented as:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_{2,t}}}}{\frac{\tau}{\lambda_{2,t}}} - e^{-\frac{\tau}{\lambda_{2,t}}} \right)$$

where $y(\tau)$ is the spot-rate curve with τ denoting the time to maturity. L_t , S_t and C_t are model factors representing the level, slope and curvature of the spot-rate curve. $\lambda_{1,t}$ governs the rate of decay of the slope whereas $\lambda_{2,t}$ governs the rate of growth and decay of the curvature factor.

Observed that the Bliss model is the same as the three-factor model if $\lambda_{1,t}$ is equals to $\lambda_{2,t}$.

Even though the three-factor model has appealing properties and it is very widely used, it still does not capture all the different shapes and forms the term structures on interest rates can assumed, especially those with multiple peaks/deeps. Thus necessitating various extensions and modifications of the base model. Below we consider some of these models and we start by describing the two factor model. The two factor model is based on the idea that the first two principal components explains most of the variations in interest rates.

5.1.2 Two-Factor Model

The two-factor model is a restriction rather than an extension of the three-factor model. The restriction is motivated by studies (see Litterman and Scheinkman (1991)) that shows that interest rates variations can be explained by its first three principal component factors. The three principal components can also be interpreted as level, slope and curvature according to how they affects the term structures of interest rates, and hence treated as a factor model.

However, the first and second principal components (i.e. the level and the slope) captures most of the variance in the interest rates while the third principal component (the curvature factor) contribution to the variance is negligible. The third principal component can therefore be ignored thereby leaving us with a two factor model.

The two-factor Nelson Siegel model represents the spot-rate curve in its dynamic form as:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t} \right)$$

where $y(\tau)$ is the spot-rate curve with τ time to maturity.

The argument that the first two principal components explains all the variations in interest rates and thus a two-factor model can efficiently be used to forecast the term structure of interest rates, was put forward by Diebold, Piazzesi, and Rudebusch (2005). Diebold, Piazzesi and Rudebusch in the same paper also pointed out that the two factor-model might be inadequate to fit the entire term structure.

An extension of the Nelson-Siegel three-factor model by adding a second curvature factor with a separate decay parameter, was first introduced by Svensson (1994). The additional curvature gives the model more flexibility and hence a better in-sample-fit.

Björk and Christensen (1999) also extended the Nelson-Siegel three-factor model by adding a second slope factor that decays at a much faster rate than the slope factor in the original three-factor model. These extensions, in their dynamic representations are referred to as four-factor models.

5.1.3 Four-Factor Model

The four-factor model adds an additional factor to the Nelson-Siegel three-factor model. The additional factor can either be a second curvature- (as in Svensson (1994)), or second slope factor (as in Björk and Christensen (1999)). The extra factor gives the model a greater flexibility, thus enabling it to capture the various shapes and multi-modalities the term structures of interest rates can attain.

We start by studying one of the most popular four-factor Nelson Siegel term structure model, i.e. the extension by Svensson (1994).

Nelson Siegel Svensson Model

Svensson (1994) extended the Nelson Siegel model by adding a second curvature factor with a factor loading component that has a different decay parameter. This model is often referred to as *The Nelson-Siegel-Svensson Model*. The Nelson-Siegel-Svensson model is one of the most popular models used by central banks all around the world to model, estimate and forecast the term structures of interest rates. Svensson pointed out that, the additional curvature term will increase the flexibility and in-sample fit of the original Nelson Siegel model.

The Nelson-Siegel-Svensson model for the spot-rate curve, in its dynamic form, is given by:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_{1,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_{2,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{2,t}}} - e^{-\frac{\tau}{\lambda_{2,t}}}}{\frac{\tau}{\lambda_{2,t}}} \right)$$

where $y(\tau)$ is the spot-rate curve with τ time to maturity. L_t , S_t , $C_{1,t}$ and $C_{2,t}$ are the factors and $\lambda_{1,t}$ governs the rate of decay of the slope and the first curvature factor loading components while $\lambda_{2,t}$ governs the rate of growth and decay of the second curvature factor loading component.

Observed that the added component mainly affects medium-term maturities. The introduction of an additional medium-term component makes it possible and more easy to fit interest rates term structures shapes with more than one local maximum or minimum along various maturities.

The main drawback of the Nelson-Siegel-Svensson is that it is highly nonlinear, which makes the estimation of the model parameters very difficult.

Another weakness in Svensson extension is what is called the *multicollinearity problem*. In general, multicollinearity arises when two or more parameters cannot be identify separately because there exist a high level of dependence between them. Multicollinearity in the Nelson-Siegel-Svensson model occurs when $\lambda_{1,t}$ and $\lambda_{2,t}$ assumes similar values, which makes it impossible to identify $C_{1,t}$ and $C_{2,t}$ separately.

One way to overcome the collinearity problem is to make sure that the medium-term components are different when $\lambda_{1,t}$ is approximately equals to $\lambda_{2,t}$. This leads us to the *Adjusted Svensson* model proposed by Michiel De Pooter (June 5, 2007).

Adjusted Nelson-Siegel-Svensson Model

Michiel De Pooter (2007) proposed scaling the second curvature factor in the Nelson-Siegel-Svensson model so as to eliminate its multicollinearity problem. The adjusted model have the same limiting properties as the Svensson extension, but the second curvature factor loading decays at a much faster rate than the first curvature factor loading. To achieved this faster decay rate, Michiel De Pooter(2007) increases the time to maturity in the numerator of the second curvature factor loading by a factor of two.

In its dynamic form, the *Adjusted Nelson-Siegel-Svensson model* represents the spot-rate curve with time to maturity τ as below:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_{1,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}} \right) + C_{2,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{2,t}}}}{\frac{\tau}{\lambda_{2,t}}} - e^{-\frac{2\tau}{\lambda_{2,t}}} \right)$$

Björk and Christensen added a second slope factor to the three-factor Nelson-Siegel model that decays at a much faster rate than the slope factor in the original three-factor model. This was achieved by doubling the time to maturity in the second slope factor loading by a factor of two. Note that the time to maturity in both the numerator and the denominator are increased by this factor.

Björk and Christensen (1999) Four-Factor Model

Björk and Christensen (1999) proposed an extension of the three-factor Nelson-Siegel model to increase its flexibility by adding a fourth factor that mainly affects interest rates with short-term maturities. This additional factor resembles the slope component of the three-factor Nelson-Siegel model. Thus, the slope of the term structure of interest rates is now captured by a weighted sum.

Björk and Christensen model the spot rate curve, in its dynamic form, as:

$$y(\tau) = L_t + S_{1,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}} \right) + S_{2,t} \left(\frac{1 - e^{-\frac{2\tau}{\lambda_{2,t}}}}{\frac{2\tau}{\lambda_{2,t}}} \right)$$

It is worth mentioning the comment made by *Aruoba (2006 b)* on the in-sample fit of the three- and four-factor models. Aruoba reported that the improvement in the in-sample fit gain from the four factor models is not that much significant. This goes a long way to show that the simplicity in the three-factor model hides within it a power to model, estimate and forecast the term structures of interest rates that is hard to beat.

Björk and Christensen (1999) Five-Factor Model

Björk and Christensen (1999) also suggested a five-factor model. The five-factor model is similar to a combination of the Svensson extension and Björk and Christensen (1999) extension. In other words, they add both a slope and curvature to the original Nelson-Siegel model. The five-factor model was shown by Diebold, Rudebusch and Aruoba (2006 b) to have negligible effect on the in-sample fit of the three-factor model. Below we state how the five-factor model dynamics of the spot rate curve .

$$y(\tau) = L_t + S_t \left(\frac{\tau}{\lambda_{1,t}} \right) + C_{1,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} \right) + C_{2,t} \left(\frac{1 - e^{-\frac{\tau}{\lambda_{1,t}}}}{\frac{\tau}{\lambda_{1,t}}} - e^{-\frac{\tau}{\lambda_{1,t}}} \right) + S_{2,t} \left(\frac{1 - e^{-\frac{2\tau}{\lambda_{2,t}}}}{\frac{2\tau}{\lambda_{2,t}}} \right)$$

APPLICATION OF THE NELSON SIEGEL MODEL

The three factor Nelson Siegel model is used to model, fit and estimate the term structure of Swedish government bonds, for both short- and long maturities. As mentioned in chapter 5, the Nelson Siegel three-factor model, in its dynamic form is given by:

$$y(\tau) = L_t + S_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} \right) + C_t \left(\frac{1 - e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} - e^{-\frac{\tau}{\lambda_t}} \right)$$

where $y(\tau)$ is the spot-rate curve with τ denoting the time to maturity. L_t , S_t and C_t are model factors representing the level, slope and curvature of the spot-rate curve. The exponential decay parameter is denoted as λ_t .

6.1 Intuition Behind The Three Factor Nelson Siegel Model

In order to understand intuitively what the Nelson-Siegel factor loadings look like, and why the factors they load are called "Level", "Slope", "Curvature", I plotted the factor loadings as a functions of maturity as shown in Figure 6.1. The maturities chosen here have nothing to do with the tenor structure associated with the data and the the decay parameter is fixed in order to remove the nonlinearity in the model. My intension here is just to show how the various parameters affects the term structure of the Swedish government yields.

Observe that the decay parameter λ governs the exponential rate of growth of the slope factor as well as the rate of growth and decay of the curvature factor. Thus small λ produces slow decay and can better fit the yield curve at long maturities, whereas large λ produces fast decay and can therefore fit yields with maturities less than a year better.

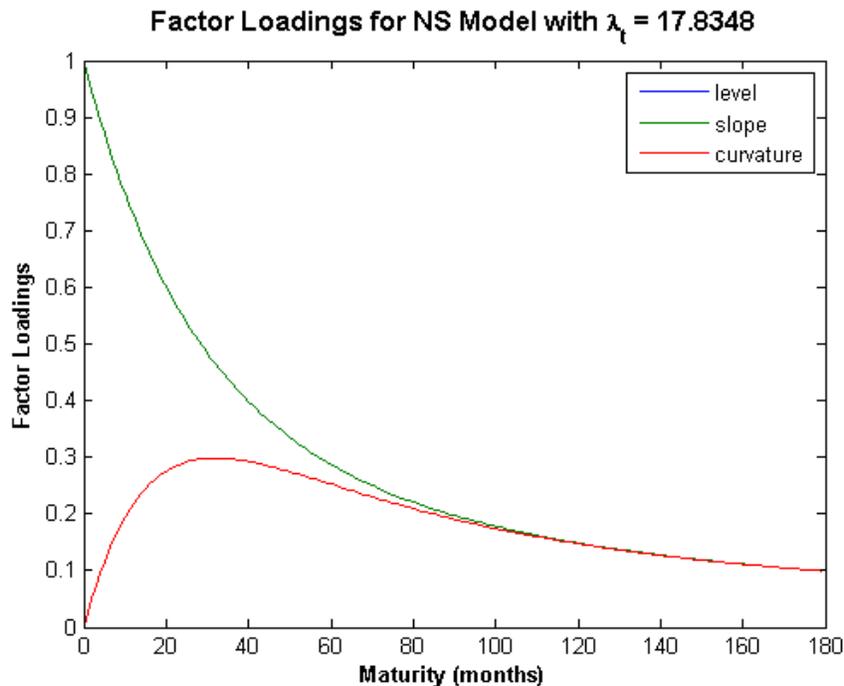


Figure 6.1. The graph shows the factor loadings of the three-factor Nelson Siegel model with the decay parameter fixed at $\lambda = 17.8348$. The level factor loading is a constant one and its unaffected by λ . The slope factor starts at one and decays at an exponential rate that is determined by λ . The curvature factor induces a hump in the model and it increases from zero to a maximum and then decreases. The rate at which it increases and decreases is also controlled by the decay parameter λ .

6.2 Data

The data used throughout this thesis are monthly averages of observed Swedish government bonds from the period January 1997 to December 2011 that was directly retrieved from the Swedish Central Banks homepage (www.riksbanken.se). Under this period, 8 different maturities were considered and used, namely one-month, three-months, six-months, one-year, two-years, five-years, seven-years and ten-years.

I restrained from converting the simple rates, (i.e rates with maturities less than a year), to effective rates. This should be necessary to obtained a yield curve for bonds of the same type and hence the same credit-risk. I instead used the rates as given by the Swedish Central Bank and constructed yield curve with both short- and long rates, as in definition 2.3.

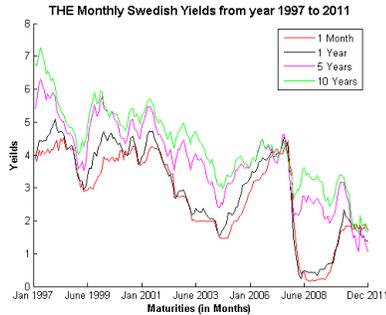


Figure 6.2. The time series of the observed monthly Swedish government yields from the period January 1997 to December 2011, for four selected maturities.

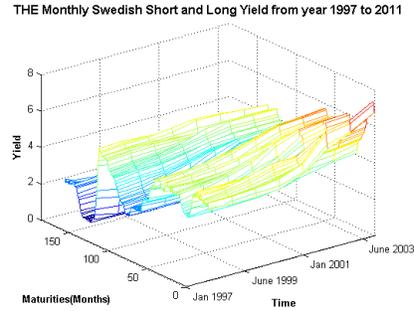


Figure 6.3. A three dimensional plot of Swedish government yields with maturities ranging from one month to ten years for the period January 1997 to December 2011.

The only modification that I made on the data is to fill in the missing rates for the one year bonds. This was done by using the matlab functions *missdata* and *id-data*. The result, thus obtained, were then compared to the *one year STIBOR rates* under the same period, and the difference was insignificant. I therefore proceeded by replacing the whole column of 1-year bonds with missing data points with the results obtained from the matlab interpolation function *missdata*. In Figure 6.2, I plotted the time series for some selected maturities in order to gain some insights on the trajectory of the yield. The maturities chosen for visualization are namely 1-month, 1-year, 5-years and 10-years.

Using the whole tenor structure, the data was then presented as a three dimensional plot as shown in Figure 6.3. From Figure 6.3, variation in the level, slope and curvature of the yield curve can be observed. Also, observed that the interest rates with longer maturities (i.e. the long-rates), varies less than the interest rates with short maturities (i.e. the short-rates).

Following *Diebold and Li (2006)*, the three factors in the Nelson Siegel model were empirically defined. From the illustrations in Figure 6.1, observed that the level-factor is loaded by a constant, one. Therefore, an increase in the level increases all yield equally, as the loading is identical for all maturities. The level factor was defined from data as that rates with the longest maturities, which in our case, are the 10-years rates.

The slope of the yield curve governs the short-term components of the term structure. The slope-factor was empirically defined as the difference between the 10-years yield and the 3-months yield.

The medium-term factor is related to the curvature of the yield curve. Empirically, the curvature factor was defined as twice the 2-years yield minus the sum of the

3-months and 10-years yield.

It is very important to note that the factors are defined as above in order to induced as little correlation as possible between them. In principle, minimal correlation between the factors is desired in order to avoid the factors been *unidentifiable* and thereby causing *collinearity issues*. Let $(\beta_l, \beta_s, \beta_c)$ represent the empirically defined factors, then their pairwise correlations are $\rho(\beta_l, \beta_s) = 0.20069$, $\rho(\beta_s, \beta_c) = -0.31891$ and $\rho(\beta_l, \beta_c) = 0.17703$. As can be observed, there exist a very weak pairwise correlation between the empirical level-, slope- and curvature factors.

6.3 Descriptive Statistics, data

A descriptive statistics of the data including the empirical level, slope and curvature defined above, are presented in Table 6.1. The sample autocorrelation functions, for displacements of one-, twelve- and thirty months are also included. From Table 6.1, we observed that the long-rates varies less than the short-rates, thus providing, an upwards sloping median term structure as shown in Figure 6.10.

Table 6.1. Descriptive Statistics for Swedish Government Yields

Maturity	Mean	Std. dev.	minimum	maximum	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
1	2.802	1.278	0.152	4.524	0.9840	0.509	-0.035
3	2.809	1.316	0.152	4.522	0.9850	0.505	-0.012
6	2.870	1.337	0.166	4.725	0.9840	0.510	0.022
12	3.022	1.358	0.257	5.085	0.9843	0.536	0.062
24	3.335	1.304	0.611	5.364	0.9726	0.532	0.208
60	3.884	1.198	1.069	6.300	0.9622	0.487	0.340
84	4.079	1.185	1.279	6.543	0.9620	0.510	0.389
120(Level)	4.291	1.178	1.677	7.268	0.9605	0.494	0.384
Slope	1.482	0.867	-0.586	3.286	0.9563	-0.083	-0.290
Curvature	-0.431	0.648	-2.230	0.926	0.8959	0.087	-0.225

The Descriptive statistics for observed monthly Swedish government yields and the empirical level, slope and curvature, for the period January 1997 to December 2011. The data based level-, slope- and curvature factors are defined as in section 6.3. The sample autocorrelation functions at displacements of 1, 12 and 30 months are also included.

The median term structure plotted in Figure 6.10, shows that the Swedish government bonds displays the main characteristic features term structures exhibits namely, upwards sloping, concave shape, persistent yield dynamics and a decreasing volatility for long maturities rates.

6.4 Estimation of model parameters

To estimate the model parameters, *state space representation* was introduced, where the *state equation* represents the *dynamic factors* in the three-factor model, i.e. the level, slope and curvature, and the *measurement equation* represents the *instantaneous rates*. The system takes the form:

$$\begin{aligned} X_t &= (I - A)\mu + AX_{t-1} + \eta_t, & \eta_t &\in N(0, Q) \\ Y_t &= ZX_t + \epsilon_t, & \epsilon_t &\in N(0, \Sigma) \end{aligned}$$

where $X_t = (L_t, S_t, C_t)$, I is an identity matrix, μ is a mean vector and A is a diagonal $AR(1)$ coefficient matrix.

In the above state space representation, we assumed that both the measurement- and state errors, i.e. η_t and ϵ_t , are independent and Gaussian.

From the above state space representation, two methods can be used to estimate the factor dynamics and the parameters in our model. These methods basically depend on whether the measurement equations and the state equations are estimated separately or together, and also on the *linearity or non-linearity* of the factor-loading matrix Z_t , which is governed by the decay parameter, λ . These methods are briefly described below.

1. First method, which was also used by Nelson and Siegel (1987), fixed the decay parameter to a specific value. The value at which the decay parameter is fixed is very important in this approach, as it affects both the slope and curvature of the term structure. I followed *Diebold and Li (2006)* and fixed the decay parameter to a value that maximizes the medium term component at a maturity of *2.5-years*, which gives a decay parameter $\lambda = 16.7291$, given our tenor structure. The *linear system* thus obtained is more convenient and simple to deal with and *ordinary least squares* method, can be applied to obtain a time-series of factor estimates.

The estimated factors are then model as independent univariate $AR(1)$ processes. I instead used the *state equation* above and model the estimated factors, which ensures *stationary ar-coefficients* and hence a stable system if started at its unconditional mean and unconditional variance.

2. In the second approach, all the parameters in our system are estimated simultaneously by *direct maximum likelihood method*. Observe that no restrictions are imposed on the decay parameter. The factor dynamics can be retrieved by using a *Kalman Filter*. This approach is more demanding and requires nonlinear optimization procedures.

In this thesis, only the first method, which *Diebold and Li (2006)* referred to as *the two-step approach*, was implemented and its application on Swedish government bonds is studied in the next section.

6.5 Analysis

I applied the two step approach on Swedish government bonds, as described in section 6.4. However, I deviated from the above approach in choosing the fixed value of the decay parameter, λ . I used the empirically defined factors as starting values and estimate the decay parameter by using the matlab function *fminsearch* which gave a time-series of decay parameters. The median value of this time-series, which was 17.8348, was taken to be the fixed decay parameter. With this value, the information matrix Z , for our model, was constructed and *ordinary least squares* used to estimate the factors. The resulting factor estimates, given the fixed decay parameter, against the empirically defined factors are shown in Figure 6.4, 6.5 and 6.6. The fit observed in Figures 6.4 through 6.6 shows that indeed the three factors in our model can be interpreted as *level*, *slope* and *curvature*

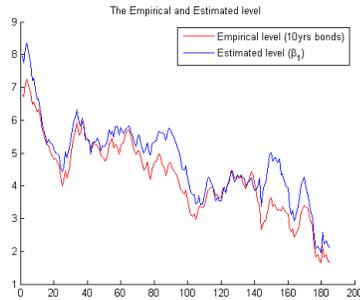


Figure 6.4. The model- against the data-based level factor.

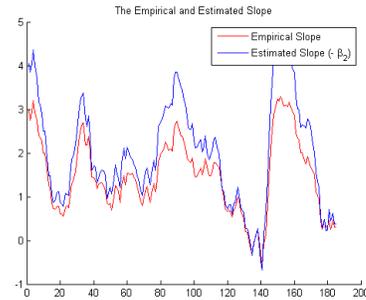


Figure 6.5. The model- against the data-based slope factor

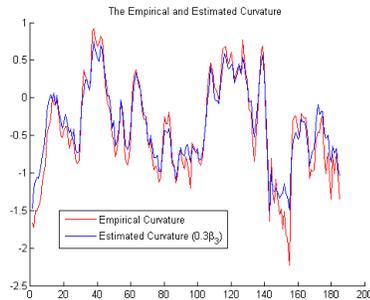


Figure 6.6. The model- against the data-based curvature.

In order to avoid collinearity, there should exist a weak negative correlation between the slope- and the curvature-factor. The correlation between the estimated slope- and curvature factor was shown to be -0.58094 . The correlations

between the empirical- and estimated-factors, for our fixed λ , was also computed and as expected, shown to be very high. If we let $(\beta_l, \beta_s, \beta_c)$ represent the empirical factors and $(\hat{\beta}_l, \hat{\beta}_s, \hat{\beta}_c)$ the estimated factor, then the correlation between them, using Swedish government bonds are $\rho(\beta_l, \hat{\beta}_l) = 0.94343$, $\rho(\beta_s, \hat{\beta}_s) = -0.98972$, $\rho(\beta_c, \hat{\beta}_c) = 0.97122$.

The descriptive statistics and the sample autocorrelations of the estimated factor are presented in Table 6.2 and in Figures 6.7, 6.8 and 6.9 respectively. From Figures 6.7, 6.8 and 6.9 we observed that the level factor is most persistent of the three and the the curvature factor is indeed the least persistent in our model.

Table 6.2. Descriptive Statistics, Estimated factors

Estimated Factor	Mean	Std. dev.	minimum	maximum	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
level	4.7590	1.2041	1.9775	8.3468	0.9491	0.3791	0.2743
Slope	-2.0078	1.2278	-4.8071	0.6699	0.9598	-0.0428	-0.3608
Curvature	-1.1996	1.7457	-5.1354	2.4900	0.9142	0.1381	-0.2766

The table shows the descriptive statistics for the estimated Nelson Siegel factors, by ordinary least squares, with fixed $\lambda = 17.8348$. The last three columns in the table shows the sample autocorrelation of the three factors with displacement of 1-, 12- and 30 months

To assess our factor estimates, we used the estimated factors to reconstruct the data and thereafter compared the fitted-yields with the actual-yields. Firstly, we plotted the average term structure for the data against the average fitted term structure as shown in Figure 6.11.

The fitted/estimated average term structure was obtained by evaluating the Nelson Siegel three-factor model at the mean values of the time-series of the estimated factors. The average term structure of the data was computed as the mean value of the Swedish government rates under study. From Figure 6.11, we observed that the actual and the fitted average yields agrees quite well.

Furthermore, we choose some maturities and compared the data with their corresponding estimates, to observe how well the fitted term structure follows the observed data. In other words, we observed an *in-sample fit* of our model. The plot of the selected maturities are shown in Figure 6.12, 6.13, 6.14, 6.15.

Figures 6.12-6.14 shows that the three-factor model can replicate the variety of shapes and forms the term structure of interest rates can take. However, the model do have problem in fitting the data when there are large variation between the rates. The model also misfits when the term structure exhibits more than one *minima or maxima*.

The *residuals* of the model were plotted in Figure 6.16. Table 6.3 we displayed the descriptive statistics of the residuals. The residuals plot in Figure 6.16 shows that the model does not perfectly fits the data, which may be due to several reasons. For example, it might be indicating that there is some *pricing error*. In Table 6.3,

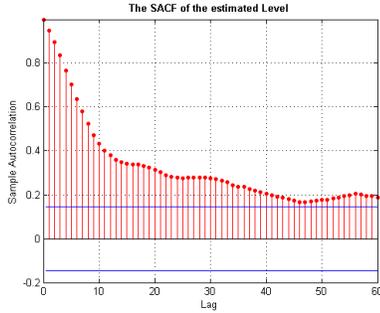


Figure 6.7. The ACF for $\hat{\beta}_{l,t}$

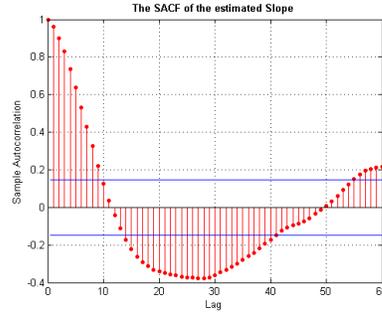


Figure 6.8. The ACF for $\hat{\beta}_{s,t}$

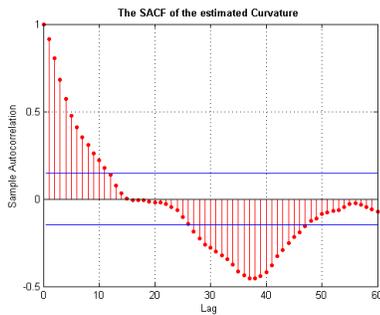


Figure 6.9. The ACF for $\hat{\beta}_{c,t}$

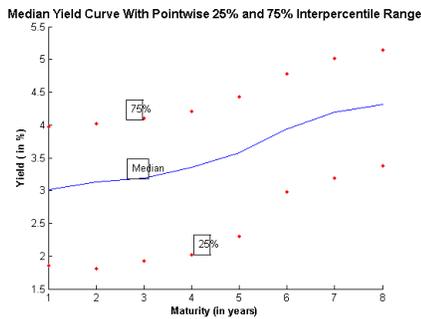


Figure 6.10. The median yield curve with 25th and 75th percentile.

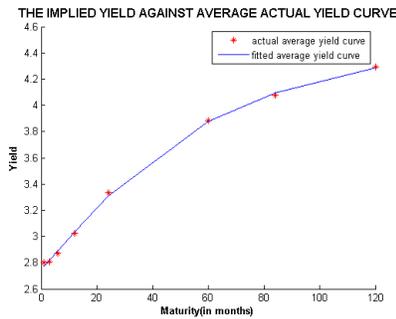


Figure 6.11. The data-based and model-based average term structure

we present the correlations at displacements of 1-, 12- and 30 months descriptive statistic of the estimated factors, given the fixed λ above. The sample autocorrelations in Table 6.3 shows that the pricing errors observed in Figure 6.16 are

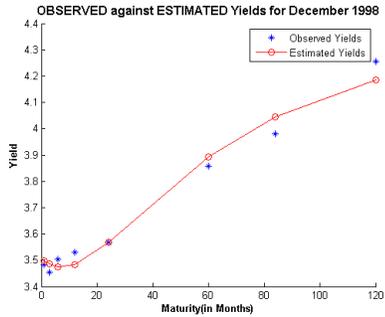


Figure 6.12. The actual and the fitted Swedish government bonds in Dec 1998.

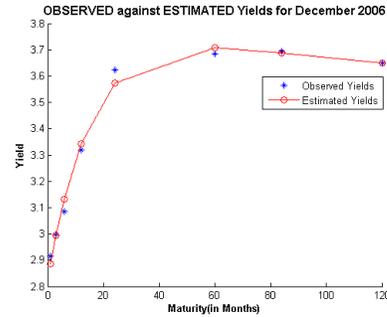


Figure 6.13. The actual and the fitted Swedish government bonds in Dec 2006.

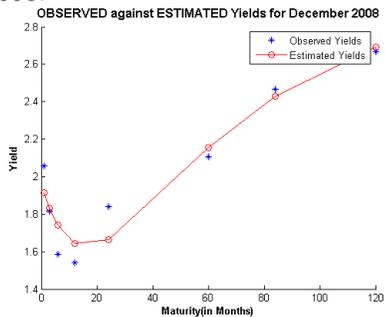


Figure 6.14. The actual and the fitted Swedish government bonds in Dec 2008.

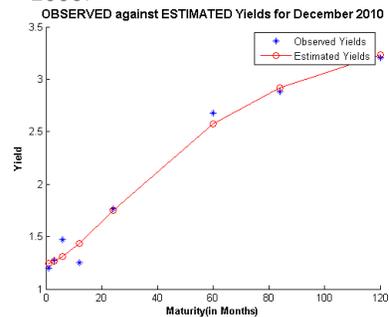


Figure 6.15. The actual and the fitted Swedish government bonds in Dec 2010.

persistent. This confirms the impossibility to model the term structure perfectly. There always exist a difference between the actual rates and the fitted rates.

6.5.1 Modeling the Estimated Factors

In the previous section, we have estimated the three factors in our model by ordinary least squares and analyzed most of its statistical properties. Observed that if we were only interested in fitting the term structure, the measurement equation in the state space representation will be sufficient.

However, the second step in the *two step approach* proposed by *Diebold and Li (2006)* suggested that we model the estimated factors as well. We therefore proceeded by modeling the estimated factors using the state equation in section 6.4. In Figures 6.17, 6.18 and 6.19 we plotted the empirical factors against the *OLS* factor estimates and the state representation factor estimates. The figure shows

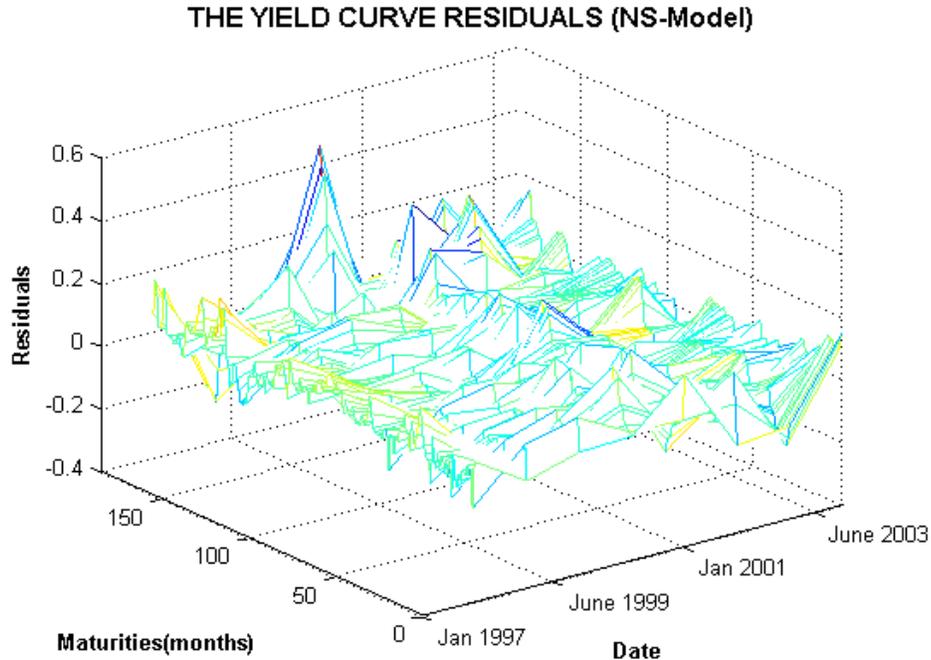


Figure 6.16. The residual of the Nelson Siegel model fitted to Swedish government bonds.

that the factors obtained by using the state representation are almost similar to those obtained by the simple *OLS* method.

However, it is important to note that modeling the estimated factors enables us to construct forecast of the term structure. For example, given a fixed decay parameter in the three-factor Nelson Siegel model, the spot-rates depends only on the model factors i.e. *level*, *slope* and *curvature*. Thus, having a forecast of the factors

$$X_{t+h|t} = (I - A)\mu + AX_{t,t} \quad (6.5.1)$$

enables us to forecast the term structure as follows

$$\hat{y}_{t+h|t}(\tau) = ZX_{t+h|t} \quad (6.5.2)$$

The estimated mean μ , ar-coefficient matrix A and the variance σ , are shown in Table 6.4.

Observed that the co-variations between the factors are not reported because we are modeling the factors as *independent univariate processes*. The residuals autocorrelation of of our state space model fitted to the estimated factors are shown

Table 6.3. Descriptive Statistics, residuals

Maturity (months)	Mean	Std	minimum	maximum	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
1	0.0275	0.0573	-0.1098	0.2004	0.6377	0.0899	-0.0958
3	-0.0114	0.0436	-0.2534	0.0585	0.6145	0.1228	-0.0131
6	-0.0219	0.0538	-0.1835	0.1625	0.5966	-0.0731	0.0200
12	-0.0147	0.1123	-0.3654	0.4697	0.7578	0.0263	-0.1350
24	0.0262	0.0828	-0.2150	0.2964	0.7348	-0.1067	0.0315
60	0.0039	0.0782	-0.1423	0.2883	0.8683	0.1948	-0.0709
84	-0.0161	0.0603	-0.1787	0.1504	0.9235	0.0525	-0.0605
120	0.0066	0.0555	-0.1362	0.1531	0.8419	-0.0201	-0.1028

The table shows the residuals of the Nelson Siegel three-factor model with $\lambda = 17.8348$ fitted to Swedish government bonds at various maturities.

Table 6.4. Estimated Dynamic Factors

Estimated Factor	Mean	AR-Coeff	Var
level	4.6271	0.9942	0.0748
Slope	-1.9436	0.9151	0.0652
Curvature	-1.0886	0.9386	0.3766

The estimated mean, AR-coefficient, variance and the prediction error of the estimated Nelson Siegel factors modeled as independent factors using the state space representation in section 6.4

in Figure 6.20, 6.21 and 6.22. From these figures, we could observed that the auto-correlations are very small, thus indicating that the model captures the conditional means of the level-, slope and curvature factors.

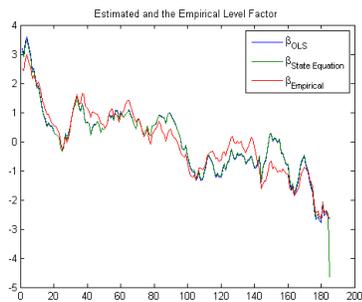


Figure 6.17. The empirically defined level plotted against the *OLS* and state equation representation

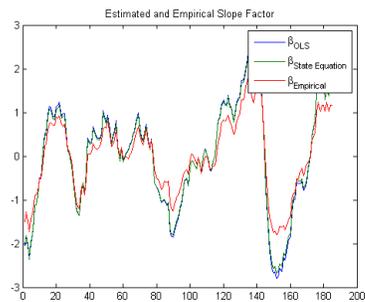


Figure 6.18. The empirically defined slope plotted against the *OLS* and state equation representation

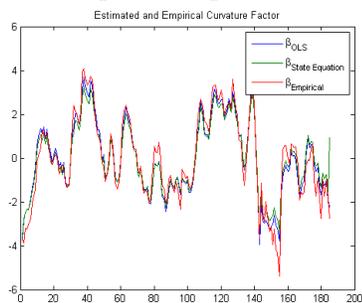


Figure 6.19. The empirically defined curvature plotted against the *OLS* and state equation representation

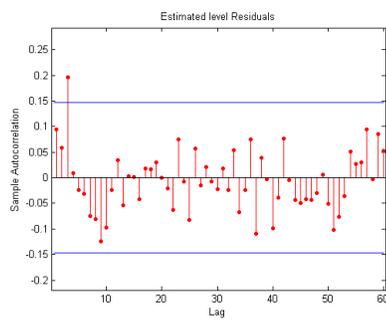


Figure 6.20. The residual ACF for our state space model fitted to $\hat{\beta}_{1,t}$.

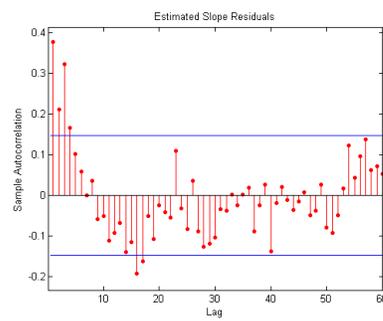


Figure 6.21. The residual ACF for our state space model fitted to $\hat{\beta}_{2,t}$.

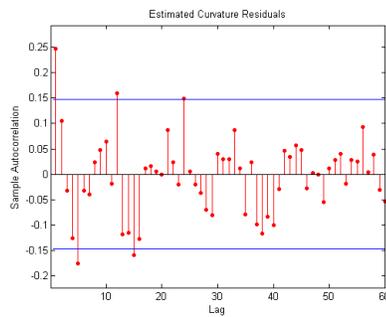


Figure 6.22. The residual ACF for our state space model fitted to $\hat{\beta}_{3,t}$.

6.6 Conclusion

In this thesis, I have shown that the term structure of Swedish government bonds can be fitted by applying the dynamic representation of the three-factor Nelson Siegel model. This was achieved by applying *Diebold and Li (2006)* two step approach, where the model factors, i.e *the level*, *the slope* and *the curvature*, were first estimated using *ordinary least squares* and thereafter model as *stable independent AR(1) processes*. I have therefore shown that the, with only three unknown parameters, the *two-step approach* can be used to estimate the term structure of Swedish government bonds. The estimates gave a good *in-sample fit* for short maturities but they display difficulties in fitting Swedish government bonds with long maturities and at dates where the term structure exhibits multiple *minima/maxima*.

I therefore instead to further study how these drawbacks can be improve upon in a masters thesis. I intend to follow *Svensson (1997)* and extend the three-factor model by adding a second curvature in the model. I also intend to used both the three-factor model and the Svensson extension to predict and forecast the term structure of Swedish government bonds. The performance of these models *in-sample* and *out-sample* fit will be contrasted. I also intend to lift the *arbitrage-free* assumption made on the Swedish government bonds throughout this thesis.

BIBLIOGRAPHY

- [1] *Interest Rate Models, Theory and Practice*, Damiano Brigo and Fabio Mercurio, August 2001.
- [2] *Time Series Analysis*, Henrik Madsen .
- [3] *Arbitrage Theory in Continuous Time*, Björk 3th Edition.
- [4] *Finance Market, Instruments and Investments*, Hans Byström, 2nd Edition.
- [5] *Modeling Non-Linear and Non-Stationary Time Series*, Hendrik Madsen and Jan Holst, December 2000.
- [6] *Stationary Stochastic Processes*, Georg Lindgren, Holger Rootzén and Maria Sandsten, September 2009.
- [7] *Derivative Pricing part I and II*, Sebastian Åberg, 2010.
- [8] *Räntebärande Instrument, Värdering och Riskhantering*, Hossien Asgharian and Lars Nordén.
- [9] *De Finansiella Marknaderna i ett Internationellt Perspektiv*, Leif Hässel, Marie Norman and Christian Andersson .
- [10] *Estimating Forward Interest Rates with the Extended Nelson and Siegel Method*, Lars E.O. Svensson, (1994).
- [11] *Interest Rate Dynamics and Consistent Forward Rate Curves*. Björk T. and Christensen B. J. (1999).
- [12] *Forecasting the term structure of government bond yields*, Francis X. Diebold and Canlin Li, (2006) .
- [13] *Estimating the Yield Curve Using the Nelson-Siegel Model, A Ridge Regression Approach*, Jan Annaert, Anouk G.P Claes, Marc J.K De Ceuster, Hairui Zhand, 2000.

-
- [14] *An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model*, Jens H.E. Christensen, Francis X. Diebold, Glen D. Rudebusch, November 2008.
- [15] *Calibrating the Nelson-Siegel-Svensson model*, M. Gilli, Stefan Grobe and E.Schumann. march 2010.
- [16] *Examining the Nelson-Siegel Class of Term Structure Models*, Michiel De Pooter, June 2007.