Eötvös Loránd University

Faculty of Science

Yield Curve Modeling

Thesis

Balázs Márton Süli

Actuarial and Financial Mathematics MSc

Quantitative Finances Major

Supervisors:

Dr. András Zempléni

associate professor

Department of Probability Theory and Statistics

Dr. Daniel Niedermayer

Solvency Analytics

Budapest, 2014

Contents

1 Introduction

The main topic of this thesis is yield curve modeling. I have tried to collect the most relevant information on that but still not to exceed the limits of an MSc thesis. The idea of a thesis about yield curve modeling has come from the swiss Solvency Analytics group. Reliable yield curve models can be very useful when calculating sensitivites and capital charges of corporate bonds within the Solvency II framework. For the thesis to be useful for Solvency Analytics, I have focused mostly on corporate bonds and I have chosen to write it in English.

The first few pages of the thesis is concentrated on concepts such as fixed income securities, risks affecting them, corporate bonds, YTM, zero-coupon yield curve, discount curve, forward curve and no-arbitrage. After those the concepts of discount function and instantenous forward rates are introduced. The next section of the thesis is about one factor short rate models. After a general description of these types of interest rate models two popular models are introduced: the Vasicek and Cox-Ingersoll-Ross models. In this section, I have relied on the knowledge I have learned at the university lectures of Dr. György Michaletzky [1] and I used similar notations. The Statistical Yield Curve Models section presents some methods to model the yield curve based on observable market prices and bond properties. It starts with a method called Coupon Stripping and after that other types of yield curve models follow such as polynomial or spline-based models and Nelson-Siegel type curves. I have relied on two books mostly: Handbook of Fixed Income Securities [2] and Interest Rate Modelling [3]. The Additional Features section presents some alternative but still popular ways to model the yield curve. They can be very useful when the construction of statistical yield curve models are not possible. The last section, Applied Methods summarizes the numerical implementations I have written to be able to fit some models to real data. I have focused on the polynomial and spline estimation models here and presented some outputs of the apllications.

I would like to thank my supervisors Dr.András Zempléni and Dr.Daniel Niedermayer for the numerous advice and help they have provided me and made the following thesis much better.

1.1 Fixed Income Securities

Fixed income securities constitute a huge and important part of the financial products universe. There are a lot of different types of them, and as financial markets became more and more complex, the evolution of these products have started to speed up. The basic fixed income securities grant given cash flows on certain future dates to the security holder. The payer of these cash flows is the issuer of the security. These are one of the most simple type of financial products, but even these are affected by several types of risks. It is in the investor's interest to have a reliable mathematical model of the products, and a good model always counts with the risks.

The most important type of risk affecting fixed income securities is called interest-rate risk. It is the risk arising from the constant change of the fixed income securities market. If interest rates increase, that means investors can expect a higher return on their new investments in the market, and this lowers the value of older ones. The market value of securities moves in the opposite direction of interest-rate movements, if rates rise the market values fall while if rates fall the market values rise. Of course this type of risk does not affect the investor who holds the security until maturity and doesn't plan to sell it before that.

Another important risk affecting fixed income securities is credit risk. The source of this type is the issuer of the security and its ability to meet its duties. There is always a possibility that the issuer can't or won't pay the amount when it must be payed, and this possibility is different for different issuers. A bigger chance of default on the issuer's side makes the security less valuable and this is usually compensated with higher yield, to have someone buying them. Among others liquidity risk, currency risk and inflation risks also affect the price of a fixed income security. To be able to precisely measure the value of an investment, and to make solid investment choices, sophisticated mathematical models are needed even for the above mentioned most simple types. In practice, there are more complex features such as embedded options, seniority restrictions or convertible bonds, which call for more complex models.

Corporate bonds refer to bonds issued by different corporations. The most important difference between governmental bonds and corporate bonds is the different credit risk associated with the security. In most cases governmental bonds are considered to contain very little credit risk, while some corporations have much bigger chance of default. This extra risk beared by the investor is compensated with an increased return on the investment usually. Another distinction between governmental bonds and corporate bonds is that corporate bonds usually contain additional features like convertibility or embedded options, and are often less liquid than governmentals. These make the modelling of corporate bonds harder mathematically, and less accurate models can be expected, than in the case of governmentals. But as more and more complex financial products arise, mathematical models become more and more sophisticated, trying to find a solution for actual financial problems.

1.2 Yield Curve

How can the investor compare the return of several investments, and choose the one that's the most approriate for his or her needs? There are different mathematical tools to measure the return of an investment. In the case of coupon bonds, one of the most popular measure of return is Yield To Maturity. Let's consider a coupon bond with n years time-to-maturity, fixed $\$C$ coupon payments at the end of every year, and the redeem of the principal which is \$100 at the end of the n years. If the market price of this coupon bond is P_C now, the yield-to-maturity (YTM) is defined as:

$$P_C = \sum_{i=1}^{n} \frac{C}{(1+YTM)^i} + \frac{100}{(1+YTM)^n}$$

According to this equation, the YTM can be calculated every time the cash flow pattern and the current market value of a security is known. Yield-to-maturity is an easily computable, but not too sophisticated measure of a bond's yield. It assumes that every cash flow is discounted with the same rate, which is not a realistic approach. In practice cash flows further in time carry more risks, and usually should be discounted with a larger discount rate than the closer ones. This is why the zero-coupon yield curve is so important in the field of modelling fixed income securities.

For a zero-coupon bond the computation of the yield is more simple. Let P_Z be the current market price of the zero-coupon bond, and the only one cash flow it contains is \$100, t years from now. Then the effective yield (r_E) of this product comes from:

$$P_Z = \frac{100}{(1+r_E)^t}$$

The compounded yield (r_C) comes from the equation:

$$P_Z = 100e^{-r_C t}$$

The zero coupon yield curve refers to the illustration of zero-coupon yields of different maturities in a given financial market. This is one of the most important tools when modelling fixed income securities.

1.3 No-arbitrage Condition

There is an essential condition when modelling fixed incomes in a given market and this is the no-arbitrage condition. Economic considerations suggest that there should be no arbitrage opportunities in well behaving financial markets. If any opportunity occurs for an arbitrage, the arbitrageur's activity will move the maket prices in a direction that extinguishes it. This phenomenon is very useful in mathematical modeling. It creates a continuous and solid connection between the prices of financial products in relating markets. In the case of fixed income securities, it is the no-arbitrage condition that ensures that the knowledge of zero coupon yields itself defines the prices of coupon bonds in that market.

In general it is possible to buy a coupon bond and sell its cash flows in tranches, like different zero coupons. Because of that, the prices of this two types of bonds should be in some kind of harmony. The price of a coupon bond should be the sum of the prices of the zero-coupons of the cash flows it contains, or else an arbitrage opportunity arises.

Now, with the use of no-arbitrage conditions let's see how the zero-coupon yield curve determines the prices of fixed income securities. It is important to bear in mind that yield curves are abstract theoretical concepts, usually adjusted to real market prices to model reality accurately. Now, suppose that in a mathematical model of a financial market, the zero coupon yield curve is known. Let $r(t)$ be the compounded yield of a zero coupon bond maturing in t years for every $t > 0$. In this model, a coupon bond with known cash flow properties can only have one specific price. With the use of the yield curve this price can easily be determined. Assume that bond B has C_i cash flows at times t_i. Then P_B, the theoretical price of bond B is defined as:

$$P_B = \sum_i C_i e^{-r(t_i)t_i}$$

This equation exhibits a strong relationship between the theoretical yield curve of a model, and the prices of securities in the financial market. If the market price of B differs from P_B an arbitrage opportunity arises in the model and that is inconsistent with the no-arbitrage condition. It is important to mention that in practice, there is an error associated with market prices and so there will always be a small difference between the theoretical price and the market price. The model is estimated usually by somehow minimizing these errors.

As shown, the knowledge of the yield curve is crucially important when pricing securities but it is an abstract concept. To have a model that describes reality well the yield curve should be approximated with the use of market prices of securities in a market. The starting set of securities of this approximation is very important. If one wants to use the information provided by the yield curve, for example to anticipate the price of a bond waiting to be issued, it is crucial that the curve is typical of that type of bond. The estimation should start from similar type of products. Since the yield curve represents a relationship between maturity and return, the types of risks (other than interest-rate risk) should be considered when collecting the starting set of bonds.

1.4 Different Types of Curves

The three most popular types of curves used to represent the term structure of interest rates are the discount curve, the zero-coupon yield curve and the instantenous forward curve. They present the same core information in different ways and since that they can be computed from each other. Let us denote the discount function with $d(t)$. It refers to the present value of 1 dollar (if dollar is the currency one wants to use) received in t years (if a year is the time unit one wants to use). The yield of a zero-coupon bond maturing in t years is written as $r(t)$. The term structure of these are called the spot or zero-coupon yield curve or shortly yield curve. The forward rates express the market's expectations of interest rates in the future. On an investment beginning at time t and ending at time T (where $t < T$) the market expects a yield equal to $f(t, T)$.

The following relationships stand among these three rates. If the spot rates are known, the discount rates can be calculated as:

$$d(t) = e^{-r(t)t}$$

and conversely:

$$r(t) = \frac{-\log d(t)}{t}. \quad (1)$$

The forward rate can be obtained from the spot rates with the use of:

$$f(t, T) = \frac{Tr(T) - tr(t)}{T - t} \quad (2)$$

The spot rate is the following forward rate trivially:

$$r(t) = f(0, t).$$

The conversion between the discount rates and the forward rates can be easily obtained using the spot curve and the mentioned equations.

There is one more curve of importance when describing the term structure of interest rates and that is the instantenous forward rate curve. The forward rates mentioned before needed two parameters: the beginning of an investment t and the end T. If the spot yield curve is differentiable it is possible to define the instantenous forward rate: $f(t)$. The instantenous forward rates represent the expected yield of an investment beginning at t and ending in $t + \Delta t$ when $\Delta t \to 0$. It describes the present market expectations of the evolution of future short rates. Using equation (2), $f(t)$ can be expressed as:

$$
\begin{aligned}
f(t) = & \lim_{\Delta t \to 0} f(t, t + \Delta t) = \\
& = \lim_{\Delta t \to 0} \frac{(t + \Delta t) r(t + \Delta t) - t r(t)}{t + \Delta t - t} = \\
& = \lim_{\Delta t \to 0} \frac{\Delta t r(t + \Delta t) + t(r(t + \Delta t) - r(t))}{\Delta t} = \\
& = r(t) + t \lim_{\Delta t \to 0} \frac{r(t + \Delta t) - r(t)}{\Delta t} = r(t) + \frac{\partial r(t)}{\partial t} t
\end{aligned}
$$

In some yield curve models it is better to estimate the instantenous forward rate curve and the spot and discount curves can be calculated from it.

2 Interest Rate Modelling

2.1 One-factor Short Rate Models

This section provides a few terms and concepts on the type of interest rate models called one-factor short rate models. Short rate models are widely used mathematical models to describe the stochastic evolution of different interest rates. They are called short rate models because the basic concept of this model is the instantenous short rate $r(t)$. It refers to the (annualized and compounded) yield that can be earned on an infinitesimally short investment at time t. It is important to emphasize that altough the notation is the same, this $r(t)$ is not nearly the same as the one mentioned in the previous chapter. Here, the t parameter refers to time in a model and not the time-to-maturity feature of a product observed now.

In the ever-changing world of financial markets the interest rate on actual short term investment opportunities can vary quickly and unpredictably in time. The rates on longer investments also change in time but they are considered much less volatile. In one-factor short rate models the instantenous short rate's evolution is defined by a stochastic process, usually an Ito-process under the risk-neutral measure:

$$dr(t) = a(t)dt + b(t)dW(t)$$

In this equation $a(t)$ and $b(t)$ are time-dependent coefficients of the stochastic process, and $W(t)$ is a Wiener-process under the risk-neutral measure.

There is an other interest-rate process in this model and that is the forward rate. The forward rate is the expected (annualized and compounded) instantenous interest rate of time T, at time t. Of course T must be higher or equal than t. In the model the forward rate is a stochastic process with two time parameters:

$$d_t f(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t) \tag{3}$$

Here the $\alpha(t,T)$ and $\sigma(t,T)$ coefficients are the functions of two time parameters.

The short rate model also incorporates two value processes of financial products. One of them is the value process of a zero coupon bond. For convenience the face value of this zero coupon bond is set to be \$1. Let's consider a \$1 zero coupon bond maturing at time T. The value of this bond at time t is denoted by $P(t,T)$. Since the values of bonds depend on the level of interest rates and expectations of these, the evolution of $P(t,T)$ is also described with a stochastic process in the model:

$$d_tP(t,T) = P(t,T)(m(t,T)dt + v(t,T)dW(t))$$

Here the $m(t,T)$ and $v(t,T)$ coefficients are functions of two time dependent parameters. It is important to mention that this mathematical model presumes that there is a bond maturing at every time t.

Another participant in this model is the value of a bank deposit $B(t)$. In opposite of the bond where present value is derived from a future cash flow, the bank deposit's value origins from the past. Bank deposits pay the timely short interest rate at every time t and are doing it compoundedly. That's why the evolution of the value of a bank deposit depends on the evolution of the short rate. The connection between the two is:

$$dB(t) = r(t)B(t)dt$$

The value of the deposit at time t can also be expressed as:

$$B(t) = B(0)e^{\int_0^t r(s)ds}, \tag{4}$$

where $B(0)$ is the value of the deposit at $t = 0$.

2.2 Connections

There are certain connections between these processes. From Equation (4) it is apparent that with the knowledge of $B(0)$ the short rate process determines the value process of the bank deposit. There is a connection between the value of the bond and the forward rate process. The present value of a bond should be the discounted value of the future cash flow. At time t the expectations of future interest rate levels are the $f(t,s)$ values where s is larger or equal than t. A fair valuation of the bond in this model should be:

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \tag{5}$$

This is the appropriate value here as the expected future value of a bank deposit of size $P(t,T)$ at time t is exactly one dollar at time T:

$$B_{exp}(T) = B(t)e^{-\int_t^T f(t,s)ds} = P(t,T)e^{\int_t^T f(t,s)ds} = e^0 = 1$$

If it wouldn't be one dollar an arbitrage opportunity would arise.
Expressing the forward rate from Equation (5):

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

Using these strong connections between these two processes the knowledge of one of them results that the other one can be expressed too. Assuming that the forward rate is known in the form presented in Equation (3), so the coefficient functions $\alpha(t,T)$ and $\sigma(t,T)$ are known. It can be shown that the coefficients of the bond value's process $m(t,T)$ and $v(t,T)$ can be expressed using $\alpha(t,T)$ and $\sigma(t,T)$ functions as follows:

$$m(t,T) = f(t,t) - \tilde{\alpha}(t,T) + \frac{1}{2}\tilde{\sigma}^2(t,T)$$

and

$$v(t,T) = -\tilde{\sigma}(t,T)$$

The $\tilde{\alpha}(t,T)$ and $\tilde{\sigma}(t,T)$ functions are integrals of $\alpha(t,T)$ and $\sigma(t,T)$:

$$\tilde{\alpha}(s,T) = \int_s^T \alpha(s,u)du$$

and

$$\tilde{\sigma}(s,T) = \int_s^T \sigma(s,u)du$$

This defines the coefficients of the bond's value process perfectly. The other way of expressing one from the other is also available. Assume that the stochastic process of the bond value is known, so the coefficient functions $m(t,T)$ and $v(t,T)$ are known. Using this information and the connection between the two processes, the coefficients $\alpha(t,T)$ and $\sigma(t,T)$ of the forward rate can be expressed:

$$\alpha(t,T) = v(t,T)v_T(t,T) - m_T(t,T)$$

and

$$\sigma(t,T) = -v_T(t,T),$$

where:

$$m_T(t,T) = \frac{\partial}{\partial T} m(t,T)$$

and

$$v_T(t,T) = \frac{\partial}{\partial T} v(t,T).$$

These equations define a two-way connection between the bond value process and the forward rate process.

Now let's consider the short rate process and its possibilities. It can be shown that assuming the knowledge of the forward rate's evolution the short rate process can be expressed. The short rate process's $a(t)$ and $b(t)$ coefficient functions take the following form:

$$a(t) = \alpha(t,t) + \int_0^t \alpha_T(s,t)ds + f_T(0,t) + \int_0^t \sigma_T(s,t)dW(s)$$

and

$$b(t) = \sigma(t,t)$$

In these expressions the T in the lower corner means the function's derivative by its second variable similarly to $m_T(t,T)$ and $v_t(t,T)$ mentioned above.

But what if one knows the short rate process and wants to express the forward rate from that? It is provable that it's impossible to derive the forward rate or bond value processes merely from the knowledge of the short rate process only, because an equivalent martingale measure can't be obtained. The forward process contains an additional information and that is the market's expectations of the future. It is needed to observe the market prices of bonds and find the appropriate market price of risk process, to be able to derive the forward rate process and the bond value from the short rate.

2.3 Two Popular Short Rate Models

The following two models are among the most popular one factor short rate models.

Vasicek: The Vasicek-model uses a mean-reverting stochastic process with a constant diffusion coefficient to model the short rate process's evolution:

$$dr(t) = (k - \Theta r(t))dt + \sigma dW^*(t),$$

where $t \geq 0$, $k \geq 0$, $\Theta > 0$, $\sigma > 0$.
$W^*(t)$ refers to a Wiener-process under the $\mathbb{P}^*$ probability measure. The $\mathbb{P}^*$ probability measure is the one where the bank deposit is the numeraire process. This measure can be obtained by observing the market prices of bonds and using the market price of risk process with Girsanov's theorem. Under this measure the bond's value can be expressed as:

$$P(t,T) = B(t)\mathbb{E}_{\mathbb{P}^*}\left(\frac{1}{B(t)}\middle|\mathcal{F}(t)\right) = \mathbb{E}_{P^*}\left(e^{-\int_t^T r(u)du}\middle|\mathcal{F}(t)\right) \tag{6}$$

CIR: The Cox-Ingersoll-Ross model constructs the short rate process from the sum of squared, independent Ornstein-Uhlenbeck processes [13] . This way the result is a non-negative stochastic process. Let us begin from the Ornstein-Uhlenbeck processes: $X_1, X_2, \ldots, X_d$ for a positive integer d. Each of them can be described as:

$$dX_j(t) = -\frac{\beta}{2}X_j(t)dt + \frac{\sigma}{2}dW_j(t),$$

where the following stand: $\beta > 0$, $\sigma > 0$ and $W_j(t)$-s are independent Wiener-processes. Now the short rate process is:

$$r(t) = \sum_{j=1}^{d} X_j^2(t)$$

The stochastic evolution of the short rate can be expressed as:

$$dr(t) = \Big(\alpha - \beta r(t)\Big)dt + \sigma\sqrt{r(t)}dW^*(t),$$

where

$$\alpha = d\frac{\sigma^2}{4}$$

and

$$dW^*(t) = \sum_{j=1}^{d} \frac{X_j(t)}{\sqrt{r(t)}} dW_j(t).$$

This last equation relies on Levy's theorem.
In the Cox-Ingersoll-Ross model the dynamics of the zero coupon bond is just like in the Vasicek-model:

$$P(t,T) = B(t)\mathbb{E}_{\mathbb{P}^*}\left(\frac{1}{B(t)}\middle|\mathcal{F}(t)\right) = \mathbb{E}_{P^*}\left(e^{-\int_t^T r(u)du}\middle|\mathcal{F}(t)\right)$$

2.4 Yield Curve Calibrating

Short rate models like these are very useful when one wants to analyze the evolution of interest rates in time, but are they consistent with usual yield curve shapes observable in the markets? Let's see the Vasicek model as an example.

According to Equation (6) the spot discount function can be evaluated at every future points of time. The discount function $d(t)$ is really the function of bond prices $P(0,t)$

evaluated now. Using Equation (6) it is the following in the Vasicek-model:

$$d(t) = \mathbb{E}_{P^*}\left(e^{-\int_0^t r(u)du}\middle|\mathcal{F}(0)\right)$$

It is easy to see that with the knowledge of the parameters of the model (k, Θ and σ) and $r(0)$ the discount function can be constructed. Also the zero-coupon yield curve, which is denoted by $z(t)$ can be computed from the discount function easily. It takes the following parametric form:

$$z(t) = r(0)e^{-\Theta t} - \frac{k}{\Theta}\big(e^{-\Theta t} - 1\big) - \frac{\sigma^2}{2}\frac{1}{\sigma^2}\big(e^{-\Theta t} - 1\big)^2$$

This is a combination of $e^{-\Theta t}$ and $e^{-2\Theta t}$ functions of time. With the three parameters of $r(t)$ it isn't always possible to match the shape of current yield curves. This is a shortcoming of the Vasicek-model and other parametric one-factor short rate models. If one wants an interest rate model that can match the actual observable yield curves some improvements has to be made.

Extended Vasicek-model: An improved version of the original constant parametered Vasicek model is the Extended Vasicek model. In order to be able to calibrate the model to actual yield curves one of the constant parameters of the Vasicek-model is made to be time-varying. The dynamics of the extended model is:

$$dr(t) = \big(k(t) - \Theta r(t)\big)dt + \sigma dW^*(t),$$

where $k(t)$ is an arbitrary deterministic function of time, and Θ and σ are positive constants. Deriving the actual zero coupon yield curve from this model results in:

$$z_2(t) = r(0)e^{-\Theta t} + \int_0^t k(u)e^{-\Theta(t-u)}du + \frac{\sigma^2}{2}\frac{1}{\sigma^2}\big(e^{-\Theta t} - 1\big)^2$$

It is possible to find a $k(t)$ function that is appropriate in a sense, that the $z_2(t)$ function of the interest rate model will match the market zero coupon yield curve, derived from observable prices of bonds.

A general method of calibrating interest rate models to market yield curves can be shown through the CIR-model. Let $z(t)$ be a process that is used for modelling short rate dynamics in a CIR model:

$$dz(t) = \big(\alpha - \beta z(t)\big)dt + \sigma\sqrt{z(t)}dW^*(t)$$

The short rate process now is a sum of $z(t)$ and a deterministic function $y(t)$:

$$r(t) = z(t) + y(t)$$

Here the value of a bond, $P(t,T)$ takes the following form:

$$P(t,T) = P_{CIR}(t,T)e^{-\int_t^T y(u)du}$$

The deterministic $y(t)$ function can be found as follows:

$$y(t) = f(0,t) - f_{CIR}(0,t)$$

This approach is suitable to other kinds of short rate models as well. The types of yield curve models that can be derived from interest rate models, like the ones mentioned above called consistent. The need for model consistency depends on the goal of the modelling. In some cases it is very important to be consistent with an interest rate model, while there are other cases as well where other features are more desirable. The following section introduces models where consistency isn't among the main requirements.

3 Statistical Yield Curve Models

There are several methods for modelling the yield curve, based on observable market prices of securities. Some of them are more theoretical while others are often used in practice. Different situations appeal to different yield curve models. Each model has its own advantages and disadvantages. They can differ for example in tractability, consistency with stochastic interest rate models or in the type of rates they estimate. The discount curve, the spot curve and the forward curve can also be the goal of a method. Once one of this three curves is estimated, the others can be easily calculated from the one known. In the followings, yield refers to the annualized and continuously compounded yield of a security.

3.1 Coupon Stripping

Let us start with an easy model which is mostly theoretical. The goal is to estimate a discount curve that is consistent with the market prices of securities of that kind. In most of the situations one has to start from a set of coupon bond prices and properties and derive the discount curve from that. A method called 'Coupon Stripping' or 'Bootstrapping' [2] is an easy way to obtain the discount function's value on specific dates though it requires very special and unrealistic starting set of information. This approach can be used only if the following are given: a set of coupon bonds with cash flows like that on every cash flow date there is exactly one maturing coupon bond.

Let's consider an example just to illustrate the method. Assume that we look at 10 bonds in the market with market prices P_i and annual coupon payments of size C_i and the i-th bond matures i years from now:

i	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
C_i	4	7	3.5	4.25	2	7	6.5	3	4.5	3.25
P_i	101.74	107.01	98.5	99.68	87.93	108.89	101.6	76.71	83.36	72.18

The goal is to estimate the discount function's value at bond maturity dates. The first step uses the first period only. The discount function's value is:

$$d(1) = \frac{P_1}{100 + C_1} = \frac{101.74}{104} = 0.98$$

The next step uses the value resulted from the step before and the next bond's properties.

$$d(2) = \frac{P_2 - C_2 d(1)}{100 + C_2} = \frac{100.15}{107} = 0.94$$

Following this method every new discount factor can be calculated with the properties of the bond maturing at that date and the past discount factor values:

$$d(n) = \frac{P_n - \sum_{i=1}^{n-1} C_n d(i)}{100 + C_n}$$

The resulting discount function of this method is:

d(1)	**d(2)**	**d(3)**	**d(4)**	**d(5)**	**d(6)**	**d(7)**	**d(8)**	**d(9)**	**d(10)**
0.98	0.94	0.87	0.84	0.79	0.73	0.64	0.58	0.52	0.48

It is useful to set **d(0)** as 1 because the present value of money available now is its face value.

With a linear algebraic approach the coupon stripping method can be presented as follows. Let P be the vector of coupon bond prices maturing on different incremental dates: $t_i \quad i \in \{1, \ldots, n\}$ and C be the cash-flow matrix such as $C(i, j)$ denotes the i-th bond's cashflow on date t_j. From now on $C(i, j)$ is denoted by $C_i(t_j)$ for convenience. C is a square matrix due to the restrictions on the dataset. The discount factors on each t_i date are denoted by $d(t_i)$ and $D = (d(t_1), \ldots, d(t_n))$ is a vector comprised of them.

In an arbitrage-free solution the bonds' prices should be equal to the discounted present values of their cash flows, so

$$P_i = d(t_1)C(i,t_1) + \cdots + d(t_n)C(i,t_n)$$

should stand for every $i \in \{1,\ldots,n\}$. Now by solving the $CD = P$ linear system D is obtained and so the discount values on the defined t_i dates.

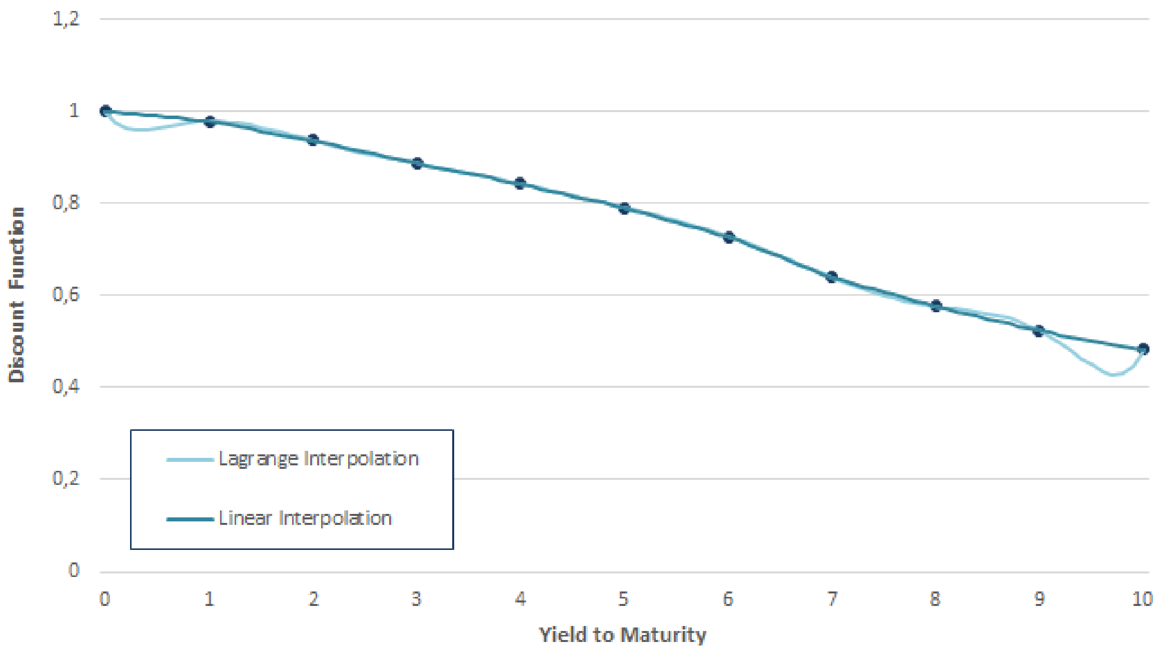
As it was mentioned before this approach is mostly theoretical and rarely used in practice. One of the main problems is that a starting set of information as it is required for the Coupon Stripping method is very unrealistic in practice. There is rarely available a representative set of similar type of bonds maturing exactly one years (or any other given period) from each other in the market.

3.2 Interpolation

To have a discount value for every t in a time interval, one can consider the discount curve linear between the dates where it's values are known. With this a continuous but not smooth (in most cases the first derivative is not continuous) yield curve is estimated. Another way to fit a continuous curve onto the points known can be using polynomial approximation techniques. In spite of the fact that a perfect fit on the vertices can be obtained by a polynomial function with the use of Lagrange approximation, it is not suitable in practical use. Although the estimated curve is infinitely differentiable and connects the points this approach doesn't result in realistic yield curves. This is because between the initial points and usually on the short and long ends of the time interval very unrealistic shapes can occur.

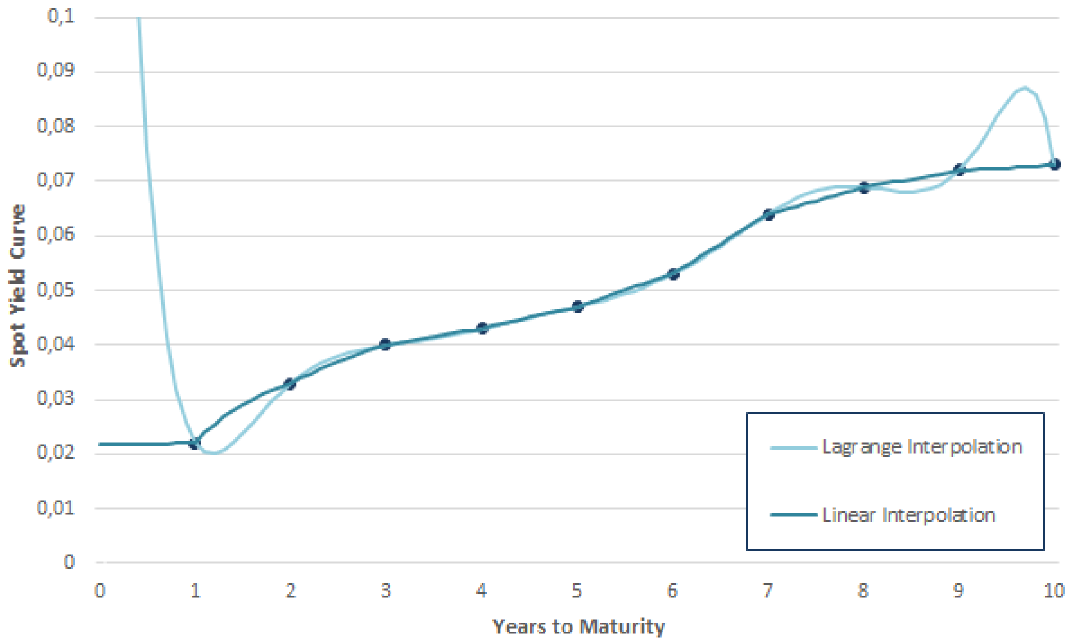
Figure 1 shows the discount factor values obtained by the coupon stripping method example before. The dark blue marks represent the estimated discrete discount factor values. The dark blue line is a linear interpolation while the light blue is a Lagrange polynomial interpolation. It can be seen that the Lagrange interpolation results in some extra curvatures at both ends of the sample.

Figure 1: Linear and Lagrange interpolation of the discount function.



While the difference between the two interpolation method is not that big in the case of the discount function, when computing the spot yield curve it becomes much more significant. The spot yield cuve is computed from the values of the discount curve using the connection mentioned in Equation (1). Figure 2 shows the resulting curves. The dark blue marks represent the yields computed from the original results of the coupon stripping. The dark blue is the yield curve estimated from the linearry interpolated discount function while the light blue is the one derived from the Lagrange interpolated discount function.

Figure 2: Yield curves calculated from the interpolated discount functions.



The instantenous forward curve approximation derived from the spot yield curve is presented in Figure 3. This curve is computed by dividing the time interval between every year into ten equal small intervals. Then using Equation (2) to calculate $f(t_i)$ as:

$$f(t_i) = f(t_i, t_{i+1}),$$

for every $i \in \{0, 1, \ldots, 99\}$ where this represents the new time lattice.

Because of the absence of a continuous first derivative in the case of linear interpolation, this approximation of the instantenous forward curve becomes very ragged as shown in Figure 3 . It is a good example of why a smooth curve is required. While the forward curve computed from the Lagrange-interpolation is smooth, it has apparently too much curvature which makes the results unrealistic.

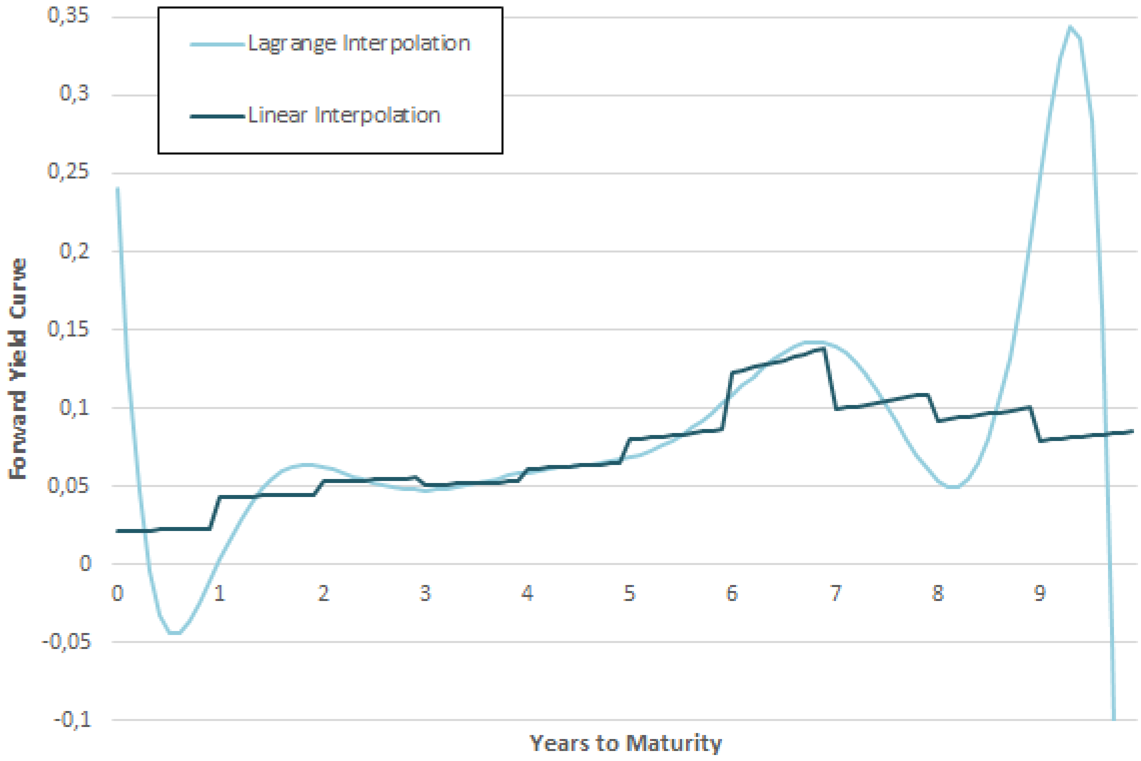


Figure 3: Forward curves derived from the spot curves.

3.3 Including Errors

The Coupon Stripping method doesn't count with errors in the prices of securities. In reality the bond market prices rarely equal the theoretical prices obtained (sum of the discounted cash flows). There can be smaller or larger differences arising from several reasons like liquidity, the cash flow structure or tax effects. It is reasonable to include an error term in the arbitrage-free equation:

$$P_i = d(t_i)C(i, t_1) + \cdots + d(t_n)C(i, t_n) + e_i,$$

where e_i refers to the error term, the difference between the theoretical and quoted market price of the i-th bond. This approach results in more realistic models of the yield curve but requires more sophisticated estimation methods.

Suppose a set of coupon bond prices and properties are known, now without the restrictions stated before. Let P be the vector comprised of the market prices of the bonds, $C(i,j)$ the cashflow matrix and D the vector of the discount values. In the Coupon Stripping case the cashflow matrix is a squared matrix due to the strict initial conditions. Without those restrictions C has usually much more columns than rows and has a lot of zero values. Let ε be the vector of the e_i error terms as mentioned above. Now the following holds:

$$P = CD + \varepsilon$$

One way to estimate D is by minimizing the error terms. Without any other restrictions on D this can be done by Ordinary Least Squares (OLS) estimation. The task is to find D^* where

$$D^* = \min_{D} \left\{ \varepsilon^T \varepsilon \quad | \quad \varepsilon = P - CD \right\}$$

The solution of the OLS regression is

$$D^* = (C^T C)^{-1} C^T P$$

It is important to mention that $C^T C$ is not always an invertible matrix, but with the use of a pseudo-inverse this problem is solvable. Although this is an elegant approach in theory, it is not a good choice in practice. Even with carefully chosen starting data very unrealistic shapes can occur as a result if one wants to somehow produce a continuous function out of the estimated values. This method doesn't place any restrictions on the $d(t)$ function and it is a source of some problems. One of them is that estimated values on similar maturities often don't have similar values which results in a ragged curve. That is not realistic at all.

3.4 Parameterised Curves

Using OLS technique without any restrictions on the $d(t)$ discount function is not a good way of modelling the yield curve. To obtain better results it is beneficial to look for more specific curves only. A popular approach is to look for curves which are parameterised functions of time-to-maturity. For example if a model's goal is to estimate the $d(t)$ discount curve one should look for a $d(t; a, b, ...)$ parameterised function of time-to-maturity with parameters $a, b, \ldots$ etc. These are called parameterised yield curve models. A big advantage of using parameterised yield curve models is that they result in real curve and not just some set of individual estimated discount factors. An other advantage is that much less parameters are needed to be estimated than in the simple OLS regression case where there was a parameter for every cashflow time.

Parameterised yield curve models can be linear or non-linear. Linear curves can be expressed as the linear combination of some basis elements and so they are easy to optimise for a best fit. Non-linear curves can not be expressed like that. In linear models it is much easier to obtain a best fit because only the linear coefficients should be estimated and OLS regression provides a good way of doing this. For a given set of $\theta_i(t)$ basis functions where $i = 1, 2, \ldots, K$, the function looked for can always be expressed as:

$$d(t) = \sum_{i=1}^{K} \lambda_i \theta_i(t) \tag{7}$$

Now a vector of parameters $(\lambda_1, \ldots, \lambda_K)$ determines a $d(t)$ function so one has to estimate these parameters only to obtain a yield curve.

With the use of $\Lambda = (\lambda_1, \ldots, \lambda_K)$ and Θ as a matrix like that $\Theta(i, j) = \theta_j(t_i)$ where $i = 1, \ldots, n$ and $j = 1, \ldots, K$, the vector $D = (d(t_i), \ldots, d(t_n))$ can be expressed as:

$$D = \Theta\Lambda$$

According to the basic problem the following also stands:

$$P = CD + \varepsilon$$

Now with the use of the notation:

$$\widehat{D} = C\Theta, \tag{8}$$

the linear problem can be expressed as:

$$P = \widehat{D}\Lambda + \varepsilon \tag{9}$$

The solution can be easily obtained with the use of OLS regression by minimalizing $\varepsilon^T\varepsilon$. It is:

$$\Lambda^* = (\widehat{D}^T\widehat{D})^{-1}\widehat{D}^T P \tag{10}$$

Once the Λ vector of the λ_i coefficients are estimated, the discount function, $d(t)$ can be obtained with the use of Equation (7) .

3.5 Polynomial Estimation

One of the easiest ways to model the yield curve is to look for a polynomial function that fits the data most accurately but still has the characteristics of realistic yield curves. High degree polynomials would fit the data probably very well but they usually result in unrealistic shapes. The most popular way is to find a cubic polynomial that fits the data well. Polynomial yield curve estimations are linear yield curve models since every n-th degree polynomial can be expressed as the linear combination of the x^j power functions for $j \in \{0, \ldots, n\}$. Using the x^j power functions as a basis, the discount curve is:

$$d(t) = \sum_{j=0}^{n} \lambda_j x^j(t)$$

With those basis functions Θ's elements are: $\Theta(i,j) = x^j(t_i)$.
The solution can be obtained using Equation (8), Equation (9) and Equation (10). While in some rare cases polynomial yield curve estimations lead to quick and good solutions

they are not the best choice most of the times. A more appropriate model is spline approximation.

3.6 Spline Yield Curve Models

Using splines to model the yield curve can lead to good results. Spline yield curve models are linear in the coefficients so they are easy to fit. Splines are piecewise polynomial functions with some restriction on the derivatives. It is called knot points where the different polynomial parts meet. If the spline's domain is a closed interval the ends are sometimes called knot points too. This section considers the ends as knot points too. A k-th degree spline is a function that is a k-th degree polynomial between the knot points and $k-1$ times differentiable everywhere. Because of that a spline of order m with n knot points needs only $n+m-1$ parameters to be fully determined. The polynomial on the first interval is defined by $m+1$ parameters and all the other parts need only one more since their derivatives must be the same in the knot points. There are $n-2$ parts other than the first, so this gives $m+1+n-2=m+n-1$. As mentioned spline models are linear in the coefficients so first let's introduce the basis functions. A good choice is to use basis splines or B-splines. It is possible to define the basis splines recursively. If a defined set of knot points is known: $\{P_1, \ldots, P_n\}$, the definition of B-splines of order zero is the following:

$$B_{0,i}(t) = \begin{cases} 1 & \text{if } P_i \leq t < P_{i+1} \\ 0 & \text{else} \end{cases}$$

In this definition the lower index i refers to the i-th basis spline available for definition on the set of knot points in an incremental order. Starting from the basis spilnes of order zero, it is possible to define B-splines of higher order recursively:

$$B_{m,i}(t) = \frac{t-P_i}{P_{i+m}-P_i} B_{m-1,i}(t) + \frac{P_{i+m+1}-t}{P_{i+m+1}-P_{i+1}} B_{m-1,i+1}(t),$$

where the lower index m refers to the order (also called degree) of the spline.

$B_{m,i}(t)$ is a spline function of order m that is zero outside the $[P_i, P_{i+m}]$ interval. Because of these on a set of n knot points $n-m-1$ basis splines of order m are available.

In Appendix 1, I give an example of B-splines on a set of knot points.

As mentioned every m-th degree spline with n knot points on a closed interval needs $n+m-1$ parameters to be determined. So for example a cubic spline with knot points $\{P_1, \ldots, P_n\}$ needs $n+2$ parameters and so $n+2$ basis functions to determine the spline. There are $n-4$ B-splines that belong to the knot points so there are 6 more basis functions required. To get those 6 more 'out of interval' knot points should be defined such as: $P_{-2} < P_{-1} < P_0 < P_1 < \cdots < P_n < P_{n+1} < P_{n+2} < P_{n+3}$. This new set of knot points determines $n+2$ B-splines. Now the basis can be this $n+2$ basis splines restricted to the interval $[P_0, P_n]$. The $d(t)$ discount function can be expressed now as the linear combination of the basis functions:

$$d(t) = \sum_{i=-2}^{n-1} \lambda_i B_{3,i}(t)$$

Let's use the notations $\Lambda = (\lambda_{-2}, \lambda_{-1}, \ldots, \lambda_{n-1})$ and B as a $k \times (n+3)$ sized matrix like that $B(i,j) = B_{3,j}(t_i)$, where there is k cash flow times. If D denotes the vector $(d(t_1), \ldots, d(t_k))$ the following equation stands:

$$D = B\Lambda$$

Now denoting $\widehat{D} = CB$ and minimizing the term $\varepsilon^T \varepsilon$ using OLS techniques, the solution is:

$$\Lambda^* = (\widehat{D}^T \widehat{D})^{-1} \widehat{D}^T P$$

This is called the solution of the unconstrained problem. If $d(0) = 1$ is required it is called the constrained problem. The constrained problem gives a better fit on the short end of the yield curve by fixing the first value to 1. This is a reasonable requirement in accordance to practice.

By denoting the vector of the B-splines values at 0 as W, like that:

$$W = (B_{3,-2}(0), B_{3,-1}(0), \ldots, B_{3,n-1}(0))$$

and set $w = 1$, the constraint becomes the following additional requirement in the problem:

$$W\Lambda = w$$

Using this condition the problem can be expressed as:

$$\Lambda^{**} = \min_{\Lambda} \left\{ \varepsilon^T \varepsilon \middle| \varepsilon = P - \widehat{D}\Lambda, W\Lambda = w \right\}$$

And the solution of this problem is:

$$\Lambda^{**} = \Lambda^* - (\widehat{D}^T \widehat{D})^{-1} W^T (W (\widehat{D}^T \widehat{D})^{-1} W^T)^{-1} (W\Lambda^* - w).$$

The results of spline methods like these rely heavily on the selection of the knot points. One suggestion made in the literature is to try to position the knot points so there are similar number of observations between them. This produces better fits or more realistic curves usually but sometimes there are exceptions when it is better to try an other starting set of knot points.

3.7 Smoothing Conditions

The OLS method used for spline approximations minimized the following criterion with changing the Λ parameter:

$$c_1 = (P - CB^T\Lambda)^2$$

This c_1 term refers to the goodness-of-fit of the curve only in consideration of the distance between the fitted curve and the observed prices. A smaller c_1 value means the observed values are closer to the fitted curve. On one hand this a desirable property but on the other hand it can carry some adverse effects on other important properties of the yield curve such as smoothness. It is showed by Díaz and Skinner that significant liquidity and tax induced errors occur when modelling corporate yield curves compared to Treasury yield curve models [9]. The universe of corporate bonds are less heterogenous by nature than governmentals. Sometimes it is better to have a yield curve with less curvature than to try to overfit the values. It can be reasonable to use some further criterion that controls the smoothness of the estimated curve. These are called Smoothing Criterions and there are several different ones. The total curvature of an estimated $d(t)$ curve between P_0 and P_n can be measured as:

$$c_2 = \int_{P_0}^{P_n} \left(\frac{\partial^{m-1} d(t)}{\partial t^{m-1}}\right)^2 dt$$

for an m-th order spline approximation.
An m-th order spline is called natural if the following is true:

$$\frac{\partial^{m-1} d(t)}{\partial t^{m-1}}\bigg|_{P_0} = \frac{\partial^{m-1} d(t)}{\partial t^{m-1}}\bigg|_{P_n} = 0$$

The natural smoothing criterion is the following:

$$\Lambda^* = \min_{\Lambda}\{c_2 | c_1 < S\}$$

This criterion minimizes the curvature of a fit while keeping c_1's value under S. A bigger S value places more importance on smoothness and less on the close fit of the values while a smaller S emphasizes the fit more. The solution of the natural smoothing criterion problem is a $(2m-3)$-th order natural spline with knots at the data points. That is why it is called the natural smoothing criterion.

Fisher, Nychka and Zervos's studies show that better fits can occur if the forward rate curve's curvature is restricted and not the discount function's [4]. The following is the smoothing spline criterion:

$$c_3 = \Theta \int_0^T \Big(\frac{\partial^2 h(t|\Lambda, \bar{\xi})}{\partial t^2}\Big)^2 dt + (P - Cd(t|\Lambda, \bar{\xi}))^2.$$

where h is an invertible function of the yield curve, $\bar{\xi}$ is the set of knot points: $\bar{\xi} = (P_{-3}, P_{-2}, \ldots, P_{n+3})$ and Θ is called the smoothing parameter. Now it is possible to use the curvature for example of the yield curve, the forward rate curve or the discount curve as a basis for the approximation. Θ carries the relative importance between the curvature and the fit in this approach. A large Θ lays more weight on the first part of the criterion. An extension is to let Θ vary by time. This method was proposed by Waggoner [11]. It can be reasonable since yield curves usually have much more curvature in the short end than in the long. Using the time-varying Θ parameter, the criterion becomes:

$$c_3 = \int_0^T \Theta(t) \Big(\frac{\partial^2 h(t|\Lambda, \bar{\xi})}{\partial t^2}\Big)^2 dt + (P - Cd(t|\Lambda, \bar{\xi}))^2.$$

3.8 Non-linear Models

There are some popular non-linear yield curve models. The most widely used is the Nelson-Siegel model. This is a four parameter model that is used to estimate the forward rate curve. It is very popular thanks to the relatively low parameter number and easy computation. In spite of its small parameter number it can capture most of the yield curve shapes. The disadvantage of the Nelson-Siegel model lies in precision. It is not a good choice if very accurate estimations are required or when one is trying to model a complex yield curve. The four parameters are: $\beta_0, \beta_1, \beta_2, k$. The forward rate curve estimation is:

$$f(t) = \beta_0 + (\beta_1 + \beta_2 t)e^{-kt}$$

From this expression of the forward rate curve the spot curve can be written as:

$$r(t) = \beta_0 + \left(\beta_1 + \frac{\beta_2}{k}\right)\frac{1 - e^{-kt}}{kt} - \frac{\beta_2}{k}e^{-kt}.$$

The short rate $r(0)$ equals $\beta_0 + \beta_1$ and the long rate is $\lim_{t\to\inf} r(t) = \beta_0$. The other two parameters β_2 and k can adjust the height and location of a hump in the curve.
An extension of the original Nelson-Siegel model is the Svensson model. It is sometimes called Nelson-Siegel-Svensson model since it is the original model with two more parameters:

$$f(t) = \beta_0 + (\beta_1 + \beta_2 t)e^{-k_1 t} + \beta_3 t e^{-k_2 t}$$

Wiseman [12] introduced a $2(n+1)$ parameter model to estimate the forward curve. It consists of $n+1$ exponential decay terms with β_i coefficients:

$$f(t) = \sum_{i=0}^{n} \beta_i e^{-k_i t}$$

4 Additional Features

4.1 Corporate Bond Valuation using Credit Spread

There was a very important criterion of the statistical yield curve models of the previous section and that is the availability of data. All of the models are based on the information a carefully gathered dataset of similar bonds contain. If for example one wants to obtain a yield curve to use for pricing a bond, the starting set of information should consist of other bonds with similar credit risk, liquidity, currency and country of the bond that ought to be priced. That way the resulting yield curve provides a useful tool for pricing the bond. If the information basis of the model is not chosen well it will not provide an appropriate tool for valuation. But what if there aren't enough observable information or simply there aren't enough time and resources to use a statistical yield curve model? Situations can occur where only a quick valuation is needed and there isn't enough time to carefully select the best datasets and precisely adjust statistical models. And it can also happen that there is too little information available on a financial product or issuer company. There are ways to estimate a yield curve and to evaluate a security without the need to construct statistical yield curve models.

The second most important type of risk affecting fixed income securities is credit risk (after interest-rate risk). Credit risk, more importantly the perceived credit risk of investors, is an important factor that shapes the yield curve and affects the price of a security affected by credit risk. When a company's perceived credit risk grows, that means investors think that default is more probable than before. Investors can analyze individually the credit risk of a company but the vast majority relies on the opinion of big credit rating agencies. The biggest three are Standard and Poor's, Moody's and Fitch Group. If a company becomes downgraded by one of these companies, it means its creditworthiness has worsened and this implies that a larger yield margin is expected by investors to buy the bonds' of that company. Skinner and Díaz (2001) showed that it is reasonable to pool bonds by broad rating agency categories when constructing statistical yield curve models. Doing this doesn't affect the results significantly in a bad way. They also found that corporate yield curve models contain significant liquidity and tax-induced errors and thus they are less reliable than Treasury yield curve estimates. [9]

There are some rates in the markets considered (credit) risk-free. The U.S. Treasuries's yields as well as the LIBOR swap rates are usually the most popular when referring to risk free rates. They contain nearly zero credit risk and that's why they are usually used as benchmark rates. A not too precise approach is that corporate bonds differ from risk free securities like U.S. Treasuries only in their perceived credit risk. Of course it is not true since for example the U.S. Treasuries are usually much more liquid products than some corporate bonds, but this approach can provide an easy and quick way with relatively good results when estimating a yield curve or pricing products. If credit risk is the only difference in a model from the risk-free securities it is reasonable to use the risk-free yield curve as a starting point and to transform it somehow to cope with the extra credit risk. A big advantage of this approach is that precise statistical yield curves for risk-free interest rates like U.S. Treasuries or LIBOR swap rates are usually available. Now one only has to use some transformations to take in account the credit risk. One of the easiest ways is to use a paralell shift of the original risk free curve of to get the yield curve of the credit risky product. The difference between the yield of a credit risky product and a risk-free product of the same maturity is called credit spread.

4.2 Pricing with CDS

If there are CDSs' available for a company, this provide a way of pricing and modelling the yield curve of bonds of that company. Credit Default Swaps are contracts that provide insurance against the risk of default by a particular company [5]. The buyer of the CDS pays money periodically to the seller and the seller pays money if the company (or government) defaults. It is called credit event and the details are fixed in the contract. The CDS has a face or notional value. In the case of a credit event there are two popular procedures. One is that if a credit event happens the CDS buyer gives the secured product to the CDS seller while the seller pays the face value to the buyer. The other case is a cash settlement payed to the buyer and the amount is usually determined by a calculation agent or auction mechanism.

The periodical payments (usually annual or semiannual) the buyer pays to the seller is computed using the face value and the credit default swap spread (or CDS rate):

$$P_{CDS} = FV \times CDS_{spread}$$

Here P_{CDS} refers to the annual cost of the CDS, FV denotes Face Value and CDS_{spread} is the credit default swap spread. The CDS spread of CDS rate is the factor that contains the perceived credit risk of the company by the market. If market participants think a company is becoming more credit risky the CDS rate rises while the opposite happens if the company becomes more stable in the eyes of the public.

Credit Default Swaps have a maturity also. That is the time horizon when the contract holds. Until maturity the buyer will pay the payments to the seller and the seller provide security. The only case the contract terminates before maturity is a credit event. If that occurs the buyer doesn't have to continue paying the periodical payments. CDS spreads can be different on different time horizons and this way there is a time structure of CDS spreads too. While it looks obvious that companies with lower credit ratings usually have higher CDS spreads the relation between CDS maturity and the spread of the same company is not that easy to see. Research on Eurobonds and domestic bonds from EU-countries shows that the amount of the spread and maturity have a positive relationship in the case of investment grade debt. [6]

Bond pricing with the use of CDS rates relies on no arbitrage considerations. If an investor buys a bond and a CDS with maturity equal to the bond's maturity and face value equal to the bond's face value this portfolio theoretically becomes a risk-free portfolio. Then this portfolio should have a return equal to the return on a risk-free product with the same maturity. Otherwise arbitrage opportunities would arise as the arbitrageur buys the cheaper product and sells the more expensive contemporarily.

This no-arbitrage consideration provides an approach of constructing yield curves and pricing credit risky bonds. The credit spread on bonds should be equal to the CDS spread of the same bond and maturity. Using this, the yield curve of a credit risky product should be a benchmark risk-free yield curve plus the term structure of CDS spreads of the product. Figure 4 shows an arbitrary example of a risk-free yield curve, a credit spread time structure and the yield curve affected by this credit spread.

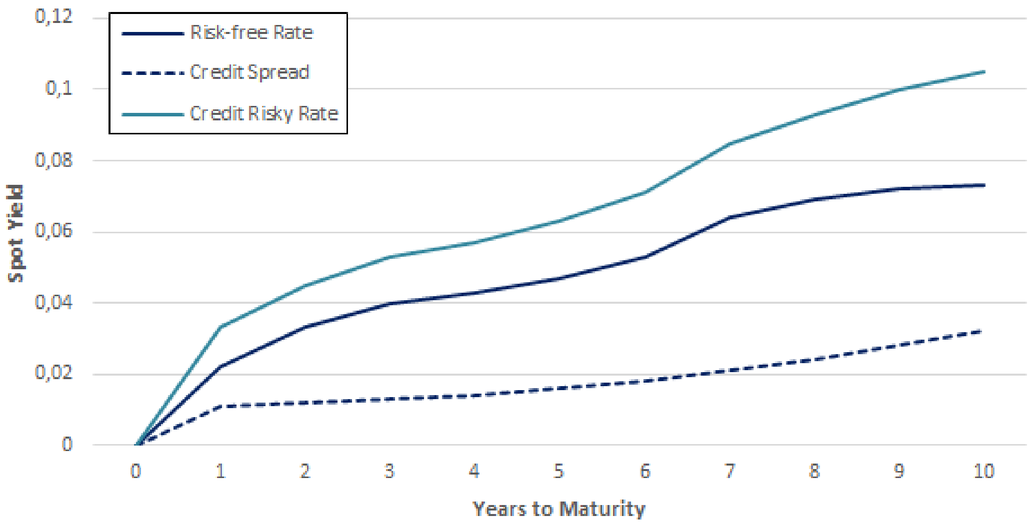


Figure 4: An example of the risk-free yield curve, the credit spread and the yield curve affected by credit risk as the sum of the first two.

4.3 Corporate Bond Spreads

There are other ways to estimate the credit risk of a product too without the use of CDS spreads. Credit Default Swaps are not always available in the market, so credit spread should be estimated from something else. If a bond with similar maturity and credit properties is observable in the market, one can derive the credit spread from that. An example is a bond of the same company or a very similar company (same size, geopraphical region, market sector, management stlye, leverage...etc.). Now the credit spread of this bond can be derived from the market price, risk-free yield curve and properties of that bond. The credit spread will be the value the risk-free yield curve should be shifted with in order to reproduce the observable market price. It is a parallel shift of the risk-free yield curve upwards. If CS denotes the credit spread, $r_f(t)$ is the risk-free interest rate, P_c is the market price of the coupon bond that matures n years from now and pays $\$C$ coupon annually and a principal of \$100 at maturity, the following stands:

$$P_c = \sum_{t=1}^{n} Ce^{-(r(t)+CS)t} + 100e^{-(r(n)+CS)n}$$

If the credit spread is obtained this way, it can be used to construct the yield curve and price of the product where it is needed. The yield curve can be constructed two ways. One of them is to simply use a paralell shift similar to the computed credit spread on the risk-free curve. It is a quick and easy way to have a new yield curve that can be used to price the product. The other way is to estimate a time structure of credit spreads and then add this to the risk-free yield curve. The time structure can be obtained by evaluating bond spreads on different maturities. Of course one should start the process from similar type of bonds in credit properties.

4.4 Liquidity

An other important factor that affects the prices and yields of financial products is liquidity. There are different definitions and measures of liquidity. Generally a highly traded product is called liquid. Liquid products are easy to buy and easy to sell in short period of time without the deterioration of their value. They can be converted to cash quickly. Illiquid products cannot be sold quickly without some waste in their value. In stock markets one way to measure liquidity is observing the bid-ask spread of a product. A relatively big spread means a smaller incoming cashflow when one wants to sell the product quickly so it is considered illiquid while a small spread refers to a liquid product.

Liquidity carries an extra value for investors. Illiquid products usually have higher yields than liquid ones for compensation. U.S. Treasuries are one of the most liquid financial products available so they are usually used as a benchmark for valuation. Even among Treasuries liquidity carries an extra value, this causes the slight difference between the yield to maturities of on-the-run and off-the-run treasuries. The most lately issued treasuries are called on-the-run and the ones that are in the market for more time called seasoned or off-the-run. New issues are usually traded more than old ones, that causes the higher liquidity of them. Figure 5 shows the slight difference between the yields of U.S. Treasuries due to the liquidity of the product. On-the-run securities have a bit lower yield to maturity than seasoned ones:

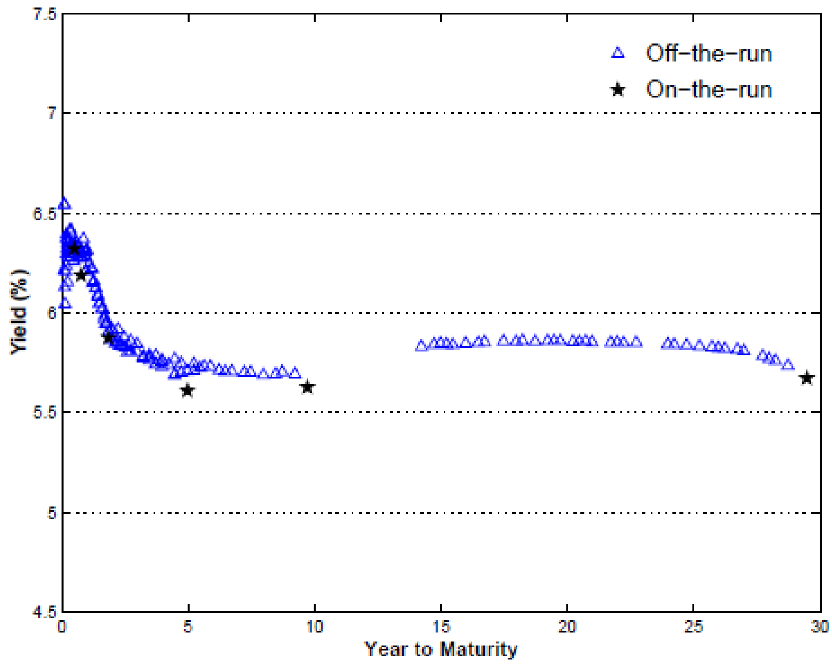


Figure 5: Yield to maturities' of on-the-run and off-the-run treasuries. [10]

When using this as the risk-free rate for calculating credit spreads, the results do not only represent the risk of default of the issuer, it also incorporates the yield premium arising from the much less liquidity on the secondary markets of corporate bonds. In a research, Longstaff [7] et al. (2005) shows that the default component explains only 50% of the spread between the yields of Aaa/Aa rated bonds and treasuries. In the case of Baa rated bonds it is 70%. The remaining part is mostly explained by the liquidity factor.

When one wants to estimate yield curves appropriate for evaluating a bond the mentioned ways starting from a risk-free yield curve can all lead to relatively good results. However, Feldhütter suggests that the swap rate is the best as a benchmark risk-free rate for corporate bond credit spread estimations. [8] The research also concludes that corporate bonds spreads are a better estimator of credit spread then CDS-spreads since the latter can be disconnected from the credit component at times. One example according to them is the Greek CDS and bond spread at 24 June, 2010. While CDS spread has hit record high, the bond spreads didn't reacted that sensitively probably because of the support from the European Central Bank bond buying program. Since it is possible to buy CDSs without owning the bond of the subject company (this is called Naked CDS contract) there can be speculative motivations that drive the prices. Even if the bonds' cash flows are protected some way and thus the values of them the credit event can happen and CDS owners can make profit from that. Figure 6 shows the Greek CDS and bond spread around 24 June, 2010.

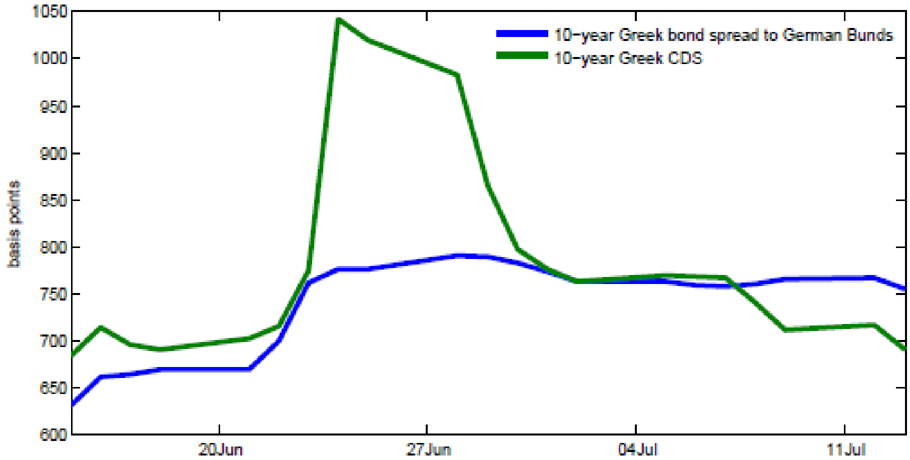


Figure 6: Historical data of the Greek CDS and bond spread values [8]

5 Applications

5.1 Data Analysis

The data I have used is corporate bond data from Bloomberg. The plan was to focus on CHF denominated bonds, altough the models are appropriate for any other currencies of course. Swiss Bond Index was the starting point of collecting a representative sample of bonds. The index consisted of 1276 bonds on the 30th September 2014. I used Microsoft Excel and its Bloomberg extension to obtain the appropriate data.

From the known ISIN values I requested the following informations of the bonds:

- **NAME:** name of the issuer company
- **PX_LAST:** last quoted market price of the bond
- **COUPON:** coupon rate
- **CPN_FREQ:** frequency of the coupon payments
- **COUNTRY_FULL_NAME:** country of the company
- **INDUSTRY_GROUP:** industry group
- **RTG_MOODY:** Moody's rating of the issuer company

The bold notations refer to the Bloomberg variables used in Excel. The data shows the states of 14th, November 2014. It is observable in the spreadsheet 'ISINs from SBI' in the attached Excel-file, titled data.xlsm. Since the bond index doesn't contain all bonds of the companies included I needed some more work on that. First I separated the bonds with the industry group 'Banks' as a focus of the analysis. From the original 1276 bonds in the bond index 634 was in the 'Banks' industry group. I further narrowed the results by centering the attention only on companies rated Aaa by Moody's. This way the set of 634 products decreased to 269. The next step was to collect all other bonds of the issuer companies of these 269 products. I used the BOND_TO_EQY_TICKER function to transform the ISIN to the Equity Ticker and after that the BOND_CHAIN function to list all bonds available of the company the ticker denotes. It is in spreadsheet 'All bonds of Aaa Banks'. After using a simple VBA-code to position the results I got 49998 bonds in a column. Of course there were a lot of repetitions which I later corrected. For example if there was a company with more than one bond in the Bond Index all of the company's bonds appeared more than one times at this point. When finished removing duplicates with Excel 7086 different bonds remained.

After that I requested the same data mentioned for the bonds: name, last price, ...etc. In more than half of the 7068 bonds the result was not perfect because some or all of the information wasn't available from Bloomberg. I cleared the data from the imperfect ones and it resulted in 3029 different bonds with all the important information needed. The only type where I left missing data is the Moody's rating category since that type is not used in calculations directly. The spreadsheet 'All bonds of Aaa Banks 2' contains the results. Figure 7 shows the first few elements of the resulting dataset.

	A	B	C	D	E	F	G	H	I
1			NAME	PX_LAST	MATURITY	COUPON	CPN_FREQ	COUNTRY_FULL_NAME	RTG_MOODY
2	RD675300 Corp	BYLAN 3.500 11/14/14	BAYERISCHE LANDESBANK	99,98	2014.11.14.	3,5	0	GERMANY	#N/A N/A
3	RD673205 Corp	BYLAN 3.800 11/14/14	BAYERISCHE LANDESBANK	99,98	2014.11.14.	3,8	0	GERMANY	#N/A N/A
4	RD675305 Corp	BYLAN 4.300 11/14/14	BAYERISCHE LANDESBANK	99,99	2014.11.14.	4,3	0	GERMANY	#N/A N/A
5	EF682388 Corp	BYLAN 4.000 11/14/14	BAYERISCHE LANDESBANK	100	2014.11.14.	4	1	GERMANY	WR
6	EI871962 Corp	BNG 1.000 11/17/14	BK NEDERLANDSE GEMEENTEN	100,06	2014.11.17.	1	2	NETHERLANDS	Aaa
7	ED693240 Corp	KFW 4.500 11/17/14	KFW	100	2014.11.17.	4,5	2	GERMANY	Aaa
8	EJ936772 Corp	NDASS 0.000 11/17/14	NORDEA BANK FINLAND NY	100	2014.11.17.	0,2236	4	FINLAND	Aa3
9	EJ679073 Corp	SHBASS 0.000 11/17/14	SVENSKA HANDELSBANKEN NY	100	2014.11.17.	0,4011	4	SWEDEN	Aa3
10	EI893289 Corp	PFZENT 0.125 11/18/14	PFANDBRIEF SCHW KANTBK	99,995	2014.11.18.	0,125	1	SWITZERLAND	Aaa
11	EC741830 Corp	PFZENT 3.250 11/18/14	PFANDBRIEF SCHW KANTBK	100,005	2014.11.18.	3,25	1	SWITZERLAND	#N/A N/A
12	EC741830 Corp	PFZENT 3.250 11/18/14	PFANDBRIEF SCHW KANTBK	100,005	2014.11.18.	3,25	1	SWITZERLAND	#N/A N/A
13	EI049111 Corp	DEKA 0.000 11/18/14	DEKABANK	100	2014.11.18.	0,603	2	GERMANY	#N/A N/A
14	EH012174 Corp	CFF 4.375 11/19/14	CIE FINANCEMENT FONCIER	100,005	2014.11.19.	4,375	1	FRANCE	Aaa
15	EI051573 Corp	CM 3.300 11/19/14	CANADIAN IMPERIAL BANK	100,006	2014.11.19.	3,3	2	CANADA	Aa3
16	ED702347 Corp	BYLAN 4.000 11/20/14	BAYERISCHE LANDESBANK	100,0255	2014.11.20.	4	1	GERMANY	Aaa

Figure 7: The first few elements of the dataset after collecting and clearing it.

I have compared two types of linear yield curve models: polynomial approximation and spline estimation. To be able to use the Ordinary Least Squares method I needed a cash-flow matrix of the selected bonds. I have made a VBA code to construct a cashflow matrix from a selection of bonds. The code is in the attached data.xlsm file. The process starts with collecting the maturities of the bonds into an array. Once it is done a new array is made that contains all the dates when there is a cash flow from any one bonds of the starting set. The date when the data is observed is included here even if there is no payment. This date has to be an input in a cell and the code reads it from there. This much bigger array is constructed with the use of the maturities and coupon frequency data. I used a bubble sort algorithm to sort these dates in an incremental order and the algorithm also unifies repeating dates. The cash flow matrix's column will present the cash flows of the products on these dates. The rows will be associated with the bonds. I have set the matrix's elements to zero and after that filled with the appropriate cash flows. In the case of every bond the algorithm started with the principal payment and then moved backwards by the coupon payment frequencies to fill the selected cells with the appropriate coupon payment amount. The selection of bonds I analyzed is Aaa rated bonds of companies from Switzerland. Figure 8 shows an excerpt from the Cash-Flow matrix result of the algorithm. The code also constructs three more arrays on different spreadsheets. The vector of the last quoted prices, the incremental cash flow dates in a column and the vector of the maturities. I found it easier to gather this four important type of arrays from different spreadsheets when it came to Python implementation.

	A	B	C	D	E	F	G	H	I	J
1		2014.11.14	2014.11.18	2014.12.03	2014.12.15	2014.12.20	2015.01.25	2015.01.30	2015.02.23	2015.03.15
2	PFZENT 0.125 11/18/14	0	100,125	0	0	0	0	0	0	0
3	PFZENT 2.250 12/20/14	0	0	0	0	102,25	0	0	0	0
4	PSHYPO 3.000 01/30/15	0	0	0	0	0	0	103	0	0
5	PSHYPO 1.500 02/23/15	0	0	0	0	0	0	0	101,5	0
6	PFZENT 2.625 03/15/15	0	0	0	0	0	0	0	0	102,625
7	PFZENT 0.375 03/16/15	0	0	0	0	0	0	0	0	0
8	PSHYPO 2.500 04/10/15	0	0	0	0	0	0	0	0	0
9	PSHYPO 0.250 06/19/15	0	0	0	0	0	0	0	0	0
10	PSHYPO 2.670 06/19/15	0	0	0	0	0	0	0	0	0
11	PSHYPO 3.625 06/19/15	0	0	0	0	0	0	0	0	0
12	PFZENT 2.500 06/30/15	0	0	0	0	0	0	0	0	0
13	PFZENT 2.000 09/15/15	0	0	0	0	0	0	0	0	0
14	PSHYPO 1.875 09/28/15	0	0	0	0	0	0	0	0	0
15	PSHYPO 1.000 10/15/15	0	0	0	0	0	0	0	0	0
16	PSHYPO 3.240 10/15/15	0	0	0	0	0	0	0	0	0
17	PSHYPO 3.680 10/15/15	0	0	0	0	0	0	0	0	0
18	PFZENT 3.250 11/02/15	0	0	0	0	0	0	0	0	0
19	PSHYPO 1.125 12/03/15	0	0	1,125	0	0	0	0	0	0
20	PFZENT 2.250 12/15/15	0	0	0	2,25	0	0	0	0	0
21	PSHYPO 0.250 01/25/16	0	0	0	0	0	0,25	0	0	0
22	PFZENT 2.500 03/30/16	0	0	0	0	0	0	0	0	0
23	PSHYPO 1.125 04/25/16	0	0	0	0	0	0	0	0	0

Figure 8: A part of the cash flow matrix constructed from Aaa rated bonds from Switzerland.

5.2 Modeling in Python

After gathering and managing the data to obtain the appropriate cashflow matrix and arrays of prices, maturities and cashflow dates I continued by constructing the models in Python. This software was suggested by my Swiss co-supervisor, as they use it in their daily work. I have tried two different methods mentioned in sections (3.5) and (3.6) to fit a yield curve to the data: polynomial and spline approximations. Both are linear yield curve models and I have used the OLS method when minimizing the error terms.

I used the following packages under Python: `numpy` to handle the array datatype and `pandas` to import the data from the excel file into Python. I also used the `math` and `matplotlib` packages. After importing the data from the excel file I transformed the arrays of dates to numbers that represent the true annualized time to maturites from 14th November. (Excel stores dates as numbers that represent days from 1st January, 1900.)

For the polynomial model I have made a function:
`polynomial (cf, prices, dates, deg, con)` , with five arguments:

- **cf:** The cashflow matrix as a numpy array datatype.
- **prices:** The market prices of the bonds as a numpy array datatype.
- **dates:** The cashflow dates in an ascending order as a numpy array datatype.
- **deg:** The degree of the polynomial estimation.
- **con:** A parameter to decide whether a constrained (con=1) or an unconstrained (con=0) model is to be computed.

The return of the function is an array with the estimated discount function values on the cashflow dates. I used this to present the results in a diagram but the code can be easily modified to return for example the estimated vector of the λ_i parameters or the discount function's values on different time to maturity values. The algorithm is based on the OLS method mentioned in the 'Statistical Yield Curve Models' chapter of this thesis.

The function I wrote for spline fitting is `spline (matu, cf, prices, dates, deg, con)` . It has one more argument than the polynomial function and that is an array of the maturities already imported from the excel file. This vector is used when computing the knot points. I have written a subfunction `knotpoints (matu, deg)` that computes the vector of the knot points. There is three way the knot points can be calculated and the user can choose which way he or she prefers. The first is fully automated, the number of knot points is the square root of the number of observations rounded down to the nearest integer. The points are located so that there are approximately equal number of observations between every point. This approach is suggested by J. Huston McCulloch [14]. The second way is that the user inputs the number of knot points and then these are located automatically similarly to the first case. The third is fully manual. The location of the knot points are typed in manually one by one. Either way has been chosen the extension of the knot points is automatically made by the code using parameter *deg*. After the knotpoints are generated the `spline` function estimates the spline yield curve model with OLS technique and after that returns the arrray of the discount function's values on cashflow dates.

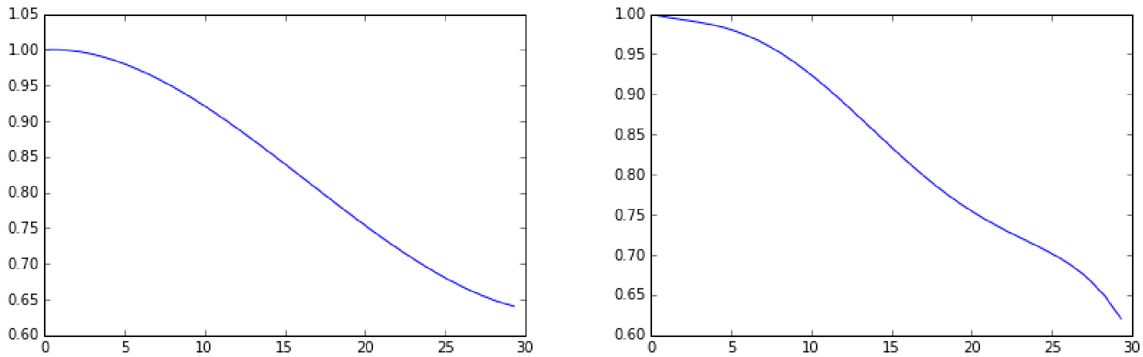
I have made two small functions: `yieldcurve (discount, dates)` and `forwardcurve (yieldcurve, dates)`. These are short codes that generate the spot and instantenous forward curves from the discount curve values or from the spot curve. The returns are the corresponding curve's values on the cashflow dates.

5.3 The results

I have used Aaa rated bonds of Swiss companies in the Bank industry to fit the statistical yield curves. A total of 175 bonds were used. I have compared the results by the outlook of the curves and the sum of squared errors of the fittings.

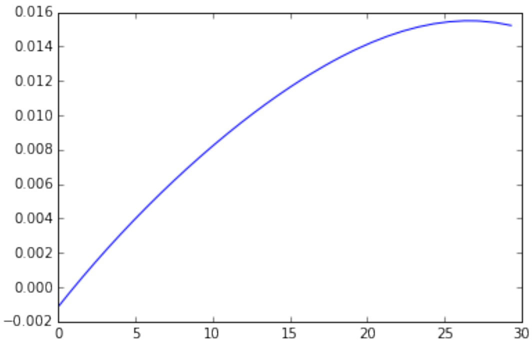
Polynomial Model: The discount function was modeled as a third and as a fifth degree polynomial function. The constrained versions (where a constraint ensured that $d(0) = 1$) resulted in much better short ends especially for the yield and forward curves. Of course this raised the error term a bit as constraints usually do. Figure 9 presents the resulting discount functions. The yield curves computed from the discount functions are presented by Figure 10.

Figure 9: Polynomial modeling of the discount function.

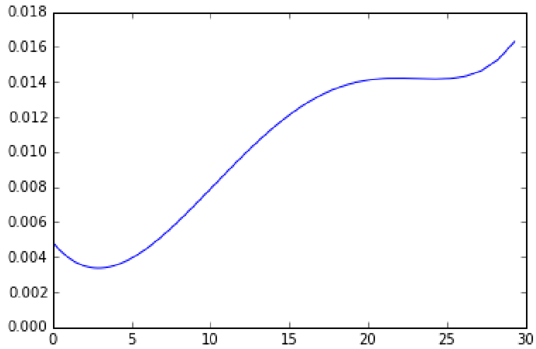


(a) 3rd degree polynomial. (b) 5th degree polynomial.

The yield curves computed from the discount functions are presented by Figure 10.



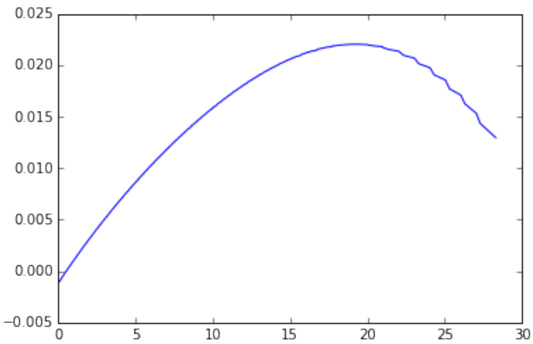
(a) Yield curve derived from the 3rd degree polynomial.



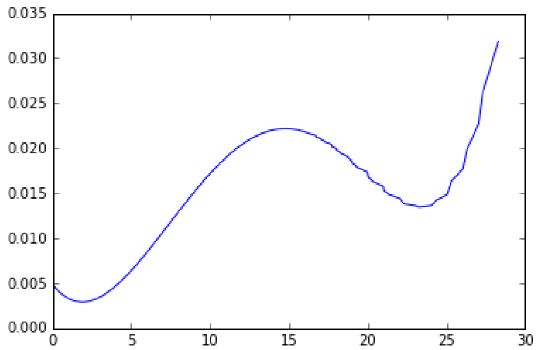
(b) Yield curve derived from the 5th degree polynomial.

Figure 10: Yield curves computed from the discount functions.

The forward curves calculated from the yield curves are presented by Figure 11.



(a) Forward curve of the 3rd degree discount function.

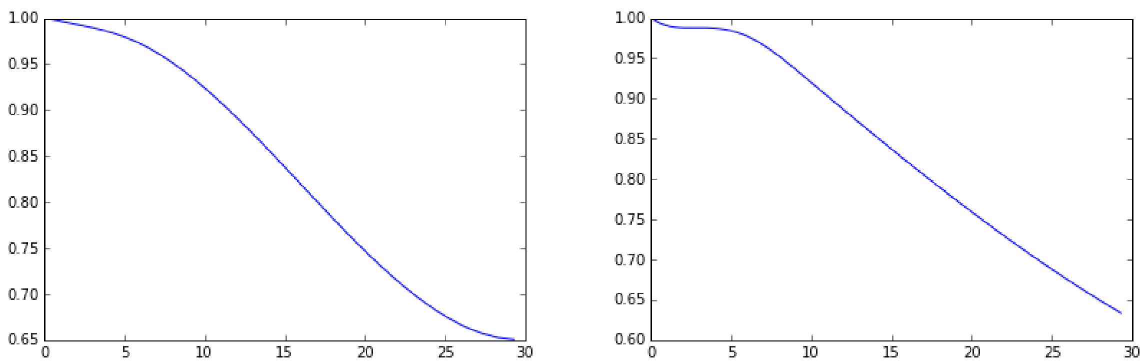


(b) Forward curve of the 5th degree discount function.

Figure 11: Forward curves calculated from the spot yield curves.

The sum of squared errors in the third degree case is 187.31. When fitting the fifth degree polynomial the error decreased to 155.12. On these figures it is apparent how some extra curvature in the discount function escalates when computing the yield and the forward curves. Altough higher degree polynomials produce smaller errors, sometimes the results' curvature may be considered too big to be realistic.

Spline Model: On the same data I examined two cubic spline models. Both of them with the constraint $d(0) = 1$ because those produced much more realistic results. The difference was the number of the knot points. In the first case a model with 4 knot points was approximated while in the second case a model with six knot points. The knot points were located automatically and the extensions were made automatically also with the use of the **knotpoints** function. Figure 12 shows the estimated discount functions.



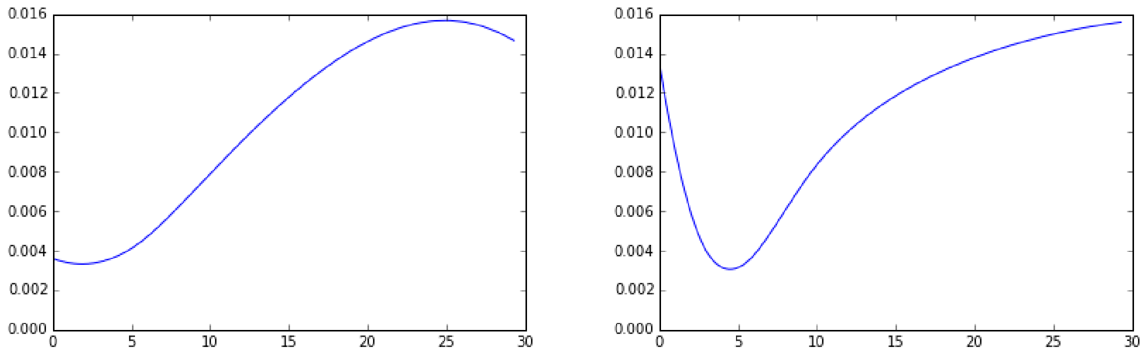
(a) Cubic spline with 4 knot points.

(b) Cubic spline with 6 knot points.

Figure 12: Cubic spline estimations of the discount function.

The yield curves estimated from the discount functions above are presented in Figure 13.

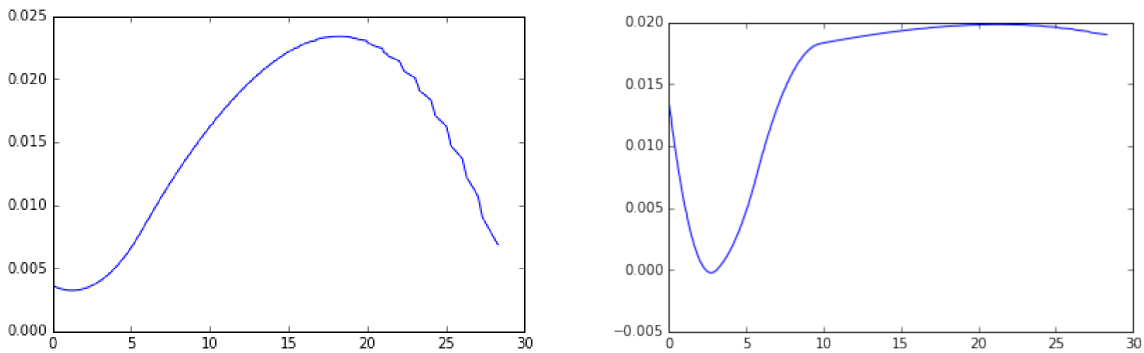
The forward curves are presented in Figure 14.



(a) Yield curve computed from the spline with 4 knot points.

(b) Yield curve computed from the spline with 6 knot points.

Figure 13: Yield curves computed from the spline discount function estimations.



(a) Forward curve of the spline with 4 knot points.

(b) Forward curve of the spline with 6 knot points.

Figure 14: Forward curves estimated from the yield curves.

The sum of squared errors of the fittings are the following. In the case of the cubic spline with 4 knot points the error was 170.56. The sum of squared errors in the case of 6 knot points was 130.27. By comparing the outlook of the curves and the error values in this four cases I conclude that the best fit was the cubic spline with 4 knot points on this set of bonds. Altough the one with 6 knot points resulted in smaller errors, the shapes of the yield curve and forward curve are not realistic.

6 Summary

I found that linear yield curve models are quite easy to implement and to work with in applications. They are not slow even when working with much bigger observations number and produced sensible results. The algorithm that slowed down the modeling is the part that computes the cashflow matrix from the list of the bonds in Excel. I plan to improve that, for example by changing the bubble sort algorithm to a quicksort, and by restructuring the algorithm. The excel file with the VBA function and the Python code saved as an ipython notebook file can be downloaded from:

https://dl.dropboxusercontent.com/u/100401660/data.xlsm

and

https://dl.dropboxusercontent.com/u/100401660/yieldcurve.ipynb

I have plans to implement some nonlinear models in Python, such as Nelson-Siegel or Svensson. And also to compare the different types of models with different statistical methods.

7 Appendix

7.1 Appendix 1: An example of B-splines

The following graphs show the B-splines belong to the set of knot points: $\{0, 2, 5, 8, 11, 14, 19\}$.

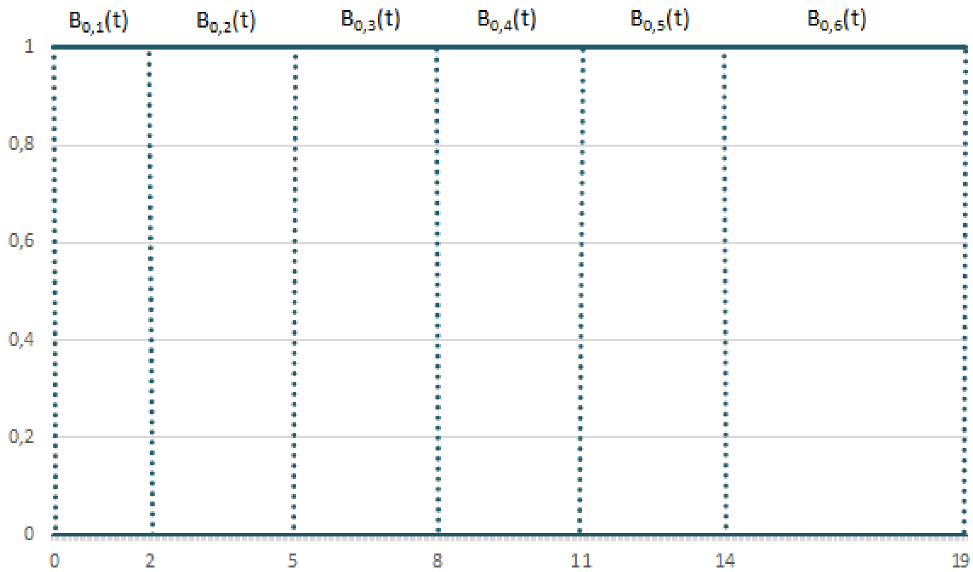


Figure 15: B-splines of order zero.

The basis splines of order one and two:

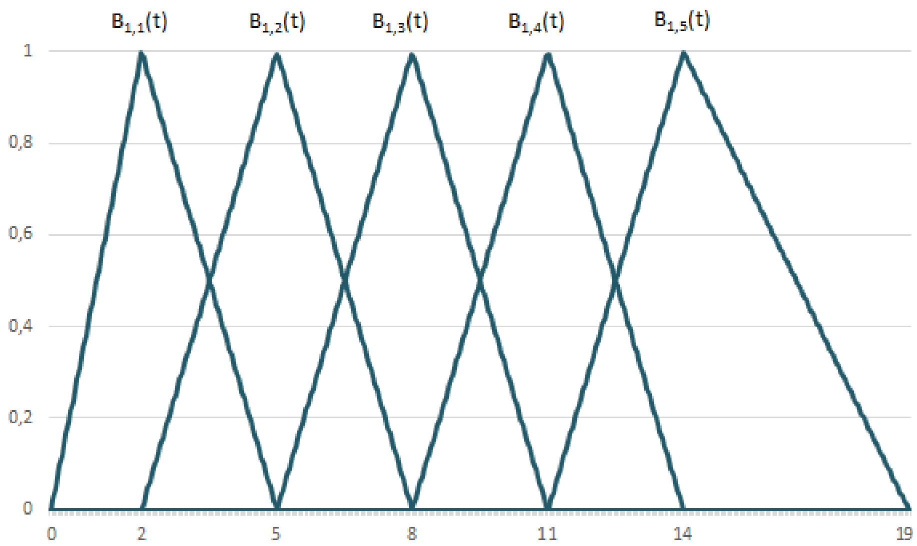


Figure 16: B-splines of order one.

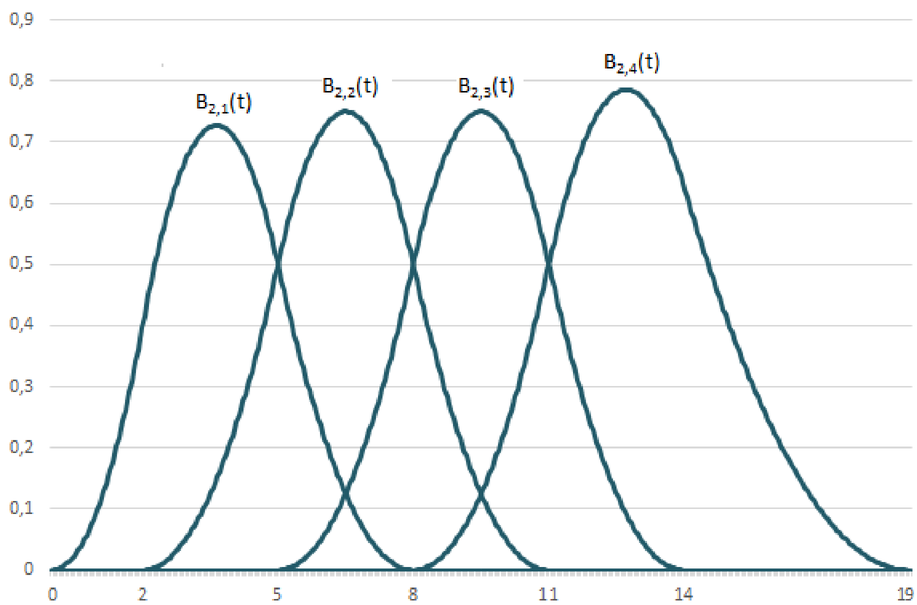


Figure 17: B-splines of order two.

The basis spline of order three:

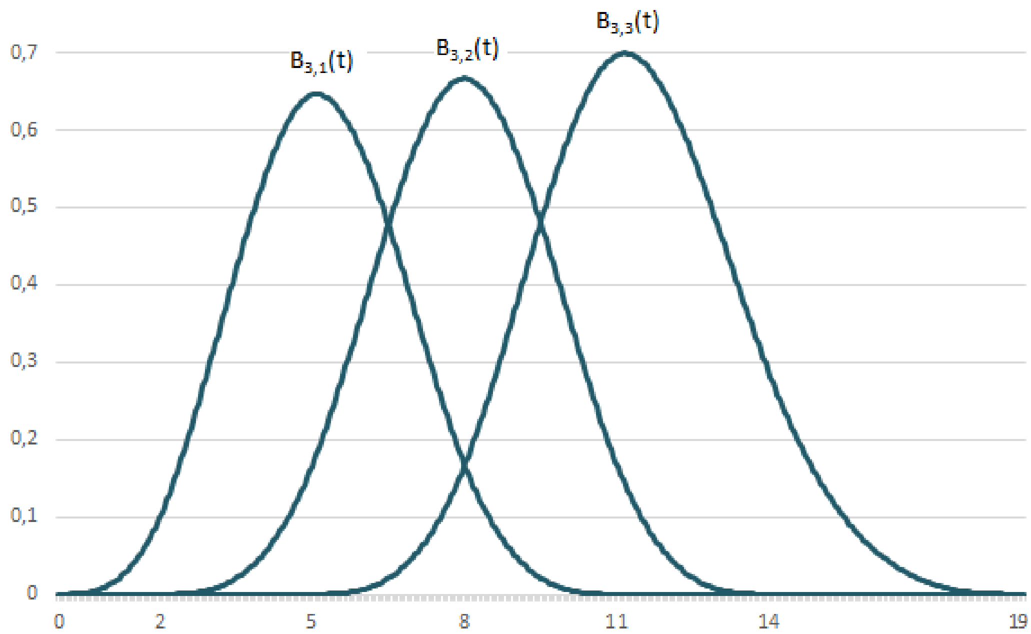


Figure 18: B-splines of order three.

References

[1] Mitzhaletzky György: *Interest Rate Models (university lecture, 2014)*

[2] Frank J. Fabozzi, Steven V. Mann: *Handbook of Fixed Income Securities (McGraw-Hill, 2005)*

[3] Jessica James, Nick Webber: *Interest Rate Modelling (John Wiley and Sons, 2000)*

[4] Mark Fisher, Douglas Nychka, David Zervos: *Fitting the term structure of interest rates with smoothing splines, (FEDS 95-1, 1994)*

[5] John C. Hull: *Options, futures and other derivatives (6th edition, 2005)*

[6] Stefan Trück, Matthias Laub, Svetlozar T. Rachev: *The Term Structure of Credit Spreads and Credit Default Swaps - an empirical investigation (2004)*

[7] Francis A. Longstaff, Sanjay Mithal, Eric Neis: *Corporate Yield Spreads: Default Risk of Liquidity? (2005)*

[8] Peter Feldhütter: *Where to look for credit risk? - The corporate bond market or the CDS market? (Annual Financial Market Liquidity Conference presentation, 2014)*

[9] Antonio Díaz, Frank Skinner: *Estimating Corporate Yield Curves (2001)*

[10] Jerry Yi Xiao: *Term Structure Estimations for U.S. Corporate Bond Yields*

[11] Daniel F. Waggoner: *Spline methods for extracting interest rate curves from coupon bond prices, 1997*

[12] Julian D. A. Wiseman: *The exponential yield curve model, 1994*

[13] Cox, J.C., J.E. Ingersoll, S.A. Ross: *A Theory of the Term Structure of Interest Rates. (Econometrica 53: 385–407, 1985)*

[14] J. Huston McCulloch: *The Tax-adjusted Yield Curve (Journal of Finance, 1975)*