COMENIUS UNIVERSITY IN BRATISLAVA
Faculty of Mathematics, Physics and Informatics
Department of Applied Mathematics and Statistics

Extraction of Nelson-Siegel Factors from Bond Prices

Master Thesis

9.1.9 Applied Mathematics
Economic and Financial Mathematics

Supervisor:
Mgr. Juraj Katriak

Author:
Bc. Róbert Berec

Bratislava 2010

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
Katedra aplikovanej matematiky a štatistiky

EXTRAKCIA NELSON-SIEGELOVÝCH FAKTOROV Z CIEN DLHOPISOV

Diplomová práca

9.1.9 Aplikovaná matematika
Ekonomická a finančná matematika

Vedúci diplomovej práce:
Mgr. Juraj Katriak

Diplomant:
Bc. Róbert Berec

Bratislava 2010

I declare this thesis was written on my own, with the only help provided by my supervisor and the refered-to literature.

Bratislava, April 23, 2010 ....

Róbert Berec

Acknowledgment

I would like to express my deep sense of gratitude to Mgr. Juraj Katriak for his invaluable help and guidance during the elaboration of this thesis. I would like to also thank to my family and Zuzka for their support during my study.

Názov práce:	Extrakcia Nelson-Siegelových faktorov z cien dlhopisov
Pracovisko:	Katedra aplikovanej matematiky a štatistiky, FMFI UK v Bratislave
Autor:	Bc. Róbert Berec
Vedúci DP:	Mgr. Juraj Katriak
Dátum:	23.04.2010
Kľúčové slová:	bonds, spot rate curve, Nelson-Siegel model

Abstrakt: Cieľom tejto diplomovej práce je v úvode teoreticky spracovať základný Nelson-Siegelov model a jeho dve modifikované prístupy, ktoré budeme v praktickej časti porovnávať. Túto praktickú časť vykonáme na sade nemeckých štátnych dlhopisov a buďe spočívať v porovnaní týchto troch metód, doplnených o dve váhované alternatívy, v schopnosti aproxímacie výnosovej krivky na jednej strane a v takzvanej out-of-sample analýze na strane druhej. Druhé spomínané porovnanie je prínosné najmä tým, že odráža schopnosť modelov teoreticky oceniť dlhopisy, ktoré neboli zahrnuté do odhadovania, a následne máme možnosť túto teoretickú cenu porovnať s reálnymi trhovými hodnotami. V praktickej časti taktiež analyzujeme a popisujeme extrahované parametre z cien dlhopisov pre jednotlivé metódy.

Thesis name: Extraction of Nelson-Siegel Factors From Bond Prices

Department: Department of Applied Mathematics and Statistics, FMFI UK in Bratislava
Author: Bc. Róbert Berec
Supervisor: Mgr. Juraj Katriak
Date: 23.04.2010
Keywords: bonds, spot rate curve, Nelson-Siegel model

Abstract: The aim of this thesis is to theoretically discuss the basic Nelson-Siegel model and two modified approaches in the first part of this work and compare them in the practical part. The set of German government bonds is used in practical part and will consist of comparing the ability of these three methods, added with two weighted alternatives, in the approximation of the yield curve on the one hand and the so-called out-of-sample analysis on the other. The latter comparison is particularly useful because it reflects the ability of models to theoretically price the bonds that were not included in the estimation and then we are able to compare these theoretical prices with real market values. We also analyze and describe the extracted parameters from bond prices for individual methods in practical part of this thesis.

Contents

Preface

Methods for modeling and estimating the spot rate curves represent multifarious area in today's financial mathematics.

In the real world we are not able to detect the spot rates for all periods because we are limited with the number of bonds issued in a particular country. The Nelson-Siegel model overcomes this barrier and allows us to capture the entire behaviour of the spot rate curve with only four parameters which reduces the storage costs of storing historical data. Nelson and Siegel (1987) demonstrated that their proposed model is capable of capturing many of the typically observed shapes that the spot rate curve assumes over time. As we will discuss later these are one of reasons why many central banks in developed countries currently use this model for estimating the spot rate curves.

Accurate estimates of the current term structure of interest rates are of crucial importance in many areas of finance. Equally important is the ability to correctly calculate the price of bonds which were not involved in the estimation. This requirements can be tested using the in-sample fit and the out-of-sample fit, respectively.

In addition to the original Nelson-Siegel approach we examine also the modified Diebold-Li approach which fix the parameter lambda on prespecified value and one approach outlined by Krishnan et al. (2008) that estimates the parameters directly from bond prices.

Chapter 1

Introduction to Bond Markets

In this chapter some basic definitions and problems of the bond theory associated with this thesis will be explained. Primarily we will focus on bonds categorization, bond pricing, yields, term structure of interest rates and bootstraping.

1.1 Bonds

First of all let us define what a bond is. According to Cairns (2004) a bond is a securitized form of loan, where the buyer of the bond lends the issuer an initial price P in return for a predetermined[1] sequence of payments in certain times. Depending on forms of this payments we distinguish several types of bonds, for example *zero-coupon bond* has only one payment. We will discuss later in this chapter about other types of bonds.

Bonds are mostly issued by national governments (*government bonds*) and the names differ from country to country. In the UK, government bonds are called *gilt-edged securities* or *gilts* for short, German federal government bonds issued with maturities of up to 30 years are called *bunds* and US government debt obligations are called *treasuries.*

Bonds are also issued by institutions other than national governments, such as regional governments, international institutions, banks and companies (the latter giving rise to the name *corporate bonds*). Bonds with identical characteristics issued by different issuers does not imply they will have the same price. For example, consider two zero-coupon bonds, one issued by federal government of Germany and one by Goldman Sachs. The bond issued by the company will probably trade at a lower price than the government bond because the market makers[2] will take into consideration the possibility of default of the Goldman

[1] Except index linked or floating rate bonds, where only the next coupon is known.

[2] A market maker is a company, or an individual, that quotes both a buy and a sell price in a financial instrument or commodity held in inventory, hoping to make a profit on the bid/offer spread.

Sachs. It is generaly assumed that government bonds of countries such as USA or Germany are default free, whereas corporate bonds are subject to varying degrees of default risk depending upon the financial health of the issuing company.

As we have mentioned above, we can divide bonds into several groups according to types of payments, most common are:

- Zero-coupon bonds
- Fixed-rate bonds
- Floating-rate bonds
- Index-linked bonds

A *zero-coupon bond* is a bond bought at a price lower than its face value, with the face value repaid at the time of maturity. It does not make periodic interest payments, or have so-called "coupons", hence the term zero-coupon bond.

A *fixed-rate bond* is a long term debt paper that carries a predetermined interest rate. The interest rate is known as coupon rate and interest is payable at specified dates for the entire term of bond.

Floating-rate bonds are bonds that have a variable coupon, equal to a money market reference rate, like LIBOR or federal funds rate, plus a spread which remains constant. At the beginning of each coupon period, the coupon for this period is calculated by taking the fixing of the reference rate for that day and adding the spread.

Index-linked bond is bond in which payment of income on the principal is related to a specific price index, often the Consumer Price Index. This feature provides protection to investors by shielding them from changes in the underlying index.

In this thesis we will focus primarily on zero-coupon bonds and fixed-rate bonds.

1.2 Fixed-Rate Bonds

The structure of a default-free, fixed-rate bond market can generally be characterized as follows. We pay a price P for a bond in return for a stream of payments at certain times (annually, semi-annually, quarterly,...). Symbolically the cash-flow of payments can be written as follows: C, C, ..., $C+F$, where C is a bond's coupon and F is a face value (typically 100).

The face value of a bond—also referred to as its redemption value, maturity value, par value, or principal—is the amount that the issuer agrees to repay the bondholder on the maturity, or redemption, date, when the debt ceases to exist and the issuer redeems the bond, Choudhry (2005). A bond's coupon C is the periodic interest payment made to owners during the life of the bond and is represented as a fraction of the face value:

$$C = cF,$$

where c, $c \in (0,\ 1)$, is fixed (constant) coupon rate for all periods.

The price $P_t(\tau)$ at time t of bond maturing in τ years is equal to discounted values of its future cash-flows. Using discrete compounding interest we get the following price of fixed-rate bond $P_t(\tau)$ at time t maturing at $\tau = t_n$ with coupon payments C in times t_i, Fabozzi (2005):

$$P_t(\tau) = \sum_{i=1}^{n} \frac{C}{(1 + r_t(t_i))^{t_i}} + \frac{F}{(1 + r_t(\tau))^{\tau}}, \tag{1.1}$$

and using continuous compounding interest we get:

$$P_t(\tau) = \sum_{i=1}^{n} C e^{-r_t(t_i)t_i} + F e^{-r_t(\tau)\tau}. \tag{1.2}$$

In equations 1.1 and 1.2 $r_t(t_i)$ represents today's (at time t) annualized *yield* for t_i years, C is coupon and F is a face value of the bond. It is very important to notice that we can observe only bond prices, not the yields on the markets.

Equation 1.1 calculates the so-called *dirty price* of a bond. Bond prices are quoted (e.g. on Bloomberg or Reuters) as *clean prices*, the price of a coupon bond not including any *accrued interest*. Accrued interest is an accounting convention that treats coupon interest as accruing every day a bond is held. Therefore we need to add accrued interest to the clean price obtained from datacenters to get the dirty price. The difference between the clean and dirty price can be seen on figure 1.1. The turning point in dirty price development is the 11th March 2008 when coupon of a bond is paid. On coupon payment dates the dirty price equals clean price of a bond.

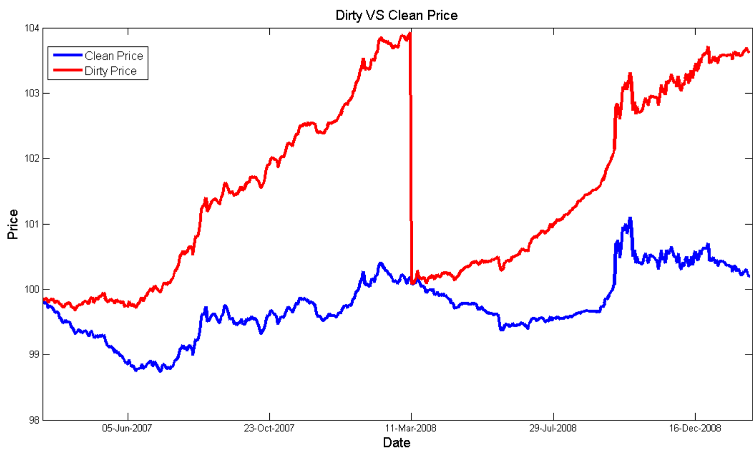


Figure 1.1: Evolution of clean and dirty prices for coupon bond over time.

1.3 Zero-coupon Bonds

A zero-coupon bond is the simplest fixed-income security. Zero-coupon bonds are very similar to fixed-rate bonds but have no coupon payments during its lifetime, i.e. $C = 0$. The rate earned on a zero-coupon bond is also referred to as the *spot interest rate*. Therefore evaluating price of zero-coupon bond $P_t(\tau)$ at time t maturing in time $t_n = \tau$ is straightforward:

$$P_t(\tau) = \frac{F}{(1 + r_t(\tau))^\tau}, \tag{1.3}$$

and using continuous compounding interest we obtain:

$$P_t(\tau) = Fe^{-r_t(\tau)\tau}. \tag{1.4}$$

Note that value of $P_\tau(\tau) = F$ for all τ and arbitrage considerations also indicate that $P_t(\tau) \leq F$ for all τ. Starting with section 1.5 the symbol $P_t(\tau)$ will denote the price of a zero-coupon bond (if not stated explicitly otherwise).

1.4 Yield to Maturity

The yield to maturity or internal rate of return on any investment is the interest rate that will make the present value of the cash flows equal to the price (or initial investment).

By substituting y for all rates in 1.1, we get:

$$P_t(\tau) = \sum_{i=1}^{n} \frac{C}{(1 + r_t(t_i))^{t_i}} + \frac{F}{(1 + r_t(\tau))^\tau} = \sum_{i=1}^{n} \frac{C}{(1 + y)^{t_i}} + \frac{F}{(1 + y)^{t_n}}, \tag{1.5}$$

where C and F have the same values. The yield y that solves this equation is called yield to maturity.

Analogically we can derive yield to maturity using continuous compounding interest:

$$P_t(\tau) = \sum_{i=1}^{n} Ce^{-r_t(t_i)t_i} + Fe^{-r_t(\tau)\tau} = \sum_{i=1}^{n} Ce^{-y\ t_i} + Fe^{-y\ \tau}. \tag{1.6}$$

It is important to notice that yield to maturity is not the same as interest rate for this maturity. This is true only for zero-coupon bonds. More insight can be found in Melichercik et al. (2005), Choudhry (2005) or Fabozzi (2005).

1.5 Spot Rates

Cairns (2004), Fabozzi (2005) define spot rate $r_t(\tau)$ at time t for maturity at time τ as the yield to maturity of zero-coupon bond with face value $F = 1$:

$$r_t(\tau) = -\frac{\log \frac{P_t(\tau)}{F}}{\tau} = -\frac{\log P_t(\tau)}{\tau}, \tag{1.7}$$

that is:

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

In some literature spot rates are also called zero-coupon rates, e.g. Diebold and Li (2006) and Bolder et al. (2004). By compounding multiple spot rates with different maturities we can construct *spot rate curve*, also called *zero-coupon yield curve* or *term structure of interest rates*. This curve reflects the dependency between yield and the maturity of the zero-coupon bond. This is the curve we are mainly interested in.

1.6 Forward Rates

According to Cairns (2004), Fabozzi (2005) the *forward rate* $f_t(\tau_1,\ \tau_2)$ at time t (continuously compounding) which applies between times τ_1 and τ_2 ($0 \leq \tau_1 < \tau_2$) is defined as:

$$f_t(\tau_1,\ \tau_2) = \frac{1}{\tau_2 - \tau_1} \log \frac{P_t(\tau_1)}{P_t(\tau_2)}$$

The forward rate arises within the terms of a *forward contract*. Under such a contract we agree at time t that we will invest \$1 at time τ_1 in return for $e^{(\tau_2-\tau_1)f_t(\tau_1,\ \tau_2)}$ at time τ_2. The *instantaneous forward rate* is defined as limit value of $f_t(\tau_1)$ by letting the maturity of such a forward contract go to τ_1:

$$f_t(\tau_1) = \lim_{\tau_2 \to \tau_1} f_t(\tau_1,\ \tau_2).$$

The instantaneous forward rate curve (we will denote it also as forward rate curve) reflects the relationship bewteen time and instantaneous forward rates on infinitesimal maturity forward contracts.
Given the forward rate curve, we cen determine the spot rate on a zero-coupon bond maturing at τ, denoted by $r_t(\tau)$, by taking the equally weighted average over the forward rates, Nawalkha et al. (2005):

$$r_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u)\mathrm{d}u. \tag{1.8}$$

1.7 STRIPS

STRIPS (*Separate Trading of Registered Interest and Principal of Securities*) are zero-coupon bonds that have been created out of coupon bonds by market makers rather than by the government.

The STRIPS program allows an investor to split a note or bond into a series of zero-coupon securities, one corresponding to each coupon payment (*coupon STRIPS*) and the principal payment of the underlying security (*principal STRIP*). For example, stripping a thirty-year bond would generate thirty-one individual zero-coupon securities: thirty *coupon STRIPS* maturing every year to the maturity date of the bond and one *principal STRIP* maturing on the maturity date of the bond. The final coupon STRIPS and principal STRIP are treated as distinct securities.

It is common to find principal and coupon STRIPS maturing at the same time, both representing presumably identical default-free cash flows, quoted with not equal bid-ask spreads. Evidence suggests that these disparities are real (not quotation errors) and are mainly caused by different liquidities of coupon and principal strips. This phenomena is studied in Daves and Ehrhardt (1992). Further discussion about STRIPS problematics can be found in Tuckman (2002) and Sack (2000).

1.8 Bootstrapping

In many bond markets only very few zero-coupon bonds are issued and traded[1] (all bonds issued as coupon bonds will eventually become a zero-coupon bond after their next-to-last payment date). Usually, such zero-coupon bonds have a short maturity. To acquire knowledge of the market spot rates for longer maturities we have to extract information from the prices of traded coupon bonds.

Following Fabozzi (2005), Nawalkha et al. (2005), Munk (2003), Choudhry (2005) and Melichercik et al. (2005) the bootstrapping method consists of iteratively extracting spot rates using a sequence of increasing maturity zero-coupon and coupon bond prices. This method requires the existence of at least one bond that matures at each bootstrapping date.
To illustrate this method, suppose we are in time $t = 0$ and we have three bonds displayed in table 1.1. Coupons are paid annually and face value of each bond is $F = 100$.

Bond	Time to Maturity (years)	Coupon	Price
1	1	0%	\$ 90.7
2	2	3%	\$ 97.4
3	3	5%	\$ 99.6

Table 1.1: Bond data for bootstrap example.

First bond in our table is zero-coupon bond, thus the calculation of corresponding spot rate, using 1.7, is straightforward:

$$r_0(1) = -\frac{\ln \frac{P_0(1)}{F}}{1} \doteq 9.76\%.$$

To calculate $r_0(2)$ from the price of second bond we have to first discount the coupon payment with $r_0(1)$ and then derive the spot rate $r_0(2)$:

$$P_0(2) = c\,F\,e^{-r_0(1)\ 1} + F(1+c)\,e^{-r_0(2)\ 2} = 0.03\,100\,e^{-0.0976} + 100\,(1+0.03)\,e^{-2\ r_0(2)},$$

and $r_0(2)$ then equals:

$$r_0(2) = -\frac{\log \frac{97.4-3\ e^{-0.0976}}{103}}{2} \doteq 4.21\%.$$

[1] For example in our dataset discussed in chapter 3 there were 7 zero-coupon bonds, 60 coupon bonds and 33 coupon strips in total on the 28th January 2002.

Finally we calculate $r_0(3)$ from the price of third bond. Analogically we can write down the equation for $P_0(3)$:

$$\begin{aligned}P_0(3) &= c\ F\ e^{-r_0(1)\ 1} + c\ F\ e^{-r_0(2)\ 2} + F(1+c)\ e^{-r_0(3)\ 3} = \\ &= 0.05\ 100\ e^{-0.0976} + 0.05\ 100\ e^{-2\ 0.0421} + 100\ (1+0.05)\ e^{-3\ r_0(3)},\end{aligned}$$

and for $r_0(3)$ we get:

$$r_0(3) = -\frac{\log \frac{90.4687}{105}}{3} \doteq 4.97\%.$$

The results are sumarized in table 1.2, where the times to maturities are related with corresponding spot rates.

Time to Maturity (years)	Spot rate
1	9.76%
2	4.21%
3	4.97%

Table 1.2: The resulting values of spot rates.

This example illustrates the basic method of bootsrapping which extracts the spot rates from bond prices. Bootstrapping technique is conceptually neat but may not work so well in practice. Problems arise when coupon payment dates and maturities do not coincide. In chapter 3 we present a modified approach which partly overcomes this issue.

1.9 Discrete and Continuous Interest

Although we have used both discrete and continuous compounding interest in this chapter, following parts of this thesis calculates only with continuous compounding interest.

Continuous compounding interest calculations are both easier to display and more comfortable to manipulate in programs. Diebold and Li (2006), Nelson and Siegel (1987) also employ continuous compounding interest so to be coherent with their work we also adopt this notation.

Chapter 2

Nelson-Siegel Model

In this chapter we describe the original Nelson-Siegel model introduced by Nelson and Siegel (1987), the modified Nelson-Siegel approach proposed by Diebold and Li (2006) and a modified approach suggested by Krishnan et al. (2008). Optimization issues of these models are also discussed in this chapter.

2.1 Nelson-Siegel Model

Nelson and Siegel (1987) suggest to fit the forward rate curve at a given date with a class of approximating functions[1]. The functional form they advocate consist of the product between a polynomial and an exponential decay term. Diebold and Li (2006) in their work also use this functional form. This approximating function can be also derived by assuming that forward rates follow a second order differential equation with two equal real roots:

$$\ddot{f} + \alpha_1 \dot{f} + \alpha_2 f = 0. \tag{2.1}$$

Solving this equation gives us the Nelson-Siegel function for forward rates:

$$f_t(\tau) = \beta_{0t} + \beta_{1t} e^{-\lambda_t \tau} + \beta_{2t} \lambda_t \tau e^{-\lambda_t \tau}, \tag{2.2}$$

where β_{0t}, β_{1t}, β_{2t} and λ_t are parameters.

The spot rate curve (yield as a function of maturity) can be easily derived from 2.2 using 1.8:

[1] This approach is very popular thanks to its convenient and parsinomious three-component exponential approximation. BIS (2005) reports that currently nine out of thirteen central banks which report their curve estimation methods to the Bank for International Settlements use either the Nelson-Siegel or its modifications

$$
\begin{aligned}
r_t(\tau) = \frac{1}{\tau}\int_0^\tau f_t(m)\mathrm{d}m &= \frac{1}{\tau}\int_0^\tau \beta_{0t} + \beta_{1t}e^{-\lambda_t m} + \beta_{2t}\lambda_t m e^{-\lambda_t m}\mathrm{d}m = \\
&= \frac{1}{\tau}\left(\beta_{0t}\tau + \beta_{1t}\frac{1-e^{-\lambda_t\tau}}{\lambda_t} + \beta_{2t}\int_0^\tau \lambda_t m e^{-\lambda_t m}\mathrm{d}m\right) = \\
&= \frac{1}{\tau}\left(\beta_{0t}\tau + \beta_{1t}\frac{1-e^{-\lambda_t\tau}}{\lambda_t} + \beta_{2t}(-e^{-\lambda_t\tau} + \frac{1-e^{-\lambda_t\tau}}{\lambda_t})\right) = \\
&= \beta_{0t} + (\beta_{1t}+\beta_{2t})\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - \beta_{2t}e^{-\lambda_t\tau},
\end{aligned}
\tag{2.3}
$$

where again β_{0t}, β_{1t}, β_{2t} and λ_t are parameters. It is important to notice that for given λ_t both 2.2 and 2.3 are linear in parameters β_{0t}, β_{1t} and β_{2t}.

For purposes of fitting spot rate curves Nelson and Siegel (1987) parametrize the model 2.3 in the following form:

$$
r_t(\tau) = a_t + b_t\ \frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} + c_t\ e^{-\lambda_t\tau}, \tag{2.4}
$$

where $a_t = \beta_{0t}$, $b_t = (\beta_{1t}+\beta_{2t})$ and $c_t = -\beta_{2t}$.

On the other hand Diebold and Li (2006) parametrize model 2.3 in slightly different way:

$$
r_t(\tau) = \beta_{0t} + \beta_{1t}\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} + \beta_{2t}(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}). \tag{2.5}
$$

The limiting value of $r_t(\tau)$ as τ goes to infinity is β_{0t} and as τ goes to zero is $(\beta_{0t}+\beta_{1t})$:

$$
\lim_{\tau\to\infty} r_t(\tau) = \beta_{0t},
$$

thus β_{0t} is the long rate. Second limit can be calculated as:

$$
\begin{aligned}
\lim_{\tau\to 0+} r_t(\tau) &= \lim_{\tau\to 0+} \beta_{0t} + (\beta_{1t}+\beta_{2t})\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - \beta_{2t}e^{-\lambda_t\tau} \\
&= \beta_{0t} - \beta_{2t} + (\beta_{1t}+\beta_{2t})\lim_{\tau\to 0+}\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} \\
&=^{L'H} \beta_{0t} - \beta_{2t} + (\beta_{1t}+\beta_{2t})\lim_{\tau\to 0+}\frac{\lambda_t e^{-\lambda_t\tau}}{\lambda_t} \\
&= \beta_{0t} - \beta_{2t} + (\beta_{1t}+\beta_{2t}) \\
&= \beta_{0t} + \beta_{1t},
\end{aligned}
\tag{2.6}
$$

i.e. the short rate is $\beta_{0t}+\beta_{1t}$. The limits for forward rates function $f_t(\tau)$ are the very same as for the spot rates function $r_t(\tau)$.

To assess the range of shapes available for $r_t(\tau)$, let us fix short term spot rate at $(\beta_{0t}+\beta_{1t}) = 0$, long term spot rate at $\beta_{0t} = 1$ and $\lambda_t = 1$. The function 2.4 then becomes a function of only one parameter β_{2t}:

$$
r_t(\tau) = 1 - (1-\beta_{2t})\frac{(1-e^{-\tau})}{\tau} - \beta_{2t}e^{-\tau}. \tag{2.7}
$$

On figure 2.1 we can see all the different shapes as humps, S-shapes, and monotonic curves of $r_t(\tau)$ in equation 2.7 for parameter β_{2t} from iterval -12 to 12 in equal increments of 2. This wide range of shapes allow us to capture the relation between yield and term to maturity without resorting to more complex models involving more parameters.

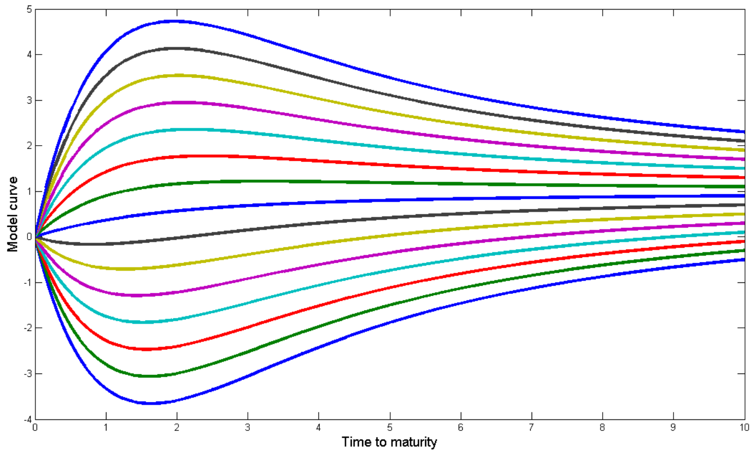


Figure 2.1: Yield curve shapes.

Nelson and Siegel (1987) suggest that the shape flexibility can be seen also in different way by interpreting the parameters of the model 2.2 as measuring the strengths of the short-term, medium-term and long-term components of the forward rate curve. β_{0t} contributes to the long-term component, the contribution of the short-term component is β_{1t} and β_{2t} indicates the contribution of the medium-term component.
This situation is better ilustrated on figure 2.2, where the long-term factor is a constant that does not decay to zero in the limit.
The medium-term component is the only function within this model that starts out at zero (thus not short-term) and decays to zero (and is therfore not long-term).
The short-term component has the fastest decay of both functions within the model that dacays monotonically to zero.
To sum up, with appropriate choices of weights for these components we can generate a large variety of forward rate curves with monotonic and humped shape, Nelson and Siegel (1987).

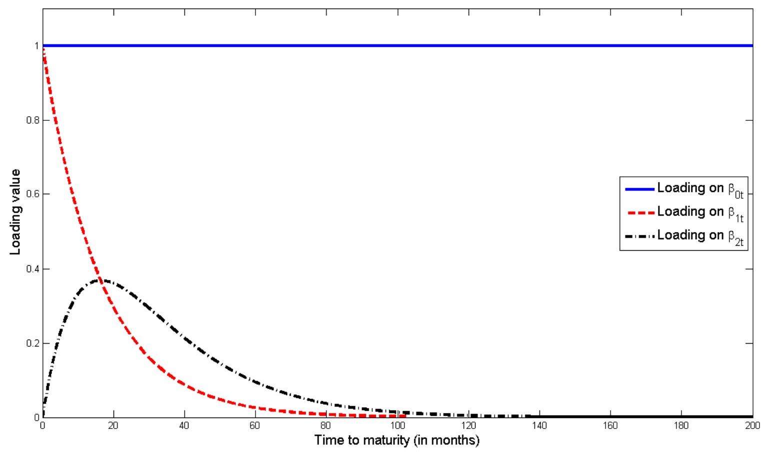


Figure 2.2: Components of the forward rate curve.

Since both this thesis and also Nelson and Siegel (1987) are mainly focused on the fitting of the spot rate curve it is more desirable to examine the development of loadings on parameters a_t, b_t and c_t in model 2.4. This situation is displayed on the figure 2.3 and the difference with conditions discussed in the previous paragraph is obvious. Parameters b_t and c_t have very similar behavior and thus we can expect higher correlation between this two parameters over time[1].

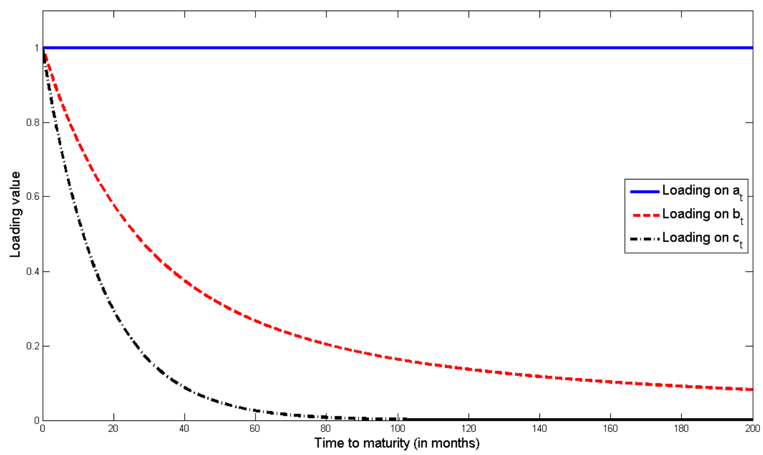


Figure 2.3: Components of the spot rate curve.

As we have already shown, loadings $(1 - e^{-\lambda\tau})/(\lambda\tau)$ and $e^{-\lambda\tau}$ have similar monotonically decreasing shape, so if we were to interpret b_t and c_t, then their loadings would be forced to be very similar.

[1]For more insight see Appendix 2.

Using parametrization proposed by Diebold and Li (2006) we overcome this drawback and can interpret β_{0t}, β_{1t}, and β_{2t} as three factors. The loading on β_{0t} is 1, a constant that does not decay to zero in the limit and may be viewed as a long-term factor. The loading on β_{1t} is $(1 - e^{-\lambda_t \tau})/(\lambda_t \tau)$, a function that starts at 1 but decays monotonically and fast to 0; thus it may be viewed as a short-term factor. The loading on β_{2t} is $(1 - e^{-\lambda_t \tau})/(\lambda_t \tau) - e^{-\lambda_t \tau}$ which starts at 0 (hence not short-term), increases, and then decays to zero (hence not long-term) and thus it may be viewed as a medium-term factor. This situation is ilustrated on figure 2.4.

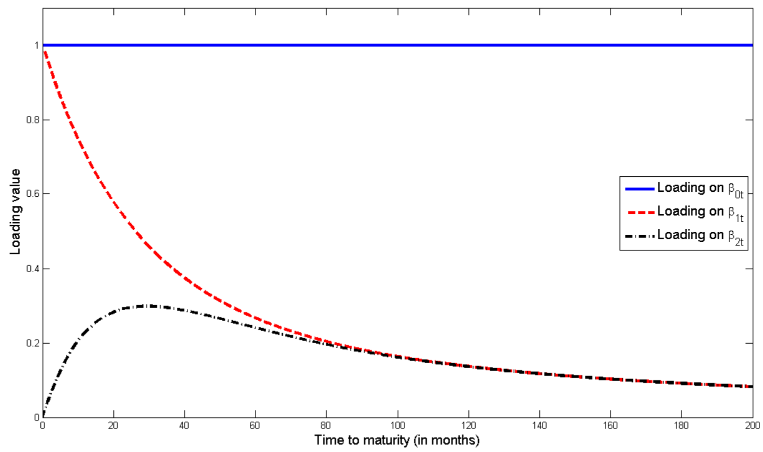


Figure 2.4: Components of the spot rate curve in extended Nelson-Siegel model.

Diebold and Li (2006) point out that long-term, medium-term and short-term factors may be also interpreted in terms of level, slope and curvature but we have to fix the parameter λ_t at 0.0609[1].

The long term factor β_{0t} governs the spot rate curve level; note that an increase in β_{0t} increases all yields equally, as the loading is identical at all maturities; thereby changing the level of the spot rate curve.

The short-term factor β_{1t} is closely related to the spot rate curve slope, which Diebold and Li (2006) define as the ten-year yield minus the three-month yield, i.e. $r_t(120) - r_t(3) = -0.78\beta_{1t} + 0.06\beta_{2t}$[2]

The medium-term factor β_{2t} is closely related to the spot rate curve curvature, which Diebold and Li (2006) define as twice the two-year yield minus the sum of ten-year and three-month yields, i.e. $2r_t(24) - r_t(3) - r_t(120) = 0.0004\beta_{1t} + 0.37\beta_{2t}$. Note that an increase in β_{2t} will have minor effect on very short or very long yields, which load minimaly on it, but will increase medium-term yields which load more heavily on it, therby increasing spot rate curve curvature.

[1]The reasons will be discussed subsequently in more detail.

[2]Using equation 2.5 where $\lambda_t = 0.0609$.

2.1.1 Model Estimation

Suppose we have M spot rates r_{τ_i} for different maturities $\tau_1,\ \tau_2, ...,\ \tau_M$. Using the model 2.5 we seek the optimal parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and λ_t in terms of best fitting the given spot rates $r_t(\tau_i)$, i.e.:

$$\begin{pmatrix} r_t(\tau_1) \\ r_t(\tau_2) \\ \vdots \\ r_t(\tau_M) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} - e^{-\lambda_t\tau_1} \\ 1 & \frac{1-e^{-\lambda_t\tau_2}}{\lambda_t\tau_2} & \frac{1-e^{-\lambda_t\tau_2}}{\lambda_t\tau_2} - e^{-\lambda_t\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_t\tau_M} & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_t\tau_M} - e^{-\lambda_t\tau_M} \end{pmatrix} \begin{pmatrix} \beta_{0t} \\ \beta_{1t} \\ \beta_{2t} \end{pmatrix}, \tag{2.8}$$

and we can rewritte this system into matrix form:

$$\boldsymbol{r_t = X_{\lambda_t}\beta_t}, \tag{2.9}$$

where $\boldsymbol{r}_t$ is M-dimensional vector, $\boldsymbol{X}_{\lambda_t}$ is Mx3 matrix and $\boldsymbol{\beta}_t$ is 3-dimensional vector:

$$\boldsymbol{r}_t = \begin{pmatrix} r_t(\tau_1) \\ r_t(\tau_2) \\ \vdots \\ r_t(\tau_M) \end{pmatrix}, \quad \boldsymbol{X}_{\lambda_t} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} - e^{-\lambda_t\tau_1} \\ 1 & \frac{1-e^{-\lambda_t\tau_2}}{\lambda_t\tau_2} & \frac{1-e^{-\lambda_t\tau_2}}{\lambda_t\tau_2} - e^{-\lambda_t\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_t\tau_M} & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_t\tau_M} - e^{-\lambda_t\tau_M} \end{pmatrix}, \quad \boldsymbol{\beta}_t = \begin{pmatrix} \beta_{0t} \\ \beta_{1t} \\ \beta_{2t} \end{pmatrix}.$$

For original Nelson-Siegel model 2.4 we get the very same results but the matrix X_{λ_t} has slightly different form:

$$\boldsymbol{X}_{\lambda_t} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} & e^{-\lambda_t\tau_1} \\ 1 & \frac{1-e^{-\lambda_t\tau_2}}{\lambda_t\tau_2} & e^{-\lambda_t\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_t\tau_M} & e^{-\lambda_t\tau_m} \end{pmatrix}, \quad \boldsymbol{\beta}_t = \begin{pmatrix} a_t \\ b_t \\ c_t \end{pmatrix}. \tag{2.10}$$

To find the optimal values of parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and λ_t in system 2.8 means to solve the subsequent problem:

$$\min_{\lambda_t,\beta_{0t},\beta_{1t},\beta_{2t}} \sum_{i=1}^{M}\left(\beta_{0t}+\beta_{1t}\,\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}+\beta_{2t}\,(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}-e^{-\lambda\tau_i})-r_t(\tau_i)\right)^2 =$$
$$= \min_{\lambda_t,\boldsymbol{\beta}_t}\left(\boldsymbol{X_{\lambda_t}\beta_t}-\boldsymbol{r_t}\right)^T\left(\boldsymbol{X_{\lambda_t}\beta_t}-\boldsymbol{r_t}\right). \tag{2.11}$$

To solve the system 2.9 using $\boldsymbol{X}_{\lambda_t}$ and $\boldsymbol{\beta}_t$ from 2.10 Nelson and Siegel (1987) suggest to employ the fact that for any given $\widetilde{\lambda} > 0$ the best-fitting values of the parameters a_t, b_t and c_t can be then computed using linear least squares:

$$\boldsymbol{\beta_t} = (\boldsymbol{X}^{\boldsymbol{T}}_{\widetilde{\boldsymbol{\lambda_t}}}\boldsymbol{X}_{\widetilde{\boldsymbol{\lambda_t}}})^{-1}\boldsymbol{X}^{\boldsymbol{T}}_{\widetilde{\boldsymbol{\lambda_t}}}\boldsymbol{r}$$

Repeating this procedure over a grid of values for $\widetilde{\lambda}$ produces the overal best-fitting values of a_t, b_t, c_t and λ_t. Large values of λ_t correspond to rapid decay in the regressors and therefore will be able to fit excessive curvature at short maturities well while being unable to fit excessive curvature over longer maturity ranges. Small values of λ_t produce slow decay in the regressors that can fit curvature at longer maturities but they will be unable to follow extreme curvature at short maturities. To illustrate this trade-off follow figure 2.5.

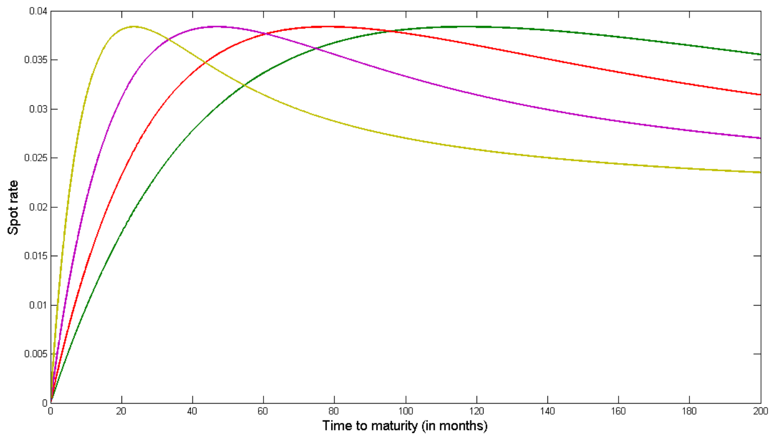


Figure 2.5: Spot rate curve sensitivity to λ_t.

Instead of searching over a grid of values for λ_t to solve the system 2.8 Diebold and Li (2006) propose to fix parameter λ_t at a prespecified value $\widetilde{\lambda}$, which lets us estimate the values of betas for each time t using ordinary least squares. Diebold and Li (2006) argue that doing so enhances not only simplicity and convenience, but also numerical trustworthiness by enabling us to replace hundreds of potentially challenging numerical optimizations with trivial least-squares regressions. The question is, what is the appropriate value of $\widetilde{\lambda}$. Parameter λ_t determines the maturity at which the loading on the medium-term, or curvature, factor obtains its maximum. According to Diebold and Li (2006) two or three year maturities are commonly used in that regard, so they simply picked the average, i.e. 30 months. The λ_t value that maximizes the loading on the medium-term factor at 30 months is $\widetilde{\lambda} = 0.0609$. The corresponding value calculated using years, i.e. value that maximizes the loading on the medium term-factor at 2.5 years, is $\widetilde{\lambda} = 0.7173$:

$$\arg\max_{\lambda} \frac{1 - e^{-2.5\lambda}}{2.5\lambda} - e^{-2.5\lambda} = 0.7173. \tag{2.12}$$

We would like to remark that the interpretation of parameters as level, slope and curvature proposed by Diebold and Li (2006) is valid only for fixed λ_t at 0.0609 (or 0.7173 using years), e.g. in definition of curvature the loading on coefficient β_{1t} is equal to 0.0004 but for different values of lambda becomes the loading on this parameter more significant (see figure 2.6).

An example of different development of fixed and non-fixed λ_t can be seen on figure 4.4 in chapter 4. It is obvious that the value of fixed lambda advocated by Diebold and Li (2006) is largely higher than non-fixed lambdas. If we were to fix the lambda on different level then we would not be able to interpret the coefficients according to Diebold and Li (2006) definitions.

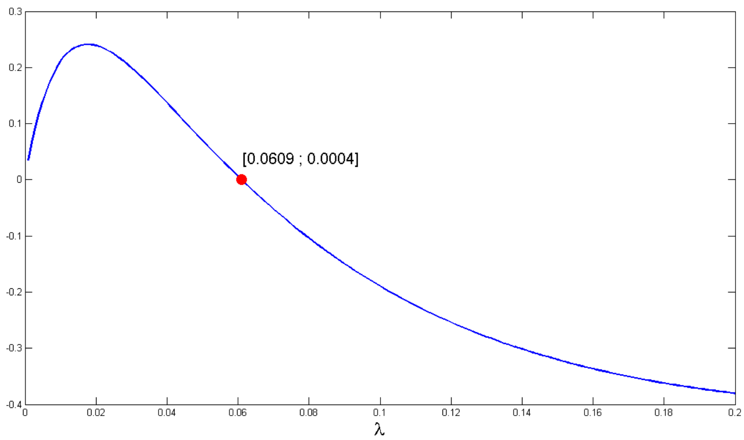


Figure 2.6: Loading on β_{1t} in definition of curvature for different λ_t.

2.2 Modified Approach

In this section we want to indroduce modified approach of parameter extraction outlined in Krishnan et al. (2008).

Instead of extraction $\boldsymbol{\beta}_t$ and λ_t from yields we propose to extract them directly from bond prices. By substituting $r_t(t_i)$ in equation 1.2 with $r_t(\tau)$ from equation 2.5 for present value of fixed rate coupon bond, i.e. for price of this bond, we get:

$$\hat{P}_t(\tau_n) = \sum_{i=1}^{n} Ce^{-(\beta_{0t}+\beta_{1t}\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}+\beta_{2t}(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}-e^{-\lambda_t\tau_i}))\tau_i}+ \\ + Fe^{-(\beta_{0t}+\beta_{1t}\frac{1-e^{-\lambda_t\tau_n}}{\lambda_t\tau_n}+\beta_{2t}(\frac{1-e^{-\lambda_t\tau_n}}{\lambda_t\tau_n}-e^{-\lambda_t\tau_n}))\tau_n}, \tag{2.13}$$

2.2.1 Estimation of Modified Approach

Suppose we have M zero-coupon and fixed-coupon bond prices $P_t(\tau_i)$ for different maturities $\tau_1,\ \tau_2, ...,\ \tau_M$. Define the following problem with parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and λ_t:

$$\min_{\beta_{0t},\beta_{1t},\beta_{2t},\lambda_t} \sum_{i=1}^{M}(\hat{P}_t(\tau_i) - P_t(\tau_i))^2, \tag{2.14}$$

and using equation 2.13 we obtain:

$$\min_{\beta_{0t},\beta_{1t},\beta_{2t},\lambda_t} \sum_{i=1}^{M}(\sum_{j=1}^{i} Ce^{-(\beta_{0t}+\beta_{1t}F_1(j)+\beta_{2t}F_2(j))\tau_j}+ \\ + Fe^{-(\beta_{0t}+\beta_{1t}F_1(i)+\beta_{2t}F_2(i))\tau_i} - P_t(\tau_i))^2, \tag{2.15}$$

where $F_1(x)$ and $F_2(x)$ equal to:

$$F_1(x) = \frac{1-e^{-\lambda_t\tau_x}}{\lambda_t\tau_x}, \qquad F_2(x) = \frac{1-e^{-\lambda_t\tau_x}}{\lambda_t\tau_x} - e^{-\lambda_t\tau_x}.$$

Model 2.8 in section 2.1.1 assumes that spot rates $r_t(\tau)$ are known. This assumption is very strict because only the bond prices are observable on the markets, thus one needs to transform them into spot rates first. We discuss this problematics in the next chaper. On the other hand the approach from this section works only with bond prices so the transformation problem is avoided.

Chapter 3

The Data

In this chapter we describe the dataset of bonds and filters used in this thesis.

We used historical end-of-day average of bid-ask price quotes for German government bonds (also known as Bunds) from the 10th Jun 1999 through the 5th September 2008 only on trading days , downloaded from the Bloomberg database into Microsoft Office Excell spreadsheets. Further adjustments were made in Mathworks software Matlab. We also downloaded the ISIN, coupon rate, coupon frequency, issue date, maturity date, description and description notes for every bond.

We focused on federal government bonds (maturity between 10 and 30 years), German federal medium-term bonds (maturity of about five years), treasury bonds (maturity of about two years) and coupon strips. The distribution of bonds by time to the maturity can be seen on figure 3.1.

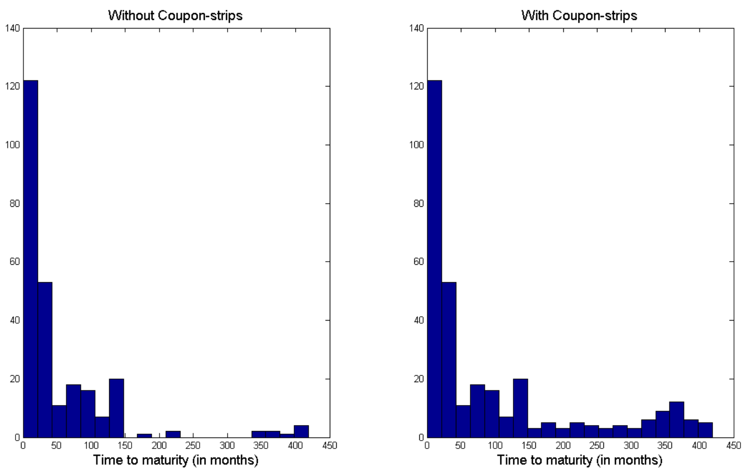


Figure 3.1: Distribution of bonds by time to the maturity.

From figure 3.1 it is clear that the biggest part of our data represents bonds with the maturity under five years. The second largest group are bonds with maturity under 15 years followed by group of bonds maturing at around 33 years. The gap between two previous groups is evident. By adding coupon strip bonds (discussed in section 1.7) to our dataset we are able to fill this gap.

Another view on the bonds in our dataset can be found on figure 3.2 where we take into account also the timeline, i.e. the ranges of maturities available for estimation over our sample. The date is shown on the horizontal axis, the remaining maturity (in months) is shown on the vertical axis.

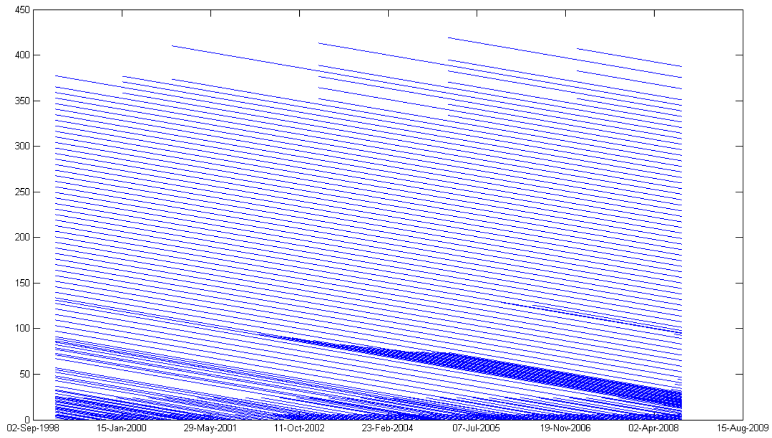


Figure 3.2: Outstanding bonds.

An immediate issue that arises is determining the set of securities to be included in the estimation. The bond securities outstanding at any point in time can differ in many dimensions, including their liquidity and their callable features. Our goal is to use a set of securities that are similar in terms of their liquidity and that do not have special features (such as being callable) that would affect their prices. In other words, we would ideally have securities that only differ in terms of their coupons and maturities.

3.1 Filtering

To achieve requirements mentioned in previous paragraph, we include in the estimation all outstanding bonds, with the following exceptions:

1. We exclude all securities with option-like features (e.g. callable bonds).
2. We exclude all securities with less than three months to maturity, since the yields on these securities often seem to behave oddly. This behavior may partly reflect the lack of liquidity for those issues.

This filtering eliminates nearly all distrubancies between bonds and allows us to estimate the spot rate curve more accurately. Although even after this elimination we can find further anomalies that should be avoided; for example see figure 3.3 which captures situation from the 1st January 2007.

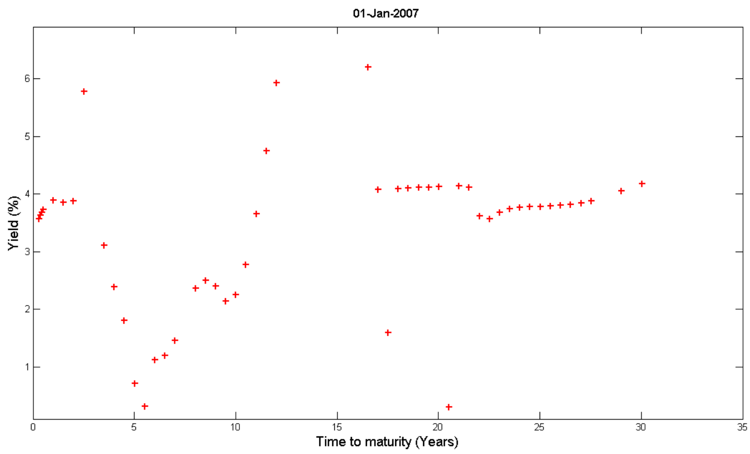


Figure 3.3: Anomaly that occures on New Year's day in 2007.

The disorder on figure 3.3 is apparent and we assume it is caused mainly by small liquidity on the markets on days like New Year's day. We also found that on such days there is much lower amount of bond quotes available. Therefore additional requirements are needed:

1. Date must be different from the 1st January.
2. Number of active[1] bonds on particular day must be larger than 75.

For standard days we get the spot rates[2] displayed on figure 3.3. The spot rates are more consistent and create the outlines of spot rate curves.

[1]Bond with quoted price is considered as active.

[2]We extracted these spot rates from bond prices using algorithm discussed in section 3.2.

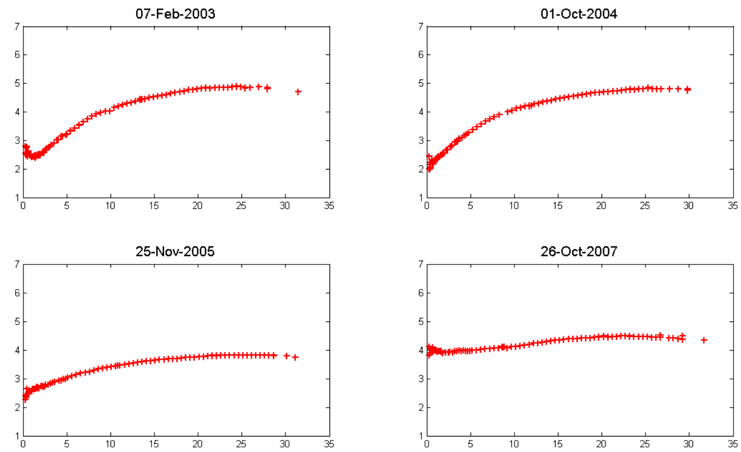


Figure 3.4: Actual spot rates for selected dates.

The number of active bonds outstanding after filtration on studied period is dispalayed on figure 3.5.

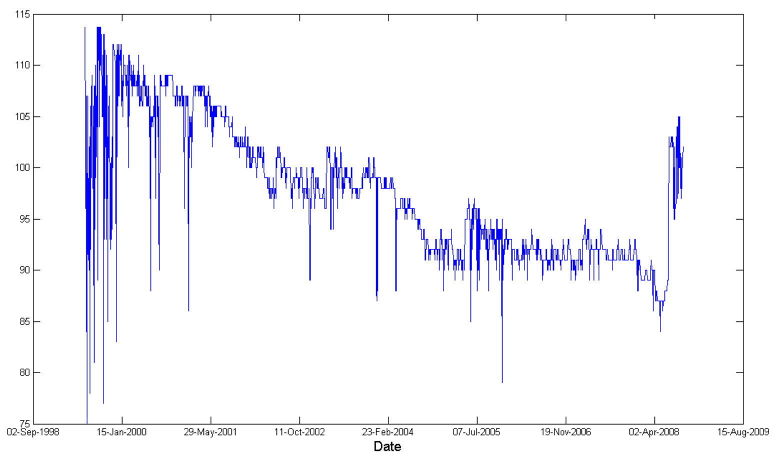


Figure 3.5: Number of active bonds on particular dates after filtration.

In figure 3.6 we provide a three-dimensional plot of estimated[1] spot rate curves over the studied period. The detailed discussion about this estimation will be in the next chapter. The high variation in short term rate is visually apparent and it is clear that the typical spot rate curve is upward sloping.

[1]Estimated spot rate curves using pNL approach discussed in chapter 4.

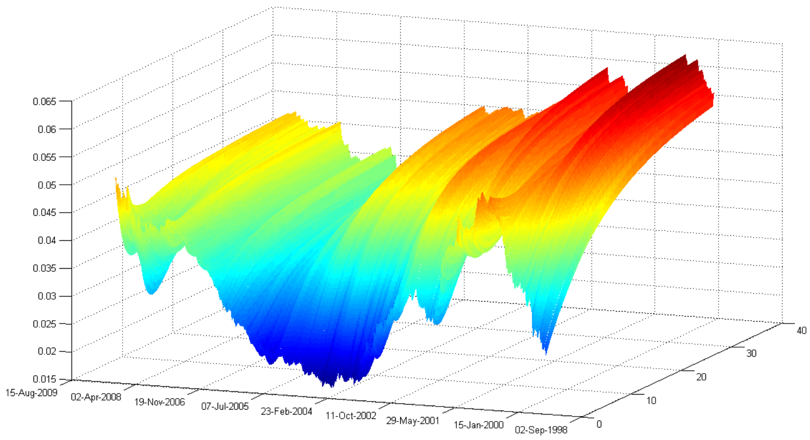


Figure 3.6: Spot rate curves, Jun 1999 - Sep 2008.

3.2 Data Transformation

In previous section we mentioned and discussed not only bond prices but also spot rates. It is necessary to clarify how the spot rates are derived from the bond prices. We follow the bootstraping method discussed in section 1.8 to extract the corresponding spot rates from prices of zero-coupon and coupon bonds. Bootstrapping technique is conceptually neat but may not work so well in practice. Problems arise when you do not have a set of bonds that mature at precise intervals. To overcome this shortcoming we have slightly modified the bootstraping method in the following way[1]:

1. We divide bonds into two groups: Zero-coupon and coupon bonds.
2. Using 1.7 we calculate spot rates from zero-coupon bonds.
3. We sort the coupon bonds according to time to maturity (from the shortest to the longest).
4. Then we gradually calculate the spot rates using bootstraping method from the second group (we utilize also spot rates calculated from the zero-coupon bonds).
    (a) If the maturity of needed spot rate is between two known maturities we calculate the weighted average of corresponding spot rates.
    (b) If the maturity of needed spot rate is too long or too short we take the nearest known spot rate.

[1]Although this algorithm works for our data there are situations in which it would have problems, e.g. if we have neither zero-coupon bonds nor coupon bonds with only one coupon payment outstanding.

5. In the end we have the set of spot rates with maturities.

This method will derive the theoretical spot rates which can precisely price every bond used in bootstraping. The weighted average mentioned in our modified procedure is computed as:

$$r_t(\tau_2) = \frac{\tau_1 - \tau_2}{\tau_1 - \tau_0}\, r_t(\tau_0) + \frac{\tau_2 - \tau_0}{\tau_1 - \tau_0}\, r_t(\tau_1), \tag{3.1}$$

what can be also rewriten as:

$$r_t(\tau_2) = \frac{r_t(\tau_1) - r_t(\tau_0)}{\tau_1 - \tau_0}\, (\tau_2 - \tau_0) + r_t(\tau_0),$$

where $r_t(\tau_2)$ represents needed spot rate for maturity τ_2; $r_t(\tau_0)$ and $r_t(\tau_1)$ are known spot rates for maturities τ_0 and τ_1 respectively[2]. Maturities τ_0, τ_1 and τ_2 satisfy the the condition $\tau_0 < \tau_2 < \tau_1$.

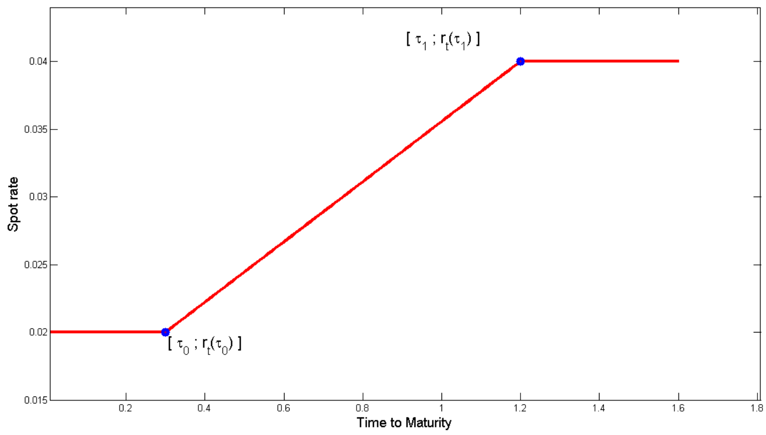


Figure 3.7: Possible positions of $r_t(\tau_2)$ (marked with red line).

To better illustrate the fourth step in our procedure we refer to figure 3.7. The red line displays possible positions of $r_t(\tau_2)$, which can be placed according to τ_2 into three areas. If the τ_2 is very short and there is available only spot rate with longer maturity, e.g. τ_0, then the spot rate $r_t(\tau_2)$ is equal to $r_t(\tau_0)$ (left-hand horizontal line). If the τ_2 is very long and there is available only spot rate with shorter maturity, e.g. τ_1, then the spot rate $r_t(\tau_2)$ is equal to $r_t(\tau_1)$ (right-hand horizontal line). And finally, if the τ_2 is between two available maturities, e.g. τ_0 and τ_1, then $r_t(\tau_2)$ is calculated using weighted average 3.1 (center sloping line).

[2]It is actually the linear interpolation.

Chapter 4

Estimation

In this chapter we present the estimation methods used to find the best values of parameters in models introduced in chapter 2 on bond data discussed in chapter 3. The results and their analysis with graphical representations are also presented in this chapter. We measure the goodness of fit of the model in sense of the mean squared errors MSE_t between theoretical and real prices of bonds $\hat{P}_t(\tau_i)$ and $P_t(\tau_i)$, respectively :

$$MSE_t = \frac{SSE_t}{m_t} = \frac{\sum_{i=1}^{m_t}\left(\hat{P}_t(\tau_i) - P_t(\tau_i)\right)^2}{m_t}, \tag{4.1}$$

where m_t is the number of bonds available at time t. In order to find appropriate statistics we make an average of SSE_t because on different dates we have different number of bonds.
We extensively utilize the mathematical software Matlab developed by MathWorks company. We employ the build in functions like $fminunc$ to find the minimum of unconstrained multivariable functions, $lsqnonlin$ to find the solution of nonlinear least squares problems, and many others.

According to chapter 2 we can distinguish three different estimation approaches:

1. Original Nelson and Siegel (1987) approach[1] to spot rates with parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t},\ \lambda_t$.
2. Modified approach of Diebold and Li (2006) to spot rates with parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and fixed $\lambda = 0.7173$.
3. Modified approach of Krishnan et al. (2008) to bond prices with parameters $\beta_{0t},\ \beta_{1t},\ \beta_{2t},\ \lambda_t$.

In chapter 3 we analysed the composition of our data and concluded that the maturities of bonds are not equally distributed over the time, i.e the shorter maturities overweight the longer maturities. By fitting the spot rates with

[1] We use parametrization from equation 2.5. See also Appendix 2.

original approach, Nelson and Siegel (1987), and modified approach, Diebold and Li (2006), we implicitly weight more the spot rates of shorter maturities over the spot rates of longer maturities, thus we fit the spot rates of longer maturities with less precision. Since prices of long-term bonds are more sensitive to changes of spot rates of longer maturities, then even small chnges of this rates can imply large changes in bond prices.
To overcome this deficiency we propose to use weights on spot rates in order to impose more importancy on spot rates with longer maturities. The weight matrix W has the form:

$$\boldsymbol{W} = \begin{pmatrix} \tau_1 & 0 & \dots & 0 \\ 0 & \tau_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tau_m \end{pmatrix}, \tag{4.2}$$

where τ_i represents time to maturity of corresponding zero-coupon bond with spot rate $r_t(\tau_i)$ which also equals to *duration*[1] of this bond.

To sum up, we identified 5 different approaches:

1. yNLSQ-W: weighted version of original Nelson and Siegel (1987) approach to spot rates
2. yNLSQ: non-weighted version of original Nelson and Siegel (1987) approach to spot rates
3. yCON-W: weighted version of modified Diebold and Li (2006) approach to spot rates
4. yCON: non-weighted version of modified Diebold and Li (2006) approach to spot rates
5. pNL: modified approach of Krishnan et al. (2008) to bond prices

4.1 yNLSQ-W and yNLSQ

The yNLSQ-W and yNLSQ approaches seek the optimal values for $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and λ_t by solving the following problem:

$$\min_{\lambda_t,\beta_{0t},\beta_{1t},\beta_{2t}} \sum_{i=1}^{m} w_{ii}\left(\beta_{0t} + \beta_{1t}\ \frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i} + \beta_{2t}\ (\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i} - e^{-\lambda\tau_i}) - r_t(\tau_i)\right)^2, \tag{4.3}$$

where w_{ii} for yNLSQ-W equals to corresponding element in weight matrix $\boldsymbol{W}$ and for yNLSQ equals to 1. The input in this case is set of spot rates $r_t(\tau_1),\ r_t(\tau_2),\ ...,\ r_t(\tau_m)$ and we use nonlinear least squares function *lsqnonlin*[2] to find the optimal values for $\beta_{0t},\ \beta_{1t},\ \beta_{2t}$ and λ_t.

[1] More insight about duration can be found in Melichercik et al. (2005), Fabozzi (2005), Cairns (2004) or Choudhry (2005).

[2] In each date we found the starting point using *fminunc* with multiple initial points. To make this process faster we calculated the gradient of minimized function, see Apendix 1.

4.2 yCON-W and yCON

The yCON-W and yCON approaches seek the optimal values for β_{0t}, β_{1t} and β_{2t} by calculating the following equation:

$$\boldsymbol{\beta_t} = \left(\boldsymbol{X_\lambda^T W X_\lambda}\right)^{-1} \boldsymbol{X_\lambda^T W r}, \tag{4.4}$$

where matrix $\boldsymbol{W}$ for yCON-W equals to weight matrix from 4.2 and for yCON equals to identity matrix I_m. In both cases λ equals to 0.7173. The input in this case is also set of spot rates $r_t(\tau_1)$, $r_t(\tau_2)$, ..., $r_t(\tau_m)$. To find the optimal values for β_{0t}, β_{1t} and β_{2t} we use the Matlab formula:

$$betas = (W * X)\backslash(W * r),$$

where *betas* is vector of β-parameters and r is vector of known spot rates.

4.3 pNL

The pNL approach seeks the optimal values for β_{0t}, β_{1t}, β_{2t} and λ_t by solving the problem:

$$\begin{aligned} \min_{\beta_{0t},\beta_{1t},\beta_{2t},\lambda_t} \sum_{i=1}^{m} (\sum_{j=1}^{i} C e^{-(\beta_{0t}+\beta_{1t}F_1(j)+\beta_{2t}F_2(j))\tau_j} + \\ + F e^{-(\beta_{0t}+\beta_{1t}F_1(i)+\beta_{2t}F_2(i))\tau_i} - P_t(\tau_i))^2, \end{aligned} \tag{4.5}$$

where $F_1(x)$ and $F_2(x)$ equal to:

$$F_1(x) = \frac{1-e^{-\lambda_t\tau_x}}{\lambda_t\tau_x}, \qquad F_2(x) = \frac{1-e^{-\lambda_t\tau_x}}{\lambda_t\tau_x} - e^{-\lambda_t\tau_x}.$$

The input to this method is set of zero-coupon and coupon bond prices $P_t(\tau_1)$, $P_t(\tau_2)$, ..., $P_t(\tau_m)$. Finding the optimal values for β_{0t}, β_{1t}, β_{2t} and λ_t is made by Matlab function $fminunc$.

4.4 Results

In this section we present the results of our estimations. First we find the optimal values for every parameter in yNLSQ-W, yNLSQ, yCON-W, yCON and pNL. The second step is calculation of the theoretical bond prices using 1.2, and the final step is the evaluation of MSE according to 4.1. Repeating this procedure for all dates we get the time series of MSE and all best-fitting coeficients for every method.

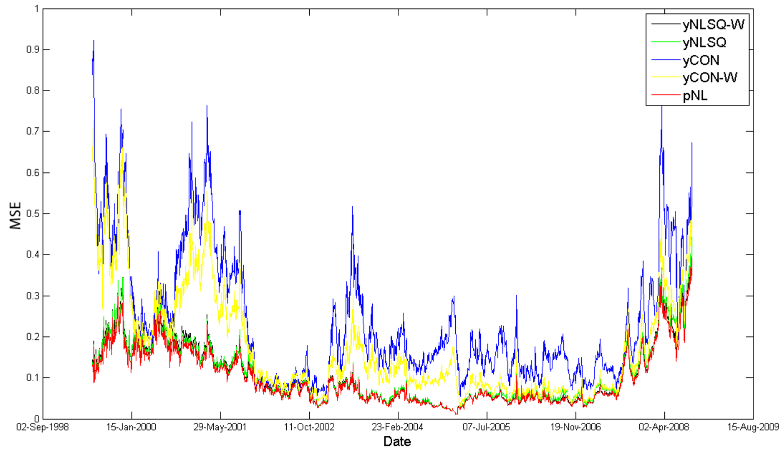


Figure 4.1: MSE over time.

The development of MSE for every method can be seen on figure 4.1. It is obvious that yCON and yCON-W gives the worst results over the whole period; what is expected because of fixation of parameter λ. But even the difference between yCON and yCON-W is substantial and making the weighted version more favourable.

The best overal performance achieved the pNL method but the distinction of yNLSQ-W and yNLSQ is unclear on the scale of figure 4.1. Therefore we added the figure 4.2 which contains the difference between MSE of yNLSQ-W and MSE of yNLSQ. The non-weighted and weighted versions give comparable results in this case.

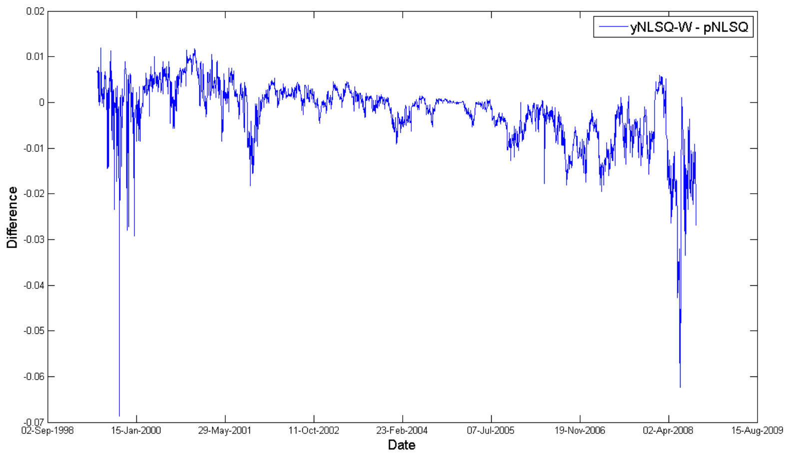


Figure 4.2: Difference between MSE of yNLSQ-W and MSE of yNLSQ.

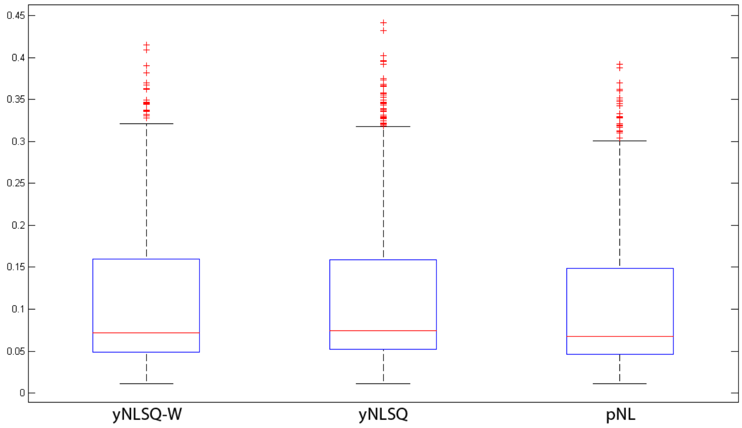


Figure 4.3: Boxplots of MSE of yNLSQ-W, yNLSQ and pNL.

Method	Min.	1st Qu	Median	3rd Qu	Max.
yNLSQ-W	0.011	0.0485	0.0718	0.1595	0.4148
yNLSQ	0.011	0.0523	0.0743	0.1587	0.4418
yCON	0.0499	0.1207	0.1789	0.3163	0.9222
yCON-W	0.0323	0.0779	0.1147	0.2432	0.7262
pNL	0.011	0.0464	0.0678	0.1484	0.3918

Table 4.1: The five-number summary of MSE for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

Figure 4.3 offers another insight onto our MSE results and exhibits graphically the five-number summary from table 4.1 into the boxplots of yNLSQ-W, yNLSQ and pNL. The lowest values of median and 3rd quartile has the pNL method at the level of 0.0678 and 0.1484, respectively. The second lowest value of median has the yNLSQL-W method and 3rd quartile the yNLSQL method. The median value of all approaches tends to be near to the 1st quartile what indicates the skewness in data, i.e. values with smaller MSE are more frequent. Overall, we can sort the studied methods according to this empirical results in the following order: pNL, yNLSQ-W and yNLSQ, yCON-W, yCON.

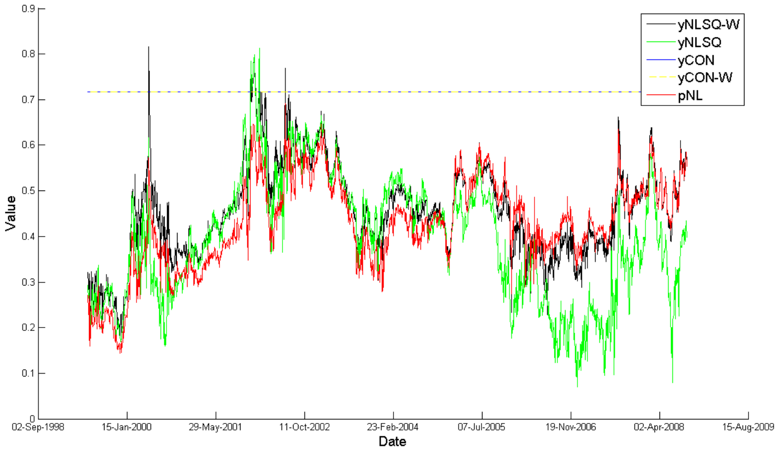


Figure 4.4: The evolution of estimated values for λ_t over time.

Method	Mean	Median	Std. Dev.	Minimum	Maximum
yNLSQ-W	0.4574	0.4543	0.1031	0.1701	0.8161
yNLSQ	0.4	0.411	0.1362	0.0698	0.813
yCON	0.7371	0.7371	0	0.7371	0.7371
yCON-W	0.7371	0.7371	0	0.7371	0.7371
pNL	0.4269	0.4255	0.0986	0.1443	0.6914

Table 4.2: Descriptive statistics of $\hat{\lambda}_t$ for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

Another important part of this section is discussion about evolution of estimated values for λ_t, β_{0t}, β_{1t} and β_{2t} over time.
The development of $\hat{\lambda}_t$ for every method is displayed on figure 4.4 (estimated values for λ_t for yCON and yCON-W are constant because of fixation of this parameter).
The values of $\hat{\lambda}_t$ for yNLSQ-W and pNL have very similar progress and fall into interval $\big(0.1443;\ 0.8161\big)$. The development of values of $\hat{\lambda}_t$ for yNLSQ method follow the previous group except interval around the 19th November 2006 where it differentiate marginaly. The descriptive statistics for estimated values of λ_t can be found in table 4.2.

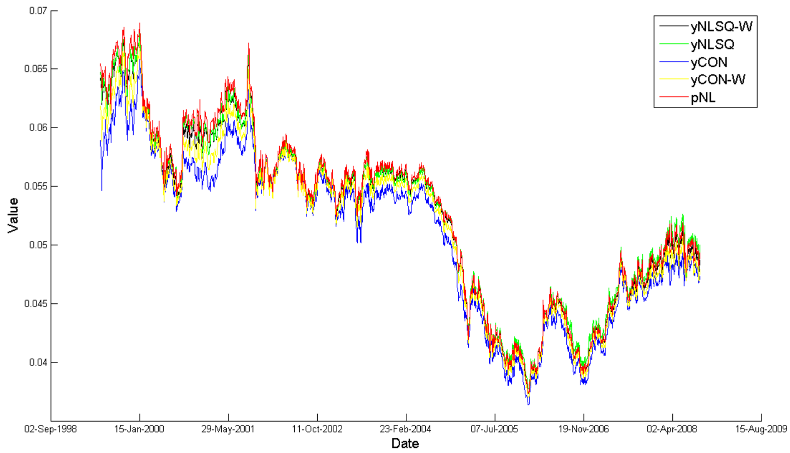


Figure 4.5: The evolution of estimated values of β_{0t} over time.

Method	Mean	Median	Std. Dev.	Minimum	Maximum
yNLSQ-W	0.0525	0.0549	0.0075	0.0373	0.0682
yNLSQ	0.0527	0.0549	0.0074	0.0377	0.0682
yCON	0.051	0.0537	0.0071	0.0363	0.0658
yCON-W	0.0518	0.0543	0.0073	0.037	0.0673
pNL	0.0529	0.0552	0.0078	0.0374	0.0689

Table 4.3: Descriptive statistics of $\hat{\beta}_{0t}$ for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

In this place we analyze the development of estimated values of β_{0t} over time for all methods. To recall the interpretation proposed by Diebold and Li (2006) for β_{0t}, it stands for the level of the spot rate curve or it's long-term rate.
We present the descriptive statistics for $\hat{\beta}_{0t}$ in table 4.3 and the development over time on figure 4.5 for all methods. Overall, the $\hat{\beta}_{0t}$ get the values from 0.0363 to 0.0689 and until the 25th October 2005 it has mainly decreasing behaviour. The extensive differences between group of yNLSQ-W, yNLSQ, yCON-W, pNL and yCON are mostly in the beginning of watched period reaching in some cases nearly 80 basis points. Estimations of methods in the former group have very similar progress over time.

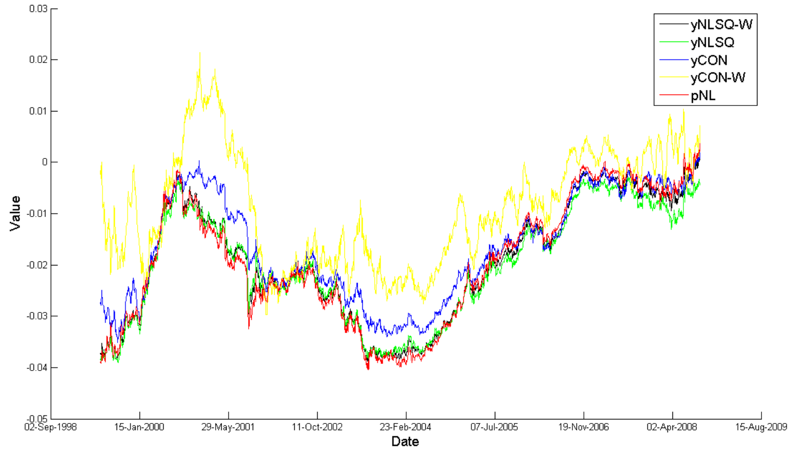


Figure 4.6: The evolution of estimated values of β_{1t} over time.

Method	Mean	Median	Std. Dev.	Minimum	Maximum
yNLSQ-W	-0.0192	-0.0196	0.0116	-0.0391	0.0015
yNLSQ	-0.0199	-0.0201	0.0108	-0.0391	-0.0029
yCON	-0.0164	-0.0173	0.0105	-0.0351	0.0019
yCON-W	-0.009	-0.0095	0.0112	-0.0297	0.0214
pNL	-0.0192	-0.0198	0.0123	-0.0404	0.0037

Table 4.4: Descriptive statistics of $\hat{\beta}_{1t}$ for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

In figure 4.6 we plot the development of $\hat{\beta}_{1t}$ and corresponding descriptive statistics are displayed in table 4.4. Majority of values fall within -0.04 and 0.0004.

We registered higher variance in estimated values between methods in this case. Both yNLSQ and yNLSQ-W exhibits nearly the same behaviour and pNL follow them with only small number of deviations.

The estimations of yCON and yCON-W approaches progress quite differently in contrast to other methods but mutually behave similarly (with different levels).

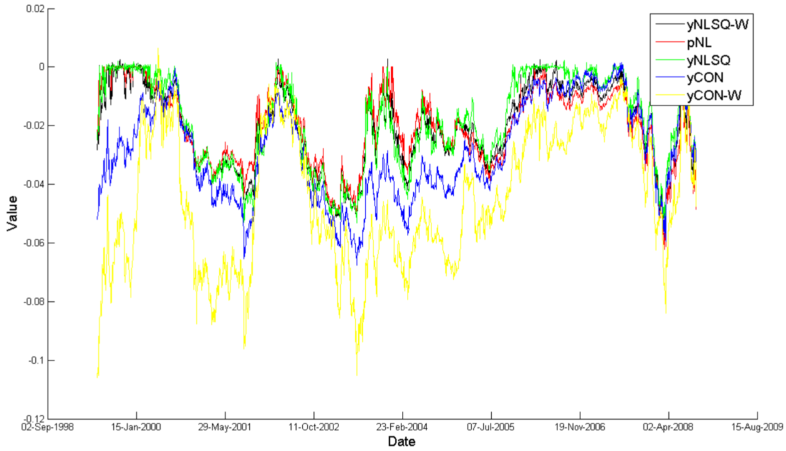


Figure 4.7: The evolution of estimated values of β_{2t} over time.

Method	Mean	Median	Std. Dev.	Minimum	Maximum
yNLSQ-W	-0.0209	-0.0202	0.0139	-0.058	0.0027
yNLSQ	-0.0192	-0.0187	0.0155	-0.0544	0.0022
yCON	-0.0291	-0.0319	0.0167	-0.0677	0.0015
yCON-W	-0.0451	-0.0492	0.0233	-0.1059	0.0065
pNL	-0.0202	-0.0187	0.0131	-0.0622	0.0001

Table 4.5: Descriptive statistics of $\hat{\beta}_{2t}$ for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

In this place we discuss and analyse the development of estimated values of the last parameter β_{2t} over time for every method. The estimated values are displayed on figure 4.7 and corresponding descriptive statistics can be found in table 4.5.

The progress of values of $\hat{\beta}_{2t}$ follow a similar pattern as in previous case. We see group of yNLSQ-W, yNLSQ and pNL developing uniformly (with only small deviations) on one side and different development of yCON and yCON-W on the other side. As for the first goup the values move in interval from -0.0622 to 0.0027 and for the second group from -0.1059 to 0.0065.

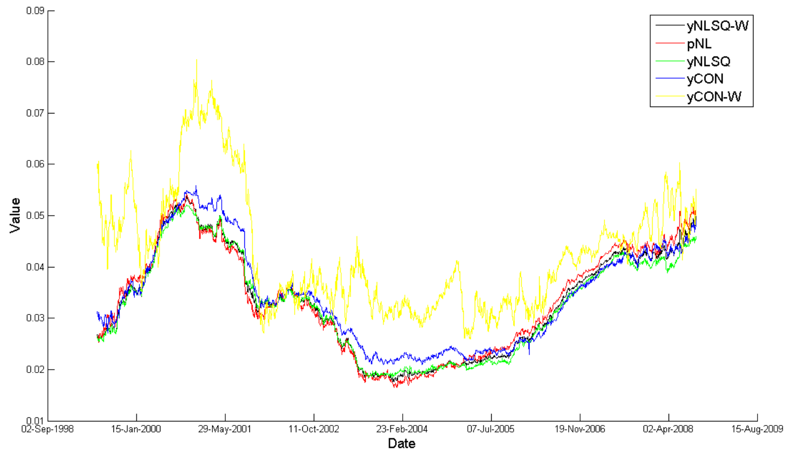


Figure 4.8: The development of the short-term rate $(\hat{\beta}_{0t} + \hat{\beta}_{1t})$ over time.

On the figure 4.8 we plot the development of the short-term rate which is calculated using equation 2.6. It is obvious that the short-term rate of yNLSQ-W, yNLSQ and pNL have similar development over the whole period and the short-term rate of yCON follows them with episodic deviations.
A combination of fixed parameter λ and weighting more on the rates on longer horizons contribute to the higher variation in short-term rates of yCON-W approach.

To provide some examples of fitted spot rate curves we include figures 4.9 and 4.10 of actual spot rates on dates from figure 3.4 and fitted spot rate curves.

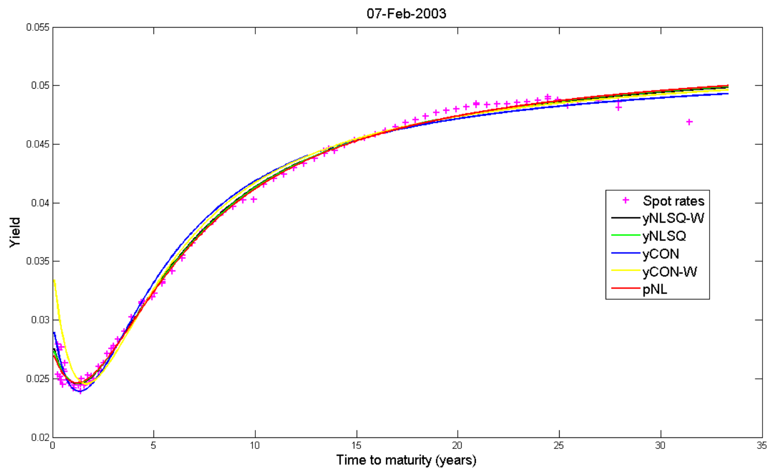


Figure 4.9: Actual spot rates and fitted spot rate curves for selected dates.

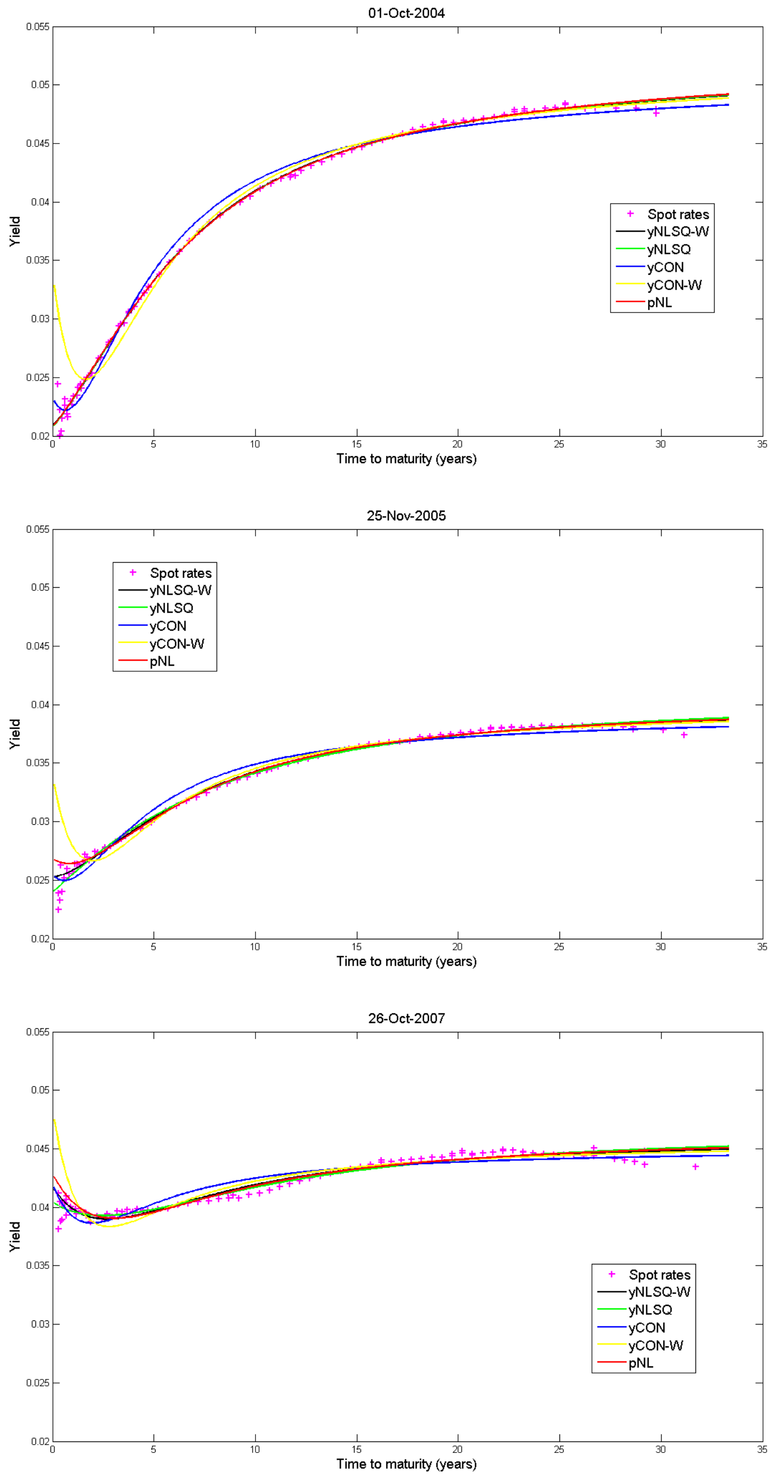


Figure 4.10: Actual spot rates and fitted spot rate curves for selected dates.

4.5 Out-of-Sample Fit

In this section we compare studied methods (yNLSQ-W, yNLSQ, yCON, yCON-W and pNL) from a different point of view. We use the so called out-of-sample fit to determine whether the estimated spot rate curve can price specific bonds (not involved in estimation) accurately, i.e. to calculate the differences between market and theoretical prices of the bonds.

At first we selected particular bonds from our dataset and set them aside. Then we estimated the parameters β_{0t}, β_{1t}, β_{2t} and λ_t from remaining bonds in order to construct the spot rate curve for every method.

Method	Min	1st Quartile	Median	3rd Quartile	Max
15Y $>$					
yNLSQ-W	0.4153	1.079	1.4539	1.8778	4.3114
yNLSQ	0.4024	1.1572	1.7323	2.4801	4.2446
yCON	0.9896	3.0242	4.0996	4.976	8.1788
yCON-W	0.6962	1.5731	2.3394	3.1184	6.0253
pNL	0.4279	1.0983	1.4407	1.8303	4.4684
10Y-15Y					
yNLSQ-W	0.1137	0.3766	0.5601	0.7898	1.6807
yNLSQ	0.064	0.2787	0.4539	0.6947	1.5092
yCON	0.1845	0.7671	1.1075	1.5429	3.394
yCON-W	0.1724	0.6707	0.9716	1.335	2.995
pNL	0.0442	0.3061	0.4682	0.6229	1.3666
Random 20					
yNLSQ-W	0.0925	0.4	0.5331	0.7144	1.8067
yNLSQ	0.0931	0.4011	0.5396	0.7186	1.7939
yCON	0.2546	0.6507	0.826	1.0636	2.7566
yCON-W	0.2085	0.5407	0.6822	0.9245	2.3566
pNL	0.0823	0.3787	0.5203	0.7086	1.7163

Table 4.6: Descriptive statistics of $RMSE$ for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL.

Afterwards using equation 1.2 and estimated spot rate curve we calculated theoretical prices $\hat{P}_t(\tau)$ of bonds we set aside. Finaly we computed the root mean squared error ($RMSE$):

$$RMSE_t = \sqrt{\frac{\sum_{i=1}^{n_t}(\hat{P}_t(\tau_i) - P_t(\tau_i))^2}{n_t}}, \tag{4.6}$$

where n_t is number of bonds we set aside and $P_t(\tau_1),\ P_t(\tau_2),\ \ldots,\ P_t(\tau_n)$ are their market prices. Repeating this procedure over the studied period we get multiple observations which can be further analysed.

The selection of bonds is an important part in this procedure which needs more discussion. We determined three different scenarios of bond selection:

1. We select all bonds with maturity over 15 years on particular date (15Y $>$).
2. We select all bonds with maturity in interval from 10 years to 20 years on particular date (10Y-15Y).
3. We randomly select 20% from all active bonds on particular date (Random 20).

The descriptive statistics of results are displayed in table 4.6 and ploted in figures 4.11 and 4.12.

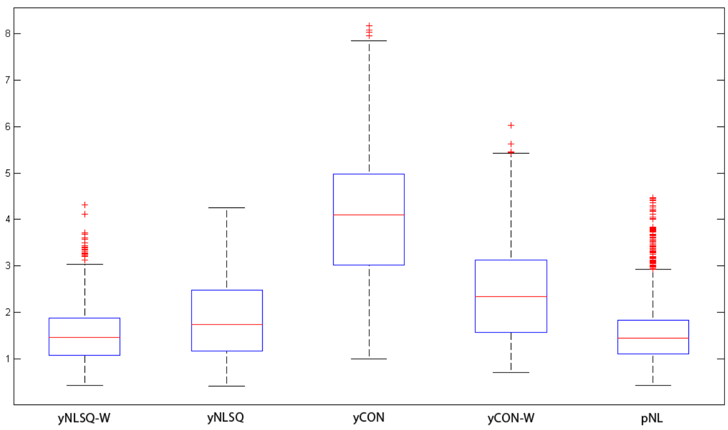
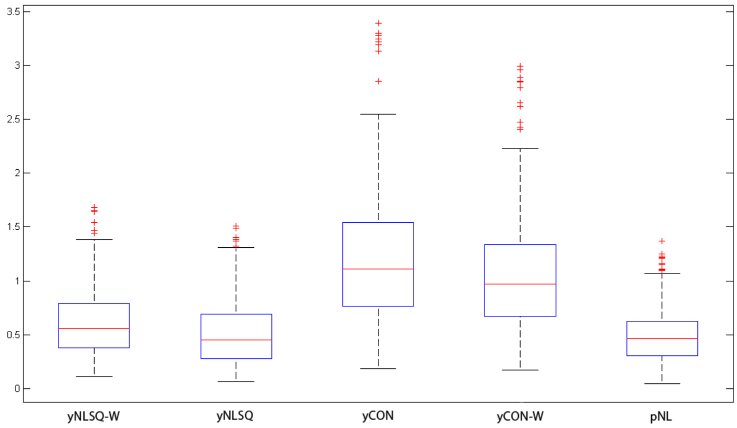


Figure 4.11: The boxplots of $RMSE$ for all methods ("15Y $>$" scenario).



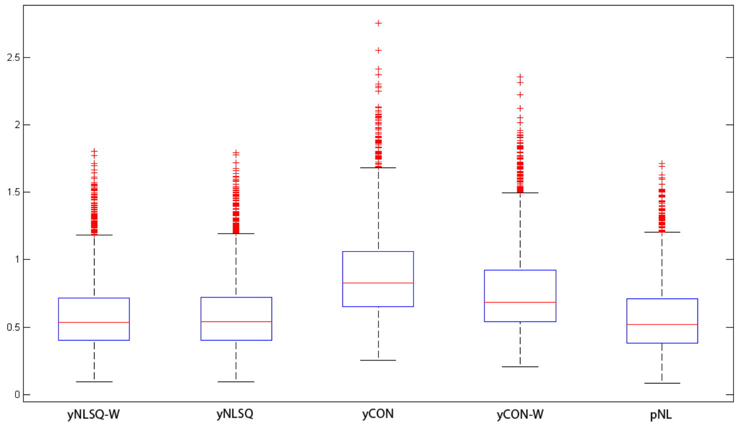


Figure 4.12: The boxplots of $RMSE$ for all methods. ("10Y-15Y" scenario on top and "Random 20" scenario on bottom).

From table 4.6 we can see that the performance of the out-of-sample fit in all scenarios the yCON and yCON-W achieved overall the worst results. Thus we do not comment them in further paragraphs but briefly discuss in the summary at the end of this section.

Obviously in case of random selection of 20% bonds the methods (yNLSQ, yNLSQ-W, pNL) performed nearly identically and the vast majority of values fall into interval from 0.082 to 1.2. This is not true for 15Y$>$ selection where yNLSQ underperformed the yNLSQ-W and pNL which again achieved very similar results. The higher variation in values of $RMSE$ is evident and is mainly caused by fact that we eliminated from estimation and are trying to price bonds with long maturities. The values fall mainly into interval from 0.4 to 3.

Finally the 10Y-15Y selection reveals little change in ranking among studied methods. The best results achieved again the pNL approach followed by yNLSQ and then by yNLSQ-W approaches. The dispersions of values are quite low in this scenario making the differences between methods less significant.

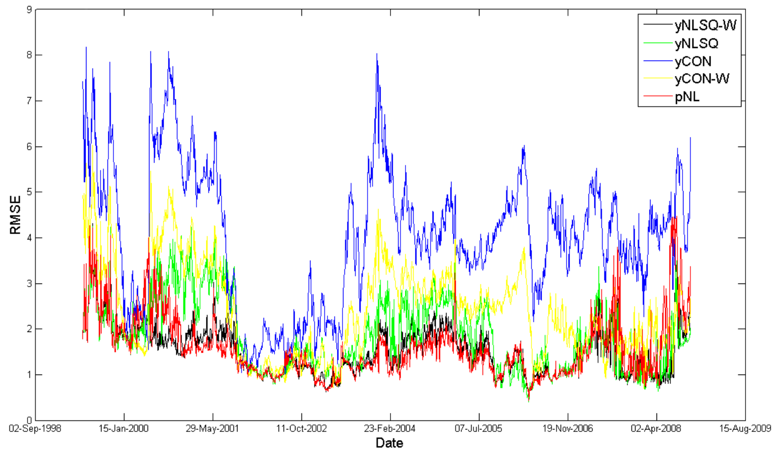


Figure 4.13: The development of $RMSE$ for all methods ("15Y $>$" scenario).

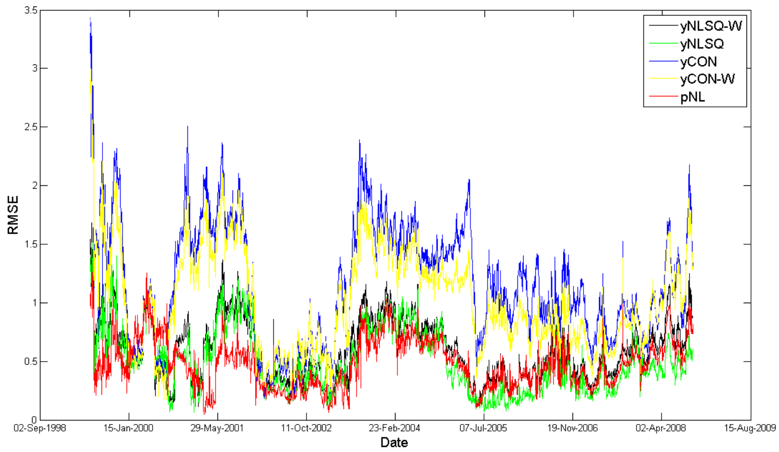


Figure 4.14: The development of $RMSE$ for all methods ("10Y-15Y" scenario).

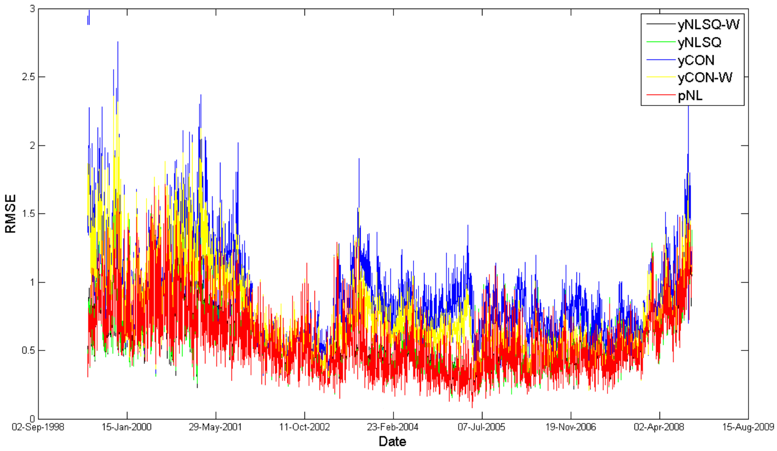


Figure 4.15: The development of $RMSE$ for all methods ("Random 20%" scenario).

We plot the development of $RMSE$ for every scenario on figures 4.13, 4.14 and 4.15. Higher volatility on figure 4.15 is caused by random selection of bonds, where the number of bonds with longer maturities fluctuate amid samples and have more significant impact on $RMSE$.

The results of the out-of-sample analysis are similar to those acquired in previous section focused on spot rate curve estimation. The pNL method obtained again overall best results in all scenarios. The yNLSQ-W method performed cometitively but in one case it was outperformed by yNLSQ approach, although the difference was marginal in comparison to the result from the first scenario. The yNLSQ method lagged behind significantly in the first scenario; nevertheless in other two scenarios it performed well. The results of yCON and yCON-W methods were poor in contrast with approaches mentioned above and the weighted version was able to price the bonds more accurately than it's non-weighted version.

According to results from this section, we can rank the methods in this order: pNL, yNLSQ-W, yNLSQ, yCON-W and yCON.

4.6 Financial Crisis

Because of financial crisis that escalated in late 2008 we cut our data on the 5th September 2008 when the Lehman Brothers filed for bankruptcy and triggered the liquidity shortfall on financial markets and scepticism between banks.
Higher demand for large emmisions of government bonds pushed their prices up and deepened differences among other less liquid bonds. The financial crisis would have distorted the results and mislead our aim to compare the studied methods.

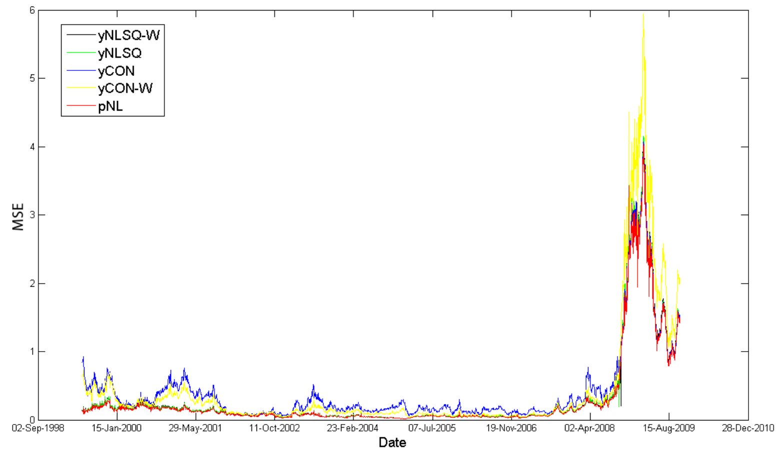


Figure 4.16: MSE over time including the financial crisis.

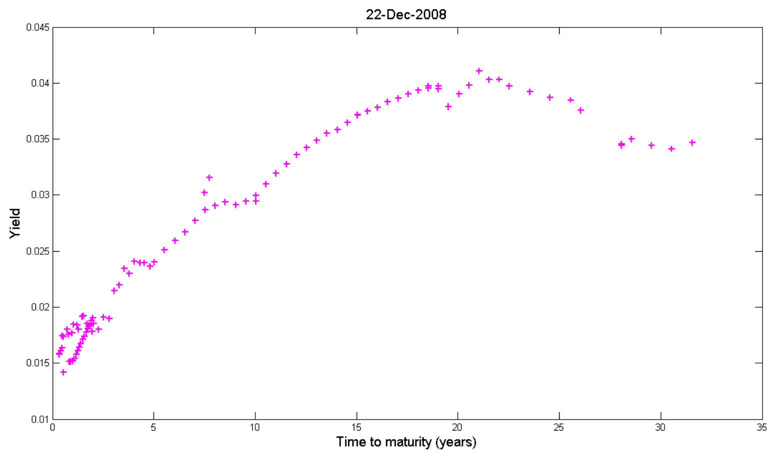


Figure 4.17: Actual spot rates on 22nd December 2008.

To illustrate this situation we widened our time horizon up to the 20th October 2009. The development of MSE can be seen on figure 4.16 and an example of disorder amid spot rates from the 22nd December 2008 is pictured on figure 4.17. The reader can find corresponding summary of MSE from the 6th September 2008 to the 20th October 2009 in table 4.7. All methods obviously performed poorly and values in table 4.7 are in large contrast with values in table 4.1.
It is evident from figure 4.17 that the spot rates create an unusual and irregular shape which can be hardly estimated with methods studied in this thesis. Thus we make no conclusions from this results and we added this section only for informative purposes.

Method	Min.	1st Qu	Median	3rd Qu	Max.
yNLSQ-W	0.2027	1.1852	1.6329	2.6726	4.148
yNLSQ	0.1978	1.1457	1.6122	2.6606	4.148
yCON	0.5801	1.1435	1.5954	2.6716	4.069
yCON-W	0.4424	1.6761	2.3202	3.5242	5.9381
pNL	0.3432	1.1332	1.5839	2.5454	4.0347

Table 4.7: The five-number summary of MSE for yNLSQ-W, yNLSQ, yCON, yCON-W and pNL from 6th September 2008 to 20th October 2010.

Conclusion

In this thesis we have theoretically discussed and practically applied to real data the Nelson-Siegel model and its two modifications.

We showed that these methods are able to capture the various shapes of the spot rate curves and that the individual parameters can be interpreted as short-term, medium-term and long-term or in the case of Diebold-Li modification with fixed lambda as the level, slope and curvature. We found the latter interpretation misleading because it does not apply for different values of lambda.

In the practical part we analyzed and filtered the data of German bonds which were then transformed into a set of spot rates. This step was necessary because the original Nelson-Siegel approach and modified Diebold-Li approach require as input a set of spot rates. The approach outlined in Krishnan et al. (2008) requires bond prices as input.
From the nature of our data we concluded that it would be appropriate to append another two methods which take into consideration the weighting of spot rates with different maturity lengths. Therefore we finally identified five different approaches: yNLSQ-W, yNLSQ, yCON, yCON-V and pNL.

We set up two tests to compare these methods: in-sample fit and out-of-sample fit. In the in-sample fit we estimated for every approach the spot rate curve in all dates and calculated the MSE as an indicator of the estimation accuracy. Based on the results we sorted the methods in the following order (from best to worst): pNL, yNLSQ-W, yNLSQ, yCON-W and yCON. We found that the pNL approach which estimates the parameters from bond prices gives the best results and can be regarded as natural benchmark because it directly minimizes the SSE. The weighted versions performed better than their non-weighted equivalents and both versions of Diebold-Li approach (yCON, yCON-W) obtained the worst outcomes. The inability to change the parameter lambda dramatically affected their flexibility to fit the spot rate curve.

In the out-of-sample fit we set aside a group of bonds, estimated spot rate curves for every method, and then we calculated the theoretical prices of bonds which we omited from the estimation. Theoretical prices were compared with their real market values and used to calculate the $RMSE$. We distinguished three scenarios of selection: we set aside the group of bonds with more than 15 years to maturity; we set aside the group of bonds with maturity between 10 and 15 years; and last we randomly set aside 20% of bonds. The results in this case were very similar to previous output and once again we sorted the methods in the following order from best to worst: pNL, yNLSQ-W, yNLSQ, yCON-W and yCON.

According to programming complexity we recomend to use the pNL method because it does not require sophisticated transformation of spot rates from bond prices. Overall, we can say that for the set of data we used appears to be the most reliable and accurate the pNL method, which estimates the parameters directly from bond prices, followed by methods in respective order: yNLSQ-W, yNLSQ, yCON-W, and finally yCON.

References

Batten, J., A., Fetherston, T., A., Szilagyi, P., G., *European fixed income markets: money, bond, and interest rate derivatives*, John Wiley&Sons, Inc., 2004

Bliss, R., *Testing term structureestimation methods*, Advances in Futures and Options Research 9, 97-231, 1997

Bolder, D., J.,Gusba, S., *Exponentials, polynomials, and fourier series: more yield curve modelling at the Banko of Canada*, Bank of Canada: Working Paper 2002-29, 2002

Bolder, D., J.,Johnson, G.,Metzler, A., *An empirical analysis of the canadian term structure of zero-coupon interest rates*, Bank of Canada: Working Paper 2004-48, 2004

Cairns. J.G.A., *Interest Rate Models, An Introduction*, Princeton University Press, 2004

Choudhry, M., *Fixed-income securities and derivatives handbook : analysis and valuation*, Bloomberg Press Princeton, 2005

Daves, P., R., Ehrhardt, M., C., *Liquidity, reconstruction, and the value of U.S. treasury STRIPS*, Working Paper, University of Tennessee, 1992

Diebold, F., X.,Li, C.,*Forecasting the term structure of government bond yields*, Journal of Econometrics, Elsevier, vol. 130(2), pages 337-364, 2006

Fabozzi, F., J., *The handbook of fixed income securities, Seventh edition*, McGraw-Hill, 2005

Fama, E., Bliss, R., *The information in long-maturity forward rates*, American Economic Review 77, 680-692, 1987

Henderson, T., M., *Fixed income strategy: the practicioner's guide to riding the curve*, John Wiley&Sons, Inc., 2003

Krishnan, C., N., V.,Ritchken, P., H., Thomson, J., B., *Predicting Credit Spreads*, http://ssrn.com/abstract=970997, 2008

Melichercik, I., Olsarova, L., Uradnicek, V., *Kapitoly z financnej matematiky*, Epos, 2005

Monetary and Economic Department, *Zero-coupon yield curves: technical documentation*, BIS Papers No.25, Bank for International Settlements, 2005

Munk, C., *Fixed income analysis: securities, pricing, and risk management*, University of Southern Denmark, 2003

Nawalkha, S., K., Soto, G., M., Beliaeva, N., A., *Interest rate risk modeling: the fixed income valuation course*, John Wiley&Sons, Inc., 2005

Nelson, C., R., Siegel, A., F., *Parsinomious modeling of yield curves*, Journal of Business 60, 473-489, 1987

Sack, B., P., *Using treasury STRIPS to measure the yield curve*, FEDS Working Paper No. 2000-42, 2000

Schich, S., T., *Estimating the German term structure*, Discussion paper 4/97, Deutsche Bundesbank, 1997

Svenson, L., E., O., *Estimating and interpreting forward interest rates: Sweden 1992-1994*, Working paper No.WP/94/114, International Monetary Fund, 1994

Tuckman, B., *Fixed income securities, tools for today's markets, Second edition*, John Wiley&Sons, Inc., 2002

Turc, J., Ungari, S., *Filtering the interest rate curve, the MENIR framework*, Societe Generale Cross Asset Research, 2009

Appendix 1

We want to minimize the following problem:

$$\min_{\lambda,\boldsymbol{\beta}} \big(\boldsymbol{y} - \boldsymbol{X_\lambda \beta}\big)^T \big(\boldsymbol{y} - \boldsymbol{X_\lambda \beta}\big), \tag{4.7}$$

and for given λ the corresponding $\boldsymbol{\beta}$ can be calculated as:

$$\boldsymbol{\beta}_\lambda = \boldsymbol{(X_\lambda^T X_\lambda)^{-1} X_\lambda^T y} \tag{4.8}$$

so we can rewrite the equation 4.7 as:

$$\begin{aligned}&\min_\lambda \big(\boldsymbol{y} - \boldsymbol{X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T y}\big)^T \big(\boldsymbol{y} - \boldsymbol{X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T y}\big) = \\ =&\min_\lambda \boldsymbol{y}^T \big(\boldsymbol{I} - \boldsymbol{X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T}\big)^T \big(\boldsymbol{I} - \boldsymbol{X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T}\big)\boldsymbol{y}.\end{aligned} \tag{4.9}$$

The expression $\big(\boldsymbol{I} - \boldsymbol{X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T}\big)$ is in fact the orthogonal projector. For orthogonal projector $\boldsymbol{P}$ we get:

$$\boldsymbol{P} = \boldsymbol{P}^T \boldsymbol{P}. \tag{4.10}$$

Applying equation 4.10 on 4.9 we obtain:

$$\min_\lambda \big(\boldsymbol{y^T y} - \boldsymbol{y^T X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T y}\big). \tag{4.11}$$

In order to make our optimization faster, we need to supply the first derivative of the objective function into the *fminunc* method in Matlab.
Since $\boldsymbol{y^T y}$ is λ-independent it's derivative is 0, thus we only need to find the first derivative of the second part of the objective function. We will write $\boldsymbol{X}$ instead of $\boldsymbol{X_\lambda}$. We have:

$$\begin{aligned}-\frac{\partial \boldsymbol{y^T X (X^T X)^{-1} X^T y}}{\partial \lambda} &= -\boldsymbol{y}^T \Big[\frac{\partial \boldsymbol{X}}{\partial \lambda} \boldsymbol{(X^T X)^{-1} X^T} + \boldsymbol{X} \frac{\partial \boldsymbol{(X^T X)^{-1} X^T}}{\partial \lambda}\Big]\boldsymbol{y} \\ &= -\boldsymbol{y}^T \Big[\frac{\partial \boldsymbol{X}}{\partial \lambda} \boldsymbol{(X^T X)^{-1} X^T} \\ &\quad + \boldsymbol{X} \frac{\partial \boldsymbol{(X^T X)^{-1}}}{\partial \lambda} \boldsymbol{X^T} \\ &\quad + \boldsymbol{X (X^T X)^{-1}} \frac{\partial \boldsymbol{X^T}}{\partial \lambda}\Big]\boldsymbol{y}.\end{aligned} \tag{4.12}$$

We utilize the knowledge of first derivative according to single variable x of an inverse matrix $\boldsymbol{A}$:

$$\frac{\partial \boldsymbol{A}^{-1}}{\partial x} = -\boldsymbol{A}^{-1}(\frac{\partial \boldsymbol{A}}{\partial x})\boldsymbol{A}^{-1}, \tag{4.13}$$

and apply it on derivative of inverse matrix in third line of equation 4.12 and acquire the desirable derivative:

$$\begin{aligned}
-\frac{\partial \boldsymbol{y^T X(X^T X)^{-1} X^T y}}{\partial \lambda} = &- \boldsymbol{y}^T \Big[\frac{\partial \boldsymbol{X}}{\partial \lambda}\boldsymbol{(X^T X)^{-1} X^T} \\
&- \boldsymbol{X(X^T X)^{-1}}(\frac{\partial \boldsymbol{X^T}}{\partial \lambda}\boldsymbol{X} + \boldsymbol{X^T}\frac{\partial \boldsymbol{X}}{\partial \lambda})\boldsymbol{(X^T X)^{-1} X^T} \\
&+ \boldsymbol{X(X^T X)^{-1}}\frac{\partial \boldsymbol{X^T}}{\partial \lambda}\Big]\boldsymbol{y}.
\end{aligned}$$

The matrix $\frac{\partial \boldsymbol{X}}{\partial \lambda}$ is equal to:

$$\frac{\partial \boldsymbol{X}}{\partial \lambda} = \begin{pmatrix}
0 & \frac{e^{-\lambda\tau_1}(1+\lambda\tau_1-e^{\lambda\tau_1})}{\lambda^2\tau_1} & \frac{e^{-\lambda\tau_1}(1+\lambda\tau_1+\lambda^2\tau_1^2-e^{\lambda\tau_1})}{\lambda^2\tau_1} \\
0 & \frac{e^{-\lambda\tau_2}(1+\lambda\tau_2-e^{\lambda\tau_2})}{\lambda^2\tau_2} & \frac{e^{-\lambda\tau_2}(1+\lambda\tau_2+\lambda^2\tau_2^2-e^{\lambda\tau_2})}{\lambda^2\tau_2} \\
\vdots & \vdots & \vdots \\
0 & \frac{e^{-\lambda\tau_M}(1+\lambda\tau_M-e^{\lambda\tau_M})}{\lambda^2\tau_M} & \frac{e^{-\lambda\tau_M}(1+\lambda\tau_M+\lambda^2\tau_M^2-e^{\lambda\tau_M})}{\lambda^2\tau_M}
\end{pmatrix}$$

Appendix 2

A different development of parameters using parametrization proposed by Nelson and Siegel (1987) and Diebold and Li (2006) can be found on figure 4.18. We used the data discussed in chapter 3. Higher correlation between estimated parameters b_t and c_t is evident.

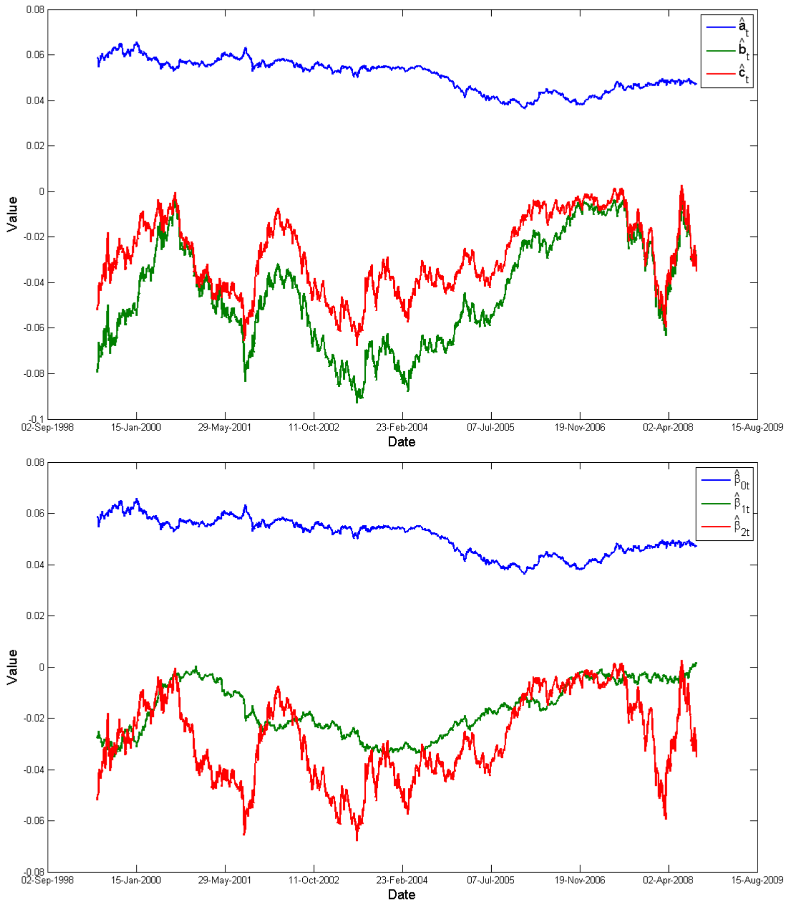


Figure 4.18: Development of estimated parameters using parametrization proposed by Nelson and Siegel (1987) (top) and by Diebold and Li (2006) (bottom).

List of Tables

List of Figures