Estimation of Bond Yield Curve by Yield Differences for Lithuanian Government Bonds

Obligacijų grąžų kreivės vertinimas naudojant grąžų skirtumus Lietuvos vyriausybės obligacijoms

Egidijus Vaičiulis

VILNIUS 2018
Obligacijų grąžų kreivės vertinimas naudojant grąžų skirtumus Lietuvos vyriausybės obligacijoms

Santrauka


Raktiniai žodžiai: obligacijos, kuponinės obligacijos, grąžų skirtumai, Nelson-Siegel grąža.

Estimation of Bond Yield Curve by Yield Differences for Lithuanian Government Bonds

Abstract

The paper considers bond pricing. Different yield curve specifications and loss functions are used to construct yield curves for pricing. This work introduces a procedure for estimation by yield difference minimization. The procedure is compared to price minimization for severa different curve specifications. The models are evaluated by changes in $MSE$ of prices.

Keywords: bonds, coupon bonds, yield differences, Nelson-Siegel yield.
Contents

1 Introduction 3
   1.1 Zero and fixed coupon bonds 3
   1.2 Relation between forward rate, yield and discount curves 4
   1.3 Lithuanian government bond data 4

2 Model specifications for bond yield curves 7
   2.1 Nelson-Siegel specification 7
   2.2 Nelson-Siegel-Svensson specification 8
   2.3 Merrill-Lynch Exponential Spline specification 8
   2.4 Applied estimation procedures 9
   2.5 Evaluation statistics 14

3 Empirical results 15
   3.1 Data processing 15
   3.2 Model fitting results 16
   3.3 Time consumption of yield difference optimization 18

4 Conclusions 19

5 References 20

A Appendices 21
1 Introduction

Bonds are a fixed income investments that allow the issuer to borrow funds at a fixed or variable interest rate. Essentially, the bond specifies future cash flows that a buyer of the bond is entitled to receive. The bonds that are modelled in this paper include zero and fixed coupon bonds. Using only such basic bonds eliminates the uncertainty of future cash flow amounts. The price of a bond is determined by discounting the individual future cash flows to the present. Discounting usually is done by constructing a forward rate, yield or discount curve for the entity issuing the bond and applying the respective rates to cash flows. This thesis compares the performance of two loss functions for modeling the forward and discount curves using several curve shape models that differ in flexibility. The loss functions are based on weighted price difference minimization and yield difference minimization.

1.1 Zero and fixed coupon bonds

The most appealing bonds to be used for yield curve construction are zero coupon bonds. Such bonds usually have a single lump sum payment at a future date called a notional amount and, therefore, only require a single yield for the price of the bond to be calculated. Coupon bonds pose the problem that a single price per coupon bond is observed in the market, but the price is affected by several yields as both the notional amount at bond maturity has to be discounted as well as each coupon payment. When there are no interest payments left between the present day and bond maturity, the coupon bond essentially becomes a zero coupon bond, however with a larger notional value as it will pay out the notional and interest for the last period upon maturing.
1.2 Relation between forward rate, yield and discount curves

The three curves are related functions and can be easily obtained from each other [Diebold and Rudebusch (2012)]. If we assume that the notional value of a zero coupon bond is 1 EUR, \( Y(t) \) denotes the yield, where \( t \) is the time to maturity, then the price \( P(t) \) given continuous compounding can be stated as:

\[
P(t) = \exp(-tY(t)).
\]  

Consequently, the forward rate curve \( F(t) \) can be calculated as follows:

\[
F(t) = -\frac{P'(t)}{P(t)}. 
\]  

From the forward rate curve we can obtain the yield curve:

\[
Y(t) = \frac{1}{t} \int_{0}^{t} F(u)du. 
\]  

Therefore, it is sufficient for a model to estimate any of the three curves.

1.3 Lithuanian government bond data

Lithuanian government regularly issues bonds. On 2017-12-12 the Lithuanian government had 66 bonds (excluding bonds up to 1 year maturity) outstanding with total face value of 3.45 billion EUR [CSDL (2017)] with maturities up to 10 years. This signifies an active market that may benefit from accurate yield curve estimation.

The used data set was retrieved from Bloomberg database on 2016-12-17 for the period from 1996-12-19 to 2016-12-16. The bond security identifier (ISIN) list was retrieved from Central Securities Depository of Lithuania [CSDL (2017)]. From the total list, at least partial price data was available for
103 bonds. Old price data was not collected consistently and was not available for full issuance of the bonds. Therefore, the data set was restricted to the period from 2012-09-07 to 2016-12-16. This period ensured at least 6 zero coupon bonds for initial value estimation for the Nelson-Siegel based models as the most extensive Nelson-Siegel type model has 6 estimated parameters. In the selected period the total amount of zero coupon bonds was 17 and 65 for coupon bonds. However, the Lithuanian bond quote data is relatively irregular as the number of quoted bonds each day varies. Within the restricted date range the number of quoted coupon bonds is at least 20 for most of the sample. Whereas there are only up to 2 zero coupon bonds quotes for half the period. The quote number distributions are given in Figures 1 and 2.

The time to maturity for zero coupon bonds was at most 2.9 years, whereas for coupon bonds it was 20 years. This suggests that zero coupon bonds cover a relatively small fraction of the yield curve and most of the curve will have to be built using coupon bond data. A sample day distribution of yields by maturity is presented in Figure 3.

The coupon bonds also had significantly different coupon rates, ranging from 0.3% to 9.95% (Figure 4) with higher coupons related to longer maturities. Taking into account the distribution of bond yields to maturity,
that would suggest that there was high uncertainty about the Lithuanian government long term yields as the coupons are significantly higher than the observed rates after bond release.

Figure 4: Coupon bond coupon rate distribution by percentages
2 Model specifications for bond yield curves

The aim of the paper is to compare the performance of yield models when estimation is done by yield difference minimization to the baseline estimation, which is done by minimizing price differences. In order to extensively test the performance, three different yield curve specifications are used that vary in their flexibility. The Nelson-Siegel specification is the most basic one with the least parameters to be estimated. The Nelson-Siegel-Svensson specification adds an additional mid term yield component, allowing more flexibility in the yield curve. The Merrill-Lynch exponential spline specification is the most flexible with significantly more terms and it also models the discount curve rather than the yield curve directly. The curve specifications are provided in the sections below.

2.1 Nelson-Siegel specification

The Nelson-Siegel specification suggests that there are three parts relating to the yield curve. A long term rate that is represented by $\beta_0$, a short term yield component represented by the term near $\beta_1$ and a medium term component represented by the term near $\beta_2$ (Figure 4). The yield for a zero coupon bond with remaining time to maturity $t$ is given by Guirreri (2010); Consiglio and Guirreri (2011):

$$\hat{Y}(t) = \beta_0 + \beta_1 \left[ \frac{1 - \exp(-t/\lambda_1)}{t/\lambda_1} \right] + \beta_2 \left[ \frac{1 - \exp(-t/\lambda_1)}{t/\lambda_1} - \exp(-t/\lambda_1) \right], \quad (2.1)$$

where $\beta_0, \beta_1, \beta_2$ and $\lambda_1$ are coefficients estimated during fitting. The $\beta_0, \beta_1, \beta_2$ coefficients determine the magnitude of each of the terms, while the $\lambda_1$ coefficient determines the timing and is also referred to as the decay coefficient. The larger the $\lambda_1$ value, the faster the short and mid term components approach zero.
2.2 Nelson-Siegel-Svensson specification

The Svensson (1994) extension adds an additional mid term yield component to the Nelson-Siegel specification to allow for a secondary hump or trough in the yield curve. It is done to capture a wider range of yield curve shapes at the cost of 2 additional parameters. With the Svensson extension the yield for a zero coupon bond with remaining time to maturity $t$ is given by Guirreri (2010); Consiglio and Guirreri (2011):

$$
\hat{Y}(t) = \beta_0 + \beta_1 \left[ \frac{1 - \exp(-t/\lambda_1)}{t/\lambda_1} \right] \\
+ \beta_2 \left[ \frac{1 - \exp(-t/\lambda_1)}{t/\lambda_1} - \exp(-t/\lambda_1) \right] \\
+ \beta_3 \left[ \frac{1 - \exp(-t/\lambda_2)}{t/\lambda_2} - \exp(-t/\lambda_2) \right],
$$

(2.2)

where $\beta_0, \beta_1, \beta_2, \beta_3, \lambda_1$ and $\lambda_2$ are coefficients estimated during fitting.

2.3 Merrill-Lynch Exponential Spline specification

The Merill-Lynch Exponential Spline specification specifies the discount curve instead of the yield curve. This specification was chosen due to its high flexibility in the number of terms that are used to construct the discount curve and relatively fast calculation time. The discount factor for a zero coupon
A bond with time to maturity equal to $t$ is given by Bolder et al. (2005):

$$
\hat{D}(t) = \sum_{k=1}^{9} z_k e^{-\alpha kt},
$$

(2.3)

where $z_1,...,z_9$ and $\alpha$ are coefficients estimated during fitting. The Yield curve can be derived by the following equation:

$$
\hat{Y}(t) = -(\ln(\hat{D}(t))/t),
$$

(2.4)

### 2.4 Applied estimation procedures

The estimation procedures have to differ for price and yield difference minimization due to the fact that for yield difference method an individual yield has to be assigned to each bond cash flow that results in the market price for each bond.

Two optimization algorithms are used for the estimation. One method is Differential Evolution (DE) Ardia et al. (2016), that is used only to find the best set of starting values for the Merrill-Lynch exponential spline specification. The algorithm is essentially creating a random population of coefficient sets. Then randomly combining the population members using random weights to form new sets and after a certain number of such mutations the best set is chosen. The algorithm is based on a random generator and therefore does not necessarily yield the same result every time, even when the same parameters are used.

Therefore, for any final optimization a quasi-Newton algorithm based on Broiden, Fletcher, Goldfarb and Shanno (BFGS) R Core Team (2015) is used. The method is based on constructing a Hessian approximation of the function being optimized and taking small iterative steps in the direction of the largest improvement in the function value according to the approximated
Hessian matrix. The method performance is significantly impacted by the initial values as the number of iterations required to reach a stationary point decreases if the initial values are close.

**Procedure for Nelson-Siegel and Nelson-Siegel-Svensson specification optimization by price difference minimization:**

1. Obtain initial coefficient values for the estimation. The initial values for the first day of estimation are equal to the coefficient estimates from only zero coupon bonds using an existing R implementation [Guirreri (2010); Consiglio and Guirreri (2011)]. For all next days, previous day final coefficient estimates are used.

2. The following price difference is minimized using BFGS method to obtain final coefficients:

   \[ \sum_{i \in I} (P_i - \hat{P}_i)^2, \]

   where \( I \) is the set of all bonds with market price quotes on the estimation day; \( P_i \) is the market price quote of bond \( i \);

   \[ \hat{P}_i = \sum_{k \in K_i} n_{i,k} \exp(-t_{i,k} \hat{Y}(t_{i,k})); \]

   \( K_i \) is the index of all future cash flows of bond \( i \);

   \( t_{i,k} \) is the time to maturity of cash flow \( k \) of bond \( i \);

   \( n_{i,k} \) is the amount of future cash flow \( k \) of bond \( i \);

   \( \hat{Y}(t_{i,k}) \) is the result of the Nelson-Siegel or Nelson-Siegel-Svensson specification with current iteration coefficients for time to maturity \( t_{i,k} \).

**Procedure for Merrill-Lynch specification optimization by price difference minimization:**

1. Obtain initial coefficient values for the estimation. The initial values for the first day of estimation are equal to the set of coefficients obtained from
differential evolution optimization algorithm that result in the lowest price difference:

$$\sum_{i \in I} (P_i - \hat{P}_i)^2,$$

(2.5)

where $I$ is the set of all bonds with market price quotes on the estimation day;

$P_i$ is the market price quote of bond $i$;

$\hat{P}_i = \sum_{k \in K_i} n_{i,k} \exp(-t_{i,k}\hat{Y}(t_{i,k}))$;

$K_i$ is the index of all future cash flows of bond $i$;

$t_{i,k}$ is the time to maturity of cash flow $k$ of bond $i$;

$n_{i,k}$ is the amount of future cash flow $k$ of bond $i$;

$\hat{Y}(t_{i,k})$ is the result of Merrill-Lynch specification with tested differential evolution coefficient set for time to maturity $t_{i,k}$.

For all next days, previous day final coefficient estimates are used.

2. The following price difference is minimized (same as Nelson-Siegel optimization) using BFGS method to obtain final coefficients:

$$\sum_{i \in I} (P_i - \hat{P}_i)^2.$$

Procedure for all specification optimization by yield difference minimization:

1. Obtain initial coefficient values for the estimation. The initial values for the first day of estimation are equal to the corresponding specification initial values in optimization by price difference. For all next days, previous day final coefficient estimates for the corresponding specification are used.

2. Using initial coefficients assign yields to all $i$ and $k$ so that:

$$\hat{y}_{i \in I, k \in K_i} = \hat{Y}^{\text{initial}}(t_{i,k}),$$

(2.6)
where $I$ is the set of all bonds with market price quotes on the estimation day;

- $K_i$ is the index of all future cash flows of bond $i$;
- $t_{i,k}$ is the time to maturity of cash flow $k$ of bond $i$;
- $\hat{Y}_{\text{initial}}(t_{i,k})$ is the result of Nelson-Siegel, Nelson-Siegel-Svensson or Merrill-Lynch specification with initial coefficient values for time to maturity $t_{i,k}$.

3. Perform shift of assigned $\hat{y}_{i,k}$:

$$\hat{y}_{i,k}^{\text{shift}} = \hat{y}_{i,k} + c_i,$$

(2.7)

where $c_i$ is a constant for each bond that satisfies the condition:

$$P_i = \sum_{k \in K_i} n_{i,k} \exp(-t_{i,k}\hat{y}_{i,k}^{\text{shift}}),$$

(2.8)

where $P_i$ is the market price quote of bond $i$;

- $n_{i,k}$ is the amount of future cash flow $k$ of bond $i$;
- $t_{i,k}$ is the time to maturity of cash flow $k$ of bond $i$.

4. Optimize $\hat{y}_{i,k}^{\text{shift}}$ values to minimize:

$$\sum_{i \in I, k \in K} t_{i,k} n_{i,k} (\hat{y}_{i,k}^{\text{shift}} - \hat{Y}_{\text{opt}}(t_{i,k}))^2,$$

(2.9)

where:

$\hat{Y}_{\text{opt}}(t_{i,k})$ is the Nelson-Siegel, Nelson-Siegel-Svensson or Merrill-Lynch specification with coefficients obtained from BFGS algorithm that minimize the yield difference above for a given set of $\hat{y}_{i,k}^{\text{shift}}$. The initial values for the first day of estimation are equal to the coefficient estimates from only zero coupon bonds using an existing R implementation [Guirreri (2010); Consiglio and Guirreri (2011)]. For all next days, previous day final coefficient estimates are used.
5. Final coefficients are equal to those obtained by BFGS minimization for the best values of $\hat{y}_{i,k}^{shift}$.

The function minimized in step 4 is weighted by time to maturity and cash flow amount. This weighting is chosen for three reasons. First being that the yields become relatively high for very short maturities as any change in price suggests a large yield. This happens due to prices being discrete values as only a limited set of decimal places is recorded. Secondly, yields assigned to longer maturity cash flows have a more significant impact on the price. To compensate for these two properties the yields are weighted by time to maturity that assigns lower significance for the low maturity cash flow yields. Thirdly, for each bond, only 1 price is observed. In cases when it is a zero coupon bond the only cash flow of the bond comprises all of the bond price. However, for coupon bonds the price can consist of many cash flows and the part in the market price is determined by the yield and amount of the cash flow. As yields are being estimated by the model, the weighting is done only by the cash flow amount.
2.5 Evaluation statistics

The aim is to evaluate the shape fit of the yield curve, rather than the closeness of fit of the specifications on the day of estimation. Additionally, comparing by MSE of price differences would obviously result in superiority of the price difference method as it is exactly the function minimized. Therefore, the statistic that will be used to evaluate the estimation is the change of price MSE between the day of estimation and the next available price quote date when using the estimation day coefficients. This will test whether the fitted yield curve suggests the correct change in yield from the initial estimation day to the next price quote day.

The statistic should ideally yield slightly negative values due to the fact that bond prices converge to the bond cash flow amount as maturity decreases and consequently the discount factor approaches 1.

**Definition for statistic calculation**

From the optimization procedures we obtain optimal coefficients for each of the three model specifications for each day which are then used to construct the optimal yield curves for each model. For each specification the following is calculated for each day $w$:

$$
\Delta MSE_w = \frac{1}{I_w} \sum_{i \in I} \left( P_{i,w+1} - \sum_{k} n_{i,k,w+1} \exp(-t_{i,k,w+1} \hat{Y}_w^{opt}(t_{i,k,w+1})) \right)^2 - \frac{1}{I_w} \sum_{i \in I} \left( P_{i,w} - \sum_{k} n_{i,k,w} \exp(-t_{i,k,w} \hat{Y}_w^{opt}(t_{i,k,w})) \right)^2,
$$

(2.10)

where $\Delta MSE_w$ denotes the change in mean squared error of the specification for day $w$;

$\frac{1}{I_w}$ is a weight that is equal to 1 divided by the number of unique bonds with quoted prices on both day $w$ and $w + 1$;
$P_{i,w}$ is the quoted market price of bond $i$ on day $w$;

$n_{i,k,w}$ is amount of $k^{th}$ future cash flow of bond $i$ as at day $w$;

t_{i,k,w}$ is the time to maturity of the $k^{th}$ future cash flow of bond $i$ as at day $w$.

3 Empirical results

3.1 Data processing

The bond price quote data from Bloomberg database was retrieved in the form of clean prices at market close. Clean bond prices do not include the accrued interest of the bond for the next coupon payment. These prices are used due to the property that they do not fluctuate significantly at the time of a bond coupon payment. Using such prices means that the cash flow of an upcoming bond coupon payment has to be calculated at each day separately as the coupon cash flow reaches 0 at the time of its maturity. These upcoming coupon payment cash flows were generated assuming a actual/actual coupon accrual scheme.

The time to maturity was calculated assuming that 1 year equals 365.25 days. Time was processed without separation into working and non-working days.

When calculating bond yields to maturity from observed bond prices, part of bonds had yields above 100%. All such bonds were close to their maturity, usually 1 week to maturity. Because such yields are caused due to bond prices being tracked to limited number of decimal places, the bond price quotes that suggest higher than 100% yields were removed from the dataset.
3.2 Model fitting results

The model fittings were done for the Nelson-Siegel specification, Nelson-Siegel-Svensson specification and Merill-Lynch specification with \( k = 5 \) and \( k = 9 \). The additional Merill-Lynch model with \( k = 5 \) was introduced due to high fluctuations in quoted bond price counts, which resulted in over parametrization by the specification with \( k = 9 \) in part of the sample.

The fittings by the price and yield difference minimization resulted in similar fittings for a significant part of the sample. Figures 7 and 8 show an example fit of the Nelson-Siegel model for two selected days, where one shows a relatively close fit by both procedures and a divergence in the second.

Figure 7: Nelson-Siegel fitting results for sample day 2012-10-30

Figure 8: Nelson-Siegel fitting results for sample day 2013-12-19

Table 1 gives the summary statistics for the fitted models. The average \( \Delta \text{MSE} \) for yield difference models is lower than price for all specifications and the Merill-Lynch models have an expected negative average, which would suggest a better fit to the bond data. It has to be noted, that due to over parametrization on some days, the Merill-Lynch specifications were fitted on a lower number of days from the sample than the Nelson-Siegel specifications. However, when yield models are compared to the price models by the percentage of times that the \( \Delta \text{MSE} \) was lower, only the Nelson-Siegel and Nelson-Siegel-Svensson specifications are higher for the yield difference mod-
Table 1: Model fitting summary

<table>
<thead>
<tr>
<th>Specification</th>
<th>Yield Diff</th>
<th>Price Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta MSE$</td>
<td>StDev</td>
</tr>
<tr>
<td>NS</td>
<td>0.00494</td>
<td>0.00321</td>
</tr>
<tr>
<td>NSS</td>
<td>0.00338</td>
<td>0.00433</td>
</tr>
<tr>
<td>ML, $k = 5$</td>
<td>-0.00538</td>
<td>0.02255</td>
</tr>
<tr>
<td>ML, $k = 9$</td>
<td>-0.00282</td>
<td>0.01517</td>
</tr>
</tbody>
</table>

Table 2: Kurtosis of $\Delta MSE$

<table>
<thead>
<tr>
<th>Specification</th>
<th>Yield Diff Kurtosis</th>
<th>Price Diff Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>161.712</td>
<td>36.911</td>
</tr>
<tr>
<td>NSS</td>
<td>58.079</td>
<td>287.73</td>
</tr>
<tr>
<td>ML, $k = 5$</td>
<td>170.172</td>
<td>188.275</td>
</tr>
<tr>
<td>ML, $k = 9$</td>
<td>946.411</td>
<td>287.729</td>
</tr>
</tbody>
</table>

This is due to the fact that Merill-Lynch models fitted by yield differences have higher negative values than price differencing (Figure 13). This makes the price difference models more consistent at predicting changes, although in some cases the yield differencing gives a significantly better prediction. The T-values presented in Table 1 are given assuming a normal distribution. However, the $\Delta MSE$ distributions of the models are not normally distributed. The distributions are leptokurtic (Table 2), which would suggest that the T-values are a conservative approximation. This still supports that the Nelson-Siegel model estimated by yield differences was significantly more accurate than the price difference model. For other models the differences are not significant and for Merill-Lynch specification with $k = 9$ the yield difference model is inconsistent.
Siegel-Svensson models the errors occur in a similar pattern, whereas for the Merill-Lynch model the errors are completely uncorrelated between yield and price minimization models (Table 3). This suggests that Merill-Lynch models are influenced by different factors when estimated by yield and price difference minimization.

### 3.3 Time consumption of yield difference optimization

The applied optimization procedure for yield differences does have significant disadvantages when compared to price differencing that arise from its nature. The main disadvantage is the time it takes to minimize the yield differences. The method requires to assign yields to each cash flow that may deviate from the fitted curve by different amounts. In the case of zero coupon bonds, both the price and yield difference difficulty is the same as there is only 1 yield per bond and it is observed through market price quote. However, for coupon bonds yield differences increases as it adds an additional optimization parameter for each coupon payment. This significantly increases the time required to reach a stationary value for the yield difference method and during the model fittings for the sample the time required for 1 day computation differed by up to a factor of $10^3$ for some days, where there were only coupon bond quotes observed. This factor is significantly reduced for days
with coupon bonds available as the zero coupon bonds provide the basis of the yield curve and the optimization only slightly adjusts the coefficients from the initial values.

4 Conclusions

The goal of this study was to compare yield curve estimation by yield differences to price differences in terms of predicting the next day change in yields. For this purpose Nelson-Siegel, Nelson-Siegel-Svensson and Merill-Lynch yield models were implemented for Lithuanian government bonds.

Yield difference minimization showed significantly more accurate results for the Nelson-Siegel specification with lower variance than the price difference minimization. However, for other models the accuracy was not significantly different or was inconsistent.

The introduced yield difference minimization method had significantly higher time costs when estimating curves while using mainly coupon bonds. However, the estimation time decreases significantly when more zero coupon bonds are observed.
5 References


Appendices

Figure 9: $\Delta MSE$ for Nelson-Siegel specification by price differences

Figure 10: $\Delta MSE$ for Nelson-Siegel specification by yield differences

Figure 11: $\Delta MSE$ for Nelson-Siegel-Svensson specification by price differences

Figure 12: $\Delta MSE$ for Nelson-Siegel-Svensson specification by yield differences
Figure 13: $\Delta MSE$ for Merrill-Lynch, $k = 5$ specification by price differences

Figure 14: $\Delta MSE$ for Merrill-Lynch, $k = 5$ specification by yield differences

Figure 15: $\Delta MSE$ for Merrill-Lynch, $k = 9$ specification by price differences

Figure 16: $\Delta MSE$ for Merrill-Lynch, $k = 9$ specification by yield differences