

termstrc: A Package for Term Structure Estimation with R

Robert Ferstl Josef Hayden

Abstract

Zero-coupon yield curves and credit spread curves are important inputs for various financial models, e.g. pricing of securities, risk management, monetary policy issues. Since zero-coupon rates are rarely directly observable, they have to be estimated from market data for existing coupon bonds. The literature broadly distinguishes between parametric and spline-based methods. We implement two widely-used term structure estimation procedures, i.e. the parametric Nelson and Siegel approach and the Svensson approach.

Moreover, we implement the traditional way of credit spread calculation, where individually estimated zero-coupon yield curves are subtracted from a risk-free reference curve. Goodness-of-fit tests are provided to compare the results of the different estimation methods. We illustrate the usage of our functions by practical examples with data from European and CEE government bonds, and European corporate bonds.

Keywords: term structure, interest rates, R.

1. Introduction

The R package **termstrc** consists of methods for estimating the term structure of interest rates from market data.

In this section we provide the basic concepts of bond pricing which are required for the estimation of the term structure of interest rates. Section 2 explains the ideas behind the two used models. Details about the notation and the optimisation procedure are presented in section 3. Section 4 includes a complete step-by-step example.

Since zero-coupon rates are rarely directly observable, they have to be estimated from market data for existing coupon bonds. For a coupon bond the following information is typically observable on the market: the maturity date m , the clean price p_c and the cashflows c_t (coupons and redemption payment) which occur at $t = 1, \dots, n$. An investor who wants to buy the bond has to pay the dirty price p_d which consists of the quoted market price (clean price) p_c and the accrued interest a .

The term structure of interest rates can usually be described in three ways: the discount function, the spot rate function and the forward rate function.

The bond pricing equation under continuous compounding is the present value of all cashflows.

$$p_c + a = \sum_{t=1}^n c_t e^{-s_t m_t} \quad (1)$$

The spot rate s_t is the yield-to maturity for a t -period zero-coupon bond. m_t is the time to maturity of the t -th cash flow. A plot of the yields against times to maturity produces the zero-coupon yield curve.

The yield-to-maturity is the solution for y in the following equation.

$$p_c + a = \sum_{t=1}^n c_t e^{-y m_t} \quad (2)$$

An equivalent formulation of the bond price equation makes use of the discount factors $d_t = \delta(m_t) = e^{-s_t m_t}$. The continuous discount function $\delta(\cdot)$ is formed by interpolation of the discount factors.

$$p_c + a = \sum_{t=1}^n c_t \delta(m_t) \quad (3)$$

The implied j -period forward rate is calculated as

$$f_{t|j} = \frac{js_j - ts_t}{j - t} = s_j + (s_j - s_t) \frac{t}{j - t} \quad (4)$$

The instantaneous forward at the time t equals

$$f_t = \lim_{\Delta t \rightarrow 0} \left[s_{t+\Delta t} + (s_{t+\Delta t} - s_t) \frac{t}{\Delta t} \right] = s_t + t \frac{\partial s}{\partial t} \quad (5)$$

The spot rate is the integral of the instantaneous forward rate.

2. The Nelson/Siegel and Svensson method

Assuming that a function between the spot rates and the maturities exists, we can estimate the term structure of interest rates. [Nelson and Siegel \(1987\)](#) proposed a parsimonious model of the yield curve. They model the forward rate with a second order differential equation.

$$f(m, \mathbf{b}) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} \exp\left(-\frac{m}{\tau_1}\right) \quad (6)$$

The corresponding spot rate curve is as follows.

$$s(m, \mathbf{b}) = \beta_0 + \beta_1 \frac{1 - \exp(-\frac{m}{\tau_1})}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1 - \exp(-\frac{m}{\tau_1})}{\frac{m}{\tau_1}} - \exp(-\frac{m}{\tau_1}) \right) \quad (7)$$

The choice of the parameters $\mathbf{b} = (\beta_0, \beta_1, \beta_2, \tau_1)$ enables us to construct various possible shapes of the spot rate and the instantaneous forward rate curve. The constant term β_0 is the limit of the instantaneous forward rate when the time to maturity tends to infinity. The first exponential term $\beta_1 \exp(-\frac{m}{\tau_1})$ can be interpreted as a short-term contribution. It will decrease monotonically to zero with time to maturity when β_1 is positive. The second exponential term $\beta_2 \frac{m}{\tau_1} \exp(-\frac{m}{\tau_1})$ generates a hump-shape, which is realized in the medium term, because it starts at zero and decays to zero.

The following code replicates figure 2 on page 477 in [Nelson and Siegel \(1987\)](#), see Figure 1.

```
m <- seq(0,10,0.01)      # time to maturity

beta_1 <- 1               # parameters
tau_1 <- 1
beta_2 <- 1
beta_0 <- 1

par(lwd=2)

# beta_1
```

```

plot(m,beta_1*exp(-m/tau_1),type="l",col=1,
     xlab="Time to maturity",
     ylab="Model curves",ylim=c(0,2.1),lty=4)

# beta_2
lines(m,beta_2*(m/tau_1)*exp(-m/tau_1),lty=2,col=2)

# forward rate curve
lines(m,beta_0+beta_1*exp(-m/tau_1)+beta_2*(m/tau_1)*exp(-m/tau_1),col=3)

abline(h=1,lty=3,col=4)

legend("topright",legend=c(expression(beta[1]*exp(-m/tau[1])),
  expression(beta[2]*(m/tau[1])*exp(-m/tau[1])),
  expression(beta[0]),
  expression(f(m,b))),
  lty=c(4,2,3,1), col=c(1,2,4,3))

```

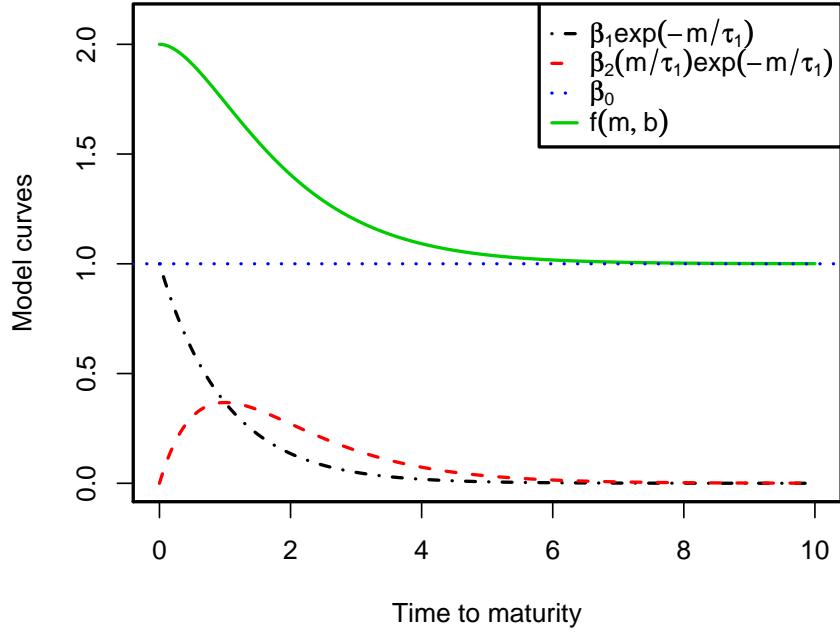


Figure 1: Decomposition of the Nelson/Siegel forward rate function

Svensson (1994) extended the functional form of the forward rate curve by two additional parameters which allows for a second hump-shape.

$$f(m, \mathbf{b}) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} \exp\left(-\frac{m}{\tau_1}\right) + \beta_3 \frac{m}{\tau_2} \exp\left(-\frac{m}{\tau_2}\right) \quad (8)$$

The spot rate is calculated as

$$s(m, \mathbf{b}) = \beta_0 + \beta_1 \frac{1 - \exp(-\frac{m}{\tau_1})}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1 - \exp(-\frac{m}{\tau_1})}{\frac{m}{\tau_1}} - \exp(-\frac{m}{\tau_1}) \right) + \beta_3 \left(\frac{1 - \exp(-\frac{m}{\tau_2})}{\frac{m}{\tau_2}} - \exp(-\frac{m}{\tau_2}) \right), \quad (9)$$

with a parameter vector $\mathbf{b} = (\beta_0, \beta_1, \beta_2, \tau_1, \beta_3, \tau_2)$.

The impact of the parameters can be described as follows (see [Bolder and Streliski 1999](#), p.7):

- β_0 is the asymptotic value of the forward rate (spot rate) function. The curve will tend towards the asymptote as the time to maturity approaches ∞ ($\beta_0 > 0$).
- β_1 determines the starting (short-term) value of the curve in terms of deviation from the asymptote (the sum β_0 and β_1) is the vertical intercept. Moreover, it defines the basic speed with which the curve tends toward its long-term trend.
- τ_1 specifies the position of the first hump or the U-shape on the curve ($\tau_1 > 0$).
- β_2 determines the magnitude and direction of the hump. If $\beta_2 > 0$ a hump will occur at τ_1 , whereas $\beta_2 < 0$, a U-shaped value will occur at τ_1 .
- τ_2 specifies the position of the second hump or the U-shape on the curve ($\tau_2 > 0$).
- β_3 analogously to β_2 determines the magnitude and direction of the second hump.

The following code replicates figure 1 on page 8 in [Bolder and Streliski \(1999\)](#). It shows the decomposition of the forward rate curve, see Figure 2.

```
m <- seq(0,15,0.01)      # time to maturity

beta_1 <- 1               # parameters
tau_1 <- 1
beta_2 <- 1
beta_0 <- 1
tau_2 <- 2
beta_3 <- 2

par(lwd=2)

# beta_1
plot(m, beta_1*exp(-m/tau_1), type="l", col=1,
     xlab="Time to maturity",
     ylab="Model curves", ylim=c(0,2.4), lty=4)

# beta_2
lines(m, beta_2*(m/tau_1)*exp(-m/tau_1), lty=2, col=2)

# beta_3
lines(m, beta_3*(m/tau_2)*exp(-m/tau_2), lty=6, col=6)

# forward rate curve
lines(m, beta_0+beta_1*exp(-m/tau_1)+beta_2*(m/tau_1)*exp(-m/tau_1) +
     beta_3*(m/tau_2)*exp(-m/tau_2), col=3)

# beta_0
```

```
abline(h=1,lty=3,col=4)

legend("topright",legend=c(expression(beta[1]*e^(-m/tau[1])),
  expression(beta[2]*(m/tau[1])*e^(-m/tau[1])),
  expression(beta[0])),
  expression(beta[3]*(m/tau[2])*e^(-m/tau[2])),
  expression(f(m,b))),
  lty=c(4,2,3,6,1), col=c(1,2,4,6,3))
```

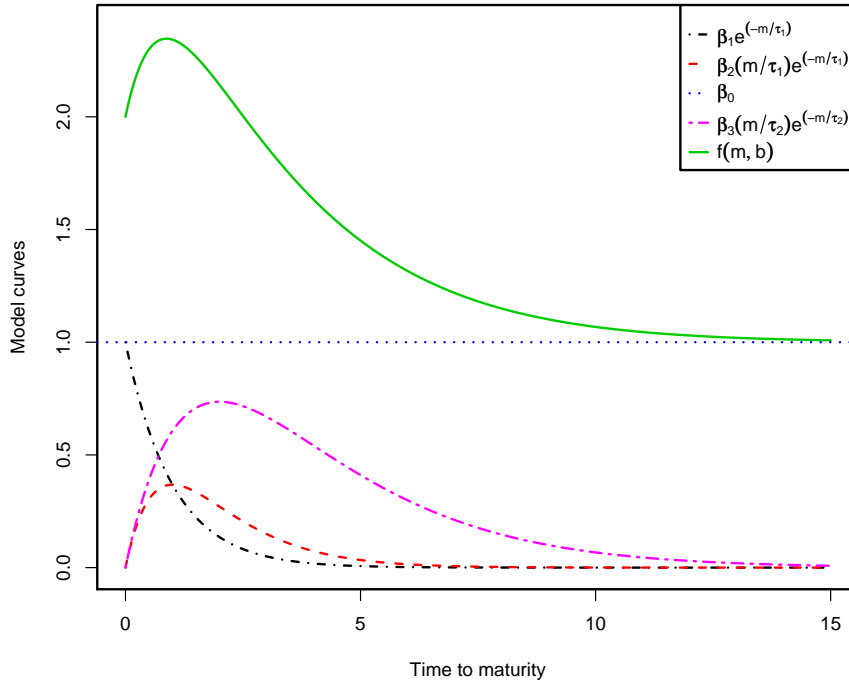


Figure 2: Decomposition of the Svensson forward rate function

Credit-spread estimation

The credit-spread estimation in the package **termstrc** is limited to the usual way of calculating a credit spread curve, i.e subtracting the yield-curve of a reference country $s_{ref}(\mathbf{m}, \mathbf{b})$ from the yield-curves $s_j(\mathbf{m}, \mathbf{b})$ of other countries .

$$cs_j(\mathbf{m}) = s_j(\mathbf{m}, \mathbf{b}) - s_{ref}(\mathbf{m}, \mathbf{b}) \quad (10)$$

$cs_j(\mathbf{m})$ credit-spread between country j and reference country ref
 $s_j(\mathbf{m}, \mathbf{b})$ spot-rate curve of country j with maturity vector \mathbf{m}
 $s_{ref}(\mathbf{m}, \mathbf{b})$ spot-rate curve of the reference country

The spot-rate functions are defined as in equation (7) and (9). For a more advanced method of credit-spread estimation see (Jankowitsch and Pichler 2004).

3. Estimation procedure

In this section we describe in detail the optimisation problem which has to be solved when estimating the term structure of interest rates from market data.

3.1. Notation

Maturity matrix \mathbf{M}

$$\mathbf{M}_{[n \times m]} = \{m_{ij}\} \quad (11)$$

The number of rows n is determined through the number of cashflows of the bond j with the longest maturity. For each bond j exists a column with the corresponding cashflow dates. Dates after the maturity of the bond j are filled up with zeros till the maturity date of the bond with the longest maturity. One element m_{ij} of the matrix refers, therefore, to the maturity date of the i -th cashflow of the j -th bond.

Cashflow matrix \mathbf{C}

$$\mathbf{C}_{[n \times m]} = \{c_{ij}\} \quad (12)$$

The cashflow matrix is defined analogously to the maturity matrix. One element c_{ij} of the matrix refers to the cashflow of the i -th cashflow of the j -th bond. Note, that the last cashflow of a each bond includes the redemption payment.

Discount factor matrix \mathbf{D}

$$\mathbf{D}_{[n \times m]} = \{d_{ij}\} \quad (13)$$

The discount factor matrix is also defined analogously to the maturity matrix. One element d_{ij} of the matrix refers to the discount factor associated with the i -th cashflow of the j -th bond. An individual discount factor d_{ij} is constructed as

$$d_{ij} = e^{-m_{ij}s(m_{ij},b)},$$

where $s(m_{ij},b)$ is the Nelson/Siegel or Svensson spot rate function defined in equation (7) and (9).

Clean price vector \mathbf{p}^c

$$\mathbf{p}_{[1 \times m]}^c = \{p_j^c\} \quad (14)$$

p_{c_j} is the quoted price of the j -th bond ($j = 1 \dots m$).

Accrued interest vector \mathbf{a}

$$\mathbf{a}_{[1 \times m]} = \{a_j\} \quad (15)$$

Different conventions for the calculation of accrued interest are used in the market. A basic form for the j -th bond is as follows.

$$a_j = \frac{\text{number of days since last coupon payment}}{\text{number of days in current coupon period}} \cdot \text{coupon}_j \quad (16)$$

Dirty price vector \mathbf{p}^d

$$\mathbf{p}_{[1 \times m]}^d = \{p_j^d\} \quad (17)$$

The dirty price vector is the sum of the clean price vector and the accrued interest vector and consists of the dirty prices of all bonds j .

$$\mathbf{p}^d = \mathbf{p}^c + \mathbf{a}$$

Weights vector \mathbf{w}

$$\mathbf{w}_{[1 \times m]} = \{w_j\} \quad (18)$$

Whereas w_j is the weight for bond j with Duration d_j :

$$w_j = \frac{\frac{1}{d_j}}{\sum_{i=1}^m \frac{1}{d_i}}$$

The duration for a bond j is a weighted average of the time to cashflows.

$$d_j = \frac{\mathbf{C}_{[n \times j]} (\mathbf{D}_{[n \times j]} \cdot \mathbf{M}_{[n \times j]})^t}{\mathbf{C}_{[n \times j]} (\mathbf{D}_{[n \times j]})^t} \quad (19)$$

Whereas (\cdot) denotes a element by element multiplication and $()^t$ the transposed matrix (vector).

3.2. Objective function

In the objective function we minimise the sum of weighted pricing errors. The weights are used to avoid heteroscedasticity.

$$\mathbf{b}_{opt} = \min_b \left((\boldsymbol{\iota}_{[1 \times n]} [\mathbf{C} \cdot \mathbf{D}] - \mathbf{p}^d)^2 \mathbf{w} \boldsymbol{\iota}_{[m \times 1]} \right) \quad (20)$$

The element by element multiplication of the cashflow matrix \mathbf{C} with the discount factor matrix \mathbf{D} returns a matrix with the present values of all cash flows. Multiplying this with a unit vector $\boldsymbol{\iota}$ from the left results in the vector of fair bond prices.

The parameter vector is subject to the constraints $\beta_0 > 0, \tau_1 > 0, \tau_2 > 0$.

3.3. Start parameters for the numerical optimisation

The selection of good start parameters is important to find a global minimum of the objective function.

[Bolder and Streliski \(1999\)](#) point out that the estimation difficulties are caused by the spot rate function (forward rate function), which is linear in the betas, however, non-linear in the taus. They present two different approaches for the generation of start parameters. The first one is based on an enumeration algorithm. A matrix with different start parameter sets is generated. For each set the optimization is performed and the one with the best solution is chosen as start parameter set for the final optimisation. The second approach divides the parameters into linear and non-linear. For the optimisation one group is held constant, whereas the other group is varied. This method improves the speed of convergence.

[Ferenczi \(2005\)](#) proposed an optimisation algorithm which is based on the HCP Algorithm of [Novak and Ritter \(1996\)](#). The advantage of this algorithm is, that it converges to the global minimum of the objective function.

4. Examples

In this section we present a detailed description of the **termstrc** package. We describe the included data sets, the available functions and show the results of the demo examples.

4.1. Data sets

The package **termstrc** includes two data sets in list format. The first one (**eurobonds**) consists of Austrian, German, Italian and Hungarian government bond data. The second one (**corpbonds**) includes data of bonds from different rating categories (AAA, AA+, AA, AA-, A+, A, A-, BBB+, BBB, BBB-). For each bond the ISIN-code, maturity date, start date, coupon rate, quoted price, accrued interest, cashflows, cashflow dates and the date stamp for the quoted price are available.

The structure of the data sets can be explored using the function **str**:

```
| str(eurobonds)
```

```
List of 4
```

```
$ AUSTRIA:List of 8
..$ ISIN      : chr [1:14] "AT0000383690" "AT0000383864" ...
..$ MATURITYDATE:Class 'Date'  num [1:14] 13614 21014 13893 ...
..$ STARTDATE  :Class 'Date'  num [1:14] 9962 10057 10241 ...
..$ COUPONRATE : num [1:14] 0.0575 0.0625 0.0500 0.0413 0.0400 ...
..$ PRICE      : num [1:14] 104 135 105 105 104 ...
..$ ACCRUED    : num [1:14] 3.48 2.16 4.21 3.47 1.38 ...
..$ CASHFLOWS  :List of 3
.. ..$ ISIN: chr [1:115] "AT0000383690" "AT0000383690" ...
.. ..$ CF : num [1:115] 5.75 105.75 6.28 6.23 6.24 ...
.. ..$ DATE:Class 'Date'  num [1:115] 13249 13614 13346 13710 ...
..$ TODAY      :Class 'Date'  num 13102
$ GERMANY:List of 8
..$ ISIN      : chr [1:29] "DE0001134468" "DE0001134922" ...
..$ MATURITYDATE:Class 'Date'  num [1:29] 16972 19726 13517...
..$ STARTDATE  :Class 'Date'  num [1:29] 6014 8769 9865 9976 10046 ...
..$ COUPONRATE : num [1:29] 0.0600 0.0625 0.0600 0.0600 0.0650 ...
..$ PRICE      : num [1:29] 121 132 104 105 139 ...
..$ ACCRUED    : num [1:29] 2.48 5.45 5.23 2.25 2.44 ...
..$ CASHFLOWS  :List of 3
....
....
....
```

4.2. Functionality of the package

The complete functionality of the package is provided by the function **termstrc_estim**, that is designed as follows:

```
| termstrc_estim(countries, bonddata, maturity_spectrum, method, fit,
|               weights, startparam, control)
```

Table 4.2 gives a description of the input parameters.

The first element of the **countries** vector is used as the reference country for the credit-spread calculation. A specific start parameter algorithm has not been implemented. Therefore, a start parameter set for each country has to be provided by the user. The start parameter matrix for the countries Austria, Germany and Italy may look like this:

Input	Description	e.g.
<code>countries</code>	countries vector	<code>c("AUSTRIA", "ITALY")</code>
<code>bonddata</code>	data set	<code>eurobonds</code>
<code>maturity_spectrum</code>	maturity spectrum	"all" or e.g. <code>c(1,20)</code>
<code>method</code>	Nelson/Siegel or Svensson	"Nelson/Siegel" od. "Svensson"
<code>fit</code>	Optimisation method	"prices" or "yields"
<code>weights</code>	Durationweighting	"none" or "duration"
<code>startparam</code>	start parameter vector	<code>b <- matrix(rep(1,8),nrow=2)</code>
<code>control</code>	control parameter for <code>nlminb</code>	<code>control=list(eval.max=1000)</code>

Table 1: Description of inputs for the function `termstrc_estim`

	<code>beta0</code>	<code>beta1</code>	<code>beta2</code>	<code>tau1</code>
GERMANY	0.02547394	-0.01216259	-0.02547394	1
AUSTRIA	0.02611532	-0.01136742	-0.02611532	1
ITALY	0.02578871	-0.01520725	-0.02578871	1

The elements und sub-lists of the returned list of the function `termstrc_estim` are shown in table 4.2. You may use this code to explore the structure of the returned list:

```
library(termstrc)
demo(euro01)
str(myres)
```

List element	Content
maturity_spectrum	Includes the chosen maturity spectrum
method	Includes the chosen estimation method
fit	Includes the chosen objective function
weights	Includes the type of optimisation, i.e "none" or "duration"
n_countries	the number of countries (rating categories) used for the optimisation
cashflows	the cashflow matrix for all specified countries (sub-list)
maturities	the maturity matrix for all specified countries (sub-list)
dirty_prices	the dirty prices for all specified countries(sub-list)
estimated_prices	the estimated prices for all specified countries (sub-list)
yields	the calculated yields for all bonds of the specified countries (sub-list)
opt_result	the optimal parameter vector for the specified countries (sub-list)
spotrates	the spotrates calculated with the optimal parameter vector (sub-list)

Table 2: content of the returned list of the function `termstrc_estim`

S 3 Methods

The package uses the R - S3 class definition. For the class `termstrc_singlecurve` exists a `summary`, `print` and `plot` method.

The `summary` method prints goodness of fit test for the price and yield deviation (root mean square error and average absolute error) and convergence information of the used solver (`nlminb`). The `print` method prints the optimal parameter set for the specified countries or rating categories. Different plots for the zero-coupon- yield curves of the countries and spread curves are provided by the `plot` method.

4.3. Demos

In this section we present a step by step guidance on how to use the package to estimate the term-structure. We use the included data sets.

Step 1: Load the package and the desired data set

```
library(termstrc)
data(eurobonds)
```

Step 2: Set the parameters of the function `termstrc` as required

```
countries <- c("GERMANY", "AUSTRIA", "ITALY")
bonddata <- eurobonds
maturity_spectrum <- "all"
method <- "Nelson/Siegel"
fit <- "prices"
weights <- "duration"
control <- list(eval.max=100000)

b <- matrix(c(0.02547394, -0.012162592, -0.02547394, 1,
              0.02611532, -0.011367422, -0.02611532, 1,
              0.02578871, -0.015207250, -0.02578871, 1),
            nrow=3, ncol=4, byrow=TRUE)

rownames(b) <- countries
colnames(b) <- c("beta0", "beta1", "beta2", "tau1")
```

Step 3: Assign the function

```
myres <- termstrc_estim(countries, bonddata, maturity_spectrum,
                      method, fit, weights, startparam=b, control)
```

Step 4: Use the S3 print method to get the optimised parameter set

```
print(myres)
```

```
-----
Parameters for single-curve estimation:
```

```
Method: Nelson/Siegel
Fitted: prices
Weights: duration
```

```
-----

      GERMANY    AUSTRIA    ITALY
beta_0 0.04153832 0.04228676 0.04599300
beta_1 -0.01699020 -0.01813795 -0.02089901
beta_2 -0.01326502 -0.01208279 -0.01874076
tau_1   2.20819586  2.48874374  2.36039421
```

Step 5: Use the S3 summary method to get the goodness of fit test and the convergence information

```
summary(myres)
```

Goodness of fit tests:

	GERMANY	AUSTRIA	ITALY
RMSE-Prices	0.2852092109	0.0606758664	0.1588754369
AABSE-Prices	0.1629764768	0.0416447015	0.1041235559
RMSE-Yields	0.0004808922	0.0003668975	0.0007122635
AABSE-Yields	0.0004177424	0.0002694297	0.0006036436

Convergence info

GERMANY "relative convergence (4)"
 AUSTRIA "relative convergence (4)"
 ITALY "relative convergence (4)"

Step 6: Use the S3 plot method to generate different plots of the termstructure and the spread-curves. By default the zero-coupon-yield curve is plotted separately for each specified country and for all countries together. Moreover the spread-curves are plotted.

```
| plot(myres)
```

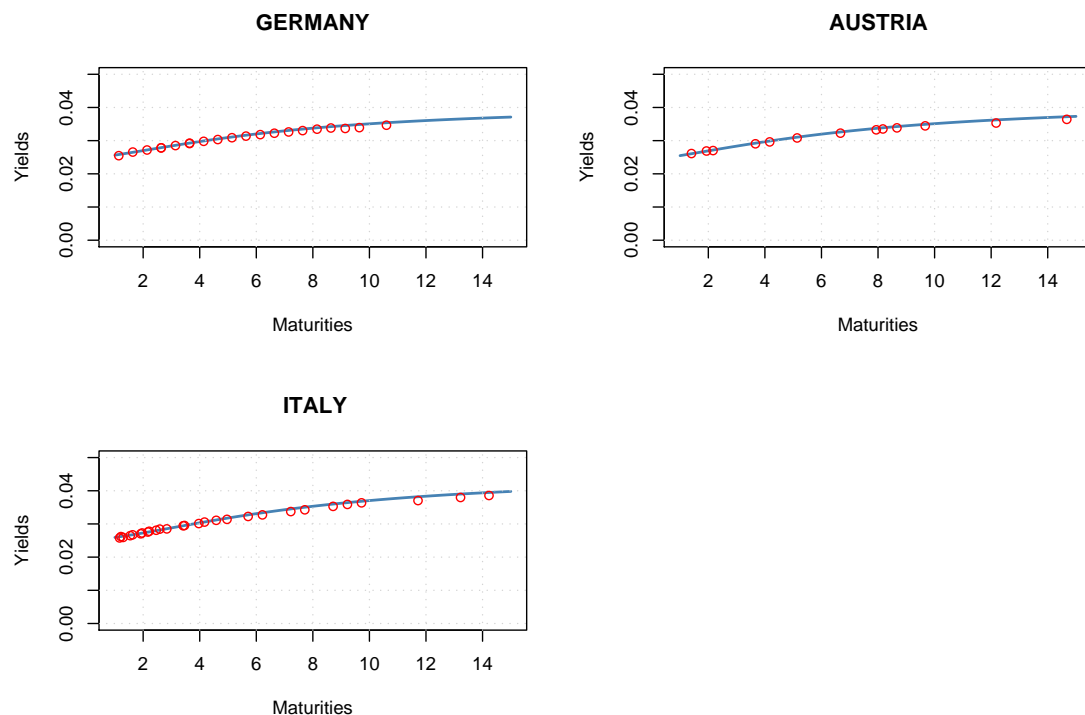


Figure 3: Zero-coupon yield curves

The presented example is easily accessible using the demo function in R. Use `demo(package="termstrc")` to explore the included demo examples.

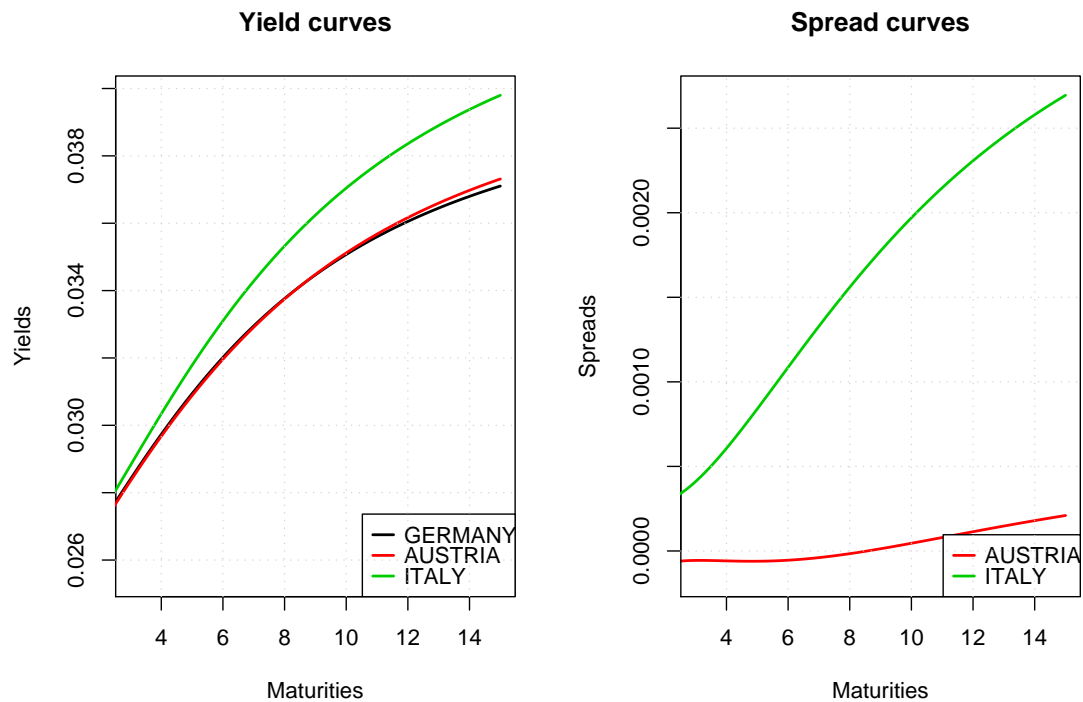


Figure 4: Spread curves

5. Discussion

Acknowledgements

The authors are grateful to the students of the finance management science lab class of 2005.

List of Tables

1	Description of inputs for the function <code>termstrc_estim</code>	9
2	content of the returned list of the function <code>termstrc_estim</code>	10

List of Figures

1	Decomposition of the Nelson/Siegel forward rate function	3
2	Decomposition of the Svensson forward rate function	5
3	Zero-coupon yield curves	12
4	Spread curves	13

References

Bolder D, Streliski D (1999). “Yield Curve Modelling at the Bank of Canada.” *Technical Reports 84*, Bank of Canada.

- Ferenczi I (2005). *Globale Optimierung unter Nebenbedingungen mit dünnen Gittern*. Master's thesis, Technische Universität München - Zentrum Mathematik. URL <http://www.mathematik.uni-wuerzburg.de/~ferenczi/tagungen.html>.
- Jankowitsch R, Pichler S (2004). "Parsimonious Estimation of Credit Spreads." *The Journal of Fixed Income*, **14**(3), 49–63. 10598596.
- Nelson C, Siegel A (1987). "Parsimonious Modeling of Yield Curves." *The Journal of Business*, **60**(4), 473–489.
- Novak E, Ritter K (1996). *Global Optimization Using Hyperbolic Cross Points*, pp. 19–33. Kluwer Academic Publishers.
- Svensson LE (1994). "Estimating and Interpreting Forward Interest Rates: Sweden 1992 -1994." *Technical Reports 4871*, National Bureau of Economic Research, Inc.