ESQUANT
Statistical Consulting

drpr package

Neil Diamond, B.Sc.(Hons), Ph.D., A.Stat.

ESQUANT Statistical Consulting

February 22, 2016

Contents

List of Tables

1 Introduction

The ERA (WA) has set out a series of steps using Excel to calculate the Debt Risk Premium (DRP), see Attachment 3 of (ERA, 2015). In calculating the DRP, the ERA (WA) used three techniques:

- Gaussian Kernel
- The Nelson-Siegel Methodology
- Nelson-Siegel-Svensson Model

For all three techniques, the ERA (WA) give detailed instructions as to how to calculate the cost of debt and how to combine the results to give a final figure.

2 Reproduction of ERA Excel calculations

The ERA Excel calculations have been reproduced in a series of workbooks. The workbooks have been named according to the Table number in the ERA report. The calculations have been done for the 20 day period beginning 2 January 2015 and ending 30 January 2015. I've assumed that the data selection steps have been followed, that is the input data for Table 145 of ERA(2015) are available.

In presenting the results, I have given comparisons to the results calculated in R using the **drpr** package, outlined in section 4.

2.1 Table 145

Given a series of remaining terms to maturity, amount issued in Australian dollars, and yields, this workbook gives the Gaussian Smoothed yield according to the RBA Methodology (Arsov et al. 2013). The workbook is intended to be used four times with Target Tenors in cell K1 of 3, 5, 7, and 10 years. The output, to be inserted into the workbook corresponding to Table 146, is the Target Tenor yield, given in cell K9, and the Effective Term to Maturity, given in the last cell of column J.

The workbook is named `Table145.xlsx`. The results for 3, 5, 7, and 10 years are in `Table145Tenor3.xlsx`, `table145Tenor5.xlsx`, `Table145Tenor145Tenor7.xlsx`, and `Table145Tenor10.xlsx`, respectively.

The results are given in Table 1. There are no differences between the Excel and R results to four decimal places.

Tenor	Effective Term (Excel)	Target Tenor Yield (Excel)	Effective Term (R)	Target Tenor Yield (R)
3	3.9903	3.9762	3.9903	3.9762
5	5.2397	4.2135	5.2397	4.2135
7	6.6277	4.4960	6.6277	4.4960
10	8.5236	4.3652	8.5236	4.3652

Table 1: Output from Table145

2.2 Table 146

Taking the results from Table 145, in this table the results are converted to Annualized yields, and then the extrapolation of the results is carried out, in order to remove the bias of the Gaussian smoothing method. The workbook is named `Table146.xlsx` and the annualized 10 year yield is given in cell F2, and the extrapolated 10 year annualized yield is given in cell F4.

The results are given in Table 2.There are only marginal differences to four decimal places.

2.3 Table 147

This table gives the starting value for the decay factor λ and is given in `Table147.xlsx`. The input workbook is `Table147Input.xlsx` and the output workbook is `Table147Output.xlsx`. The starting value is 0.7173 with both Excel and R.

	Excel	R
Extrapolated 10 Year yield (Semi-Annual Basis)	4.2633	4.2633
Extrapolated 10 Year yield (Annualized)	4.3088	4.3087

Table 2: Output from Table146

2.4 Table 148

The starting value for the decay factor, 0.7173, is input into cell A1 of `Table148.xlsx`. Columns B and C are the remaining terms to maturity and yield values, respectively. Columns D and E are the first and second loadings, corresponding to the Nelson-Siegel model with λ given by 0.7173. Multiple linear regression, with an intercept, is applied with columns D and E as the independent variables and Column C as the dependent variable. Cells G17, G18, and G19 give the starting values for beta1, beta2, and beta3, respectively.

The input workbook is `Table148.xlsx` and the output workbook is `table148Output.xlsx`. The starting values are given in Table 3. Both Excel and R have the same starting values.

	Excel	R
beta1	5.5258	5.5258
beta2	-1.4936	-1.4936
beta3	-4.0384	-4.0384
lambda	0.7173	0.7173

Table 3: Start Values (Output from Table148)

2.5 Table 149

With the starting values, this workbook calculates the residual sum of squares. Solver is then used to minimise the residual sum of squares by changing the parameter values. The input workbook is given in `Table149Input.xlsx` and the output workbook is given in `Table149Output.xlsx`. Note that the parameter estimates are given in cells E1 to E4, given in Table 4. There are slight differences between Excel and R[1].

	Excel	R
beta1	7.0060	7.0128
beta2	-3.9174	-3.9232
beta3	0.0085	-0.0078
lambda	0.1292	0.1293

Table 4: Nelson-Siegel Parameter Estimates (Output from Table149)

2.6 Table 150 and 151

Taking the parameter estimates from Table 149, in this table the semi-annualized yields at 3, 5, 7, and 10 years are calculated and then converted to annualized yields. The workbook is given in `Table150and151.xlsx` and the 10 year annualized yield is given in cell D3.

The results are given in Table 5. The results differ only in the fourth decimal place.

[1] There are also some slight differences in the parameter estimates if the package is run on a Windows or Linux computer rather than on a Mac (on which the package was developed). In particular, the estimated beta3 parameter on Window or Linux is positive. This has almost no effect on the subsequent results.

	Excel	R
Ten Year Yield (semi annual basis)	4.8094	4.8095
Ten Year Yield (Annual Basis)	4.8672	4.8673

Table 5: Output from Table150and151

2.7 Table 152

The starting value for the first decay factor and second decay factors are input into cells A1 and A2 of `Table152.xlsx`. The ERA have suggested last years values. However, Diebold and Rudebusch suggest the values that maximise the loadings at 2.5 and 5 years term to maturity. Note that the ERA are using "inverse" decay factors-accordingly the start values should be 1.3941 and 2.7882 respectively.

Columns B and C are the remaining terms to maturity and yield values, respectively. Columns D, E, and F are the first, second, and third loadings, corresponding to the Nelson-Siegel-Svensson model with λ_1 given by 1.3941 and λ_2 given by 2.7882. Multiple linear regression, with an intercept, is applied with columns D, E, F as the independent variables and Column C as the dependent variable. Cells H17, H18, H19, and H20 give the starting values for beta1, beta2, beta3, and beta4, respectively.

The input workbook is `Table152.xlsx` and the output workbook is `table152Output.xlsx`. The starting values are given in Table 6. Both Excel and R have the same starting values.

	Excel	R
beta1	6.8368	6.8368
beta2	-5.8047	-5.8047
beta3	5.0232	5.0232
beta4	-7.8727	-7.8727
lambda1	1.3941	1.3941
lambda2	2.7882	2.7882

Table 6: Start Values (Output from Table152)

2.8 Table153

With the starting values, this workbook calculates the residual sum of squares. Solver is then used to minimise the residual sum of squares by changing the parameter values. The input workbook is given in `Table153Input.xlsx` and the output workbook is given in `Table153Output.xlsx`. Note that the parameter estimates are given in cells E1 to E4, given in Table 7. There are some differences between Excel and R. This is not that surprising given the overparameterisation of the Nelson-Siegel-Svensson model and the differing termination criteria used by Excel and R when fitting the model[2].

	Excel	R
beta1	6.2120	6.2072
beta2	-6.2120	-6.2072
beta3	26.8364	85.8643
beta4	-26.0640	-84.7073
lambda1	1.3581	1.4431
lambda2	1.6235	1.5252

Table 7: Nelson-Siegel-Svensson Parameter Estimates (Output from Table153)

[2] Again, there are differences when the package is run on a Mac or Windows/Linux. The differences in the parameter estimates do not impact the ten year cost of debt calculation.

2.9 Table154and155

Taking the parameter estimates from Table 153, in this table the semi-annualized yields at 3, 5, 7, and 10 years are calculated and then converted to annualized yields. The workbook is given in `Table154and155.xlsx` and the 10 year annualized yield is given in cell D3.

The results are given in Table 8. There are differences between Excel and R.

	Excel	R
Ten Year Yield (semi annual basis)	4.7032	4.8268
Ten Year Yield (Annual Basis)	4.7585	4.8851

Table 8: Output from Table154and155

2.10 Table156

Finally, the three estimates of the 10 year annualized yield are averaged to give an overall estimate, and then the average of the 10 year swap rate is annualized and subtracted from the overall estimate, to give the final 10 year DRP estimate in cell B24 of the workbook given in `Table156.xlsx`. The differences between Excel and R are quite small.

Average Swap rate (Semi-Annual basis)	2.9546	2.9546
Annualized Swap Average (Annual Basis)	2.9764	2.9764
10 year final cost of debt estimate	4.6448	4.6870
10 year DRP	1.6684	1.7106

Table 9: Output from Table156

3 Critique of the ERA (WA) Methodology

In this section some strengths and weaknesses of the ERA methodology are given.

1. It was relatively easy to follow the steps given by the ERA. However, it would probably be better to have a VBA routine which only requires the user to supply the data and the various calculations are done with the press of a button.

2. The workbooks do not contain column headers and this may cause errors. For example, in Table 145 if the Yield and Maturity were interchanged the results would be incorrect. By adding column errors, this risk would be eliminated. Similarly, it would be helpful to add labels next to important cells in the workbooks.

3. A major weakness of all three procedures is that no adjustment is made for the credit ratings of the bonds. It is possible to make these adjustments and this should improve the reliability of the results.

4. The Nelson-Siegel-Svensson Parametric Form used by the ERA is (dropping subscripts t)

$$y(\tau) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\tau/\lambda_1}}{\tau/\lambda_1} \right) + \beta_2 \left(\frac{1 - e^{-\tau/\lambda_1}}{\tau/\lambda_1} - e^{-\tau/\lambda_1} \right) + \beta_3 \left(\frac{1 - e^{-\tau/\lambda_2}}{\tau/\lambda_2} - e^{-\tau/\lambda_2} \right).$$

This is the original form in which it was published by Svensson. However, just as the Nelson-Siegel model is usually written as

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

rather than as it was introduced by Nelson and Siegel (1987), the The Nelson-Siegel-Svensson equation is now usually written as (see, for example, Diebold and Rudebusch, 2013, p. 106.)

$$y(\tau) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \beta_2 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \right)$$

However, in **drpr** I have retained, at this stage, the parameterisation used by the ERA.

5. The Nelson-Siegel-Svensson model is over-parameterised, leading to multi-collinearity, and is difficult to fit. Based on this, I suggest that it not be used in the DRP calculation.

4 R Package drpr

An R package has been built to undertake the above calculations in R. The package has a vignette which shows how the package can be used. The package:

- implements Gaussian Smoothing according to the ERA method, and fits Nelson-Siegel and Nelson-Siegel-Svensson curves;
- provides print, summary, and plot methods;
- calculates the DRP based on the three methods;
- for the Nelson-Siegel and Nelson-Siegel-Svensson models, a new re-parameterisation is used to impose the usual constraints for those models;
- adjusts for differences in credit ratings, if desired.

4.1 Parameterisation of Nelson-Siegel model

The Nelson-Siegel model can be written as:

$$\begin{aligned} y(\tau) &= \beta_1 + \beta_2 \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \\ &= \beta_1 \left(1 - \frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + (\beta_1+\beta_2) \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right). \end{aligned}$$

The usual constraints are that $\beta_1 \geq 0$, $\beta_1 + \beta_2 \geq 0$, and $\lambda > 0$. Let

$$\begin{aligned} \delta_1 &= \log(\beta_1) \\ \delta_2 &= \log(\beta_1+\beta_2) \\ \gamma &= \log(\lambda) \end{aligned}$$

Then

$$y(\tau) = \exp(\delta_1) \left(1 - \frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau} \right) + \exp(\delta_2) \left(\frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau} \right) + \beta_3 \left(\frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau} - e^{-\exp(\gamma)\tau} \right).$$

Once the parameters δ_1, δ_2, β_3, and γ are estimated, the original parameters can be calculated using

$$\begin{aligned} \beta_1 &= \exp(\delta_1) \\ \beta_2 &= \exp(\delta_2) - \exp(\delta_1) \\ \beta_3 &= \beta_3 \\ \lambda &= \exp(\gamma). \end{aligned}$$

4.2 Parameterisation of Nelson-Siegel-Svensson model

The Nelson-Siegel-Svensson model can be written as:

$$\begin{aligned} y(\tau) &= \beta_1 + \beta_2 \left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau} \right) \\ &= \beta_1 \left(1 - \frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} \right) + (\beta_1+\beta_2) \left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau} \right) + \beta_3 \left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau} \right). \end{aligned}$$

The usual constraints are that $\beta_1 \geq 0$, $\beta_1 + \beta_2 \geq 0$, $\lambda_1 > 0$, and $\lambda_2 > 0$ with $\lambda_2 > \lambda_1$. Let

$$\begin{aligned} \delta_1 &= \log(\beta_1) \\ \delta_2 &= \log(\beta_1+\beta_2) \\ \gamma_1 &= \log(\lambda_1) \\ \gamma_2 &= \log(\lambda_2 - \lambda_1) \end{aligned}$$

Then

$$
\begin{aligned}
y(\tau) &= \exp(\delta_1)\left(1-\frac{1-e^{-\exp(\gamma_1)\tau}}{\exp(\gamma_1)\tau}\right)+\exp(\delta_2)\left(\frac{1-e^{-\exp(\gamma_1)\tau}}{\exp(\gamma_1)\tau}\right)+ \\
&\quad \beta_3\left(\frac{1-e^{-\exp(\gamma_1)\tau}}{\exp(\gamma_1)\tau}-e^{-\exp(\gamma_1)\tau}\right)+\beta_4\left(\frac{1-e^{-(\exp(\gamma_1)+\exp(\gamma_2))\tau}}{(\exp(\gamma_1)+\exp(\gamma_2))\tau}-e^{-(\exp(\gamma_1)+\exp(\gamma_2))\tau}\right).
\end{aligned}
$$

Once the parameters δ_1, δ_2, β_3, γ_1, and γ_2 are estimated, the original parameters can be calculated using

$$
\begin{aligned}
\beta_1 &= \exp(\delta_1) \\
\beta_2 &= \exp(\delta_2)-\exp(\delta_1) \\
\beta_3 &= \beta_3 \\
\beta_4 &= \beta_4 \\
\lambda_1 &= \exp(\gamma_1) \\
\lambda_2 &= \exp(\gamma_1)+\exp(\gamma_2).
\end{aligned}
$$

4.3 Adjusting for Credit ratings

When adjusting for credit ratings, is is assumed that the curves for BBB-, BBB, and BBB+ bonds are parallel. Summaries and DRPs and based on weighted averages according to the relative proportion of BBB-, BBB, and BBB+ bonds in the sample used for fitting.

The usual convention is take BBB bonds as the base level. The Nelson-Siegel model, adjusting for credit ratings is then:

$$
y(\tau)=\beta_1+\beta_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)+\beta_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}-e^{-\lambda\tau}\right)+\beta_5\text{BBB-}+\beta_6\text{BBB+}
$$

where BBB- and BBB+ are dummy variables for BBB- and BBB+ bonds, respectively. The constraints are $\beta_1 \geq 0$, $\beta_1+\beta_2+\beta_6 \geq 0$, $\lambda > 0$, $\beta_5 \geq 0$, and $\beta_6 \leq 0$.

However it is somewhat easier to fit the model using BBB+ as the base level. The model is

$$
y(\tau)=\beta_1^*+\beta_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)+\beta_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}-e^{-\lambda\tau}\right)+\beta_5^*\text{BBB}+\beta_6^*\text{BBB-}
$$

where BBB and BBB- are dummy variables for BBB and BBB- bonds, respectively. The constraints are $\beta_1^* \geq 0$, $\beta_1^*+\beta_2 \geq 0$, $\lambda > 0$, $\beta_5^* \geq 0$, $\beta_6^* \geq 0$, with $\beta_6^* \geq \beta_5^*$, where

$$
\begin{aligned}
\beta_1^* &= \beta_1+\beta_6 \\
\beta_5^* &= -\beta_6 \\
\beta_6^* &= \beta_5-\beta_6
\end{aligned}
$$

and hence

$$
\begin{aligned}
\beta_1 &= \beta_1^*+\beta_5^* \\
\beta_5 &= \beta_1^*+\beta_5^*+\beta_6^* \\
\beta_6 &= -\beta_5^*.
\end{aligned}
$$

Let

$$
\begin{aligned}
\delta_1 &= \log(\beta_1^*) \\
\delta_2 &= \log(\beta_1^*+\beta_2) \\
\gamma &= \log(\lambda_1) \\
\theta_1 &= \log(\beta_5^*) \\
\theta_2 &= \log(\beta_6^*-\beta_5^*).
\end{aligned}
$$

Then

$$y(\tau) = \exp(\delta_1)\left(1 - \frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau}\right) + \exp(\delta_2)\left(\frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau}\right) + \beta_3\left(\frac{1-e^{-\exp(\gamma)\tau}}{\exp(\gamma)\tau} - e^{-\exp(\gamma)\tau}\right) + \exp(\theta_1)\text{BBB} + (\exp(\theta_1)+\exp(\theta_2))\text{BBB-}.$$

Once the parameters δ_1, δ_2, β_3, γ, θ_1, and θ_2 are estimated, the original parameters can be calculated using

$$\begin{aligned}
\beta_1^* &= \exp(\delta_1) \\
\beta_2 &= \exp(\delta_2) - \exp(\delta_1) \\
\beta_3 &= \beta_3 \\
\lambda &= \exp(\gamma) \\
\beta_5^* &= \exp(\theta_1) + \exp(\theta_2) \\
\beta_6^* &= -\exp(\theta_1)
\end{aligned}$$

The same method is used for the Nelson-Siegel-Svensson model.

5 References

Diebold, F.X. and Rudebusch, G.D. (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach.* Princeton University Press, New Jersey.

ERA (2015). Mid-West and South West Gas Distribution Systems Access Arrangements 2015-2019: Consultation on Proposed Amendments to the Final Decision.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Edition. Springer: New York.

Nelson C.R. and Siegel, A.F. (1987). 'Parsimonious Modeling of Yield Curves', *The Journal of Business*, **60**(4), pp 473–489.

Svennson, L.E.O. (1995). 'Estimating Forward Interest Rates with the Extended Nelson-Siegel Method,' *Sveriges Riksbank Quarterly Review*, **3**, pp 13–26.