Forecasting the yield curve: art or science?

Molenaars, Tomas K. and Reinerink, Nick H. and Hemminga, Marcus A.

RiskCo BV, Utrecht, The Netherlands, RiskCo BV, Utrecht, The Netherlands, RiskCo BV, Utrecht, The Netherlands

5 February 2015

Forecasting the yield curve: art or science?

Tomas K. Molenaars, Nick H. Reinerink, Marcus A. Hemminga*

RiskCo BV
Utrecht
The Netherlands

* **Corresponding author**
Dr. Marcus A. Hemminga
RiskCo BV
Maliebaan 22
3581 CP Utrecht
The Netherlands
Tel: +31 30 2059994
E-mail: marcus.hemminga@riskco.nl
Web site: http://www.riskco.nl/

Abstract

The objective of our work is to analyze the forecast performance of the dynamic Nelson-Siegel yield curve model and, for comparison, the first order autoregressive (AR(1)) model applied to a set of US bond yield data that covers a large timespan from November 1971 to December 2008. As a reference we take the random walk model applied to the yield data. For our analysis, we make use of a simple parameter representing the relative forecast performance to compare forecasting results of different methods. Our findings indicate that none of the yield curve models convincingly beats the random walk model. Furthermore, our results show that deriving conclusions on basis of model testing for a limited time period is inadequate.

This paper is a condensed version of the paper by Molenaars et al., Forecasting the yield curve - Forecast performance of the dynamic Nelson-Siegel model from 1971 to 2008 (2013). It will be published in the Dutch periodical 'De Actuaris' (The Actuary) 22-4, March 5, 2015.

Introduction

A yield curve (i.e., the term structure of interest rates) represents the relationship between interest rates and the remaining time to maturity. Forecasting of the yield curve will provide important information for monetary policy, as it is a basis for investment and saving strategies. In this view, the development of models for forecasting yield curves is of fundamental importance to banks and financial institutions, such as life insurers and pension funds.

For modeling the zero-coupon yield curve Diebold and Li (2006) constructed forecasting models based on the Nelson-Siegel model (Nelson and Siegel, 1987) and tested the forecast performance using US Treasuries bond yields. This dynamic Nelson-Siegel model (De Pooter, 2007; Christensen et al., 2009) utilizes a set of exponential components whose contributions are analyzed as a function of time. This method, in fact, is based on modeling the yield curve using its shape. It was found that this approach forecasts well, especially for a 6 and 12-month forecast horizon. This success has given rise to the popularity of the dynamic Nelson-Siegel model in forecasting studies of yield curve. However, the question is: how well does this model perform over a large time period?

To tackle this problem, we use a simple parameter representing the relative forecast performance with respect to the random walk model to facilitate the interpretation of the forecasting quality. We systematically examine the dynamic Nelson-Siegel model and the AR(1) model using the US Treasuries bond yields for an extensive historic data set ranging from November 1971 to December 2008. This data set is provided by Robert Bliss and covers the period from November 1971 (1971:11) to December 2008 (2008:12) with maturities 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months.

Theory and methodology

The models that we use in the forecasting procedures are summarized in Table 1. In the case of the dynamic Nelson-Siegel model, the yield curve is fitted with the following equation:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right). \qquad (1)$$

Here we have four time-dependent parameters, which can be interpreted as follows: the shape parameter λ_t governs the exponential decay rate and parameters $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ represent the contribution of the so-called long-term component, short-term component and medium-term component, respectively. Eq. (1) is not linear in λ_t, hence for every time *t* we should estimate the parameters by a nonlinear fit. However, we follow the approach of Diebold and Li (2006), by fixing $\lambda_t = \lambda$. This avoids potentially challenging numerical optimizations. Doing this enables us to estimate the remaining parameters $\beta_{i,t}$ by ordinary least-squares regression. The resulting times series for these parameters are modeled subsequently using the AR(1) model.

In the forecasting procedures with the dynamic Nelson-Siegel model in Eq. (1), the AR(1) forecast for the parameters $\beta_{i,t}, i = 1,2,3,$ can be written as:

$$\hat{\beta}_{i,t+h} = \hat{a}_{i,h} + \hat{b}_{i,h}\beta_{i,t}, \qquad (2)$$

where $\hat{a}_{i,h}$ and $\hat{b}_{i,h}$ are the estimated parameters and *h* is the forecast horizon. Assuming a constant value for λ, the forecasted yield curve at time *t*+*h* is given by:

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{1,t+h} + \hat{\beta}_{2,t+h}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \hat{\beta}_{3,t+h}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right). \tag{3}$$

To evaluate the out-of-sample performance of a forecasting procedure, we calculate the root-mean-square-error (RMSE), given by

$$\mathrm{RMSE}_{model}(\tau) = \sqrt{\frac{1}{T-t_0}\sum_{t=t_0}^{T}\left(\hat{y}_t(\tau) - y_t(\tau)\right)^2}, \tag{4}$$

where $\hat{y}_t(\tau)$ is the forecasted yield of the model, $y_t(\tau)$ is the yield from the data, and $[t_0, T]$ is the interval of times for which we make the forecasts. The smaller the RMSE, the better the forecast quality of the model.

To be able to systematically compare the quality of the huge number of forecasting results of the models, we "compress" them in terms of a the relative forecast performance parameter *F*. This parameter is defined as the relative difference in forecast error of the model with respect to the RW model:

$$F_{model} = \frac{\sum_\tau \mathrm{RMSE}_{RW}(\tau) - \sum_\tau \mathrm{RMSE}_{model}(\tau)}{\sum_\tau \mathrm{RMSE}_{RW}(\tau)}, \tag{5}$$

where $\sum_\tau \mathrm{RMSE}_{RW}(\tau)$ and $\sum_\tau \mathrm{RMSE}_{model}(\tau)$ sum over the RMSE values at all maturities τ of the random walk model and fitting model, respectively.

We take the random walk model as our bench mark, as it has the most simple no-change forecast, to provide a minimum standard on predictive accuracy for each model. Positive values of *F* denote a better forecast of the model as compared to the random walk model; negative values indicate a reduced performance. By definition, the relative forecast performance of the random walk model is 0.

Results and discussion

Our forecasting results are presented in Fig. 1A and B, which show the relative forecast performance *F* of the models NS and AR, respectively, as a function of time at different forecast horizons h. In this figure on the horizontal axis, the starting dates are shown for the various forecast periods. For example, 1994 (see arrow) reflects the forecast study carried out by Diebold & Li (2006). This point indicates a forecast period from January 1994 up to and including December 2000 (from 1994:1 to 2000:12, i.e., 84 months).

The advantage of using *F* is that it enables us to easily compare the forecasting results of different models applied to a large yield data set. However, a disadvantage is that valuable information about the effect of different maturity values τ is lost. Nevertheless, Fig. 1 demonstrates that the relative forecast performance offers an excellent way to analyze the overall trends in the forecasts at different forecast horizons.

Since the forecasting result of the dynamic Nelson-Siegel model depends on the value of λ, its effect on F_{NS} is investigated for different values of λ for a forecast horizon of 6 months. This result is presented in Fig. 2. As can be seen, taking other values for λ does not make much difference, except for $\lambda = 0.03$, which delivers poor forecasts in most cases. Again, this a another demonstration of the usefulness of the relative forecast performance parameter.

From the results shown in Fig. 1, a couple of interesting observations can be made.

(1) In comparing F_{NS} and F_{AR} in Fig. 1, it can be seen that only for about 20% of the monthly data points between 1982 and 2002 F_{NS} performs better than F_{AR} (see the periods 1993-1995 and 2000-2002). This suggests that there is no convincing advantage in using the more advanced and complicated dynamic Nelson-Siegel model over a simple AR(1) model. This can be understood, because there are a couple of inherent weaknesses in using the dynamic Nelson-Siegel model.

Firstly, one can argue that the Nelson-Siegel curve (Eq. (1)) does not properly fit the yield curve at all dates (for a fixed value of λ). In fact, the Nelson-Siegel model imposes a functional form to the yield curve. If the yield curve does not fit to this form, the Nelson-Siegel model will result in inferior forecasts. It is well known that adding a fourth term to the Nelson-Siegel equation (the Svensson extension (Svensson, 1995)), which allows for a second "hump/trough", delivers a better yield curve fit. Although there is no fundamental economic theory that supports this Nelson-Siegel-Svensson equation, it is extensively used by Central Banks (BIS, 2005; Gilli et al., 2010). Conversely, in the four-term Nelson-Siegel-Svensson equation more parameters need to be fitted, increasing the risk of fitting noise arising from parameter correlation and multiple local optima (Hawkins, 2004; Gilli et al., 2010).

Secondly, in the estimation of the β-parameters, it is assumed that λ is fixed. However, it is questionable whether the Nelson-Siegel equation with a fixed λ will perform well in all cases. In Fig. 1, we have used a constant value of λ of 0.0609 (in month^{-1}) that is optimized by Diebold and Li (2006) [1] for the result at 1994. The findings in Fig. 2 reveal that the effect of varying λ is small, thus the value of λ will not affect the main conclusions obtained from Fig. 1. Even so, the assumption of a fixed λ may be a source for the low overall relative forecast performance of the dynamic Nelson-Siegel model as compared to the forecast performance of the AR(1) model.

(2) The most striking point in Fig. 1 is that for almost all monthly data points the relative forecast performance F is negative, demonstrating that none of the models AR and NS can convincingly beat the random walk model. Thus the most simple random walk forecasting model performs the best.

(3) Finally, our results clearly show that deriving conclusions on basis of model testing for a limited time period is inadequate.

Acknowledgments

We thank Dr. Robert Bliss (Wake Forest University, Winston-Salem, USA) for kindly providing the US yield data and Dr. Michiel De Pooter (Federal Reserve Board of Governors, Washington, USA) for helpful discussions.

---

[1] In this paper it is argued that the value of λ_t that maximizes the medium-term component in Eq. (1) at exactly 30 months is $\lambda_t = 0.0609$. This statement is incorrect. The medium-term component has a bump shape with a maximum at $\lambda_t \tau = 1.793$. From this relationship, it can be seen that $\lambda_t = 0.0609$ actually corresponds to 29.44 months.

References

BIS, 2005, Zero-coupon yield curves: technical documentation. Bank for International Settlements, Basel, Switzerland, BIS Papers 25.

Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2009. An arbitrage-free generalized Nelson-Siegel term structure model. Econometrics Journal 12, C33-C64.

De Pooter, M., 2007, Examining the Nelson-Siegel class of term structure models - In-sample fit versus out-of-sample forecasting performance. Tinbergen Institute, Discussion Paper 2007-043/4.

Diebold, F.X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337-364.

Gilli, M., Grosse, S., Schumann, E., 2010. Calibrating the Nelson-Siegel-Svensson model, COMISEF, Working Paper WPS-031.

Hawkins, D.M., 2004. The problem of overfitting. J. Chem. Inf. Comput. Sci. 44, 1-12.

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

Svensson, L.E.O., 1995. Estimating forward interest rates with the extended Nelson and Siegel method. Quarterly Review, Sveriges Riksbank 3, 13-26.

Table 1

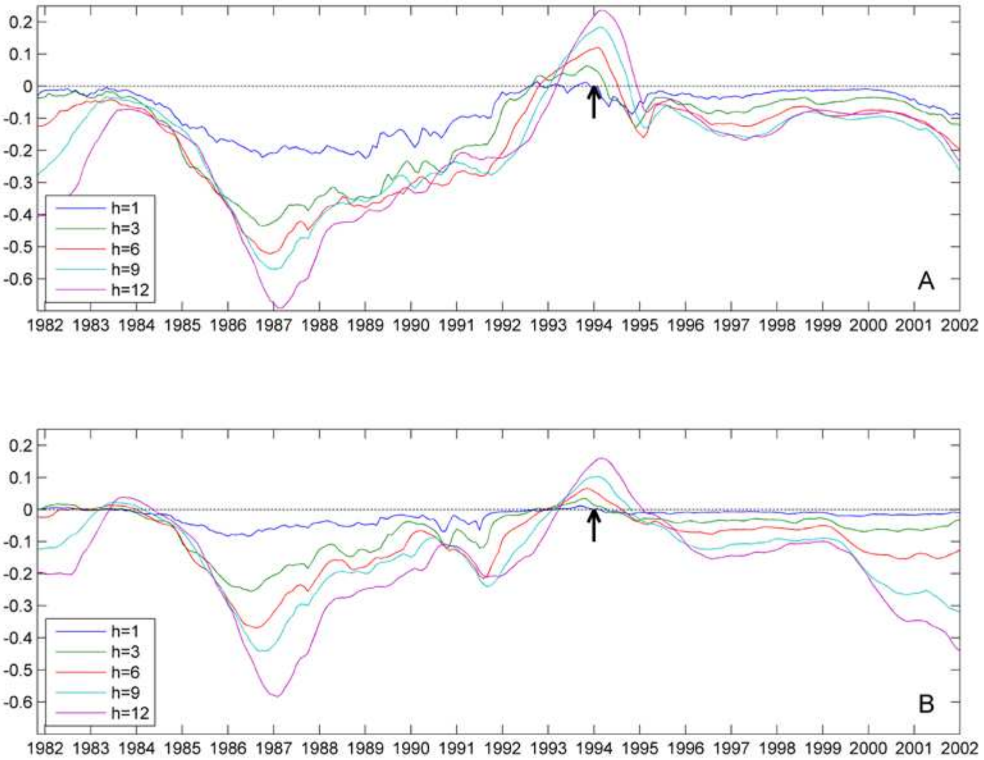
Models used in the forecasting procedures. The random walk model (RW) and first order autoregressive (AR(1)) model (AR) are applied directly to the yield data. In the dynamic Nelson-Siegel (NS) model (Eq. (1)), the AR(1) model is applied to the β-parameters from the yield curve fit. In comparing the different forecasting procedures, the random walk model is taken as a bench mark.

Abbreviation	Model type
RW	Random walk model on the yield data
AR	AR(1) model on the yield data
NS	Dynamic Nelson-Siegel model, Eq. (1) and AR(1) on the β-parameters

Figure legends

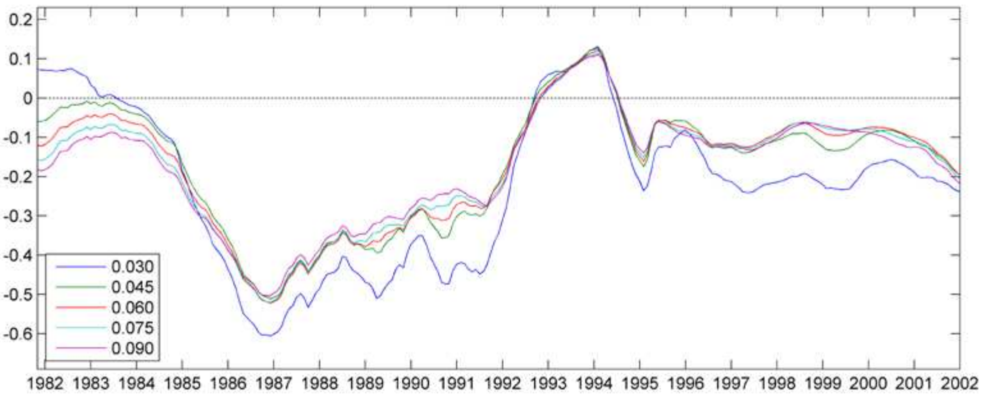
Fig. 1. Relative forecast performance F of the models NS (F_{NS}, A) and AR (F_{AR}, B) (see Table 1) for forecast horizons h 1, 3, 6, 9, and 12 months. Parameter λ is fixed at a value of 0.0609. The arrow at the year 1994 reflects the results of the forecast study carried out by Diebold and Li (2006).

Fig. 2. Effect of λ on the relative forecast performance F of the NS model for a forecast horizon of 6 months.



A
B
h=1
h=3
h=6
h=9
h=12
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002

Fig. 1



0.030
0.045
0.060
0.075
0.090
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002

Fig. 2