

## Advanced Monetary Economics

F.C. Bagliano - May 2009

Notes on:

**F.X. Diebold and C. Li**, “Forecasting the term structure of government bond yields”, *Journal of Econometrics*, 2006

# 1 Purpose of the paper

The paper presents a novel methodology for forecasting the yield curve of government bonds using the Nelson-Siegel (1987) method for deriving a continuous curve with desirable properties from a finite set of observed bond prices at different maturities. The Nelson-Siegel (N-S) technique is interpreted as a “factor model” and each N-S component is given an interpretation in terms of “level”, “slope” and “curvature” effects.

A simple time-series model for the evolution over time of such factors is estimated and used to forecast the factors out-of-sample. Using monthly US data for the 1985-2000 period, the forecasting performance of this methodology is compared with alternative methods for forecasting yields and found superior at least over a one-year forecasting horizon.

# 2 Fitting the yield curve

Given a limited set of observed bond prices at a point in time, several techniques can be used to extract a continuous yield curve, delivering the yield to maturity on bonds with any maturity. In order for this derived curve to be useful for financial analysis and forecasting, it must possess a set of desirable properties, such as flexibility (i.e. it must be able to reproduce the various observed shapes of the yield curve, such as upward-sloping, downward-sloping, hump-shaped), parsimony (i.e. it must usefully summarize the features of the term structure using a limited set of parameters) and economic interpretability.

Perhaps the most widely used method for fitting the yield curve is the three-component exponential method of Nelson and Siegel (1987), on which the Diebold-Li paper is based.

## 2.1 Preliminaries

Focusing on the simpler case of pure discount bonds (paying 1 unit at the maturity date to the holder, with no coupon payments in the intervening

periods), let's recall the theoretical relationships among three basic concepts: the *discount curve*, the *forward rate curve* and the *yield curve*. In what follows, time is continuous, interests are continuously compounded, and  $\tau$  denotes the maturity of the discount bonds (measured in months in the empirical analysis of the paper).

Denoting the yields to maturity on  $\tau$ -period discount bonds at time  $t$  as  $y_t(\tau)$ , the discount curve gives the prices of the bonds  $P_t(\tau)$  as a function of maturity as

$$P_t(\tau) = e^{-\tau y_t(\tau)} \quad \text{“discount curve”}$$

from which

$$y_t(\tau) = -\frac{1}{\tau} \log P_t(\tau)$$

The relationship between the yields to maturity and the implicit instantaneous forward rates  $f_t(\tau)$  gives the forward rate curve<sup>1</sup>

$$\begin{aligned} f_t(\tau) &= \frac{dy_t(\tau)}{d\tau} \tau + y_t(\tau) \\ &= \left( -\frac{P'_t(\tau)}{P_t(\tau)} \frac{1}{\tau} - \frac{y_t(\tau)}{\tau} \right) \tau + y_t(\tau) \\ &= -\frac{P'_t(\tau)}{P_t(\tau)} \quad \text{“forward rate curve”} \end{aligned}$$

and the yield curve

$$y_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u) du \quad \text{“yield curve”}$$

## 2.2 The Nelson-Siegel method and its interpretation

At a given date  $t$ , at which a set of yields on bonds with different maturities is available, the Nelson-Siegel method fits a smooth continuous curve of the following three-component exponential form:

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

with the following related instantaneous forward rate curve:

$$f_t(\tau) = \beta_{1t} + \beta_{2t} e^{-\lambda_t \tau} + \beta_{3t} \lambda_t \tau e^{-\lambda_t \tau}$$

---

<sup>1</sup>Equivalently, the instantaneous forward rate may be seen as the rate of decay of the discount function  $d(\tau) \equiv P_t(\tau) = e^{-\tau y_t(\tau)} = e^{-\int_0^\tau f_t(u) du}$ .

The properties of the curves are defined by the four parameters  $\beta_1, \beta_2, \beta_3$  and  $\lambda$ . The parameter  $\lambda$  governs the exponential decay of the two functions in brackets as maturity  $\tau$  goes from 0 to  $\infty$ , with higher values of  $\lambda$  determining a faster decay.<sup>2</sup> Given a value for  $\lambda$ , the three terms in the curve for  $y_t(\tau)$  can be interpreted as a parameter  $\beta_i$  ( $i = 1, 2, 3$ ) multiplied by a function of maturity  $\tau$  as follows:

1. the first component is simply  $\beta_1$  and does not depend on  $\tau$ . It is interpreted as a *long-term* component capturing the “**level**” of the yield curve. In fact, letting  $\tau \rightarrow \infty$

$$y_t(\infty) = \beta_{1t}$$

2. in the second term, a decay function  $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$  (monotonically going to 0 as  $\tau \rightarrow \infty$ ) is applied to the parameter  $\beta_2$ . Since the size of this component declines as maturity increases, it is interpreted as a *short-term* component. In fact, as  $\tau \rightarrow 0$  we have

$$y_t(0) = \beta_{1t} + \beta_{2t}$$

that, combined with the previous property, gives

$$y_t(\infty) - y_t(0) = -\beta_{2t}$$

The parameter  $\beta_2$  has therefore the interpretation of (minus) the “**slope**” of the yield curve.

3. in the third term, the parameter  $\beta_3$  is multiplied by a decay function  $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$  which starts at 0 for  $\tau = 0$ , increases for intermediate values of  $\tau$  up to a maximum (whose position is determined by the chosen value for  $\lambda$ ) and then decreases to 0 for  $\tau \rightarrow \infty$ . Therefore, this component has a smaller size for both short and long maturities and a larger size for medium-run maturities: it is then interpreted as a *medium-term* component. Moreover, the parameter  $\beta_3$  is closely related to a (conventional) measure of the curvature of the yield curve, given by (with  $\tau$  measured in months and for a chosen  $\lambda = 0.0609$ )

$$2 y_t(24) - (y_t(3) + y_t(120)) = 0.00053 \beta_{2t} + 0.37 \beta_{3t}$$

Hence, the parameter  $\beta_3$  captures the “**curvature**” of the yield curve.

---

<sup>2</sup>The parameter  $\lambda > 0$  captures the speed of convergence of the spot and instantaneous forward rates to their values for  $t \rightarrow \infty$  (the “consol” rate). A lower  $\lambda$  shifts the hump of the curve towards shorter maturities, thereby accelerating convergence to the consol rate.

Overall, the Nelson-Siegel functional form for the yield curve is interpreted by Diebold and Li as a three-factor (statistical) model for the term structure with the parameters  $\beta_i$  capturing the relevant “**factors**” affecting the shape of the yield curve at each date  $t$ :

$$\begin{aligned}\beta_{1t} &\rightarrow \text{level factor} \\ -\beta_{2t} &\rightarrow \text{slope factor} \\ \beta_{3t} &\rightarrow \text{curvature factor}\end{aligned}$$

The effects of a (unit) change in each factor  $\beta_i$  on the yields at different maturities is given by the values of the “decay” functions, or “**loadings**” (displayed in Figure 1 of the paper) and are consistent with the above interpretation of the factors as level, slope and curvature.

### 3 Modelling and forecasting the yield curve

Diebold and Li propose a multi-step empirical strategy for forecasting the yield curve out-of-sample using a time-series of cross-sections of yields. The data set consists, for each month  $t$  between January 1985 and December 2000, of yields on discount bonds of 17 different maturities, ranging from 3 months to 10 years. The steps of the methodology are the following:

1. **Factor estimation.** For each month  $t$ , the N-S continuous function for  $y_t(\tau)$  is fitted to the 17 available yields (imposing  $\lambda = 0.0609$ , which implies a maximum value for the loading of the medium-term factor at a maturity of 30 months) by means of the following *OLS* regression:

$$y_t(\tau) = \beta_{1t} + \beta_{2t} \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_{3t} \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) + u_t(\tau) \quad (1)$$

where the 17 yields are regressed on a constant and the two regressors in brackets, with  $u_t(\tau)$  representing “pricing errors”. From those regressions, time series for the estimated  $\hat{\beta}_{1t}$ ,  $\hat{\beta}_{2t}$ ,  $\hat{\beta}_{3t}$  and  $\hat{u}_t(\tau)$  are obtained.

2. **Factor modelling and forecasting.** To obtain forecasts of the yield curve over  $h$ -period horizons on the basis of information available at month  $t$ ,  $\hat{y}_{t+h|t}(\tau)$ , the three factors are modelled as autoregressions of the form (for  $i = 1, 2, 3$ )

$$\hat{\beta}_{it} = c_i + \gamma_i \hat{\beta}_{i,t-h} + \varepsilon_{it} \quad (2)$$

so that the values of factors  $\beta_i$  forecast at  $t$  for month  $t+h$  are obtained as

$$\hat{\beta}_{i,t+h|t} = \hat{c}_i + \hat{\gamma}_i \hat{\beta}_{i,t}$$

As an alternative, forecasts can be obtained from a multivariate model of the form

$$\hat{\beta}_t = \mathbf{c} + \mathbf{\Gamma} \hat{\beta}_{t-h} + \boldsymbol{\varepsilon}_t$$

where  $\hat{\beta} = (\beta_1 \beta_2 \beta_3)'$  and  $\mathbf{\Gamma}$  is a  $3 \times 3$  matrix, so that

$$\hat{\beta}_{t+h|t} = \hat{\mathbf{c}} + \hat{\mathbf{\Gamma}} \hat{\beta}_t$$

3. **Yield curve forecasting.** Once forecasts of the factors are obtained, the yield curve is forecast out-of-sample over a  $h$ -period horizon as

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

Forecasts are constructed recursively. Starting from the sample 1985(1)-1993(12), the autoregressive models for the factors (2) are estimated and used to produce factor forecasts from 1994(1) to 2000(12) at the 1-month, 6-month and 1-year horizons; then, forecasts of the yield curve are constructed from (3). The estimation sample for the factor models (2) is then updated by adding one observation at a time (becoming 1985(1)-1994(1), 1985(1)-1994(2), and so on) and new forecasts of factors and yields are obtained for the same horizons over the remaining part of the period (being 1994(2)-2000(12), 1994(3)-2000(12), and so on).

4. **Forecast evaluation.** Finally, forecasts are evaluated and compared with those obtained by alternative models (such as a simple random walk for the yield on each maturity, *VAR* models for the levels or the changes in the yields, *VECM* models with one or two common stochastic trends) using various statistics, including the *root mean squared error* for each forecasting horizon  $h$  and each maturity  $\tau$ :

$$RMSE(h, \tau) = \sqrt{\frac{\sum_{i=0}^{T-(t+h)} [y_{t+i+h}(\tau) - \hat{y}_{t+i+h|t+i}(\tau)]^2}{T - (t+h) + 1}}$$

where  $T = 2000(12)$  and  $t = 1993(12)$ .

## 4 Conclusion

Diebold and Li present a reinterpretation of the Nelson-Siegel method for yield curve fitting as a three-factor model, capturing movements in the level, slope and curvature of the curve, and assess its out-of-sample forecasting performance. The results show that such a parsimonious model outperforms several competitors in forecasting exercises over the one-year horizon.