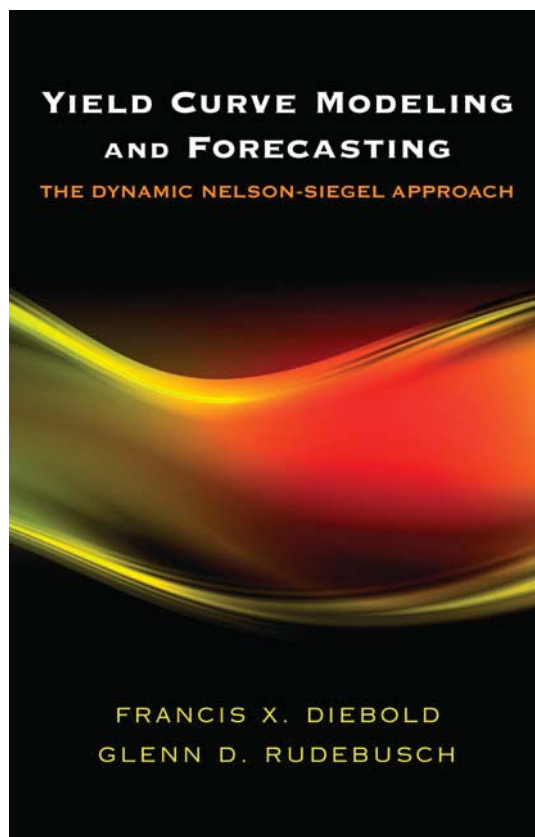


Book review



© 2013, Princeton University Press

Yield Curve Modeling and Forecasting—The Dynamic Nelson–Siegel Approach, by Francis X. Diebold and Glenn D. Rudebusch, Princeton University Press (2013). ISBN 978-0691146805.

Diebold and Rudebusch have written an original and useful book on affine term structure modelling and estimation. It is clearly written, and the detailed appendices make the approach ready to ‘plug and play’. It is important because it attempts to marry the complementary statistical estimation approach to yield curve modelling with the arbitrage-free approach. As Diebold and Rudebusch say, their model

blends two important and successful approaches to yield curve modelling: the [...] empirically based and the no-arbitrage theoretically based one. Yield curve models in both traditions are impressive successes, albeit for different reasons. Ironically, both approaches are equally impressive failures, and for the same reasons, swapped. That is, models in the [statistical] tradition fit and forecast well, but they [...] may admit arbitrage

possibilities, while the models in the arbitrage-free tradition are theoretically rigorous [...], but they may fit and forecast poorly.†

Their approach, the authors continue, ‘bridges the divide with a [statistically]-inspired model that enforces the absence of arbitrage.’‡ In this sense, their contribution is in the mould of recent work by *D’Amico et al. (2010)* and *Adrian et al. (2013)*.

The best way to understand to what extent the promise is fulfilled is to retrace the conceptual steps that the authors follow.

They start from the ‘static’ Nelson–Siegel (1987) model, which is probably the best-known yield curve fitting tool in the industry, in research and among central bankers. In its simplest formulation, the time- t and T -expiry instantaneous forward rates, f_t^T , are given by

†Diebold and Rudebusch (2013), p. 76.

‡ *ibid*, p. 76.

$$f_t^T = \beta_1^t + \beta_2^t e^{-\lambda\tau} + \beta_3^t \lambda\tau e^{-\lambda\tau} \quad (1)$$

with $\tau = T - t$.

This is all well known, but note that in equation (1) there is a time dependence for β_1^t , β_2^t and β_3^t but not for λ . In the Nelson–Siegel approach, the parameter, λ , is used to obtain the best fit to the yield curve alongside the other parameters, but, for a reason that will become apparent in the following, in the Diebold–Rudebusch approach is it kept constant in yield maturities.

Despite its unashamedly *ad hoc* nature, the Nelson–Siegel specification enjoys some promising features, which suggest an intriguing, albeit at this stage rather tenuous, link with financial reality. First of all, the limiting values for the bond prices ($\lim_{T \rightarrow t} P_t^T = 1$ and $\lim_{(T-t) \rightarrow \infty} P_t^T = 0$) are recovered for any choice of the parameter set $\{\beta_1^t, \beta_2^t, \beta_3^t, \lambda\}$. Second, it is easy to show that

$$\lim_{T \rightarrow t} (y_t^T)^{NS} = \lim_{T \rightarrow t} f_t^T = r_t \quad (2)$$

which says that the limit of the yields produced by the Nelson–Siegel model approaches the short rate, as it should, as the residual maturity τ goes to zero. Finally, the limit of the infinite-maturity yield, $\lim_{(T-t) \rightarrow \infty} (y_t^T)^{NS}$, is a constant, given by

$$\lim_{(T-t) \rightarrow \infty} (y_t^T)^{NS} = \beta_1^t \quad (3)$$

This chimes well with the predictions of affine models, which produce yield curves that tend to a long-term constant.

Finally, the Nelson–Siegel curve, despite its simplicity, is empirically observed to fit a wide variety of yield curves surprisingly well. Readers interested in a glimpse of the mathematical explanation of why such a simple functional form can perform so well are referred to the original paper by Nelson and Siegel (1987) and to the discussion in Diebold and Rudebusch (2013).[†]

So far, despite the asymmetry in the superscript notation, we have taken λ to be a fitting parameter on the same footing as the $\{\beta\}$ parameters. Let’s set, however, set the λ parameter to a fixed value, whose interpretation will be discussed later. Then, we are left with the three parameters (β_1^t , β_2^t and β_3^t) to fit the time- t yield curve. Since yield curves change day after day, so will the optimal fitting parameters $\{\beta_1^t, \beta_2^t, \beta_3^t\}$ —but not λ .

If we look at the fitting exercise in this light, the time- t yield, y_t^T ,[‡] can be seen as the result of multiplying three *fixed* quantities, 1, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$ and $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$, by three *time-varying quantities*, β_1^t , β_2^t and β_3^t . In this interpretation, the coefficients $\{\beta_1^t, \beta_2^t, \beta_3^t\}$ (but not λ !) have become dynamic *variables*, and 1, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$ and $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$ have become *parameters*.

This interpretation could seem a bit contrived were it not for the fact that the three new fixed ‘parameters’, 1, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$ and

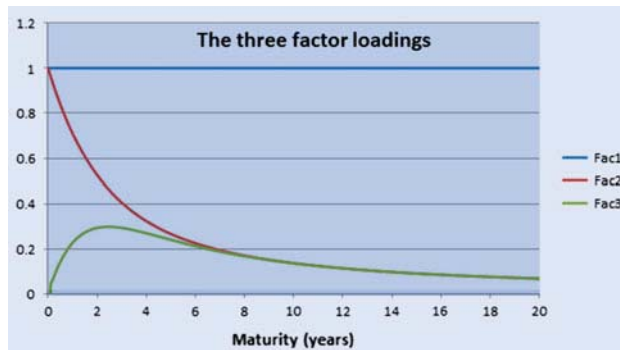


Figure 1. The factor loadings (the ‘parameters’) 1, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$ and $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$ for a lambda of 0.7308.

$\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$, lend themselves to an interesting interpretation, suggested by figure 1.§

The first loading is just a constant. This means that any shock to the first ‘variable’, β_1^t , will be transmitted equally to all the yields.

The second loading decays rapidly, with a ‘half-life’ of approximately 2.15 years. Therefore, shocks to the second ‘variable’, β_2^t , will mainly affect the short maturity yields, and will have a negligible effect on the longest maturities.

Finally, the third loading has a humped shape, with a shallow maximum around 2.5 years. So, shocks to the third ‘variable’, β_3^t , will have no effect at the very short end, a decaying effect at the long end and a maximum impact for intermediate-maturity yields (in the 2-to-5-year region).

Diebold and Rudebusch (2103)¶ give another interpretation to the three loadings. Note that β_1^t moves yield in a parallel fashion (and, therefore, changes the *level* of the yield curve); that β_2^t changes the *slope* of the yield curve and that β_3^t creates a hump in the yield curve (and, therefore, changes its *curvature*). The suggestive choice of words (‘level’, ‘slope’ and ‘curvature’) points to an obvious similarity between shocks to the variables, $\{\beta_1^t, \beta_2^t, \beta_3^t\}$ and shocks to the three first principal components.

When seen in this light, the rather artificial decomposition of the Nelson–Siegel snapshot model begins to look very interesting. Note however: (i) that the ‘slope factor’ has the same sign for all yields, while the second principal component moves yield on opposite sides of the ‘pivot point’ in opposite directions; (ii) that the first true principal component is not a constant as a function of yield maturity (it actually resembles a combination of the first and the third factors, β_1^t and β_3^t) and (iii) that the Nelson–Siegel loadings are not orthogonal, as true principal components should be (they cannot be, as they are all positive).

As Diebold and Rudebusch show, this decomposition lends itself well to statistical estimation. They also show that it can be turned into an affine arbitrage-free model. In a nutshell, the strategy Diebold and Rudebusch (2013) employ to do so is the following.||

§ For readers trying to reproduce the results in Diebold and Rudebusch (2013, p. 28), it should be noted that their decay constant ($\lambda = 0.0609$) takes months, not years, as the unit of time—and, therefore, the dimensions of their λ are months⁻¹, not years⁻¹.

¶ See p. 29.

|| See sections 3.2 to 3.3, pp. 62–78.

[†] See p. 25.

[‡] In the following we drop the superscript *NS* to denote the yields produced by the Nelson–Siegel model.

They start from three variables, which they would like to behave like the variables $\{\beta_1^t, \beta_2^t, \beta_3^t\}$ discussed above (which they rename x_i). They would like these variables to display a mean-reverting behaviour, with a reversion-speed matrix, \mathcal{K} , a reversion-level vector, θ , and a volatility matrix, Σ :

$$dx = \mathcal{K}(\theta - x) dt + \Sigma dz \tag{4}$$

Shocks to the short rate (which, as equation (2) shows, is the limit of the yield of vanishingly short maturity) are then just given by the sum of shocks to the first and second variables:

$$dr_t = d\beta_1^t + d\beta_2^t. \tag{5}$$

Next Diebold and Rudebusch look at the family of arbitrage-free affine models in the taxonomy provided by Duffie and Kan (1996), and pick the family member that most closely resembles the dynamic factors described above. More precisely, if the reversion-speed matrix, \mathcal{K} , is chosen to be

$$\mathcal{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix} \tag{6}$$

and if, inspired by equation (5), the mapping from the state variables to the short rate is given by[†]

$$r_t = x_1^t + x_2^t, \tag{7}$$

then we know that the bond price, P_t^T , is given by

$$P_t^T = e^{A_t^T + (B_t^T)^T x_t} \tag{8}$$

with A_t^T a constant depending on t and T

$$(B_t^T)^T = \left[1 \frac{1-e^{-\lambda\tau}}{\lambda\tau} \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right] \tag{9}$$

and

$$\tau = T - t. \tag{10}$$

The expression for the arbitrage-free yields turns out to be of the form

$$y_t^T = x_1^t + x_2^t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + x_3^t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) + \alpha_t^T. \tag{11}$$

with α_t a constant vector depending on t and T .

A few observations are in order.

First note that, in order to pick out the strongest family resemblance between the NS-inspired dynamic model described above and the arbitrage-free affine models in the Duffie and Kan (1996) family tree, the first row of the reversion-speed matrix, \mathcal{K} , is zero. This means that the first factor (the first ‘pseudo-principal-component’) experiences no mean reversion at all and is a unit-root process. Given the hardly detectable reversion speed of the true first principal component uncovered, at least in the \mathbb{P} measure, by statistical studies, this is not an unreasonable choice.

Second, we want to see to what extent the analogy of the state variables with principal components can be pushed. More precisely, we want to see if we recover at any point in time the relationship $y = Vx^{pc}$ that we know applies to true principal components.

For simplicity, let’s deal with just three yields, of which we choose the first one to be the short rate. We can always construct

from the Diebold–Rudebusch model three vectors, u_1, u_2 and u_3 , given by

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} \\ \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} \\ \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} \end{bmatrix}, \quad u_3 = \begin{bmatrix} \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} - e^{-\lambda\tau_3} \end{bmatrix}. \tag{12}$$

If the parallel with principal components is to be pursued, one should have

$$y_t = Ux_t = [u_1 \ u_2 \ u_3] x_t. \tag{13}$$

In the Diebold–Rudebusch model this relationship is always exactly satisfied, as one easily obtains

$$U = \begin{bmatrix} 1 & 1 & 0 \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ 1 & \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} & \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} - e^{-\lambda\tau_3} \end{bmatrix}.$$

Furthermore, the first element of the y vector (the short rate) is always exactly recovered as $y_1 = r_t = x_1^t + x_2^t$. But, thanks to equation (11), the identity also automatically applies to the other yields, for which we have

$$y_i \simeq \left[1 \frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} \frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i} \right] \begin{bmatrix} x_1^t \\ x_2^t \\ x_3^t \end{bmatrix}, \quad i = 2, 3 \tag{14}$$

We have reached the conclusion that in the Diebold–Rudebusch approach the two separate possible routes of obtaining yields (either as $y_t = Ux_t$ or as $y_t^i = -\frac{1}{T_i-t} \log P_t^{T_i}$) always exactly converge. Simple as this may seem, ensuring this convergence is not trivial, and requires some careful handling.

In sum: the interpretation of the variables $\{x\}$ as principal components is not skin-deep, but captures some important features that one would like to find in a principal component-inspired model. Is this enough?

There is no question that the specification of the reversion-speed matrix (equation (6)) from the Duffie and Kan (1996) taxonomy is the one that most closely resembles (actually, perfectly coincides with) the dynamic Nelson–Siegel model. However, the precise functional form of this model was not derived on the basis of some profound theoretical derivation. Rather, it came from a clever transformation of a ‘snapshot’ model, whose variables ended up lending themselves to a serendipitous interpretation as quasi-principal components.

This prompts an observation. The reversion-speed matrix, \mathcal{K} , is chosen in the \mathbb{Q} measure. Now, if we want to interpret the variables $\{x\}$ as quasi-principal components, their \mathbb{P} -measure dynamics should be statistically determinable. In particular, for the interpretation to hold, their Σ volatility matrix should be (close to) diagonal, and the \mathbb{P} -measure reversion-speed matrix, $\mathcal{K}^{\mathbb{P}}$, should be whatever comes out from the econometric analysis. (As Diebold and Rudebusch make clear in the opening chapters, we do know a lot about the \mathbb{P} -measure behaviour of the principal components—such as their degree of persistence—and this is precisely one of the features that makes their approach more appealing than latent-variable models).

Now, once the reversion-speed matrix under \mathbb{Q} (6) is chosen, and once the $\mathcal{K}^{\mathbb{P}}$ matrix turns out to be whatever we estimate, the affine market price of risk, $\Lambda = \Lambda_0 + \Lambda_1 x$, is fully specified.

[†]This means that, using the notation in Chapter Affine General, $u_r = 0$, and $g^T = [1, 1, 0]$.

However, it would be an incredible fluke if the ‘implied’ market price of risk turned out to bear more than a passing resemblance with the market price of risk that can be estimated from excess return studies.

Indeed, from Fama and Bliss onwards we know that the slope of the yield curve ‘contains’ a lot, if not all, of the return-predicting factor. (See also Adrian *et al.* (2013) and Cochrane and Piazzesi (2005).) But if this is the case, and if the second state variable of the Diebold–Rudebusch approach is to be interpreted as a ‘yield curve slope’, then the matrix Λ_1 should have the form

$$\Lambda_1 = \begin{bmatrix} 0 & \lambda_{21}^1 & 0 \\ 0 & \lambda_{22}^1 & 0 \\ 0 & \lambda_{23}^1 & 0 \end{bmatrix}. \quad (15)$$

It is easy to see that there is a potential incompatibility between (i) the \mathbb{Q} -measure reversion-speed matrix, \mathcal{K} , Diebold and Rudebusch choose, (ii) a market price of risk matrix, Λ_1 , that reflects some well-established facts about excess returns, (iii) a statistically-determined reversion-speed matrix $\mathcal{K}^{\mathbb{P}}$, and (iv) the interpretation of the state variables as (quasi) principal components. Once the market price of risk is made to depend in a *specified* (affine) manner on the state variables, the modeller loses her ability to assign the nature of the variables *and* their \mathbb{Q} -measure reversion-speed matrix, \mathcal{K} . (This may well be one of the reasons why latent-variables affine term structure models are currently very popular.)

This tension is not unique to the Diebold–Rudebusch approach, and resurfaces in all the approaches that Saroka (2014) dubs ‘specified-variables’, and in the constraints highlighted in Adrian *et al.* (2011). As the work in Rebonato *et al.* (2014) shows, the problem is not solved (if anything, is made more acute) if one moves from quasi- to exact principal components. To the extent that one of the stated uses to which these models may be put is prediction (see the opening quote), this may create problems, unless one only works in the \mathbb{Q} measure.

For all this, the model presented by Diebold and Rudebusch is very appealing, useful, simple and clearly presented. In its effort to identify and interpret the state variables, rather than leaving them as latent placeholders, it helps the understanding

and the intuition of the reader. In short, it constitutes a very valuable and original addition to the new generation of affine yield curve models which try to marry rigorous no-arbitrage pricing, statistical information and what we ‘now know’ about risk premia and excess returns. As usual, Princeton University Press have produced this slim book to the high quality standards readers have come to expect from this publisher.

References

- Adrian, T., Crump, R.K. and Moench, E., Pricing the term structure with linear regressions, 2013, Federal Reserve Bank of New York, Staff Report No 340, August 2008, Revised 2013. Available online at: http://www.newyorkfed.org/research/staff_reports/sr340.pdf (accessed 15 October 2014).
- Cochrane, J.H. and Piazzesi, M., Bond risk premia. *Am. Econ. Rev.*, 2005, **95**(1), 138–160.
- Duffie, D. and Kan, R., A yield-factor model of interest rates. *Math. Financ.*, 1996, **6**, 379–406.
- D’Amico, S., Kim, D.H. and Wei, M., Tips from TIPS: The informational content of treasury inflation protected security prices. Working Paper, Finance and Economics Discussion Series, Federal Reserve Board, Washington, DC, 2010. Available online at: <http://www.federalreserve.gov/pubs/feds/2010/201019/201019pap.pdf> (accessed 13 November 2014).
- Joslin, S., Singleton, K.J. and Zhu, H., A new perspective on Gaussian dynamic term structure models. *Rev. Financ. Stud.*, 2011, **24**, 926–970.
- Nelson, C.R. and Siegel, A.F., Parsimonious modeling of yield curves. *J. Bus.*, 1987, **60**, 473–489.
- Rebonato, R., Saroka, I. and Putyatin, V., A principal-component-based affine term structure model. Working Paper, submitted for publication in *Appl. Math. Finance*, 2014. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2451130.
- Saroka, I., Arbitrage-free modelling of affine principal components of the yield curve. Unpublished Thesis for MSc in Mathematical Finance, Oxford University, 2014.

Riccardo Rebonato
University of Oxford

© 2015, Riccardo Rebonato