

Interconnectedness of Sovereign Yield curves

Badics Milan Csaba, Balazs Kotro

Draft version

Abstract

Abstract: This paper is examining the linkages between the whole tenor structure of the yield curves of 12 sovereigns from all over the globe. The curves got decomposed to level, slope and curvature factors by the Nelson Siegel model. TBC...

1 Methodology

1.1 The Nelson and Siegel framework

Among the statistical models for interest rate, the influential model designed by Diebold-Li [Diebold & Li, 2006] is widely used in market applications. This model is a dynamic extension of the Nelson-Siegel model ([Nelson & Siegel, 1987]) for the cross-section fit for the yield curve. The Nelson-Siegel model corresponds to fitting the following equation for the yield curve observed in the market on a specific date:

$$y_{it}(m_{it}) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda\tau_t}}{\lambda\tau_t} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda\tau_t}}{\lambda\tau_t} - e^{-\lambda\tau_t} \right) + \epsilon_{it} \quad (1)$$

where $y_{it}(m_{it})$ are the observed rates on a given date i and maturity t , and β_{1t} , β_{2t} , β_{3t} and τ_t are parameters. The Nelson-Siegel model is a parsimonious way of fitting the yield curve while managing to capture a part of the stylized facts in interest rate process, such as the exponential formats present in the yield curves. The parameters β_{it} have economic interpretations, where β_{1t} presents a long-term level interpretation, β_{2t} short-term components, and β_{3t} medium-term components. It may also be interpreted as decompositions of Level, Slope and Curvature of the yield curve, according to the terminology developed by [Litterman & Scheinkman, 1991]. These components may be used directly in the immunization process of interest rate portfolios.

The purpose of these models is to allow fitting, and subsequent interpolation and extrapolations of the yield curve based on a parametric structure, which concurs with other non-parametric fitting models such as smoothing-splines. Besides the parsimonious estimation, the [Nelson & Siegel, 1987] model has two additional advantages over non-parametric models. The first advantage is that the extrapolation of the curve has a better performance due to the exponential nature of this model. The second advantage is that the parameters β_{1t} , β_{2t} and β_{3t} have interpretation of level, slope and curvature compatible with the interpretation of three factors proposed by [Litterman & Scheinkman, 1991], a benchmark in literature. This makes the interpretation and comparison of the results obtained in the curve fitting easier. The extension formulated by [Diebold & Li, 2006] renders the [Nelson & Siegel, 1987] model dynamic (adjusting the several days observed for the yield curve) by means of a procedure in 3 stages:

- The Nelson-Siegel model (with τ fixed, thus making the model linear in the parameters) is fitted by Ordinary Least Squares for each date, estimating the parameters β_{1t} , β_{2t} , β_{3t} .
- The dynamics of the system is modelled by a vector autoregressive (VAR) model for the parameters β_{1t} , β_{2t} and β_{3t} , estimated in the first stage.
- Forecasts for these parameters are made through the VAR model estimated for vectors β_{1t} , β_{2t} and β_{3t} . By substituting the forecasted parameters in Nelson-Siegel model given by equation '1' it is possible to forecast future interest rate curves.

According to [Diebold & Li, 2006], this dynamic formulation has the purpose of capturing the set of existing stylized facts in the term structure of interest rates, such as the fact that while the yield curve is crescent and concave, it may also assume inverted shapes like decreasing curves and slope changes. Other stylized facts captured by [Diebold & Li, 2006] models are the high persistence in the temporal dynamics (rates with same maturity are highly dependent on the past), and the fact that persistence in the long-term rates is higher than in the short-term rates.

1.2 The Toda-Yamamoto model

The Toda-Yamamoto procedure begins from the following premise: The implementation of the classic Granger Causality test from a VAR (Vector Autoregressive) model can lead to non-stationarity problems in the series, as it is necessary to confirm the type of existing cointegration. The authors of ... point out that the “conventional” Wald test produces integrated or cointegrated causal VAR models, which would inevitably lead to obtaining spurious Granger causality relationships. However, the Toda-Yamamoto procedure drastically avoids this handicap by developing a Modified Wald test (MWALD) for restrictions on the parameters of a VAR (p) model. This test is generated on a χ_p distribution, with $p = p + d_{max}$ (or number of time lags). In Wolfe-Rufael’s words, the fundamental idea underlying this procedure is to “artificially augment the correct VAR order, p , by the maximal order of integration, say d_{max} . Once this is done, a $(p + d_{max})$ -th order of VAR is estimated and the coefficients of the last lagged d_{max} vector are ignored”. The resulting VAR $(p + d_{max})$ model is formulated in Equations (3) and (4):

$$Y_t = \alpha_0 + \sum_{i=1}^k \delta_{1i} Y_{t-i} + \sum_{j=k+1}^{d_{max}} \alpha_{1j} Y_{t-j} + \sum_{j=1}^k \theta_{1j} X_{t-j} + \sum_{j=k+1}^{d_{max}} \beta_{1j} X_{t-j} + \omega_{1t} \quad (2)$$

$$X_t = \alpha_1 + \sum_{i=1}^k \delta_{2i} Y_{t-i} + \sum_{j=k+1}^{d_{max}} \alpha_{2j} Y_{t-j} + \sum_{j=1}^k \theta_{2j} X_{t-j} + \sum_{j=k+1}^{d_{max}} \beta_{2j} X_{t-j} + \omega_{2t} \quad (3)$$

where ω_{1t} and ω_{2t} are the VAR error terms and d_{\max} is the maximum order of integration, according to the original specification of the Toda–Yamamoto procedure. Therefore, in Equation (3), causality in the sense of Granger between X and Y will be detected, provided that $\delta_{1i} \neq 0$ for every i , and, on an identical basis, Equation (4) will imply causality in the sense of Granger between X and Y , if $\delta_{2i} \neq 0$ for every i .

Once the VAR $(p + d_{\max})$ model is obtained, the implementation of the Toda–Yamamoto procedure in practice requires the realization of three differentiated steps:

- Testing each time-series to conclude the maximum order of integration d_{\max} of the variables by using, individually or jointly, the following tests: ADF (Augmented Dickey–Fuller), KPSS (Kwiatkowski–Phillips–Schmidt–Shin), and/or PPE (Phillips–Perron).
- Next, the optimal lag length (p) should be obtained based on the criteria: AIC (Akaike Information Criterion), FPE (Akaike’s Final Prediction Error), SC (Schwarz), HQ (Hannan and Quinn), and LR (Likelihood-Ratio), seeking, as much as possible, an optimal length supported by the maximum degree of unanimity between criteria.
- Finally, the Granger causality test between the variables X and Y (in both directions) is properly performed by considering that the rejection of the null hypothesis implies the existence of causality in the sense of Granger according to the Toda–Yamamoto procedure and that a reciprocal rejection would indicate a bilateral causal relationship between the analyzed variables.

1.3 Multilayer causality based network

We denote the proposed multilayer causality based networks as $\overline{\Omega} = \{G^{[1]}, G^{[2]}, \dots, G^{[L]}\}$ with L layers and N nodes, where $\{G^{[a]}\} = G(V, A^{[a]})$ is layer a of multilayer causality based networks, $V = \{1, 2, \dots, N\}$ is the set of nodes, and $A^{[a]}$ is the set of edges of layer a . On each layer, nodes represent a yield curve factor, and a directed edge indicates that there is a corresponding causality effect from the starting node to the terminal one. In our case $L=3$, and we assume that the first layer, the second layer and the third layer corresponds to level, slope and curvature layers, respectively. For any two factors $i, j \in V$, we draw a direct edge from i to j on the first (second, third) layer, if node i has a level (slope, curvature) causing effect on node j . $A^{[a]} = \{a_{ij}^{[a]}\}_{N \times N}$ is a directed binary connection matrix for all pairs of nodes i and j on layer a , where the element $a_{ij}^{[a]}$ in the matrix $A^{[a]}$ is defined as

$$\alpha_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } i \text{ has a corresponding causality effect on } j \text{ layer } \alpha \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Thus, multilayer information spillover networks are simplified to a 3 dimensional $N \times N$ adjacency matrix by mathematical notation. Considering the unpredictability of the financial system and the dynamic changes of interconnectedness among the yield curve nodes, we build time-varying multilayer causality based networks using rolling window analysis. TBC...