How Arbitrage-Free is the Nelson-Siegel Model?

Laura Coroneo, Ken Nyholm, Rositsa Vidova-Koleva

Bank of Canada Conference on Fixed Income Markets
September 2008

Introduction

In this paper we show that the Dynamic Nelson-Siegel term structure model is arbitrage-free in a statistical sense

- The Nelson-Siegel model is widely used in *practice*
  - by fixed-income managers
  - by Central Banks
  - the ECB publishes daily a Svensson-Soderlind yield curve
  - and RMA (ECB) uses a regime-switching extension of the Nelson-Siegel model in its foreign reserves management
- Four reasons for its success:
  - it is easy to estimate
  - it provides an intuitive interpretation of yields (level, slope, curvature)
  - it recovers yields for all maturities
  - and it empirically fits data well in- and out-of-sample

Introduction

However, from a theoretical view point, the Nelson-Siegel model:

- is not arbitrage-free (Bjork and Christensen (1999))
- and it does not belong to the family of affine models (Diebold, Ji and Li (2004))

The affine class of no-arbitrage term structure models:

- is the preferred choice in the academic literature
- it precludes arbitrage opportunities among yields observed at different maturities
- and it also produces good in- and out-of-sample fit to observed yields

Introduction

In this paper we address the following questions:

- how much should we worry about the Nelson-Siegel model not being arbitrage-free by construction?
- how significant (in a statistical sense) is the difference between the estimates of the Nelson-Siegel model and the an Affine Arbitrage-free model?
- in other words: what is the added value of the no-arbitrage constraints?

General Framework of a Term Structure Model

$$y_{t,\tau} = a_\tau + b_\tau X_t + \epsilon_{t,\tau},$$

- $y_{t,\tau}$ - yields at time t for maturity τ
- X_t - vector of yield curve factors (common across maturities)
- a_τ - constant
- b_τ - yield curve factor loadings
- $\epsilon_{t,\tau}$ - measurement error
- Nelson-Siegel and Affine No-Arbitrage Models impose different assumptions on a_τ, b_τ and X_t

Nelson-Siegel Model

$$y_{t,\tau} = a_\tau^{NS} + b_\tau^{NS} X_t^{NS} + \epsilon_{t,\tau}^{NS}$$

where

$$\begin{aligned} a_\tau^{NS} =& 0 \\ b_\tau^{NS} =& \left[1 \quad \frac{1-\exp(-\lambda\tau)}{\lambda\tau} \quad \frac{1-\exp(-\lambda\tau)}{\lambda\tau} - \exp(-\lambda\tau) \right] \end{aligned}$$

- τ - time to maturity
- λ - decay parameter
- X_t^{NS} = [level, slope, curvature]

Affine No-Arbitrage Model

$$y_{t,\tau} = a_\tau^{NA} + b_\tau^{NA} X_t^{NA} + \epsilon_{t,\tau}^{NA}$$

- state equation: $X_t^{NA} = \mu + \Phi X_{t-1}^{NA} + u_t$
- market price of risk: $\Lambda_t = \lambda_0 + \lambda_1 X_t^{NA}$
- short rate equation: $r_t = a_1^{NA} + b_1^{NA} X_t^{NA} + v_t$
- no-arbitrage restrictions:

$$A_{\tau+1} = A_\tau + B'_\tau \left(\mu - \Sigma\ \lambda_0\right) + \frac{1}{2} B'_\tau \Sigma \Sigma' B_\tau - A_1, \qquad a_\tau^{NA} = -\frac{A_\tau}{\tau}$$

$$B'_{\tau+1} = B'_\tau \left(\Phi - \Sigma\ \lambda_1\right) - B'_1, \qquad b_\tau^{NA} = -\frac{B_\tau}{\tau}$$

Our framework

$$y_t = a^{NA} + b^{NA}\widehat{X}_t^{NS} + \epsilon_t^{NA}, \quad \epsilon_t^{NA} \sim N(0,\Omega)$$

- relying on a yield-block-resampling approach we generate multiple yield curve data sets
- we use the factors estimated from the Nelson-Siegel model, $\widehat{X}_t^{NS}$, in the no-arbitrage affine model ...
- to estimate a^{NA} and b^{NA} that satisfy the no-arbitrage restrictions
- in this way we generate empirical confidence intervals for the no-arbitrage loadings
- we then test whether $\widehat{a}^{NA}$ and $\widehat{b}^{NA}$ are significantly different from a^{NS} and b^{NS}

Our framework

Three sets of tests are used to assess whether the NS model is compatible with the no-arbitrage restrictions

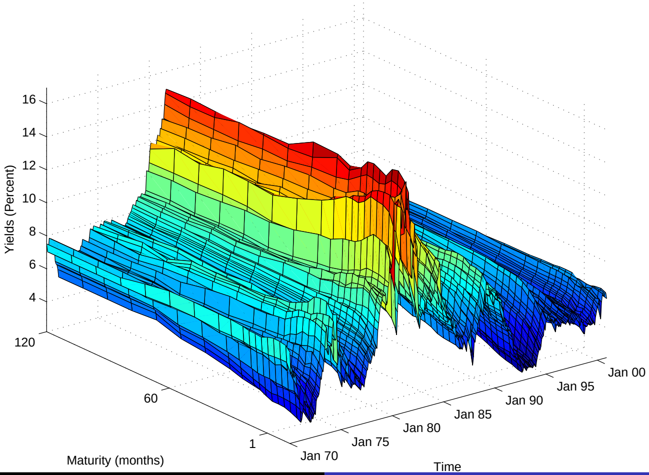
1. first using the empirical confidence intervals:

$$
\begin{aligned}
H_0^1 &: \widehat{a}_\tau^{NA} = a_\tau^{NS} = 0 \\
H_0^2 &: \widehat{b}_\tau^{NA}(1) = b_\tau^{NS}(1) \\
H_0^3 &: \widehat{b}_\tau^{NA}(2) = b_\tau^{NS}(2) \\
H_0^4 &: \widehat{b}_\tau^{NA}(3) = b_\tau^{NS}(3)
\end{aligned}
$$

2. then looking at the in-sample performance of the NS factors with and without the arbitrage constraints
3. and finally by looking at the out-of-sample forecast performance of the NS factors with and without the arbitrage constraints

Data

Figure: Zero-coupon yields



Data

- US monthly zero-coupon yields from 1970:1 to 2000:12
- Maturities 1, 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120 months (Francis Diebold's homepage)

Table: Summary statistics for selected maturities

τ	mean	std dev	min	max	$\rho(1)$	$\rho(2)$	$\rho(3)$	$\rho(12)$
3	6.75	2.66	2.73	16.02	0.97*	0.94*	0.91*	0.71*
12	7.2	2.57	3.11	15.82	0.97*	0.94*	0.91*	0.74*
24	7.46	2.44	3.78	15.65	0.98*	0.94*	0.92*	0.75*
60	7.84	2.25	4.35	15.01	0.98*	0.96*	0.94*	0.79*
84	7.99	2.18	4.35	14.98	0.98*	0.96*	0.94*	0.78*
120	8.05	2.14	4.44	14.93	0.98*	0.96*	0.94*	0.78*

Estimation

- for the Nelson-Siegel model we follow Diebold and Li (2006): fix $\lambda = 0.0609 \longrightarrow \widehat{X}_t^{NS}$
- the estimation of the affine No-arbitrage model follows (Ang, Piazzesi and Wei (2006))
  - First step:
    - VAR in the NS factors, $X_t^{NS} = \mu + \Phi X_{t-1}^{NS} + u_t$
    - OLS of the short rate (1-month yield), $r_t = a_1^{NA} + b_1^{NA} X_t^{NS} + v_t$
  - Second step:
    - min SSR to estimate the market price of risk parameters, λ_0 and λ_1
    - compute the no-arbitrage loadings $\widehat{a}^{NA}$ and $\widehat{b}^{NA}$
  - Last step:
    - to account for the two-step procedure, confidence intervals around $\widehat{a}^{NA}$ and $\widehat{b}^{NA}$ constructed using resampling

In-sample fit

Residuals from the Nelson-Siegel model

τ	mean	RMSE	MAD	$\rho(1)$	$\rho(6)$
3	0.027	0.114	0.013	0.274	0.159
12	0.046	0.122	0.015	0.586	0.127
24	-0.040	0.073	0.005	0.493	0.044
60	-0.053	0.096	0.009	0.655	0.312
120	0.002	0.140	0.020	0.699	0.345

Residuals from no-arbitrage model

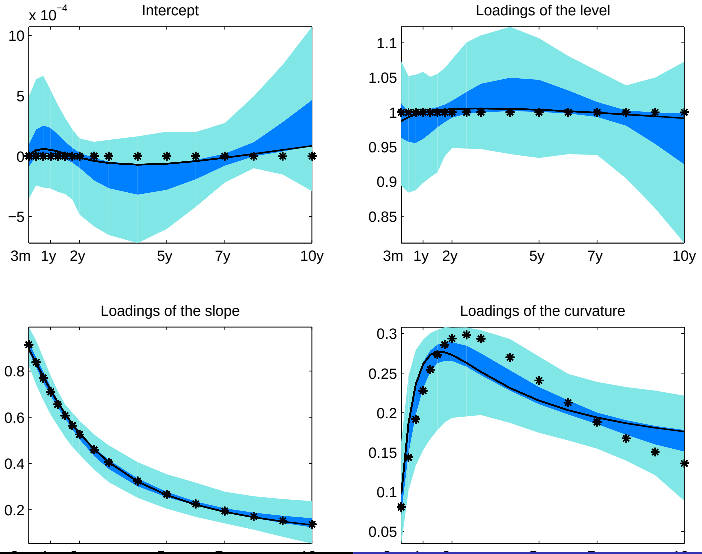
3	0.080	0.132	0.018	0.448	0.219
12	- 0.019	0.109	0.012	0.514	0.147
24	-0.041	0.071	0.005	0.491	0.134
60	0.004	0.100	0.010	0.688	0.350
120	-0.060	0.144	0.021	0.705	0.464

Test

τ	Intercept a^{NS}	Intercept $\widehat{a}^{NA}$	Level $b^{NS}(1)$	Level $\widehat{b}^{NA}(1)$	Slope $b^{NS}(2)$	Slope $\widehat{b}^{NA}(2)$	Curvature $b^{NS}(3)$	Curvature $\hat{b}^{NA}(3)$
3	0.00	0.00	1.00	0.99	0.91	0.89	0.08	0.10
12	0.00	0.01	1.00	1.00	0.71	0.72	0.23	0.26
24	0.00	0.00	1.00	1.00	0.53	0.53	0.29	0.27
30	0.00	0.00	1.00	1.01	0.46	0.46	0.30	0.26
36	0.00	-0.01	1.00	1.01	0.41	0.41	0.29	0.25
48	0.00	-0.01	1.00	1.00	0.32	0.32	0.27	0.23
60	0.00	-0.01	1.00	1.00	0.27	0.26	0.24	0.21
84	0.00	0.00	1.00	1.00	0.19	0.19	0.19	0.19
120	0.00	0.01	1.00	0.99	0.14	0.13	0.14	0.18

- Estimated $\widehat{a}^{NA}$ and $\widehat{b}^{NA}$ are close to a^{NS} and b^{NS}

Test


x 10^−4
Intercept
10
5
0
−5
3m
1y
2y
5y
7y
10y
Loadings of the level
1.1
1.05
1
0.95
0.9
0.85
3m
1y
2y
5y
7y
10y
Loadings of the slope
0.8
0.6
0.4
0.2
Loadings of the curvature
0.3
0.25
0.2
0.15
0.1
0.05

Out-of-sample forecast

Further results

- generate $h = 1$, 6, 12-steps ahead iterative forecasts
- project the yield curve factors forward using the estimated VAR parameters $\hat{X}^{NS}_{t+h|t} = \sum_{s=0}^{h-1} \hat{\Phi}^s \hat{\mu} + \hat{\Phi}^h \hat{X}^{NS}_t$
- compute out-of-sample yield forecasts for the two models, given the projected factors

$$\hat{y}^{NS}_{t+h|t} = b^{NS} \hat{X}^{NS}_{t+h|t}$$
$$\hat{y}^{NA}_{t+h|t} = \widehat{a}^{NA} + \widehat{b}^{NA} \hat{X}^{NS}_{t+h|t}$$

- Evaluation period 1994:01 to 2000:12
- The benchmark is the random walk

Out-of-sample forecast

Mean Squared Forecast Errors

	1-m ahead		6-m ahead		12-m ahead	
τ	NS	NA	NS	NA	NS	NA
3	0.91	**0.89**	0.72	**0.70**	0.64	**0.63**
12	1.01	**1.00**	**0.80**	0.81	**0.64**	0.65
24	**1.09**	1.11	0.80	0.80	0.67	0.67
30	1.04	1.04	0.80	**0.78**	0.68	**0.67**
36	0.99	**0.98**	0.80	**0.78**	0.70	**0.69**
48	0.98	0.98	0.84	**0.81**	0.76	**0.73**
60	1.10	**1.04**	0.88	**0.85**	0.81	**0.79**
84	1.08	1.08	0.91	0.91	0.87	**0.86**
120	**1.08**	1.32	**1.02**	1.08	**1.00**	1.05

MSFE relative to the random walk

Conclusion

What is the added value of the no-arbitrage constraints (for US data)?

- The factor loadings of the no-arbitrage model are not statistically different from the Nelson-Siegel ones, at a 95% confidence level
- The no-arbitrage constraints do not significantly improve the performance of the Nelson-Siegel factors

Our results suggest that the Nelson-Siegel model is compatible with the no-arbitrage constraints.