Model description: A Terminal Rate Approach to Dissecting Term Premia, Journal of Fixed Income, 2024*

Johannes Kramer
jkramer@imf.org

Ken Nyholm †
ken.nyholm@ecb.europa.eu

October 6, 2024

Abstract

We parameterize the yield curve using four factors that directly capture the term structure of term premia and the risk-free expectations curve. The model can be estimated using the code available at: `https://kennyholm.github.io/posts/matlab/`. Also the models of Joslin, Singleton, and Zhu (2011) (JSZ), Adrian, Crump, and Mönch (2013) (ACM), and Diebold and Li (2006) (DNS), can be estimated using this code. When using the code please quote: `Kramer and Nyholm, 2024, Journal of Fixed Income`.

*The code used in the paper is available for download at https://kennyholm.github.io/posts/matlab/.

†Corresponding author.

The Model

Our analytical framework is based on the dynamic Nelson-Siegel class of term structure models (among others, Diebold and Li (2006), Diebold, Rudebusch, and Aruoba (2006), and Diebold and Rudebusch (2013)). Although this modelling setup is not strictly arbitrage-free, it prevents dominant trading strategies (see, Feunou, Fontaine, Le, and Lundblad (2022)). The no-dominance criterion, which is closely related to the no-arbitrage condition, ensures that portfolios with the same price generate identical expected return, thereby preventing the creation of riskless profits through zero-cost long-short bond portfolio strategies. However, the no-dominance criterion is less stringent than the no-arbitrage condition, with the latter also excluding strategies that offer a potential profit without the risk of loss and without requiring an initial investment. The key difference therefore lies in the no-arbitrage condition's exclusion of strategies that provide an opportunity for gain, without the possibility of loss. Such pay-offs resemble strategies involving derivatives, which offer asymmetric payoff profiles.

The empirical relevance of strategies that distinguish no-dominance from no-arbitrage remains an area of investigation. If these strategies do not exist in practice, the two principles may appear indistinguishable to yield-curve practitioners. Given that our model is applied to the U.S. government market - widely regarded as efficient in its price formation - we are confident in using a model based on the no-dominance principle. For a version of the model that adheres to the no-arbitrage principle, please refer to Appendix A.

Bond prices and yields

Following the no-dominance principle suggested by Feunou, Fontaine, Le, and Lundblad (2022) for the pricing of credit-risk free bonds, the price of a 0-maturity bond is equal to

1, and recursive equations specify the pricing for bonds with residual maturities $n > 0$:

$$P_0(X_t) \equiv 1 \tag{1}$$

$$P_n(X_t) = P_{n-1}\left(g(X_t)\right) \cdot exp(-\rho' \cdot X_t) \tag{2}$$

where P denotes the price, n is the residual maturity, t counts calendar time, and the vector X holds the yield curve factors. The function $g(\cdot)$ represents the process governing the yield curve factors relevant for pricing purposes[1], and the product $\rho' X$ specifies the discount rate as a function of the yield curve factors. This can be seen from the prices from the expressions of a 1-period and a 2-period bond:

$$P_1(X_t) = P_0\left(g(X_t)\right) \cdot exp(-\rho' \cdot X_t) = 1 \cdot exp(-\rho' \cdot X_t), \tag{3}$$

$$P_2(X_t) = P_1\left(g(X_t)\right) \cdot exp(-\rho' \cdot X_t). \tag{4}$$

Given that the modelled bonds are assumed to be credit-risk free, i.e. $P_0(X_t) \equiv 1$, it is seen that (3) simply amounts to the one-period discount-factor. In (2) we need an expression for the price of a bond that is one period closer to maturity on the RHS compared to the LHS, i.e. we need $P_{t+1}(n-1)$.

Through successive substitution, the price of a n-maturity bond is:

$$P_n(X_t) = exp\left(-\sum_{j=0}^{n-1} \rho' \cdot \left(\underbrace{(g \circ g \circ \cdots \circ g)}_{j}(X_t)\right)\right) \tag{5}$$

where $g \circ g$ denotes the composition of the g-function. To ensure that bond prices satisfy the properties of: i) positivity, ii) invertibility and iii) discounting distant cash flows, the following condition is adopted:

$$g(X_t) = \Phi \cdot X_t. \tag{6}$$

[1]Drawing a comparison between the no-dominance framework and the no-arbitrage set-up, it is noted that (2) represents $\mathbb{Q}$-measure discounting, as no adjustment is made for the market-price of risk. Consequently $g(\cdot)$ represent the $\mathbb{Q}$-corresponding dynamics of the yield curve factors.

Accordingly, the expression for n-maturity yield is:

$$y_t(n) \equiv -\frac{\log\left(P_n(X_t)\right)}{n} = \frac{1}{n}\sum_{j=0}^{n-1}\rho'\left(\Phi^j\right)\cdot X_t. \tag{7}$$

With Φ defined as:

$$\Phi \equiv \begin{bmatrix} 1-a & a & 1-\gamma & 1-\gamma \\ 0 & 1 & 1-\gamma & 1-\gamma \\ 0 & 0 & \gamma & \gamma-1 \\ 0 & 0 & 0 & \gamma \end{bmatrix} \tag{8}$$

and, ρ defined as:

$$\rho = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{9}$$

it is seen that:

$$\begin{aligned} y_t(n) &= \frac{1}{n}\sum_{j=0}^{n-1}\rho'\cdot\Phi^j\cdot X_t = \rho'\cdot\sum_{j=0}^{n-1}\frac{\Phi^j}{n}\cdot X_t \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} \frac{1-(1-a)^n}{an} & 1-\frac{1-(1-a)^n}{an} & 1-\frac{1-\gamma^n}{n(1-\gamma)} & \frac{1-\gamma^n}{n(1-\gamma)}-\gamma^{n-1} \\ 0 & 1 & 1-\frac{1-\gamma^n}{n(1-\gamma)} & \frac{1-\gamma^n}{n(1-\gamma)}-\gamma^{n-1} \\ 0 & 0 & \frac{1-\gamma^n}{n(1-\gamma)} & \gamma^{n-1}-\frac{1-\gamma^n}{n(1-\gamma)} \\ 0 & 0 & 0 & \frac{1-\gamma^n}{n(1-\gamma)} \end{bmatrix}\cdot X_t \\ &= \begin{bmatrix} \frac{1-(1-a)^n}{an}, & 1-\frac{1-(1-a)^n}{an}, & 1-\frac{1-\gamma^n}{n(1-\gamma)}, & \frac{1-\gamma^n}{n(1-\gamma)}-\gamma^{n-1} \end{bmatrix}\cdot X_t. \end{aligned} \tag{10}$$

The yield equation of our model as portrayed in (10) can then be written as:

$$y_t(n) = b_n\cdot X_t + e_t, \tag{11}$$

where the loading structure is defined by:

$$b_n = \begin{bmatrix} \frac{1-(1-a)^n}{an}, & 1-\frac{1-(1-a)^n}{an}, & 1-\frac{1-\gamma^n}{n(1-\gamma)}, & \frac{1-\gamma^n}{n(1-\gamma)}-\gamma^{n-1} \end{bmatrix}. \tag{12}$$

Interpretation of the yield curve factors

The loading structure, b_n, is defined in (12), and the corresponding factor interpretations are conjectured to be:

$$X_t = \begin{bmatrix} r_t \\ C_t^* \\ \theta_t^s \\ \theta_t^c \end{bmatrix} = \left\{ \begin{array}{l} \text{short rate} \\ \text{nominal long-term natural rate of interest} \\ \text{slope of term structure of term premium} \\ \text{curvature of term structure of term premium} \end{array} \right\} \tag{13}$$

To prove this conjecture, we consider Exhibit 1 that shows the structure of the factor loadings and Exhibit 2 documenting their first derivatives. The loading for the first factor across all maturities, denoted by $b(:,1)$, exhibits a consistent decline from an initial value near unity at the 3-month maturity mark, progressively declining towards zero, as the maturity increases. This pattern is confirmed by the first row of Exhibit 2, showing that the loading is bound between 1 and 0, and that it is monotonically decreasing as $\frac{\partial}{\partial n} < 0$. These observations in conjunction with $\rho = [1, 0, 0, 0]'$, and the fact that the loadings for the remaining three factors are zero at the shortest maturity, unequivocally shows that the first factor represents the short-term interest rate. Given that the loadings of the first and second factors, $b(:,1) + b(:,2)$, sums to 1, and considering that the loading for the second factor, $b(:,2)$, has a positive slope and approaches one as the maturity increases, as seen in Exhibit 2, it becomes apparent that the second factor embodies the long-term anchor for the short-term rate, denoted by C^*. Consequently, these initial two factors, along with their respective loadings, delineate the risk-free expectations curve (y_t^{rf}): [2]

$$\begin{aligned} y_t^{rf}(n) &\equiv \frac{1}{n} \cdot \mathbb{E}_t \sum_{j=0}^{n-1} r_{t,t+j} \\ &= \frac{1-(1-a)^n}{an} \cdot r_t + 1 - \frac{1-(1-a)^n}{an} \cdot C^*. \end{aligned} \tag{14}$$

[2] In the context of deriving the yield in the cross-sectional dimension, i.e. for pricing purposes, it is noted that our model implies an AR(1) process for the short rate: $r_t = \kappa \cdot C^* + (1-\kappa) \cdot r_{t-1} = c + \beta \cdot r_{t-1}$, where $c = \kappa \cdot C^*$, and $\beta = 1 - \kappa$.

The wedge between the expectations curve and the model-fitted yield curve is then, by definition, formed by the product of the remaining two yield curve factors and their loadings, $b(:,3)$ and $b(:,4)$:

$$\begin{aligned} y_t(n) &\equiv y_t^{drf}(n) + \theta_t(n) \\ &= y_t^{rf}(n) + \left(1 - \frac{1-\gamma^n}{n(1-\gamma)}\right) \cdot \theta_t^s + \left(\frac{1-\gamma^n}{n(1-\gamma)} - \gamma^{n-1}\right) \cdot \theta_t^c + z_t, \end{aligned} \tag{15}$$

where z_t is the difference between the yield curve fitted as the sum of the risk free curve and the term structure of term premia, as defined by the second and third terms in (15).

Exhibits 1 and 2 show that the loading for the third factor is monotonically increasing ($\frac{\partial}{\partial n} > 0$ in the third row of 2) from a value of 0 (at the 1-month maturity mark) and has 1 as a limit when $n \to \infty$. This pattern is identical to that of a 'slope' factor, having little impact on short maturities with increasing impact as the maturity increases; in fact, our slope factor is similar in design to the slope factor of Nelson and Siegel (1987) as parameterized by Diebold and Li (2006), which, using n to indicate maturity, is:[3]

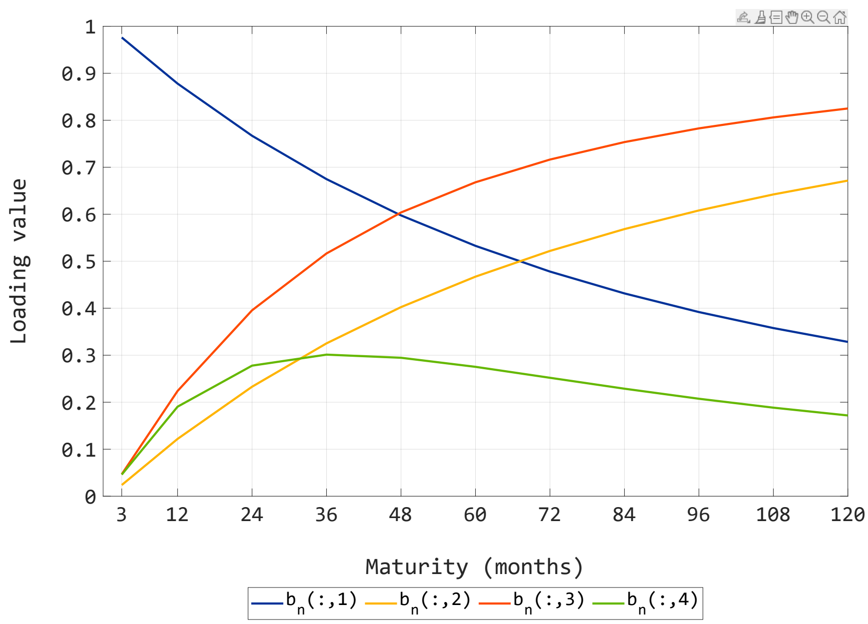
$$b^{\mathtt{NS}}_{\mathtt{slope}} = \frac{1-e^{-\lambda n}}{\lambda n}. \tag{16}$$

Disregarding the different definitions of the slope (either as the short rate minus the long rate, or the other way around), it is observed that $\lambda = (1-\gamma)$, and since for small values of λ we have that $e^{-\lambda} \approx 1 - \lambda$, we see that our model is a discrete-time version of the well-known Nelson-Siegel model with a different definition of the yield curve factors. Similarly, our fourth factor expresses the curvature component, like the curvature factor in the dynamic Nelson-Siegel model. However, noting the important difference that the slope an curvature factors in our model describe the term structure of term premia, while the dynamic Nelson-Siegel model describe the yield curve itself.

[3]Note that Diebold and Li (2006) define the slope factor as the short rate minus the long rate, whereas we use long rate minus short rate.

Exhibit 1

Loadings for the yield curve factors



Notes: The figure shows the loadings of the four factors included in the Terminal Rate Model (TRM). The loadings are plotted as functions of maturities. Factor 1 and 2 trace out the expectations component of the yield curve as a convex combination of the short rate and the equilibrium long-maturity convergence point for the short rate, labelled C^*. Factor 3 and 4 account for the slope and curvature of the term structure of term premia.

Exhibit 2

Factor loading properties

b_n	*Operation*		
	$\lim_{n\to 1}$	$\lim_{n\to\infty}$	$\frac{\partial}{\partial n}$
$b(:,1) = \frac{1-(1-a)^n}{an}$	1	0	$\frac{(1-a)^n-1}{a\,n^2} - \frac{\ln(1-a)\,(1-a)^n}{a\,n} < 0$
$b(:,2) = 1 - \frac{1-(1-a)^n}{an}$	0	1	$\frac{\ln(1-a)\,(1-a)^n}{a\,n} - \frac{(1-a)^n-1}{a\,n^2} > 0$
$b(:,3) = 1 - \frac{1-\gamma^n}{n(1-\gamma)}$	0	1	$-\frac{\gamma^n\,n\,\log(\gamma)-\gamma^n+1}{n^2\,(\gamma-1)} > 0$
$b(:,4) = \frac{1-\gamma^n}{n(1-\gamma)} - \gamma^{n-1}$	0	0	$\frac{\gamma^n\,n\,\log(\gamma)-\gamma^n+1}{n^2\,(\gamma-1)} - \frac{\gamma^n\,\log(\gamma)}{\gamma} \gtreqqless 0$

Notes: The table shows the key properties of the factor-loadings derived in (12), by showing the limits for $n \to 1$, which is the shortest possible maturity before the bond is redeemed, and $n \to \infty$ and the first derivatives. The parameters a and g determine the value loadings as $n \to 1$. As $n \to \infty$ the loadings converge to 0 or 1. The sign and functional form of the first derivative show that the loadings for factors 1-3 are monotonously increasing or decreasing functions. The derivative of loading 4 shows that this loading reaches a maximum point somewhere in the maturity spectrum, and then decreases afterwards.

Factor dynamics

Within the no-dominance framework there is no explicitly hard-wired relationship between the parameters that govern the cross-sectional behavior of yields and those that govern the time series behavior of the factors. This is in contrast to the no-arbitrage framework (see, e.g. Joslin, Singleton, and Zhu (2011)) where the market price of risk serves this role, but fully in line with the traditional agnostic approach of the Dynamic Nelson-Siegel model (see, e.g. Diebold and Li (2006)). Following the standard assumption in the literature, our yield curve factors are assumed to be appropriately modelled by a VAR(1):

$$X_t = k + K \cdot (X_{t-1} - k) + v_t \tag{17}$$

where k is the vector of means, K is the matrix of autoregressive coefficients and v_t is the residuals. X hold the yield curve factors described above.

We relying on a nominal version of Roberts (2018) where C_t^* is defined in the following way:

$$C_t^* = R_t + (xgap_t - \eta \cdot xgap_{t-1}) / \sigma, \tag{18}$$

where R_t is the 10-year nominal rate, $xgap_t$ is the output gap, and as suggested by Roberts (2018) we set $\eta = 0.75$ and $\sigma = 3.5$. This relationship expresses a simple way to estimate the equilibrium long-term rate in the economy, as the nominal 10-year risk free rate adjusted for unexpected shocks to the output gap. We then compare the 10-year term premium estimate from the TRM to those obtained by the models of Adrian, Crump, and Mönch (2013) (ACM), Joslin, Singleton, and Zhu (2011)(JSZ), and Diebold and Li (2006)(DNS).

For comparison, we use the TRM as a calibration tool to obtain estimates for the terminal rates implied by the ACM, JSZ, and DNS models. The corresponding $C_{t,ACM}^*$, $C_{t,JSZ}^*$, and $C_{t,DNS}^*$ are found in the following way:

1. The term structure of term premia is obtained by estimating Adrian, Crump, and Mönch (2013) (ACM), Joslin, Singleton, and Zhu (2011)(JSZ), and Diebold and Li

(2006)(DNS) models. This gives three matrices of term premia estimates, one for each model: $TP_{t,ACM}$, $TP_{t,JSZ}$, and, $TP_{t,DNS}$, each comprising $529-by-11$ observations.

2. The TRM is estimated using the original yield curve data such that $C^*_{t,ACM}$, $C^*_{t,JSZ}$, and $C^*_{t,DNS}$ are found as the solution to: $min_{C^*_{t,model}} \left(TP_{model} - TP_{TRM}|C^*_{t,model}\right)^2$, with $model = \{ACM, JSZ, DNS\}$

Appendix: An arbitrage free version of the model

Our purpose here is to illustrate how the standard linear modelling set-up (see, e.g., Duffie and Kan (1996), Dai and Singleton (2000), and Ang and Piazzesi (2003)) can be used to derive a discrete-time arbitrage-free version of the model presented in the text.

Within the continuous-time setting Christensen, Diebold, and Rudebusch (2011) have shown how to maintain the parametric loading structure of the Nelson and Siegel (1987), while ensuring that arbitrage constraints are fulfilled.[4]

As before, let X_t denote the vector of the modelled yield curve factors, at time t. Furthermore, let the dynamics of X_t be governed by vector autoregressive (VAR) processes of order one, under both the empirical measure, $\mathbb{P}$, and the pricing measure, $\mathbb{Q}$:

$$X_t = k^{\mathbb{P}} + \Phi^{\mathbb{P}} \cdot X_{t-1} + \Sigma^{\mathbb{P}} \epsilon_t^{\mathbb{P}}, \qquad \epsilon_t^{\mathbb{P}} \sim N(0,1) \tag{19}$$

$$X_t = k^{\mathbb{Q}} + \Phi^{\mathbb{Q}} \cdot X_{t-1} + \Sigma^{\mathbb{Q}} \epsilon_t^{\mathbb{Q}}, \qquad \epsilon_t^{\mathbb{Q}} \sim N(0,1). \tag{20}$$

with $\Sigma\Sigma' = \Omega$ being the variance of the residuals, and it is assumed that $\Sigma^{\mathbb{P}} = \Sigma^{\mathbb{Q}}$. To define our model we use:

$$\Phi^{\mathbb{Q}} = \Phi = \begin{bmatrix} 1-a & a & 1-\gamma & 1-\gamma \\ 0 & 1 & 1-\gamma & 1-\gamma \\ 0 & 0 & \gamma & \gamma-1 \\ 0 & 0 & 0 & \gamma \end{bmatrix} \tag{21}$$

As the first element in X_t is defined to be the one-period short rate, we have:

$$r_t = \rho_0 + \rho_1' X_t. \tag{22}$$

with the following constraints on (22): $\rho_0 = 0$ and $\rho_1 = [1, 0, 0, 0]'$.

We now impose absence of arbitrage on the model by introducing the unique pricing mechanism, that governs all traded assets:

$$P_{t,\tau} = \mathbb{E}_t \left[M_{t+1} \cdot P_{t+1,\tau-1} \right] \tag{23}$$

[4] See also, Krippner (2013) and Diebold and Rudebusch (2013).

The idea here is that when the bond matures at time T, its value is known with certainty, since it is default-free: the bond pays its principal value on that day, so $P_{T,0} = 1$. At any time $t + j$ before maturity, the price of the bond can therefore be found as the one-period discounted-value of the price at time $t + j + 1$, all the way back to time t. Discounting is done using the stochastic discount factor (also called the pricing kernel), which is denoted by M_t, and this quantity is assumed to be given by:

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}\right) \tag{24}$$

with

$$\lambda_t = \lambda_0 + \lambda_1 \cdot X_t, \tag{25}$$

where λ_t is of dimension (4×1) in our application, because we have four factors, λ_0 is of dimension (4×1), and λ_1 is a matrix of dimension (4×4).

It is recalled that:

$$y_{t,\tau} = -\frac{1}{\tau}\log(P_{t,\tau}), \tag{26}$$

and that we can write the yield curve expression as an affine function:

$$y_{t,\tau} = -\frac{A_\tau}{\tau} - \frac{B_\tau'}{\tau}X_t. \tag{27}$$

The bond price is therefore exponential affine in terms of A_τ and B_τ:

$$P_{t,\tau} = \exp\left(A_\tau + B_\tau' X_t\right). \tag{28}$$

To derive closed-form expressions for A_τ and B_τ, the fundamental pricing equation is invoked (24):

$$\begin{aligned} P_{t,\tau} &= \mathbb{E}_t\left[M_{t+1} \cdot P_{t+1,\tau-1}\right] && (29) \\ &= \mathbb{E}_t\left[\exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}\right) \cdot \exp\left(A_{\tau-1} + B_{\tau-1}' X_{t+1}\right)\right]. && (30) \end{aligned}$$

The expression for X_{t+1} (see equation 19) is substituted:

$$P_{t,\tau} = \mathbb{E}_t \left[\exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}\right) \cdot \exp\left(A_{\tau-1} + B_{\tau-1}'\left(k^{\mathbb{P}} + \Phi^{\mathbb{P}} X_t + \Sigma\epsilon_{t+1}^{\mathbb{P}}\right)\right)\right], \tag{31}$$

and, the terms are then separated into two groups: one to which the expectations operator should be applied, i.e. $t+1$ terms, and another group, which are known at time t:

$$\begin{aligned} P_{t,\tau} =& \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t + A_{\tau-1} + B_{\tau-1}'k^{\mathbb{P}} + B_{\tau-1}'\Phi^{\mathbb{P}} X_t\right) \\ &\cdot \mathbb{E}_t\left[\exp\left(-\lambda_t'\epsilon_{t+1}^{\mathbb{P}} + B_{\tau-1}'\Sigma\epsilon_{t+1}^{\mathbb{P}}\right)\right]. \end{aligned} \tag{32}$$

The question is then, how can we calculate the expectations part of (32):

$$\mathbb{E}_t\left[\exp\left(-\lambda_t' + B_{\tau-1}'\Sigma\right)\epsilon_{t+1}^{\mathbb{P}}\right]. \tag{33}$$

To this end, the moment generating function of the multivariate normal distribution is used. Since $\epsilon^{\mathbb{P}} \sim N(0, I)$, it is known that:

$$\mathbb{E}[\exp(a'\epsilon^{\mathbb{P}})] = \exp\left(\frac{1}{2}a' \cdot I \cdot a\right), \tag{34}$$

so, the expectation in (32) can be calculated, using $a' = (-\lambda_t' + B_{\tau-1}'\Sigma)$, as:

$$\begin{aligned} &\exp\left[\frac{1}{2}(-\lambda_t' + B_{\tau-1}'\Sigma) \cdot I \cdot (-\lambda_t' + B_{\tau-1}'\Sigma)'\right] \\ =&\exp\left[\frac{1}{2}(-\lambda_t' + B_{\tau-1}'\Sigma) \cdot I \cdot (-\lambda_t + \Sigma' B_{\tau-1})\right] \\ =&\exp\left[\frac{1}{2}\left(\lambda_t'\lambda_t - \lambda_t'\Sigma' B_{\tau-1} - B_{\tau-1}'\Sigma\lambda_t + B_{\tau-1}'\Sigma\Sigma' B_{\tau-1}\right)\right], \end{aligned} \tag{35}$$

and, since $B_{\tau-1}'\Sigma\lambda_t$ is a scalar, and for a scalar h, we know that $h = h'$, so $B_{\tau-1}'\Sigma\lambda_t = \lambda_t'\Sigma' B_{\tau-1}$. We can then write:

$$\begin{aligned} &\mathbb{E}_t\left[\exp\left(-\lambda_t' + B_{\tau-1}'\Sigma\right)\epsilon_{t+1}^{\mathbb{P}}\right] \\ =&\exp\left[\left(\frac{1}{2}\lambda_t'\lambda_t - B_{\tau-1}'\Sigma\lambda_t + \frac{1}{2}B_{\tau-1}'\Sigma\Sigma' B_{\tau-1}'\right)\right]. \end{aligned} \tag{36}$$

This term is then reinserted into (32), giving:

$$P_{t,\tau} = \exp\left(-r_t + A_{\tau-1} + B'_{\tau-1}k^{\mathbb{P}} + B'_{\tau-1}\Phi^{\mathbb{P}}X_t - B'_{\tau-1}\Sigma\lambda_t + \frac{1}{2}B'_{\tau-1}\Sigma\Sigma' B'_{\tau-1}\right). \quad (37)$$

It is recalled that $r_t = \rho_1' X_t$, and that $\lambda_t = \lambda_0 + \lambda_1 X_t$. Inserting these expressions into (37), gives:

$$\begin{aligned} P_{t,\tau} = \exp\Big(&-\rho_1' X_t + A_{\tau-1} + B'_{\tau-1}k^{\mathbb{P}} + B'_{\tau-1}\Phi^{\mathbb{P}}X_t \\ &-B'_{\tau-1}\Sigma\left(\lambda_0 + \lambda_1 X_t\right) + \frac{1}{2}B'_{n-1}\Sigma\Sigma' B'_{\tau-1}\Big). \end{aligned} \quad (38)$$

Reorganizing this expression into terms that load on X_t and terms that do not, help matching coefficients with respect to equation (28):

$$\begin{aligned} P_{t,\tau} = \exp\Big(&A_{\tau-1} + B'_{\tau-1}\left(k^{\mathbb{P}} - \Sigma\lambda_0\right) + \frac{1}{2}B'_{\tau-1}\Sigma\Sigma' B'_{\tau-1} \\ &+B'_{\tau-1}\Phi^{\mathbb{P}}X_t - \rho_1' X_t - B'_{\tau-1}\Sigma\lambda_1 X_t\Big), \end{aligned} \quad (39)$$

which is:

$$\begin{aligned} P_{t,\tau} = \exp\Bigg(&A_{\tau-1} + B'_{\tau-1}\left(k^{\mathbb{P}} - \Sigma\lambda_0\right) + \frac{1}{2}B'_{\tau-1}\Sigma\Sigma' B'_{\tau-1} \\ &+\left[B'_{\tau-1}\left(\Phi^{\mathbb{P}} - \Sigma\lambda_1\right) - \rho_1'\right]X_t\Bigg). \end{aligned} \quad (40)$$

Matching the coefficients of (40) with those of (28) establishes the recursive formulas for A_n and B_n:

$$A_n = A_{n-1} + B'_{n-1}k^{\mathbb{Q}} + \frac{1}{2}B'_{n-1}\Sigma\Sigma' B'_{n-1} \quad (41)$$

$$B'_n = B'_{n-1}\Phi^{\mathbb{Q}} - \rho_1' \quad (42)$$

with $k^{\mathbb{Q}} = k^{\mathbb{P}} - \Sigma\lambda_0$, and $\Phi^{\mathbb{Q}} = \Phi^{\mathbb{P}} - \Sigma\lambda_1$. Recall that $\rho_0 = 0$ in our model setup. Using recursive substitution, we realize that the expression for B'_n also can be written in the following way:[5]

$$B_n = -\left[\sum_{k=0}^{\tau-1}\left(\Phi^{\mathbb{Q}}\right)^k\right]' \cdot \rho_1. \tag{43}$$

which is the same expression as we obtain in the text for the loading structure under the no-dominance requirement. Hence the arbitrage-free version of the model presented in the text is obtained by adding a constant A_n/n to the yield equation of the model:

$$A_n = A_{n-1} + B'_{n-1}k^{\mathbb{Q}} + \frac{1}{2}B'_{n-1}\Sigma\Sigma' B'_{n-1}. \tag{44}$$

[5] We see this by the use of an example. For $\tau = 3$, we have:

$$\begin{aligned}
B'_1 &= -\rho'_1 \\
B'_2 &= B'_1\Phi^{\mathbb{Q}} - \rho'_1 = -\rho'_1\Phi^{\mathbb{Q}} - \rho'_1 \\
B'_3 &= B'_2\Phi^{\mathbb{Q}} - \rho'_1 = (-\rho'_1\Phi^{\mathbb{Q}} - \rho'_1)\Phi^{\mathbb{Q}} - \rho'_1 \\
&= -\rho'_1\left(\Phi^{\mathbb{Q}}\right)^2 - \rho'_1\Phi^{\mathbb{Q}} - \rho'_1 \\
&= -\rho'_1\left(\left(\Phi^{\mathbb{Q}}\right)^2 + \left(\Phi^{\mathbb{Q}}\right)^1 + \left(\Phi^{\mathbb{Q}}\right)^0\right) \\
&= -\rho'_1\left[\sum_{k=0}^{2}\left(\Phi^{\mathbb{Q}}\right)^k\right]
\end{aligned}$$

so,

$$B_3 = -\left[\sum_{k=0}^{2}\left(\Phi^{\mathbb{Q}}\right)^k\right]' \rho_1,$$

which generalises to equation (43).

References

ADRIAN, T., R. K. CRUMP, AND E. MÖNCH (2013): "Pricing the Term Structure with Linear Regressions", *Journal of Financial Economics*, 110, 110–138.

ANG, A., AND M. PIAZZESI (2003): "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables", *Journal of Monetary Economics*, 50(4), 745–787.

CHRISTENSEN, J. H. E., F. X. DIEBOLD, AND G. D. RUDEBUSCH (2011): "The affine arbitrage-free class of Nelson-Siegel term structure models", *Journal of Econometrics*, 164, 4–20.

DAI, Q., AND K. SINGLETON (2000): "Specification analysis of affine term structure models", *Journal of Finance*, 55, 1943–1978.

DIEBOLD, F. X., AND C. LI (2006): "Forecasting the term structure of government bond yields", *Journal of Econometrics*, 130(2), 337–364.

DIEBOLD, F. X., AND G. D. RUDEBUSCH (2013): *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton University Press, Princeton, New Jersey, USA.

DIEBOLD, F. X., G. D. RUDEBUSCH, AND S. B. ARUOBA (2006): "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach", *Journal of Econometrics*, 131, 309–338.

DUFFIE, D., AND R. KAN (1996): "A yield-factor model of interest rates", *Mathematical Finance*, 6, 379–406.

FEUNOU, B., J.-S. FONTAINE, A. LE, AND C. LUNDBLAD (2022): "Tractable Term-Structure Models", *Management Science*, *https://doi.org/10.1287/mnsc.2021.4214*.

Joslin, S., K. J. Singleton, and H. Zhu (2011): "A New Perspective on Gaussian Dynamic Term Structure Models", *Review of Financial Studies*, 24, 926–970.

Krippner, L. (2013): "Measuring the stance of monetary policy in zero lower bound environments", *Economics Letters*, 118(1), 135–138.

Nelson, C., and A. Siegel (1987): "Parsimonious modeling of yield curves", *Journal of Business*, 60, 473–89.

Roberts, J. M. (2018): "An Estimate of the Long-Term Neutral Rate of Interest", https://www.federalreserve.gov/econres/notes/feds-notes/estimate-of-the-long-term-neutral-rate-of-interest-20180905.htm.