A Principal-Components-based Affine Term-Structure Model with Stochastic Market-Price-of-Risk

Ng Chu Ming

University of Oxford

June 1, 2016

Abstract

We review the mathematics of affine term-structure models with principal components as state variables. Subsequently, we explore an extension of the model incorporating a stochastic market-price-of-risk in state variables that we deem attract a risk compensation.

1 Introduction

Affine Term-Structure Models (ATSMs) are popular among researchers for structural yield curve modelling due to its analytical tractability under multi-factor setups. Using specified variable approaches together with ATSMs, we are able to specify rich dynamics (through time-varying risk-neutral drifts and volatility coefficients) and impute econometric interpretations to the state variables in the model. The resultant closed-form expressions from ATSMs for zero-coupon-bond prices also greatly facilitate their pricing implementations in practice. For the reasons above (and more), ATSMs have been a mainstay in the extensive literature on bond-pricing and interest-rate derivatives.

Principal Components Analysis (PCA) is a well-known technique for dimensionality reduction when working with large, multi-dimensional datasets. Yield-curve time-series data (now widely available) typically contains spot rate data for terms ranging from 1 to 30 years in half-year steps. As such, yield curve modelling is a canonical example demonstrating the high dimensionality reduction made possible with PCA. The resultant principal components lend themselves to a seductive interpretation as 'level', 'slope' and 'curvature' of the yield curve. Furthermore, these three principal components account for a remarkable 99.9% of yield curve variation, making PCA an extremely appealing tool in term-struture modelling.

In ATSMs, the link between the real-world measure and the risk-neutral measure is usually established by the market-price-of-risk. We often make the assumption that the market-price-of-risk is an affine function of the state variables. This means that the market-price-of-risk, while stochastic, remains a *deterministic* function of the *stochastic* state variables.

In this work, we explore an extension of principal-components-based affine term-structure model (PC-based ATSM) where we attempt to specify a stochastic market-price-of-risk that is independent of the principal-component state variables.

2 Preliminaries

In this section, we provide the basic theoretical setup for affine term-structure models, principal components analysis as applied to yield curve modelling, and also the market-price-of-risk specifications in affine models. We mainly follow the terminologies in [5] and [6] and also extend some of their analysis when specifying and solving our new model.

2.1 Notation

In expressions where ambiguity might arise, we denote a vector with a decorating top arrow: $\overrightarrow{\theta}$. Matrices are decorated with an underscore: $\underline{A}$.

2.2 Affine Term-Structure Models

An N-factor affine term structure model assumes an N-dimensional vector of state variables $\overrightarrow{x_t} = (x_1(t), x_2(t), \ldots, x_N(t))^T$ with the following specifications,

1. The instantaneous short rate r_t is given by an affine function of the state variables,
$$r_t = \delta_0 + \sum_{i=1}^{N} \delta_i x_i(t) = \delta_0 + \delta^T \overrightarrow{x_t} \tag{1}$$
where $\delta^T = (\delta_1, \delta_2, \ldots, \delta_N)$ so that $\delta^T \overrightarrow{x_t}$ specifies an inner product.

2. The state variables $\overrightarrow{x_t}$ follow an affine diffusion under $\mathbb{Q}$,
$$d\overrightarrow{x_t} = \underline{K}^{\mathbb{Q}} \left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{Q}}$$
where $\overrightarrow{z_t}^{\mathbb{Q}}$ is an N-dimensional independent Brownian motion. The matrix $\underline{S}$ can be decomposed as $\underline{S} = \Sigma \, \mathrm{diag} \left(\sqrt{\alpha_j + \beta_j^T \overrightarrow{x_t}} \right)$ where α_j is a scalar, β_j is an N-dimensional vector. $\underline{K}^{\mathbb{Q}}$ and Σ are $N \times N$ matrices and $diag(\cdot)$ denotes an $N \times N$ diagonal matrix .

3. The market-price-of-risk takes the form,

$$\lambda_j(t) = \lambda_j \sqrt{\alpha_j + \beta_j^T \overrightarrow{x_t}}$$

4. The state variables $\overrightarrow{x_t}$ also follow an affine diffusion under $\mathbb{P}$,

$$d\overrightarrow{x_t} = \underline{\mathcal{K}}^{\mathbb{P}} \left(\overrightarrow{\theta}^{\mathbb{P}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}}$$

where $\overrightarrow{z_t}^{\mathbb{P}}$ is an N-dimensional independent Brownian motion.

The matrix $\underline{S}$ remains exactly the same as the one under the measure $\mathbb{Q}$. The matrix $\underline{\mathcal{K}}^{\mathbb{P}}$ is related to $\underline{\mathcal{K}}^{\mathbb{Q}}$ by $\underline{\mathcal{K}}^{\mathbb{P}} = \underline{\mathcal{K}}^{\mathbb{Q}} - \Sigma\Phi$, with the j-th row of the matrix Φ given by $\lambda_j \beta_j^T$. The vector $\overrightarrow{\theta}^{\mathbb{P}}$ is given by $\overrightarrow{\theta}^{\mathbb{P}} = (\underline{\mathcal{K}}^{\mathbb{P}})^{-1}(\underline{\mathcal{K}}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \Sigma\psi)$, with the i-th element of the vector ψ given by $\lambda_j \alpha_j$.

From the above, the link between the $\mathbb{P}$ and $\mathbb{Q}$ measures is established through the market-price-of-risk $\lambda_j(t)$, albeit through a rather complicated looking series of vector/matrix operations. In Section 2.4.1, we will show the full details of the derivation.

5. The price of zero coupon bonds is given by,

$$P_t^T = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \right]$$

and with an admissable parameterization, the solution (see [3]) is given by,

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

2.3 Principal Components in Yield Curve Modelling

The yield curve at time-t is an $N \times 1$ vector of yields $\overrightarrow{y_t} = (y_t^{T_1}, y_t^{T_2}, \ldots, y_t^{T_N})^T$. Each element $y_t^{T_i}$ is the yield of the corresponding discount bond with price $P_t^{T_i}$. Following [5], we assume the yields have the dynamics,

$$\overrightarrow{dy_t} = \overrightarrow{\mu_y} dt + \underline{\sigma} \overrightarrow{dw}^{\mathbb{Q},\mathbb{P}}$$

with

$$\underline{\sigma} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \quad \text{and} \quad E\left[\overrightarrow{dw}\overrightarrow{dw}^T\right] = \underline{\rho} dt$$

The exact forms of $\overrightarrow{\mu_y}$ corresponding to the shocks $\overrightarrow{dw}^{\mathbb{Q},\mathbb{P}}$ (representing shocks in $\mathbb{P}$ and $\mathbb{Q}$ depending on the measure we are considering) are inconsequential for our current discussion. This is because we are mainly interested in the covariance structure among the yields,

$$E\left[\overrightarrow{dy}\overrightarrow{dy}^T\right] = \underline{\sigma}\underline{\rho}\underline{\sigma}^T = \Sigma_{\text{mkt}} dt$$

Suppose we have the covariance matrix Σ_{mkt} of the N yields constructed from observing historical times series of the N yields. This exogeneous market observable can be decomposed using eigenvalue decomposition to give,

$$\Sigma_{\mathrm{mkt}} = VDV^T$$

where

$$D = \mathrm{diag}(\xi_1, \xi_2, \dots, \xi_N)$$

is a diagonal matrix of eigenvalues of Σ_{mkt} and V is an orthogonal matrix of eigenvectors of Σ_{mkt},

$$V = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \overrightarrow{v_1} & \overrightarrow{v_2} & \cdots & \overrightarrow{v_N} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

The principal components $\overrightarrow{x_t}$ are given by,

$$\overrightarrow{y_t} = \overrightarrow{u} + V\overrightarrow{x_t}$$

where $\overrightarrow{u}$ is a constant vector.

In order to satisfy the requirements of an affine model, we need the short rate to be an affine function of the state-variables (which are the principal components in this case),

$$r_t = \delta_0 + \delta^T \underbrace{\overrightarrow{x_t}}_{\text{principal components}}$$

2.3.1 Dynamics of Principal Components

With principal components explained in the previous section, we shall now specify their dynamics. We use the ATSM diffusion dynamics,

$$d\overrightarrow{x_t} = \underline{\mathcal{K}}^{\mathbb{Q}} \left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{Q}}$$

where $d\overrightarrow{z_t}^{\mathbb{Q}}$ are independent $\mathbb{Q}$-Brownian increments.

In the above, the matrix $\underline{\mathcal{K}}^{\mathbb{Q}}$ is the reversion speed matrix, the vector $\overrightarrow{\theta}^{\mathbb{Q}}$ is the reversion-level vector, and $\underline{S}$ is the diffusion matrix.

A consequence of using principal components as state variables is that the eigenvalues of Σ_{mkt} correspond to the variance explained by each of the principal components. The eigenvectors are orthogonal, therefore, the matrix $\underline{S}$ has to be a diagonal matrix of the form,

$$\underline{S} = \mathrm{diag}(\sqrt{\xi_1}, \sqrt{\xi_2}, \dots, \sqrt{\xi_N})$$

where ξ_i are the eigenvalues of the covariance matrix Σ_{mkt}.

When we specify $\underline{S}$ to be diagonal, one consequence is that it will be impossible for the matrix $\underline{\mathcal{K}}^{\mathbb{Q}}$ to also be diagonal (see the proof in [7]). This means that we will not be able to have intuitive mean-reversion dynamics where each $x_i(t)$ reverts to its own reversion level θ_i in the drift terms in the diffusion dynamics for $dx_i(t)$.

2.4 (Affine) Market Price of Risk

In the usual affine setting, the market-price-of-risk is assumed to be an affine function of the state variables,

$$\overrightarrow{\lambda_t} = \overrightarrow{\lambda_0} + \underline{\Pi}\overrightarrow{x_t}$$

where $\overrightarrow{\lambda_0}$ is a constant vector.

This definition means that the market-price-of-risk is a *deterministic* function of the *stochastic* state variables (principal components). Once the stochastic principal components take on a certain value, the market-price-of-risk assumes the same corresponding fixed value. It is important to note that the market-price-of-risk does not have its own separate and independent stochasticity.

For example, assuming we have $N = 3$ and we use the first three principal components (of 'level', 'slope', 'curvature') as state variables, then $\underline{\Pi}$ will be a 3×3 matrix. If the market-price-of-risk depends only on the slope of the yield curve, then $\underline{\Pi}$ will have the following form,

$$\underline{\Pi} = \begin{bmatrix} 0 & a & 0 \\ 0 & b & 0 \\ 0 & c & 0 \end{bmatrix}$$

so that

$$\underline{\Pi}\overrightarrow{x_t} = \begin{bmatrix} 0 & a & 0 \\ 0 & b & 0 \\ 0 & c & 0 \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} ax_2(t) \\ bx_2(t) \\ cx_2(t) \end{pmatrix}$$

As we can see, each of the $x_i(t)$ in $\overrightarrow{x_t}$ will have their market-price-of-risk as some factor times the value of the state variable $x_2(t)$ (representing the slope).

What we have specified above is that investors require compensation for bearing 'level', 'slope' and 'curvature' risk (corresponding to the principal component state variables $x_1(t)$, $x_2(t)$, $x_3(t)$ respectively). In addition, the amount of risk compensation depends solely on the 'slope' (the $x_2(t)$ state variable).

If investors only require compensation for bearing 'level' risk, then $\underline{\Pi}$ would take the form,

$$\underline{\Pi} = \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so that

$$\underline{\Pi}\overrightarrow{x_t} = \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} ax_2(t) \\ 0 \\ 0 \end{pmatrix}$$

Given this, when we adjust the drifts of $\overrightarrow{x_t}$ using $\underline{\Pi}\overrightarrow{x_t}$, only the element $x_1(t)$ in $\overrightarrow{x_t}$ will get an adjustment related to the 'slope' variable $x_2(t)$. This is what we mean when we say that investors are rewarded for 'level' risk *depending* on the 'slope' of the yield curve.

2.4.1 Moving between $\mathbb{P}$ and $\mathbb{Q}$ Measures

As mentioned earlier, the market-price-of-risk is the link that allows us to move between the $\mathbb{P}$ and $\mathbb{Q}$ measures, while preserving the affine structure in the dynamics of the state variables.

We begin by considering the $\mathbb{Q}$-measure dynamics (as stated in the affine model specification),

$$d\overrightarrow{x_t} = \underline{\mathcal{K}}^{\mathbb{Q}} \left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{Q}}$$

If we introduce the compensation for risk via the market-price-of-risk into the above, we get,

$$\begin{aligned}
d\overrightarrow{x_t} &= \underline{\mathcal{K}}^{\mathbb{Q}} \left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}} + \underline{S}\overrightarrow{\lambda_t} dt \\
&= \underline{\mathcal{K}}^{\mathbb{Q}} \left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}} + \underline{S}(\overrightarrow{\lambda_0} + \underline{\Pi}\overrightarrow{x_t}) dt \\
&= \left(\underline{\mathcal{K}}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \underline{S}\overrightarrow{\lambda_0} \right) dt - \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \right) \overrightarrow{x_t} dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}} \\
&= \left[\left(\underline{\mathcal{K}}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \underline{S}\overrightarrow{\lambda_0} \right) - \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \right) \overrightarrow{x_t} \right] dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}} \\
&= \underbrace{\left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \right)}_{\underline{\mathcal{K}}^{\mathbb{P}}} \left[\underbrace{\left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \right)^{-1} \left(\underline{\mathcal{K}}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \underline{S}\overrightarrow{\lambda_0} \right)}_{\overrightarrow{\theta}^{\mathbb{P}}} - \overrightarrow{x_t} \right] dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}} \\
&= \underline{\mathcal{K}}^{\mathbb{P}} \left(\overrightarrow{\theta}^{\mathbb{P}} - \overrightarrow{x_t} \right) dt + \underline{S} d\overrightarrow{z_t}^{\mathbb{P}}
\end{aligned}$$

where we have the relations,

$$\underline{\mathcal{K}}^{\mathbb{P}} = \underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \quad \text{and} \quad \overrightarrow{\theta}^{\mathbb{P}} = \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi} \right)^{-1} \left(\underline{\mathcal{K}}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \underline{S}\overrightarrow{\lambda_0} \right) \tag{2}$$

3 Extending a PC-based ATSM with Stochastic Market-Price-of-Risk

With the theoretical foundations set in the previous section, we shall examine how a principal-components-based ATSM can be extended to feature a stochastic market-price-of-risk.

3.1 Principal-Components-based Affine Model

In our model, we choose to use $N = 3$ principal components as our state variables, so that

$$\overrightarrow{x_t} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

where $x_1(t)$, $x_2(t)$, $x_3(t)$ correspond to the first three principal-components of 'level', 'slope' and 'curvature' respectively.

The $\mathbb{Q}$ dynamics for the model is,

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix}}_{\underline{K}^{\mathbb{Q}}} \underbrace{\left(\begin{bmatrix} \theta_1^{\mathbb{Q}} \\ \theta_2^{\mathbb{Q}} \\ \theta_3^{\mathbb{Q}} \end{bmatrix} - \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \right)}_{\left(\vec{\theta}^{\mathbb{Q}} - \vec{x}_t\right)} dt + \underbrace{\begin{bmatrix} \sqrt{\xi_1} & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 \\ 0 & 0 & \sqrt{\xi_3} \end{bmatrix}}_{\underline{S}} \begin{bmatrix} dW_1^{\mathbb{Q}}(t) \\ dW_2^{\mathbb{Q}}(t) \\ dW_3^{\mathbb{Q}}(t) \end{bmatrix} \tag{3}$$

where ξ_i are the eigenvalues of the covariance matrix Σ_{mkt}. Since the eigenvalues correspond to the variance explained in each direction of the principal components, we can see that the matrix $\underline{S}$ above correctly produces the required variance in each of the Brownian-motions $dW_1^{\mathbb{Q}}(t), dW_2^{\mathbb{Q}}(t), dW_3^{\mathbb{Q}}(t)$.

3.2 Stochastic Market-Price-of-Risk

In Section 2.4, we showed the specification of an affine market-price-of-risk where it is a deterministic (affine) function of the stochastic state variables.

In our new model, we choose to adopt a stochastic market-price-of-risk model where it is driven by its own Brownian shock. To model our stochastic market-price-of-risk, we posit that we have another *latent* state variable $\lambda(t)$ with the dynamics,

$$d\lambda(t) = \kappa_\lambda(\theta_\lambda - \lambda(t))dt + \sigma_\lambda dB^{\mathbb{Q}}(t)$$

where θ_λ is the reversion-level, κ_λ is the reversion-speed, and σ_λ is the volatility of the market-price-of-risk. The driving Brownian motion $dB^{\mathbb{Q}}(t)$ is independent of the Brownian motions $dW_i^{\mathbb{Q}}(t)$ in (3).

We assume that investors do not require a risk compensation for the stochasticity in the above market-price-of-risk specification. Hence, moving forward, we should be aware that the drift term for $d\lambda(t)$ is the same in both the $\mathbb{P}$ and $\mathbb{Q}$ measures.

Suppose (following results in [2]) investors are only compensated for 'level' risk, the stochastic market-price-of-risk should then apply only to the 'level' state-variable $x_1(t)$, with 'dependence' on the 'slope' state variable $x_2(t)$. Following the affine market-price-of-risk approach in Section 2.4, we thus modify the dynamics for $x_1(t)$ in $\mathbb{P}$ to be,

$$\begin{aligned} dx_1(t) = [\kappa_{11}(\theta_1 - x_1(t)) + \kappa_{12}(\theta_2 - x_2(t)) + \kappa_{13}(\theta_3 - x_3(t))]dt \\ + \lambda(t)\sqrt{\xi_1}x_2(t)dt + \sqrt{\xi_1}dW_1^{\mathbb{P}}(t) \end{aligned}$$

However, the explicit dependence on $x_2(t)$ above complicates the drift dynamics and is undesirable (see Chapter 27 of [5]). To solve this, we change the above term $\lambda(t)\sqrt{\xi_1}x_2(t)dt$ to $\lambda(t)\sqrt{\xi_1}dt$, removing the explicit dependence on $x_2(t)$. By doing so, we break the link between the 'slope' return-predicting factor and the market-price-of-risk, which is something we do not want.

To reconcile the issue, we re-introduce the 'dependence' via a positive statistical correlation between the Brownian shock driving the 'slope' principal component $x_2(t)$ and the Brownian shock driving $\lambda(t)$. With the above adjustments, the dynamics for $x_1(t)$ in $\mathbb{P}$ now becomes,

$$
\begin{aligned}
dx_1(t) = [\kappa_{11}(\theta_1 - x_1(t)) + \kappa_{12}(\theta_2 - x_2(t)) + \kappa_{13}(\theta_3 - x_3(t))]dt \\
+\lambda(t)\sqrt{\xi_1}dt + \sqrt{\xi_1}dW_1^{\mathbb{P}}(t)
\end{aligned}
$$

and the dynamics for the stochastic market-price-of-risk is modified to become,

$$
d\lambda(t) = \kappa_\lambda(\theta_\lambda - \lambda(t))dt + \sigma_\lambda\left(\rho dW_2^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dB^{\mathbb{Q}}(t)\right)
$$

where $dW_2^{\mathbb{Q}}(t)$ and $dB^{\mathbb{Q}}(t)$ are independent Brownian motions.

So $\lambda(t)$ is now a correlated diffusion process with $x_2(t)$ - the 'slope' principal component - and we are able to generate a positive 'dependence' between the slope of the yield curve and the market price of risk *on average*.

3.3 The Combined Model

Combining the PC-based Affine Model in Section 3.1 with the Stochastic Market-Price-of-Risk in Section 3.2, we have an Affine $\mathbb{A}_0(4)$ model (see [1] for affine model classification). The dynamics of the model in $\mathbb{P}$ is given below,

$$
\begin{aligned}
\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \\ d\lambda(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & -\sqrt{\xi_1} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}}_{\underline{\mathcal{K}}^{\mathbb{P}}} \underbrace{\left(\begin{bmatrix} \theta_1 + \frac{\sqrt{\xi_1}\theta_\lambda}{\kappa_{11}} \\ \theta_2 \\ \theta_3 \\ \theta_\lambda \end{bmatrix} - \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \lambda(t) \end{bmatrix} \right)}_{\left(\vec{\theta}^{\mathbb{P}} - \vec{x_t}\right)} dt \\
&+ \underbrace{\begin{bmatrix} \sqrt{\xi_1} & 0 & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 & 0 \\ 0 & 0 & \sqrt{\xi_3} & 0 \\ 0 & \sigma_\lambda\rho & 0 & \sigma_\lambda\sqrt{1-\rho^2} \end{bmatrix}}_{\underline{S}} \underbrace{\begin{bmatrix} dW_1^{\mathbb{P}}(t) \\ dW_2^{\mathbb{P}}(t) \\ dW_3^{\mathbb{P}}(t) \\ dB^{\mathbb{P}}(t) \end{bmatrix}}_{d\vec{z_t}^{\mathbb{P}}}
\end{aligned}
$$

Using the relations in (2) we have,

$$
\underline{\mathcal{K}}^{\mathbb{P}} = \underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\Pi
$$

$$
\begin{aligned}
\Rightarrow \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & -\sqrt{\xi_1} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix} &= \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix} \\
&- \begin{bmatrix} \sqrt{\xi_1} & 0 & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 & 0 \\ 0 & 0 & \sqrt{\xi_3} & 0 \\ 0 & \sigma_\lambda\rho & 0 & \sigma_\lambda\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

where we specified the reversion-speed matrix $\underline{\mathcal{K}}^{\mathbb{Q}}$ to have the nice form,

$$\underline{\mathcal{K}}^{\mathbb{Q}} = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}$$

The other relation in (2) provides a specification for $\overrightarrow{\theta}^{\mathbb{Q}}$,

$$\overrightarrow{\theta}^{\mathbb{P}} = \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi}\right)^{-1} \left(\underline{\mathcal{K}}^{\mathbb{Q}}\overrightarrow{\theta}^{\mathbb{Q}} + \cancelto{\emptyset}{\underline{S}\overrightarrow{\lambda_0}}\right)$$

$$\Rightarrow \overrightarrow{\theta}^{\mathbb{Q}} = \left(\underline{\mathcal{K}}^{\mathbb{Q}}\right)^{-1} \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi}\right) \overrightarrow{\theta}^{\mathbb{P}} = \left(\underline{\mathcal{K}}^{\mathbb{Q}}\right)^{-1} \underline{\mathcal{K}}^{\mathbb{P}}\overrightarrow{\theta}^{\mathbb{P}}$$

$$= \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}^{-1} \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & -\sqrt{\xi_1} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix} \begin{bmatrix} \theta_1 + \frac{\sqrt{\xi_1}\theta_\lambda}{\kappa_{11}} \\ \theta_2 \\ \theta_3 \\ \theta_\lambda \end{bmatrix}$$

Combining everything, we can specify the $\mathbb{Q}$ dynamics for our model as,

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \\ d\lambda(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}}_{\underline{\mathcal{K}}^{\mathbb{Q}}} \underbrace{\left(\overrightarrow{\theta}^{\mathbb{Q}} - \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \lambda(t) \end{bmatrix}\right)}_{\left(\overrightarrow{\theta}^{\mathbb{Q}} - \overrightarrow{x_t}\right)} dt$$

$$+ \underbrace{\begin{bmatrix} \sqrt{\xi_1} & 0 & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 & 0 \\ 0 & 0 & \sqrt{\xi_3} & 0 \\ 0 & \sigma_\lambda\rho & 0 & \sigma_\lambda\sqrt{1-\rho^2} \end{bmatrix}}_{\underline{S}} \underbrace{\begin{bmatrix} dW_1^{\mathbb{Q}}(t) \\ dW_2^{\mathbb{Q}}(t) \\ dW_3^{\mathbb{Q}}(t) \\ dB^{\mathbb{Q}}(t) \end{bmatrix}}_{d\overrightarrow{z_t}^{\mathbb{Q}}}$$

3.3.1 Solving the Combined Model

To solve the combined model in Section 3.3, we use the general framework for solving affine models in [5].

In our setup, our state variable is $\overrightarrow{x_t} = (x_1(t), x_2(t), x_3(t), \lambda(t))^T$ where we have a vector of $N = 3$ *key maturity yields* plus a 'dummy' yield corresponding to the market-price-of-risk *latent* state variable $\lambda(t)$. As such, we have $\overrightarrow{y_t} = (y_t^{T_1}, y_t^{T_2}, y_t^{T_3}, y_t^{\lambda})^T$. This is basically the class of semi-observable affine factor models $\mathbb{A}_0(4)$ described in [7].

The affine yields for our combined model are given by,

$$\overrightarrow{y_t} = \overrightarrow{c} + U\overrightarrow{x_t} \tag{4}$$

where $\overrightarrow{c}$ is a constant vector and U is a 4×4 matrix given by

$$U = \begin{bmatrix} v_{11} & v_{12} & v_{13} & 0 \\ v_{21} & v_{22} & v_{23} & 0 \\ v_{31} & v_{32} & v_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & \uparrow & 0 \\ \overrightarrow{v_1} & \overrightarrow{v_2} & \overrightarrow{v_3} & \vdots \\ \downarrow & \downarrow & \downarrow & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_2}$ as the first three eigenvectors in the matrix V from the eigen-decomposition of the market observable covariance matrix $\Sigma_{\text{mkt}} = VDV^T$.

Expanded out in full, the yields are given by,

$$\begin{bmatrix} y_t^{T_1} \\ y_t^{T_2} \\ y_t^{T_3} \\ y_t^{\lambda} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} & v_{13} & 0 \\ v_{21} & v_{22} & v_{23} & 0 \\ v_{31} & v_{32} & v_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \lambda(t) \end{bmatrix}$$

If we only consider the key maturity yields $\overrightarrow{y}_{1\ldots3} = (y_t^{T_1}, y_t^{T_2}, y_t^{T_3})^T$, they satisfy,

$$\begin{aligned} \begin{bmatrix} y_t^{T_1} \\ y_t^{T_2} \\ y_t^{T_3} \end{bmatrix} &= \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \\ \Rightarrow \overrightarrow{y}_{1\ldots3} &= \overrightarrow{c}_{1\ldots3} + V\,\overrightarrow{x}_{1\ldots3} \end{aligned} \tag{5}$$

and we are back to the same setup as a principal-components-only model. All the consistency contraints on V mentioned in [7] similarly apply.

We now specify that the first element of $\overrightarrow{y_t}$ is the short rate r_t. This is to say that,

$$y_t^{T_1} = r_t$$

where we let $T_1 \to 0$ and assume an overnight deposit rate (such as the Fed Funds rate or the EONIA index) is a suitable proxy for the short rate (which is strictly speaking unobservable).

Given the above assumption, we have,

$$\begin{aligned} r_t &= \overrightarrow{e}_1^T \overrightarrow{y_t} \\ &= \underbrace{\overrightarrow{e}_1^T \overrightarrow{c}}_{w_0} + \underbrace{\overrightarrow{e}_1^T U}_{\overrightarrow{w}_1^T} \overrightarrow{x_t} \\ &= w_0 + \overrightarrow{w}_1^T \overrightarrow{x_t} \end{aligned} \tag{6}$$

where $\overrightarrow{e_1} = (1, 0, \ldots, 0)^T$. This is the same form as required by (1) except that we used w instead of δ in the naming of variables.

The 1×4 vector $\overrightarrow{w}_1^T$ is given by,

$$
\begin{aligned}
\overrightarrow{w}_1^T &= \underbrace{\overrightarrow{e}_1^T}_{1\times 4}\underbrace{U}_{4\times 4} \\
&= [v_{11}, v_{12}, v_{13}, 0] \\
\Rightarrow \overrightarrow{w}_1 &= \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ 0 \end{bmatrix}
\end{aligned}
$$

With the definition of the short rate specified above, we know from no-arbitrage that the price of the discount bond must be given by,

$$
\begin{aligned}
P_t^T &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T (w_0 + \overrightarrow{w}_1^T \overrightarrow{x_t}) ds}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T w_0 + v_{11}x_1(t) + v_{12}x_2(t) + v_{13}x_3(t)\, ds\right)\right]
\end{aligned} \tag{7}
$$

This shows our setup correctly produces a price formula which does not involve the latent state variable $\lambda(t)$.

In what follows, all mentions of $\underline{\mathcal{K}}$ and $\overrightarrow{\theta}$ refer to the $\mathbb{Q}$ measure reversion-speed matrix and reversion-level vector.

From [1], the general solution to (7) is given by,

$$
P_t^T = \exp^{A_t^T + \left(\overrightarrow{B}_t^T\right)^T \overrightarrow{x_t}}
$$

with the vector $\overrightarrow{B}_t^T$ (which we abbreviate $\overrightarrow{B}_\tau$) and the scalar A_t^T (abbreviated A_τ) satisfying the ODEs (with $\tau = T - t$) below,

$$
\begin{aligned}
\frac{dA_\tau}{d\tau} &= -w_0 + \left(\overrightarrow{B}_\tau\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta} + \frac{1}{2}\left(\overrightarrow{B}_\tau\right)^T \underline{S}\,\underline{S}^T \overrightarrow{B}_\tau \\
\frac{d\overrightarrow{B}_\tau}{d\tau} &= -\overrightarrow{w}_1 - \underline{\mathcal{K}}^T \overrightarrow{B}_\tau \\
\overrightarrow{B}(\tau = 0) = 0, &\quad A(\tau = 0) = 0 \qquad \text{(boundary conditions)}
\end{aligned}
$$

The solution for $\overrightarrow{B}_\tau$ (and hence $\left(\overrightarrow{B}_\tau\right)^T$) is given by,

$$
\begin{aligned}
\overrightarrow{B}_\tau &= \left(e^{-\underline{\mathcal{K}}^T \tau} - I_n\right)\left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 \\
\Rightarrow \left(\overrightarrow{B}_\tau\right)^T &= \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\left[e^{-\underline{\mathcal{K}}\tau} - I_n\right]
\end{aligned}
$$

The detailed expansions for $\overrightarrow{B}_\tau$ and $\left(\overrightarrow{B}_\tau\right)^T$ for our setup is shown below,

$$\begin{aligned}
\overrightarrow{B}_\tau &= \left(e^{-\underline{\mathcal{K}}^T\tau} - I_n\right) \begin{bmatrix} \kappa_{11} & \kappa_{21} & \kappa_{31} & 0 \\ \kappa_{12} & \kappa_{22} & \kappa_{32} & 0 \\ \kappa_{13} & \kappa_{23} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}^{-1} \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ 0 \end{bmatrix} \\
\left(\overrightarrow{B}_\tau\right)^T &= \begin{bmatrix} v_{11} & v_{12} & v_{13} & 0 \end{bmatrix} \begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}^{-1} \left(e^{-\underline{\mathcal{K}}\tau} - I_n\right)
\end{aligned} \tag{8}$$

where $e^{-\underline{\mathcal{K}}\tau}$ and $e^{-\underline{\mathcal{K}}^T\tau}$ are matrix exponentials involving the matrix $\underline{\mathcal{K}}$ and its transpose.

To simplify the above expressions for $\overrightarrow{B}_\tau$ and $\left(\overrightarrow{B}_\tau\right)^T$, we assume that $\underline{\mathcal{K}}$ is diagonalizable, i.e. it has $N = 4$ distinct and real eigenvalues $\{l_i\}$. With that, we can diagonalize the reversion-speed matrix $\underline{\mathcal{K}}$ to obtain,

$$\begin{aligned}
&\underline{\mathcal{K}} = \underline{a}\,\Lambda\,\underline{a}^{-1} \\
\Rightarrow\ &\underline{\mathcal{K}}^T = \left(\underline{a}^{-1}\right)^T \Lambda\,\underline{a}^T = \underbrace{\left(\underline{a}^T\right)^{-1}}_{\underline{b}}\ \Lambda\ \underbrace{\underline{a}^T}_{\underline{b}^{-1}} = \underline{b}\,\Lambda\,\underline{b}^{-1}
\end{aligned}$$

with

$$\underline{\Lambda} = \mathrm{diag}(l_1, l_2, l_3, l_4)$$

(Note : To ensure stability, we require all the eigenvalues l_i to be positive.)

Using this assumption that $\underline{\mathcal{K}}$ can be written as $\underline{a}\,\underline{\Lambda}\,\underline{a}^{-1}$, we can write the matrix exponential $e^{-\underline{\mathcal{K}}^T\tau}$ as,

$$\begin{aligned}
e^{-\underline{\mathcal{K}}^T\tau} &= e^{-\underline{b}\underline{\Lambda}\underline{b}^{-1}\tau} \\
&= e^{-\underline{b}\underline{\Lambda}\tau\underline{b}^{-1}} \\
&= \underline{b}e^{-\underline{\Lambda}\tau}\underline{b}^{-1} \quad \text{(by up-and-down theorem)} \\
&= \underline{b}\left(e^{\underline{\Lambda}\tau}\right)^{-1}\underline{b}^{-1} \\
&= \underline{b}\begin{bmatrix} e^{-l_1\tau} & 0 & 0 & 0 \\ & e^{-l_2\tau} & 0 & 0 \\ 0 & 0 & e^{-l_3\tau} & 0 \\ 0 & 0 & 0 & e^{-l_4\tau} \end{bmatrix}\underline{b}^{-1} \\
&= \underline{b}\,\mathrm{diag}(e^{-l_i\tau})\,\underline{b}^{-1}
\end{aligned}$$

Substituting the above expression for $e^{-\underline{\mathcal{K}}^T\tau}$ into the expression for $\overrightarrow{B}_\tau$, we get,

$$\begin{aligned}
\overrightarrow{B}_\tau &= \left(e^{-\underline{K}^T\tau} - I_n\right)\left(\underline{K}^T\right)^{-1}\overrightarrow{w}_1 \\
&= \left(\underline{b}\,\mathrm{diag}(e^{-l_i\tau})\,\underline{b}^{-1} - I_n\right)\left(\underline{b\Lambda b}^{-1}\right)^{-1}\overrightarrow{w}_1 \\
&= \underline{b}\left(\mathrm{diag}(e^{-l_i\tau})\,\underline{b}^{-1} - \underline{b}^{-1}I_n\right)\left(\underline{b\Lambda}^{-1}\underline{b}^{-1}\right)\overrightarrow{w}_1 \quad (\text{since } \left(\underline{b\Lambda b}^{-1}\right)^{-1} = \underline{b\Lambda}^{-1}\underline{b}^{-1}) \\
&= \underline{b}\left(\mathrm{diag}(e^{-l_i\tau}) - \underline{b}^{-1}I_n\underline{b}\right)\underline{b}^{-1}\underline{b\Lambda}^{-1}\underline{b}^{-1}\overrightarrow{w}_1 \\
&= \underline{b}\left(\mathrm{diag}(e^{-l_i\tau}) - I_n\right)\underline{\Lambda}^{-1}\underline{b}^{-1}\overrightarrow{w}_1 \\
&= \underline{b}\,\mathrm{diag}\left[\frac{e^{-l_i\tau} - 1}{l_i}\right]\underline{b}^{-1}\overrightarrow{w}_1
\end{aligned}$$

With this, we finally see that $\overrightarrow{B}_\tau$ takes the following form,

$$\overrightarrow{B}_\tau = -\underline{b}\,\mathrm{diag}\left[\frac{1 - e^{-l_j\tau}}{l_j}\right]\underline{b}^{-1}\overrightarrow{w}_1$$

Once the vector $\overrightarrow{B}_\tau$ has been computed, the scalar A_τ can be obtained by integration. The full derivation for A_τ is given in Appendix A and the final expression is as follows,

$$\begin{aligned}
A_\tau &= \int_0^\tau \left[-w_0 + \left(\overrightarrow{B}_s\right)^T\underline{K}\overrightarrow{\theta} + \frac{1}{2}\left(\overrightarrow{B}_s\right)^T\underline{S}\,\underline{S}^T\overrightarrow{B}_s\right]ds \\
&= -w_0\tau \\
&\quad + \overrightarrow{w}_1^T\left(\underline{a}\,\mathrm{diag}\left[\frac{1 - e^{-l_i\tau}}{l_i}\right]\underline{a}^{-1}\overrightarrow{\theta} - \overrightarrow{\theta}\tau\right) \\
&\quad + \frac{1}{2}\overrightarrow{w}_1^T\underline{K}^{-1}\underline{a}\,F_s\underline{\Lambda}^{-1}\underline{a}^{-1}\overrightarrow{w}_1 \\
&\quad - \frac{1}{2}\overrightarrow{w}_1^T\underline{K}^{-1}C\underline{a}\,\mathrm{diag}\left[\frac{1 - e^{-l_i\tau}}{l_i}\right]\underline{\Lambda}^{-1}\underline{a}^{-1}\overrightarrow{w}_1 \\
&\quad - \frac{1}{2}\overrightarrow{w}_1^T\underline{K}^{-1}\underline{a}\mathrm{diag}\left[\frac{1 - e^{-l_i\tau}}{l_i}\right]\underline{a}^{-1}C\left(\underline{K}^T\right)^{-1}\overrightarrow{w}_1 \\
&\quad + \frac{1}{2}\overrightarrow{w}_1^T\underline{K}^{-1}C\left(\underline{K}^T\right)^{-1}\overrightarrow{w}_1\tau
\end{aligned}$$

4 Concise Summary of our Combined Model

To recap the results for our combined model, the following is a concise one-page summary,

1. The dynamics of the model in $\mathbb{Q}$ is as follows,

$$
\begin{aligned}
\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \\ d\lambda(t) \end{bmatrix} &= \underbrace{\begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_\lambda \end{bmatrix}}_{\underline{\mathcal{K}}} \underbrace{\left(\vec{\theta} - \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \lambda(t) \end{bmatrix} \right)}_{\left(\vec{\theta} - \vec{x_t}\right)} dt \\
&+ \underbrace{\begin{bmatrix} \sqrt{\xi_1} & 0 & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 & 0 \\ 0 & 0 & \sqrt{\xi_3} & 0 \\ 0 & \sigma_\lambda \rho & 0 & \sigma_\lambda \sqrt{1-\rho^2} \end{bmatrix}}_{\underline{S}} \underbrace{\begin{bmatrix} dW_1^{\mathbb{Q}}(t) \\ dW_2^{\mathbb{Q}}(t) \\ dW_3^{\mathbb{Q}}(t) \\ dB^{\mathbb{Q}}(t) \end{bmatrix}}_{d\vec{z_t}^{\mathbb{Q}}}
\end{aligned}
\tag{9}
$$

with the assumption that $\underline{\mathcal{K}}$ admits an eigen-decomposition as,

$$
\underline{\mathcal{K}} = \underline{a}\,\underline{\Lambda}\,\underline{a}^{-1} \Rightarrow \underline{\mathcal{K}}^T = \left(\underline{a}^{-1}\right)^T \underline{\Lambda}\,\underline{a}^T = \underbrace{\left(\underline{a}^T\right)^{-1}}_{\underline{b}} \underline{\Lambda} \underbrace{\underline{a}^T}_{\underline{b}^{-1}} = \underline{b}\,\underline{\Lambda}\,\underline{b}^{-1}
$$

and $\underline{\Lambda} = \mathrm{diag}(l_1, l_2, l_3, l_4)$ where l_i are the eigenvalues.

2. The solution to the combined model is given by,

$$
P_t^T = \exp^{A_t^T + \left(\vec{B}_t^T\right)^T \vec{x_t}}
$$

where the vector $\vec{B}_t^T \equiv \vec{B}_\tau$ and the scalar $A_t^T \equiv A_\tau$ (with $\tau = T - t$) are given by,

$$
\vec{B}_\tau = -\underline{b}\,\mathrm{diag}\left[\frac{1-e^{-l_j \tau}}{l_j}\right] \underline{b}^{-1} \vec{w}_1
$$

$$
\begin{aligned}
A_\tau = &-w_0\tau + \vec{w}_1^T \left(\underline{a}\,\mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{a}^{-1}\vec{\theta} - \vec{\theta}\,\tau \right) + \frac{1}{2}\vec{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a}\, F_\tau \underline{\Lambda}^{-1} \underline{a}^{-1} \vec{w}_1 \\
&- \frac{1}{2}\vec{w}_1^T \underline{\mathcal{K}}^{-1} C \underline{a}\,\mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{\Lambda}^{-1} \underline{a}^{-1} \vec{w}_1 + \frac{1}{2}\vec{w}_1^T \underline{\mathcal{K}}^{-1} C \left(\underline{\mathcal{K}}^T\right)^{-1} \vec{w}_1 \tau \\
&- \frac{1}{2}\vec{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a}\mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{a}^{-1} C \left(\underline{\mathcal{K}}^T\right)^{-1} \vec{w}_1
\end{aligned}
$$

with w_0 and $\vec{w}_1^T$ defined in (6). C is defined as $C = \underline{S}\,\underline{S}^T$. The matrix F_τ is given by $F_\tau = \left[\int_0^\tau e^{-\underline{\Lambda} s} \underline{a}^{-1} C \underline{a} e^{-\underline{\Lambda} s} ds\right]$.

5 Calibration of the Combined Model

Having presented the solution for our combined model, we proceed to discuss how to calibrate our new model. To obtain the data needed for calibration, we can use the US Treasury and Fed Funds rate data available for download in [4].

5.1 Choosing Eigenvalues for the Reversion-Speed Matrix

The analysis for choosing eigenvalues of the reversion-speed matrix is the same as that in [6].

We recall from the $\mathbb{Q}$ dynamics of our model in (9) that there is no interaction/dependence between the parameters driving the principal-component state variables $(x_1(t), x_2(t), x_3(t))$ and the parameters driving $\lambda(t)$.

If we drop the $\lambda(t)$ from consideration for now, we end up with exactly the same setup as in [6], which is,

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \\ dx_3(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix}}_{\underline{\mathcal{K}}_{\mathrm{PC}}} \underbrace{\left(\overrightarrow{\theta}_{\mathrm{PC}} - \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \right)}_{\left(\overrightarrow{\theta}_{\mathrm{PC}} - \overrightarrow{x_t}\right)} dt + \underbrace{\begin{bmatrix} \sqrt{\xi_1} & 0 & 0 \\ 0 & \sqrt{\xi_2} & 0 \\ 0 & 0 & \sqrt{\xi_3} \end{bmatrix}}_{\underline{S}_{\mathrm{PC}}} \underbrace{\begin{bmatrix} dW_1^{\mathbb{Q}}(t) \\ dW_2^{\mathbb{Q}}(t) \\ dW_3^{\mathbb{Q}}(t) \end{bmatrix}}_{d\overrightarrow{z_t}^{\mathbb{Q}}}$$

where in the above, we abbreviated the symbols in the system with 'PC' to avoid confusion. (Note : the values for $\sqrt{\xi_1}, \sqrt{\xi_2}, \sqrt{\xi_3}$ will come from PCA, and are the variances captured in each of the principal-component directions).

We also recall from (5) that the key maturity yields are given by,

$$\begin{aligned} \begin{bmatrix} y_t^{T_1} \\ y_t^{T_2} \\ y_t^{T_3} \end{bmatrix} &= \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \\ \Rightarrow \overrightarrow{y}_{\mathrm{PC}} &= \overrightarrow{c}_{\mathrm{PC}} + V \overrightarrow{x}_{\mathrm{PC}} \end{aligned}$$

This put us exactly in the same setup as [6], and we proceed with the calibration of our model to obtain the entries of $\underline{\mathcal{K}}_{\mathrm{PC}}$ in the same way as described in the paper.

To summarize, the problem setup is such that the only choice we need to make in the calibration of our model is to choose the eigenvalues $\{l_j\}$ of the matrix $\underline{\mathcal{K}}_{\mathrm{PC}}$. Assuming that $\underline{\mathcal{K}}_{\mathrm{PC}}$ is diagonalizable to $\underline{\mathcal{K}}_{\mathrm{PC}} = \underline{\phi}\underline{\Lambda}_{\mathcal{K}_{\mathrm{PC}}}\underline{\phi}^{-1}$ with $\underline{\Lambda}_{\mathcal{K}_{\mathrm{PC}}}$ the diagonal matrix of eigenvalues $\{l_j\}$, then it is shown in [6] that $\underline{\mathcal{K}}_{\mathrm{PC}}$ is given by,

$$\underline{\mathcal{K}}_{\mathrm{PC}} = V^T F^{-1} \underline{\Lambda}_{\mathcal{K}_{\mathrm{PC}}} F V$$

with F the 3×3 matrix with elements $[F]_{ij}$ given by,

$$F_{ij} = \frac{1}{\tau_j} \frac{1 - e^{-l_i \tau_j}}{l_i}$$

The 3-factor principal component setup can exactly recover three key maturity yields which are exogeneous market observables. However, the non-reference yields will not be exactly recovered, and the choice of our eigenvalues $\{l_j\}$ will be to minimize the discrepancy of the covariance matrix and yield curve that is recovered for these non-key maturity yields.

5.2 Calibrating the Stochastic Market-Price-of-Risk

Recall that our stochastic market-price-of-risk state variable has the following dynamics,

$$d\lambda(t) = \kappa_\lambda(\theta_\lambda - \lambda(t))dt + \sigma_\lambda \left(\rho dW_2^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dB^{\mathbb{Q}}(t)\right)$$

Our calibration process involves finding values for $\kappa_\lambda, \theta_\lambda, \sigma_\lambda$ and ρ. We follow closely the analysis in Chapter 8 of [5], whereby the market-price-of-risk is inferred from the excess returns from the strategy of being consistently long a T-maturity bond, and funding it with a very-short-maturity bond.

For example, we can look at the strategy of being long a 10-yr bond and funding it with a 1-yr bond. Using this setup, we will show that the market-price-of-risk is closely related to the Sharpe Ratio of the above strategy.

We begin the analysis by assuming that we have the following bond price dynamics,

$$\frac{dP_t^T}{P_t^T} = \mu^{\mathbb{P}}dt + \sigma^{\mathbb{P}}dW^{\mathbb{P}}(t)$$

We know from (7) that our bond price is a function of the state variables $x_1(t), x_2(t), x_3(t)$. However, since we only assign a market-price-of-risk to the state variable $x_1(t)$ by construction, we will only consider the drift terms involving $x_1(t)$ in the workings below, where we analyse the difference in drifts of the bond between the $\mathbb{P}$ and $\mathbb{Q}$ measures. As the state variables $x_2(t)$ and $x_3(t)$ do not have a market-price-of-risk, their drifts would be the same in both measures and would cancel.

Using Ito's Lemma on the bond price function (and only considering terms involving $x_1(t)$), we have,

$$\begin{aligned}
dP_t^T &= \frac{\partial P_t^T}{\partial t}dt + \frac{\partial P_t^T}{\partial x_1}dx_1(t) + \frac{1}{2}\frac{\partial^2 P_t^T}{\partial x_1^2}d\langle x_1\rangle_t \\
&= \frac{\partial P_t^T}{\partial t}dt + \frac{\partial P_t^T}{\partial x_1}\left[\mu_{x_1}dt + \sigma_{x_1}dW_1^{\mathbb{P}}(t)\right] + \frac{1}{2}\sigma_{x_1}^2\frac{\partial^2 P_t^T}{\partial x_1^2}d\langle x_1\rangle_t \\
&= \left[\frac{\partial P_t^T}{\partial t} + \mu_{x_1}\frac{\partial P_t^T}{\partial x_1} + \frac{1}{2}\sigma_{x_1}^2\frac{\partial^2 P_t^T}{\partial x_1^2}\right]dt + \sigma_{x_1}\frac{\partial P_t^T}{\partial x_1}dW_1^{\mathbb{P}}(t) \\
\Rightarrow \frac{dP_t^T}{P_t^T} &= \underbrace{\frac{1}{P_t^T}\left[\frac{\partial P_t^T}{\partial t} + \mu_{x_1}\frac{\partial P_t^T}{\partial x_1} + \frac{1}{2}\sigma_{x_1}^2\frac{\partial^2 P_t^T}{\partial x_1^2}\right]}_{\mu^{\mathbb{P}}}dt + \underbrace{\sigma_{x_1}\frac{1}{P_t^T}\frac{\partial P_t^T}{\partial x_1}}_{\sigma^{\mathbb{P}}}dW_1^{\mathbb{P}}(t)
\end{aligned} \tag{10}$$

We also know that the market-price-of-risk is given by,

$$
\begin{aligned}
\lambda &= \frac{\mu^{\mathbb{P}} - r}{\sigma^{\mathbb{P}}} \\
\Rightarrow \mu^{\mathbb{P}} &= r + \lambda \sigma^{\mathbb{P}} \\
\Rightarrow \mu^{\mathbb{P}} &= r + \frac{1}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1}
\end{aligned}
$$

We define the excess returns as,

$$
\mathbb{E}\left[\mathrm{xret}_t^T\right] = \mu^{\mathbb{P}} - r = \mathbb{E}\left[\frac{dP_t^T}{P_t^T} - r\right] = \frac{1}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1}
$$

This implies the drift of the bond in the $\mathbb{P}$ measure is given by (note that the stochastic term in (10) drops out on taking expectation below),

$$
\begin{aligned}
&\mathbb{E}\left[\frac{dP_t^T}{P_t^T} - r\right] = \frac{1}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1} \\
\Rightarrow\, &\underbrace{\frac{1}{P_t^T}\left[\frac{\partial P_t^T}{\partial t} + \mu_{x_1} \frac{\partial P_t^T}{\partial x_1} + \frac{1}{2}\sigma_{x_1}^2 \frac{\partial^2 P_t^T}{\partial x_1^2}\right]}_{\mu^{\mathbb{P}}} = r + \frac{1}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1}
\end{aligned} \tag{11}
$$

If we absorb the $\frac{1}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1}$ term into the left hand side, we get,

$$
\underbrace{\frac{1}{P_t^T}\left[\frac{\partial P_t^T}{\partial t} + \frac{\partial P_t^T}{\partial x_1}\left(\mu_{x_1} - \lambda \sigma_{x_1}\right) + \frac{1}{2}\sigma_{x_1}^2 \frac{\partial^2 P_t^T}{\partial x_1^2}\right]}_{\mu^{\mathbb{Q}}} = r \tag{12}
$$

which is nothing but an expression for the drift of the bond in the $\mathbb{Q}$ measure (which has to be equal to r).

In summary, what we have worked out in (11) and (12) above are expressions for the drifts ($\mu^{\mathbb{P}}$ and $\mu^{\mathbb{Q}}$) of the bond in the $\mathbb{P}$ and $\mathbb{Q}$ measures respectively.

Next, we consider the Sharpe Ratio of our bond strategy, which is given by,

$$
\begin{aligned}
\mathrm{SR} &= \frac{\mu^{\mathbb{P}} - r}{\sigma^{\mathbb{P}}} \\
&= \frac{\mu^{\mathbb{P}} - \mu^{\mathbb{Q}}}{\sigma^{\mathbb{P}}} \\
&= \frac{\frac{dP_t^T}{P_t^T} \frac{\partial P_t^T}{\partial x_1} \lambda \sigma_{x_1}}{\sigma_{x_1} \frac{dP_t^T}{P_t^T} \frac{\partial P_t^T}{\partial x_1}} \\
&= \lambda
\end{aligned}
$$

where in the above, we used all the relations in (10), (11) and (12).

Ultimately, we see that the market-price-of-risk equals the Sharpe Ratio of our bond strategy. This means that from the historical time-series of Sharpe Ratio values, we can back out the reversion-level, volatility and reversion-speed of $\lambda(t)$. This gives us the means of calibrating the parameters $\kappa_\lambda, \theta_\lambda, \sigma_\lambda$.

The last parameter we have to calibrate is the value for ρ. Recall that ρ is the correlation of the shocks driving $\lambda(t)$ with the shocks driving $x_2(t)$, the 'slope' principal-component.

Following [2], we can regress the excess returns of our bond strategy against the term-structure slope (where slope of the term structure is measured, for example, by the difference between five-year and three-month zero-coupon yields). From the regression, we will be able to get the R^2 value.

It is well-known that R^2 equals the squared Pearson correlation coefficient of the dependent and explanatory variable in a univariate linear least squares regression. Therefore, we will choose $\rho = R$ in the dynamics of our stochastic market-price-of-risk, and this completes its full calibration.

5.3 Calibrating the Reversion-Level Vector

To calibrate the reversion-level vector $\overrightarrow{\theta}^{\mathbb{Q}}$, we follow (again) the same strategy in [6], which is to,

1. Estimate time averages of yields or principal components using a very-long-term historical record.

2. Equate these quantities to the reversion levels $\overrightarrow{\theta}^{\mathbb{P}}$ (in the $\mathbb{P}$ measure).

3. Translate this vector $\overrightarrow{\theta}^{\mathbb{P}}$ to the $\mathbb{Q}$ measure using equation (2).

   To reiterate, the relation in (2) linking $\overrightarrow{\theta}^{\mathbb{Q}}$ and $\overrightarrow{\theta}^{\mathbb{P}}$ is,

$$\begin{aligned}
\overrightarrow{\theta}^{\mathbb{P}} &= \left(\underline{K}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi}\right)^{-1} \left(\underline{K}^{\mathbb{Q}} \overrightarrow{\theta}^{\mathbb{Q}} + \cancelto{\emptyset}{\underline{S}\overrightarrow{X}_0}\right) \\
\Rightarrow \overrightarrow{\theta}^{\mathbb{Q}} &= \left(\underline{K}^{\mathbb{Q}}\right)^{-1} \left(\underline{K}^{\mathbb{Q}} - \underline{S}\,\underline{\Pi}\right) \overrightarrow{\theta}^{\mathbb{P}}
\end{aligned}$$

For brevity, we would exclude the lengthy discussion of this calibration and refer the reader to the details in [6].

6 Conclusion and Future Work

We have given a concise review of the theoretical underpinnings of affine term-structure models, principal components analysis and market-price-of-risk as applied to bonds.

We also showed how a principal-component-based ATSM can be extended to include a new latent state variable which models a stochastic market-price-of-risk. Details on the calibration of the model were also presented.

For future work, it would be good to perform an analysis of the calibrated model and compare it against other existing models in pricing and fore-casting.

A Derivation of Solution for A_τ

For completeness, we provide a concise recap (following the derivations in [5]) of the closed-form solution for A_τ using our current symbols.

It is assumed that $\underline{\mathcal{K}}$ can be written as $\underline{\mathcal{K}} = \underline{a}\,\underline{\Lambda}\,\underline{a}^{-1}$. The integral expression for A_τ is,

$$\begin{aligned}
A_\tau &= \int_0^\tau \left[-w_0 + \left(\overrightarrow{B}_s\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta} + \frac{1}{2}\left(\overrightarrow{B}_s\right)^T \underline{S}\,\underline{S}^T \overrightarrow{B}_s \right] ds \\
&= -\int_0^\tau w_0 ds + \int_0^\tau \left(\overrightarrow{B}_s\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta}\, ds + \int_0^\tau \frac{1}{2}\left(\overrightarrow{B}_s\right)^T \underline{S}\,\underline{S}^T \overrightarrow{B}_s ds \\
&= I_1 + I_2 + I_3
\end{aligned}$$

We consider each of I_1, I_2 and I_3 in turn.

Evaluating I_1.

$$I_1 = -\int_0^\tau w_0 ds = -w_0 \tau$$

Evaluating I_2.

$$
\begin{aligned}
I_2 &= \int_0^\tau \left(\overrightarrow{B}_s\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds \\
&= \int_0^\tau \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \left[e^{-\underline{\mathcal{K}}s} - I_n\right] \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds \\
&= \int_0^\tau \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}}s} \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds - \int_0^\tau \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds \\
&= \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \int_0^\tau e^{-\underline{\mathcal{K}}s} \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \int_0^\tau e^{-\underline{a\Lambda}s\underline{a}^{-1}} \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \int_0^\tau \underline{a} e^{-\underline{\Lambda}s} \underline{a}^{-1} \underline{\mathcal{K}}\,\overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \quad \text{(by up-and-down theorem)} \\
&= \overrightarrow{w}_1^T \left(\underline{a\Lambda a}^{-1}\right)^{-1} \underline{a} \int_0^\tau e^{-\underline{\Lambda}s} \underline{a}^{-1} \left(\underline{a\Lambda a}^{-1}\right) \overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{a\Lambda}^{-1} \underline{a}^{-1} \underline{a} \int_0^\tau e^{-\underline{\Lambda}s} \underline{a}^{-1} \left(\underline{a\Lambda a}^{-1}\right) \overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{a\Lambda}^{-1} \int_0^\tau e^{-\underline{\Lambda}s} \underline{\Lambda a}^{-1} \overrightarrow{\theta}\,ds - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{a\Lambda}^{-1} \left[\int_0^\tau e^{-\underline{\Lambda}s} ds\right] \underline{\Lambda a}^{-1} \overrightarrow{\theta} - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{a} \left[\int_0^\tau e^{-\underline{\Lambda}s} ds\right] \underline{\Lambda}^{-1} \underline{\Lambda a}^{-1} \overrightarrow{\theta} - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \quad \text{(commutativity)}
\end{aligned}
$$

$$
\begin{aligned}
&= \overrightarrow{w}_1^T \underline{a} \left[\int_0^\tau \left(e^{\underline{\Lambda}s}\right)^{-1} ds\right] \underline{a}^{-1} \overrightarrow{\theta} - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \underline{a}\,\mathrm{diag}\left[\frac{1 - e^{-l_i \tau}}{l_i}\right] \underline{a}^{-1} \overrightarrow{\theta} - \overrightarrow{w}_1^T \overrightarrow{\theta}\,\tau \\
&= \overrightarrow{w}_1^T \left(\underline{a}\,\mathrm{diag}\left[\frac{1 - e^{-l_i \tau}}{l_i}\right] \underline{a}^{-1} \overrightarrow{\theta} - \overrightarrow{\theta}\,\tau\right)
\end{aligned}
$$

Evaluating I_3.

$$\begin{aligned}
I_3 &= \int_0^\tau \frac{1}{2}\left(\overrightarrow{B}_s\right)^T \underbrace{\underline{S}\,\underline{S}^T}_{C} \overrightarrow{B}_s ds \\
&= \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\left[e^{-\underline{\mathcal{K}}s} - I_n\right] C \left[e^{-\underline{\mathcal{K}}^T s} - I_n\right]\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&= \int_0^\tau \frac{1}{2}\left[\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}}s} - \overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\right]\left[C e^{-\underline{\mathcal{K}}^T s}\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 - C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1\right] ds \\
&= \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}}s} C e^{-\underline{\mathcal{K}}^T s}\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&\quad - \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C e^{-\underline{\mathcal{K}}^T s}\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&\quad - \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}}s} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&\quad + \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&= I_a + I_b + I_c + I_d
\end{aligned}$$

Consider I_d,

$$\begin{aligned}
I_d &= \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 \int_0^\tau ds \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 \tau
\end{aligned}$$

Consider I_c,

$$\begin{aligned}
I_c &= -\int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}}s} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\int_0^\tau e^{-\underline{\mathcal{K}}s} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 ds \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\underline{a}\left[\int_0^\tau e^{-\underline{\Lambda}s} ds\right]\underline{a}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1 \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1}\underline{a}\,\mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right]\underline{a}^{-1} C\left(\underline{\mathcal{K}}^T\right)^{-1}\overrightarrow{w}_1
\end{aligned}$$

Consider I_b,

$$
\begin{aligned}
I_b &= -\int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C e^{-\underline{\mathcal{K}}^T s} \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 ds \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \left[\int_0^\tau e^{-\underline{\mathcal{K}}^T s} ds\right] \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \underline{a} \left[\int_0^\tau e^{-\underline{\Lambda} s} ds\right] \underline{a}^{-1} \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 \quad \text{(since } \underline{\mathcal{K}}^T = \underline{\mathcal{K}}\text{, and up-and-down theorem)} \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \underline{a}\, \mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{a}^{-1} \underline{a}\underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 \\
&= -\frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \underline{a}\, \mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1
\end{aligned}
$$

Consider I_a,

$$
\begin{aligned}
I_a &= \int_0^\tau \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} e^{-\underline{\mathcal{K}} s} C e^{-\underline{\mathcal{K}}^T s} \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 ds \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \int_0^\tau e^{-\underline{a}\underline{\Lambda} s \underline{a}^{-1}} C e^{-\underline{a}\underline{\Lambda} s \underline{a}^{-1}} \underline{a}\underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 ds \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a} \int_0^\tau e^{-\underline{\Lambda} s} \underbrace{\underline{a}^{-1} C \underline{a}}_{M} e^{-\underline{\Lambda} s} \underline{a}^{-1} \underline{a}\underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 ds \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a} \underbrace{\left[\int_0^\tau e^{-\underline{\Lambda} s} M e^{-\underline{\Lambda} s} ds\right]}_{F_\tau = \left[m_{ij} \frac{1-e^{-(l_i+l_j)\tau}}{l_i+l_j}\right]_{ij}} \underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 \\
&= \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a}\, F_\tau \underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1
\end{aligned}
$$

Finally, we assemble all the various integrals we've computed to get A_τ as,

$$
\begin{aligned}
A_\tau &= \int_0^\tau \left[-w_0 + \left(\overrightarrow{B}_s\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta} + \frac{1}{2}\left(\overrightarrow{B}_s\right)^T \underline{S}\,\underline{S}^T \overrightarrow{B}_s\right] ds \\
&= -w_0\tau \\
&\quad + \overrightarrow{w}_1^T \left(\underline{a}\, \mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{a}^{-1} \overrightarrow{\theta} - \overrightarrow{\theta}\tau\right) \\
&\quad + \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a}\, F_s \underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 \\
&\quad - \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \underline{a}\, \mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{\Lambda}^{-1} \underline{a}^{-1} \overrightarrow{w}_1 \\
&\quad - \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} \underline{a} \mathrm{diag}\left[\frac{1-e^{-l_i\tau}}{l_i}\right] \underline{a}^{-1} C \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 \\
&\quad + \frac{1}{2}\overrightarrow{w}_1^T \underline{\mathcal{K}}^{-1} C \left(\underline{\mathcal{K}}^T\right)^{-1} \overrightarrow{w}_1 \tau
\end{aligned}
$$

References

[1] Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, **LV, NO. 5**, 2000.

[2] Gregory R. Duffee. Term premia and interest rate forecasts in affine models. *The Journal of Finance*, **LVII, NO. 1**:405–443, 2002.

[3] Darrell Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical Finance*, **6, No. 4**:379–406, 1996.

[4] Federal Reserve System. *Historical Data http://www.federalreserve.gov/releases/h15/data.htm.*

[5] Riccardo Rebonato. *Bond Pricing and Yield Curve Modelling : A Structural Approach.* 2016.

[6] Riccardo Rebonato, Ivan Saroka, and Vladyslav Putyatin. A principal-component-based affine term structure model. *SSRN*, 2014.

[7] Ivan Saroka. Affine principal-component-based term structure model. *SSRN*, 2014.