Technical Notes for the ATSM

Sunho Lee

June 25, 2024

1 Bond pricing model

1.1 JSZ normalization

The short-term interest rate r_t is

$$\underset{(1\times1)}{r_t} = \underset{(1\times d_{\mathbb{Q}})}{\iota'} \underset{(d_{\mathbb{Q}}\times1)}{X_t}, \tag{1}$$

where ι is a vector of ones. $\mathbb{Q}$-dynamics is

$$\begin{aligned} \underset{(d_{\mathbb{Q}}\times1)}{X_t} &= \underset{(d_{\mathbb{Q}}\times1)}{K_X^{\mathbb{Q}}} + \underset{(d_{\mathbb{Q}}\times d_{\mathbb{Q}})}{G_{XX}^{\mathbb{Q}}} X_{t-1} + \underset{(d_{\mathbb{Q}}\times1)}{\epsilon_{X,t}^{\mathbb{Q}}}, \\ K_X^{\mathbb{Q}} &\equiv \begin{bmatrix} \underset{(1\times1)}{k_\infty^{\mathbb{Q}}} & O_{1\times(d_{\mathbb{Q}}-1)} \end{bmatrix}' \\ G_{XX}^{\mathbb{Q}} &= Jordan(3), \\ \epsilon_{X,t}^{\mathbb{Q}} &\backsim \mathcal{N}(O_{d_{\mathbb{Q}}\times1}, \Omega_{XX}). \end{aligned} \tag{2}$$

We justify our specific form of $\mathbb{Q}$-eigenvalues in $G_{XX}^{\mathbb{Q}}$ by the below proposition.

Proposition 1. *When the eigenvalues of $G_{XX}^{\mathbb{Q}} - I_{d_{\mathbb{Q}}}$ is restricted to $\{0, \exp(-\kappa^{\mathbb{Q}}) - 1, \exp(-\kappa^{\mathbb{Q}})-1\}$, the restricted JSZ model with is observationally equivalent to the AFNS model.*

Proof. See Appendix A. □

Setting prior belief about $\kappa^{\mathbb{Q}}$ is easy, because $\kappa^{\mathbb{Q}}$ is just the DNS decay parameter. It is easy to modify our model to estimate unrestricted $\mathbb{Q}$-eigenvalues if we have reasonable prior belief. Reflecting a prior belief on the eigenvalues and estimate unrestricted eigenvalues is straight-forward. Considering a great cross-sectional fit of the DNS model, the restriction on the $\mathbb{Q}$-eigenvalues is supported empirically.

$\mathbb{P}$-dynamics is

$$\underbrace{\begin{bmatrix} X_t \\ {\scriptstyle (d_{\mathbb{Q}}\times 1)} \\ M_t \end{bmatrix}}_{F_t:d_{\mathbb{P}}\times 1} = \underbrace{\begin{bmatrix} K_X^{\mathbb{P}} \\ {\scriptstyle (d_{\mathbb{Q}}\times 1)} \\ K_M^{\mathbb{P}} \end{bmatrix}}_{K_F^{\mathbb{P}}:d_{\mathbb{P}}\times 1} + \sum_{l=1}^{p} \underbrace{\begin{pmatrix} G_{XX,l}^{\mathbb{P}} & G_{XM,l}^{\mathbb{P}} \\ {\scriptstyle (d_{\mathbb{Q}}\times d_{\mathbb{Q}})} & \\ G_{MX,l}^{\mathbb{P}} & G_{MM,l}^{\mathbb{P}} \end{pmatrix}}_{G_{FF,l}^{\mathbb{P}}:d_{\mathbb{P}}\times d_{\mathbb{P}}} \begin{bmatrix} X_{t-l} \\ M_{t-l} \end{bmatrix} + \underbrace{\begin{bmatrix} \epsilon_{X,t}^{\mathbb{P}} \\ {\scriptstyle (d_{\mathbb{Q}}\times 1)} \\ \epsilon_{M,t}^{\mathbb{P}} \end{bmatrix}}_{\epsilon_{F,t}^{\mathbb{P}}:d_{\mathbb{P}}\times 1},$$

$$\epsilon_{F,t}^{\mathbb{P}} \backsim \mathcal{N}(O_{d_{\mathbb{P}}\times 1}, \underbrace{\Omega_{FF}}_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \equiv \begin{pmatrix} \underbrace{\Omega_{XX}}_{d_{\mathbb{Q}}\times d_{\mathbb{Q}}} & \Omega_{MX}' \\ \Omega_{MX} & \Omega_{MM} \end{pmatrix}).$$

1.2 Bond price

Our guess about the bond price formula is

$$P_{\tau,t} = \exp[-a_\tau - b_\tau' X_t]. \tag{3}$$

We verify our guess with the method of undetermined coefficients. Under $\mathbb{Q}$-measure,

$$P_{\tau,t} = \mathbb{E}_t^{\mathbb{Q}}[\exp[-r_t] P_{\tau-1,t+1}]. \tag{4}$$

If we combine the above equations,

$$\exp[-a_\tau - b_\tau' X_t] = \mathbb{E}_t^{\mathbb{Q}}[\exp[-r_t - a_{\tau-1} - b_{\tau-1}' X_{t+1}]]. \tag{5}$$

By equation (2),

$$\begin{aligned} &\exp[-a_\tau - b_\tau' X_t] \\ =&\mathbb{E}_t^{\mathbb{Q}}[\exp[-r_t - a_{\tau-1} - b_{\tau-1}'(K_X^{\mathbb{Q}} + G_{XX}^{\mathbb{Q}} X_t + \epsilon_{X,t+1}^{\mathbb{Q}})]], \\ =&\mathbb{E}_t^{\mathbb{Q}}[\exp[-r_t - a_{\tau-1} - b_{\tau-1}' K_X^{\mathbb{Q}} - b_{\tau-1}' G_{XX}^{\mathbb{Q}} X_t - b_{\tau-1}' \epsilon_{X,t+1}^{\mathbb{Q}}]]. \end{aligned}$$

Expectation operator $\mathbb{E}_t^{\mathbb{Q}}$ is applied, then

$$\begin{aligned} &\exp[-a_\tau - b_\tau' X_t] \\ =&\exp[-r_t - a_{\tau-1} - b_{\tau-1}' K_X^{\mathbb{Q}} - b_{\tau-1}' G_{XX}^{\mathbb{Q}} X_t + 0.5 b_{\tau-1}' \Omega_{XX} b_{\tau-1}]. \end{aligned}$$

We substitute equation (1) into the above expression. Then,

$$\begin{aligned} &\exp[-a_\tau - b_\tau' X_t] \\ =&\exp[-a_{\tau-1} - b_{\tau-1}' K_X^{\mathbb{Q}} + 0.5 b_{\tau-1}' \Omega_{XX} b_{\tau-1} - (b_{\tau-1}' G_{XX}^{\mathbb{Q}} + \iota') X_t]. \end{aligned}$$

By the method of undetermined coefficient,

$$\therefore a_\tau = a_{\tau-1} + b_{\tau-1}' K_X^{\mathbb{Q}} - 0.5 b_{\tau-1}' \Omega_{XX} b_{\tau-1},$$

$$b_\tau = \iota + G_{XX}^{\mathbb{Q}\prime} b_{\tau-1}.$$

In reality, the bond yield is determined by

$$R_{\tau,t} = \frac{a_\tau}{\tau} + \frac{b_\tau'}{\tau} X_t + e_{\tau,t},$$
$$e_{\tau,t} \backsim \mathcal{N}(0, \sigma_{\tau,e}^2).$$

2 Affine Transformation to Principal Components

We use an affine transformation from X_t to principal components to estimate our model efficiently. For first $d_{\mathbb{Q}}$ principal components $\mathcal{P}_t$ and remaining $\mathcal{O}_t$,

$$\begin{aligned}
\begin{bmatrix} \mathcal{P}_t - c \\ \mathcal{O}_t \end{bmatrix} &:= \begin{pmatrix} W_{\mathcal{P}} \\ W_{\mathcal{O}} \end{pmatrix} R_t - \begin{bmatrix} c \\ O_{(N-d_{\mathbb{Q}})\times 1} \end{bmatrix} \\
&= \begin{pmatrix} W_{\mathcal{P}}\mathcal{A}_X - c \\ W_{\mathcal{O}}\mathcal{A}_X \end{pmatrix} + \begin{pmatrix} W_{\mathcal{P}} \\ W_{\mathcal{O}} \end{pmatrix} \mathcal{B}_X X_t + \begin{bmatrix} O_{d_{\mathbb{Q}}\times 1} \\ e_{\mathcal{O},t} \end{bmatrix}, \\
\text{where } R_t' &:= \begin{bmatrix} R_{\tau_1,t} & R_{\tau_2,t} & \cdots & R_{\tau_N,t} \end{bmatrix}, \\
\mathcal{A}_X' &:= \begin{bmatrix} a_{\tau_1}/\tau_1 & a_{\tau_2}/\tau_2 & \cdots & a_{\tau_N}/\tau_N \end{bmatrix}, \\
\mathcal{B}_X' &:= \begin{bmatrix} b_{\tau_1}/\tau_1 & b_{\tau_2}/\tau_2 & \cdots & b_{\tau_N}/\tau_N \end{bmatrix}, \\
e_{\mathcal{O},t} &\backsim \mathcal{N}(O_{(N-d_{\mathbb{Q}})\times 1}, \Sigma_{\mathcal{O}} := diag(\begin{bmatrix} \sigma_{O,1}^2 & \sigma_{O,2}^2 & \cdots & \sigma_{\mathcal{O},N-d_{\mathbb{Q}}}^2 \end{bmatrix})),
\end{aligned}$$

$W_{\mathcal{P}}$ and $W_{\mathcal{O}}$ is rotation matrices for $\mathcal{P}_t$ and $\mathcal{O}_t$, respectively. We set c to a sample mean of $\mathcal{P}_t$ to make the sample mean of latent pricing factors to zeros. Under this assumption, X_t becomes observed factors given parameters, because

$$\begin{aligned}
\mathcal{P}_t - c &= \mathcal{T}_{0,X} + \mathcal{T}_{1,X} X_t, \\
\text{where } \mathcal{T}_{0,X}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{XX}) &:= W_{\mathcal{P}}\mathcal{A}_X - c, \\
\mathcal{T}_{1,X}(: \kappa^{\mathbb{Q}}) &:= W_{\mathcal{P}}\mathcal{B}_X,
\end{aligned}$$

and

$$\begin{aligned}
X_t &= \mathcal{T}_{0,\mathcal{P}} + \mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c), \qquad (6) \\
\text{where } \mathcal{T}_{1,\mathcal{P}}(: \kappa^{\mathbb{Q}}) &:= (W_{\mathcal{P}}\mathcal{B}_X)^{-1} \equiv \mathcal{T}_{1,X}^{-1}, \\
\mathcal{T}_{0,\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{XX}) &:= -\mathcal{T}_{1,\mathcal{P}}(W_{\mathcal{P}}\mathcal{A}_X - c) \equiv -\mathcal{T}_{1,\mathcal{P}}\mathcal{T}_{0,X}.
\end{aligned}$$

The parentheses indicate deep-parameters in the latent factor state space.

Using the transformation

$$\underbrace{\begin{bmatrix} X_t \\ M_t \end{bmatrix}}_{F_t} = \underbrace{\begin{bmatrix} \mathcal{T}_{0,\mathcal{P}} \\ O_{(d_\mathbb{P}-d_\mathbb{Q})\times 1} \end{bmatrix}}_{\mathcal{T}_{0,\mathcal{F}}} + \underbrace{\begin{pmatrix} \mathcal{T}_{1,\mathcal{P}} & O_{d_\mathbb{Q}\times(d_\mathbb{P}-d_\mathbb{Q})} \\ O_{(d_\mathbb{P}-d_\mathbb{Q})\times d_\mathbb{Q}} & I_{(d_\mathbb{P}-d_\mathbb{Q})} \end{pmatrix}}_{\mathcal{T}_{1,\mathcal{F}}} \underbrace{\begin{bmatrix} \mathcal{P}_t - c \\ M_t \end{bmatrix}}_{\mathcal{F}_t},$$

we rotate our JSZ model. The rotated short-term interest rate is

$$r_t = \underbrace{\iota'\mathcal{T}_{0,\mathcal{P}}}_{\delta(:\kappa^\mathbb{Q}, k^\mathbb{Q}_\infty, \Omega_{XX})} + \underbrace{\iota'\mathcal{T}_{1,\mathcal{P}}}_{\beta(:\kappa^\mathbb{Q})'}(\mathcal{P}_t - c), \tag{7}$$

and the rotated $\mathbb{Q}$-dynamics is

$$\begin{aligned}
\mathcal{P}_t - c &= \underbrace{\mathcal{T}_{1,X}[K^\mathbb{Q}_X + (G^\mathbb{Q}_{XX} - I_{d_\mathbb{Q}})\mathcal{T}_{0,\mathcal{P}}]}_{K^\mathbb{Q}_\mathcal{P}(:\kappa^\mathbb{Q}, k^\mathbb{Q}_\infty, \Omega_{XX})} + \underbrace{\mathcal{T}_{1,X}G^\mathbb{Q}_{XX}\mathcal{T}_{1,\mathcal{P}}}_{G^\mathbb{Q}_{\mathcal{P}\mathcal{P}}(:\kappa^\mathbb{Q})}(\mathcal{P}_{t-1} - c) + \underbrace{\mathcal{T}_{1,X}\epsilon^\mathbb{Q}_{X,t}}_{\epsilon^\mathbb{Q}_{\mathcal{P},t}}, \\
\epsilon^\mathbb{Q}_{\mathcal{P},t} &\backsim \mathcal{N}(O_{d_\mathbb{Q}\times 1}, \Omega_{\mathcal{P}\mathcal{P}} \equiv \mathcal{T}_{1,X}\Omega_{XX}\mathcal{T}'_{1,X}).
\end{aligned}$$

2.1 State-Space Representation

The rotated model has a measurement equation of

$$\mathcal{O}_t = \underbrace{W_\mathcal{O}[\mathcal{A}_X + \mathcal{B}_X\mathcal{T}_{0,\mathcal{P}}]}_{\mathcal{A}_\mathcal{P}(:\kappa^\mathbb{Q}, k^\mathbb{Q}_\infty, \Omega_{XX})} + \underbrace{W_\mathcal{O}\mathcal{B}_X\mathcal{T}_{1,\mathcal{P}}}_{\mathcal{B}_\mathcal{P}(:\kappa^\mathbb{Q})}(\mathcal{P}_t - c) + e_{\mathcal{O},t}.$$

The $\mathbb{P}$-transition equation is

$$\begin{aligned}
\mathcal{F}_t = &\underbrace{\begin{bmatrix} \mathcal{T}_{1,X}[K^\mathbb{P}_X - (I_{d_\mathbb{Q}} - \sum_{l=1}^p G^\mathbb{P}_{XX,l})\mathcal{T}_{0,\mathcal{P}}] \\ K^\mathbb{P}_M + (\sum_{l=1}^p G^\mathbb{P}_{MX,l})\mathcal{T}_{0,\mathcal{P}} \end{bmatrix}}_{K^\mathbb{P}_\mathcal{F}} \\
&+ \sum_{l=1}^p \underbrace{\begin{pmatrix} \mathcal{T}_{1,X}G^\mathbb{P}_{XX,l}\mathcal{T}_{1,\mathcal{P}} & \mathcal{T}_{1,X}G^\mathbb{P}_{XM,l} \\ G^\mathbb{P}_{MX,l}\mathcal{T}_{1,\mathcal{P}} & G^\mathbb{P}_{MM,l} \end{pmatrix}}_{G^\mathbb{P}_{\mathcal{F}\mathcal{F},l}} \mathcal{F}_{t-l} + \underbrace{\begin{bmatrix} \mathcal{T}_{1,X}\epsilon^\mathbb{P}_{X,t} \\ \epsilon^\mathbb{P}_{M,t} \end{bmatrix}}_{\epsilon^\mathbb{P}_{\mathcal{F},t}}.
\end{aligned}$$

It is easy to verify that there is a sufficient degree of freedom to set parameters, $K^\mathbb{P}_\mathcal{F}$, $G^\mathbb{P}_{\mathcal{F}\mathcal{F},l}$ and the covariance of $\epsilon^\mathbb{P}_{\mathcal{F},t}$, as unrestricted parameters. Therefore, we can estimate the $\mathbb{P}$-dynamics with a unrestricted $VAR(p)$ model of

$$\mathcal{F}_t = \underbrace{\begin{bmatrix} K^\mathbb{P}_\mathcal{P} \\ K^\mathbb{P}_{M|\mathcal{F}} \end{bmatrix}}_{K^\mathbb{P}_\mathcal{F}} + \sum_{l=1}^p \underbrace{\begin{pmatrix} G^\mathbb{P}_{\mathcal{P}\mathcal{P},l} & G^\mathbb{P}_{\mathcal{P}M,l} \\ G^\mathbb{P}_{M\mathcal{P},l} & G^\mathbb{P}_{MM,l} \end{pmatrix}}_{G^\mathbb{P}_{\mathcal{F}\mathcal{F},l}} \mathcal{F}_{t-l} + \underbrace{\begin{bmatrix} \epsilon^\mathbb{P}_{\mathcal{P},t} \\ \epsilon^\mathbb{P}_{M,t} \end{bmatrix}}_{\epsilon^\mathbb{P}_{\mathcal{F},t}}, \tag{8}$$

$$\epsilon^{\mathbb{P}}_{\mathcal{F},t} \backsim \mathcal{N}(O_{d_{\mathbb{P}}\times 1}, \Omega_{\mathcal{F}\mathcal{F}} := \begin{pmatrix} \Omega_{\mathcal{P}\mathcal{P}} & \Omega'_{M\mathcal{P}} \\ \Omega_{M\mathcal{P}} & \Omega_{MM} \end{pmatrix} := \begin{pmatrix} \mathcal{T}_{1,X}\Omega_{XX}\mathcal{T}'_{1,X} & \mathcal{T}_{1,X}\Omega'_{MX} \\ \Omega_{MX}\mathcal{T}'_{1,X} & \Omega_{MM} \end{pmatrix}).$$

Given $\kappa^{\mathbb{Q}}$, there is a one-to-one mapping between $\Omega_{\mathcal{P}\mathcal{P}}$ and Ω_{XX}. Therefore, we can write the state-space equations as

$$\mathcal{O}_t = \mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k^{\mathbb{Q}}_{\infty}, \Omega_{\mathcal{P}\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c) + e_{\mathcal{O},t}, \tag{9}$$

$$\mathcal{F}_t = K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t}. \tag{10}$$

In this case, the difference equations are

$$\begin{aligned} a_\tau &= a_{\tau-1} + b'_{\tau-1}K^{\mathbb{Q}}_X - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}, \\ b_\tau &= \iota + G^{\mathbb{Q}\prime}_{XX}b_{\tau-1}, \end{aligned}$$

and also the short rate and $\mathbb{Q}$-dynamics are

$$r_t = \delta(: \kappa^{\mathbb{Q}}, k^{\mathbb{Q}}_{\infty}, \Omega_{\mathcal{P}\mathcal{P}}) + \beta(: \kappa^{\mathbb{Q}})'(\mathcal{P}_t - c),$$

$$\mathcal{P}_t - c = K^{\mathbb{Q}}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k^{\mathbb{Q}}_{\infty}, \Omega_{\mathcal{P}\mathcal{P}}) + G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_{t-1} - c) + \epsilon^{\mathbb{Q}}_{\mathcal{P},t}. \tag{11}$$

$\mathbb{Q}$-dynamics can be re-written with macro variables, for example,

$$\mathcal{F}_t = \underbrace{\begin{bmatrix} K^{\mathbb{Q}}_{\mathcal{P}} \\ K^{\mathbb{P}}_{M|\mathcal{F}} \end{bmatrix}}_{K^{\mathbb{Q}}_{\mathcal{F}}} + \sum_{l=1}^{p} G^{\mathbb{Q}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t-l} + \underbrace{\begin{bmatrix} \epsilon^{\mathbb{Q}}_{\mathcal{P},t} \\ \epsilon^{\mathbb{P}}_{M,t} \end{bmatrix}}_{\epsilon^{\mathbb{Q}}_{\mathcal{F},t}}, \tag{12}$$

$$G^{\mathbb{Q}}_{\mathcal{F}\mathcal{F},l} = \begin{cases} \begin{pmatrix} G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \\ G^{\mathbb{P}}_{M\mathcal{P},1} & G^{\mathbb{P}}_{MM,1} \end{pmatrix} & \text{if } l = 1, \\ \begin{pmatrix} O_{d_{\mathbb{Q}}\times d_{\mathbb{Q}}} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \\ G^{\mathbb{P}}_{M\mathcal{P},l} & G^{\mathbb{P}}_{MM,l} \end{pmatrix} & \text{otherwise.} \end{cases}$$

2.2 Moving to deep-parameters

Given $\{\kappa^{\mathbb{Q}}, k^{\mathbb{Q}}_{\infty}, K^{\mathbb{P}}_{\mathcal{F}}, G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}, \Omega_{\mathcal{F}\mathcal{F}}\}$, we can get the deep-parameters by

$$\begin{aligned} \Omega_{XX} &= \mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}, \qquad (13) \\ \Omega_{MX} &= \Omega_{M\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}, \\ G^{\mathbb{P}}_{XX,l} &= \mathcal{T}_{1,\mathcal{P}}G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}\mathcal{T}_{1,X}, \\ G^{\mathbb{P}}_{MX,l} &= G^{\mathbb{P}}_{M\mathcal{P},l}\mathcal{T}_{1,X}, \\ G^{\mathbb{P}}_{XM,l} &= \mathcal{T}_{1,\mathcal{P}}G^{\mathbb{P}}_{\mathcal{P}M,l}, \end{aligned}$$

$$K_X^{\mathbb{P}} = \mathcal{T}_{1,\mathcal{P}} K_{\mathcal{P}}^{\mathbb{P}} + (I_{d_{\mathbb{Q}}} - \sum_{l=1}^{p} G_{XX,l}^{\mathbb{P}}) \mathcal{T}_{0,\mathcal{P}},$$

$$K_M^{\mathbb{P}} = K_{M|\mathcal{F}}^{\mathbb{P}} - (\sum_{l=1}^{p} G_{MX,l}^{\mathbb{P}}) \mathcal{T}_{0,\mathcal{P}},$$

$$G_{XX}^{\mathbb{Q}} = \mathcal{T}_{1,\mathcal{P}} G_{\mathcal{P}\mathcal{P}}^{\mathbb{Q}} \mathcal{T}_{1,X}, \tag{14}$$

$$K_X^{\mathbb{Q}} = \mathcal{T}_{1,\mathcal{P}} K_{\mathcal{P}}^{\mathbb{Q}} + (I_{d_{\mathbb{Q}}} - G_{XX}^{\mathbb{Q}}) \mathcal{T}_{0,\mathcal{P}}. \tag{15}$$

2.3 Incorporation of the fourth observable spanned factor

Suppose the fourth factor is observable by a vector Z_t, so the pricing factor X_t is

$$X_t := \begin{bmatrix} X_{t,(1:3)} \\ Z_t \end{bmatrix}.$$

In this case,

$$\begin{aligned}
\begin{bmatrix} \mathcal{P}_t - c \\ Z_t \\ \mathcal{O}_t \end{bmatrix} &= \begin{pmatrix} W_{\mathcal{P}} \\ O_{d_z \times 1} \\ W_{\mathcal{O}} \end{pmatrix} R_t - \begin{bmatrix} c \\ -Z_t \\ O_{(N-d_{\mathbb{Q}})\times 1} \end{bmatrix}, \\
&= \begin{pmatrix} W_{\mathcal{P}} \mathcal{A}_X - c \\ O_{d_z \times 1} \\ W_{\mathcal{O}} \mathcal{A}_X \end{pmatrix} + \begin{pmatrix} W_{\mathcal{P}} \mathcal{B}_X & \\ O_{d_z \times 3} & I_{d_z} \\ W_{\mathcal{O}} \mathcal{B}_X & \end{pmatrix} X_t + \begin{bmatrix} O_{d_{\mathbb{Q}} \times 1} \\ O_{d_z \times 1} \\ e_{\mathcal{O},t} \end{bmatrix}.
\end{aligned}$$

The inverse of the transformation would be

$$\begin{bmatrix} \mathcal{P}_t - c \\ Z_t \end{bmatrix} = \begin{bmatrix} \mathcal{T}_{0,X} \\ O_{d_z \times 1} \end{bmatrix} + \begin{pmatrix} \mathcal{T}_{1,X} & \\ O_{d_z \times 3} & I_{d_z} \end{pmatrix} X_t,$$

and the transformation is

$$X_t = -\begin{pmatrix} \mathcal{T}_{1,X} & \\ O_{d_z \times 3} & I_{d_z} \end{pmatrix}^{-1} \begin{bmatrix} \mathcal{T}_{0,X} \\ O_{d_z \times 1} \end{bmatrix} + \begin{pmatrix} \mathcal{T}_{1,X} & \\ O_{d_z \times 3} & I_{d_z} \end{pmatrix}^{-1} \begin{bmatrix} \mathcal{P}_t - c \\ Z_t \end{bmatrix}.$$

Therefore, the coefficients of the transformation can be extended by augmenting the identity matrix and zeros.

3 Stochastic Discount Factor

Our market prices of risks are

$$\Omega_{\mathcal{F}\mathcal{F}}^{1/2} \underset{(d_{\mathbb{P}} \times 1)}{\lambda_t} \equiv \mathbb{E}_t^{\mathbb{P}}[\mathcal{F}_{t+1}] - \mathbb{E}_t^{\mathbb{Q}}[\mathcal{F}_{t+1}], \tag{16}$$

$$= \underbrace{K^{\mathbb{P}}_{\mathcal{F}} - K^{\mathbb{Q}}_{\mathcal{F}}}_{\lambda: d_{\mathbb{P}}\times 1} + \sum_{l=1}^{p} (\underbrace{G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l} - G^{\mathbb{Q}}_{\mathcal{F}\mathcal{F},l}}_{\Lambda_l: d_{\mathbb{P}}\times d_{\mathbb{P}}})\mathcal{F}_{t+1-l}, \tag{17}$$

$$= \begin{bmatrix} \lambda_{\mathcal{P}} + \sum_{l=1}^{p} \Lambda_{\mathcal{P}\mathcal{P},l}\mathcal{P}_{t+1-l} + \sum_{l=1}^{p} \Lambda_{\mathcal{P}M,l} M_{t+1-l} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1} \end{bmatrix}. \tag{18}$$

$$\text{where } \lambda = \begin{bmatrix} \lambda_{\mathcal{P}} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1} \end{bmatrix}, \tag{19}$$

$$\underset{(d_{\mathbb{Q}}\times 1)}{\lambda_{\mathcal{P}}} = K^{\mathbb{P}}_{\mathcal{P}} - K^{\mathbb{Q}}_{\mathcal{P}}, \tag{20}$$

$$\Lambda_l = \begin{pmatrix} \Lambda_{\mathcal{P}\mathcal{P},l} & \Lambda_{\mathcal{P}M,l} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times d_{\mathbb{Q}}} & O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{pmatrix}, \tag{21}$$

$$\underset{(d_{\mathbb{Q}}\times d_{\mathbb{Q}})}{\Lambda_{\mathcal{P}\mathcal{P},l}} = \begin{cases} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1} - G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} & \text{if } l = 1, \\ G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l} & \text{otherwise,} \end{cases} \tag{22}$$

$$\underset{(d_{\mathbb{Q}}\times(d_{\mathbb{P}}-d_{\mathbb{Q}}))}{\Lambda_{\mathcal{P}M,l}} = G^{\mathbb{P}}_{\mathcal{P}M,l}. \tag{23}$$

Next, we derive the Radon-Nikodym derivative $d\mathbb{Q}/d\mathbb{P}$ to specify a relationship between the two measures. By equation (8), risk factor $\mathcal{F}_t$ has

$$\mathcal{F}_{t+1} = K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} \tag{24}$$

in the $\mathbb{P}$-measure. We transform the above $\mathbb{P}$-dynamics into $\mathbb{Q}$-dynamics by making $\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l} - (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l})$ in the above equation. So,

$$\begin{aligned} \mathcal{F}_{t+1} &= K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l} - (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}), \\ &= (K^{\mathbb{P}}_{\mathcal{F}} - \lambda) + \sum_{l=1}^{p} (G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l} - \Lambda_l)\mathcal{F}_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}. \end{aligned}$$

By equation (17), the above equation can be written as

$$\mathcal{F}_{t+1} = K^{\mathbb{Q}}_{\mathcal{F}} + \sum_{l=1}^{p} G^{\mathbb{Q}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}.$$

By comparing the above expression with equation (12), we derive the relationship between $\epsilon^{\mathbb{P}}_{t+1}$ and $\epsilon^{\mathbb{Q}}_{t+1}$.

$$\epsilon^{\mathbb{Q}}_{\mathcal{F},t+1} = \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l},$$

$$\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} = \epsilon^{\mathbb{Q}}_{\mathcal{F},t+1} - (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}).$$

Under $\mathbb{Q}$-measure, since $\epsilon^{\mathbb{Q}}_{\mathcal{F},t+1}$ follows $\mathcal{N}(O_{d_{\mathbb{P}}\times 1}, \Omega_{\mathcal{F}\mathcal{F}})$,

$$p^{\mathbb{Q}}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|\mathcal{I}_t) = \mathcal{N}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}| - (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}), \Omega_{\mathcal{F}\mathcal{F}}) \text{ under } \mathbb{Q}\text{-measure.}$$

On the other hand,

$$p^{\mathbb{P}}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|\mathcal{I}_t) = \mathcal{N}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|O_{d_{\mathbb{P}}\times 1}, \Omega_{\mathcal{F}\mathcal{F}}) \text{ under the } \mathbb{P}\text{-measure.}$$

Radon-Nikodym derivative $d\mathbb{P}/d\mathbb{Q}$ is the one that satisfies

$$p^{\mathbb{Q}}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|\mathcal{I}_t) = \frac{d\mathbb{Q}}{d\mathbb{P}} p^{\mathbb{P}}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|\mathcal{I}_t).$$

If we substitute the above derivations into the above equation,

$$\mathcal{N}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}| - (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}), \Omega_{\mathcal{F}\mathcal{F}}) = \frac{d\mathbb{Q}}{d\mathbb{P}} \mathcal{N}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}|O_{d_{\mathbb{P}}\times 1}, \Omega_{\mathcal{F}\mathcal{F}}).$$

By the functional form of a normal density, the above equation is

$$\begin{aligned}
&(2\pi)^{-\frac{d_{\mathbb{P}}}{2}} \det(\Omega_{\mathcal{F}\mathcal{F}})^{-\frac{1}{2}} \exp[-\frac{1}{2}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}))' \Omega^{-1}_{\mathcal{F}\mathcal{F}} (\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}))] \\
=&\frac{d\mathbb{Q}}{d\mathbb{P}} (2\pi)^{-\frac{d_{\mathbb{P}}}{2}} \det(\Omega_{\mathcal{F}\mathcal{F}})^{-\frac{1}{2}} \exp[-\frac{1}{2} \epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1} \Omega^{-1}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t+1}],
\end{aligned}$$

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \exp[-\frac{1}{2} \epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1} \Omega^{-1}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t+1}] = \exp[-\frac{1}{2}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}))' \Omega^{-1}_{\mathcal{F}\mathcal{F}} (\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + (\lambda + \sum_{l=1}^{p} \Lambda_l \mathcal{F}_{t+1-l}))].$$

By equation (17), the right-hand side can be simplified as

$$\exp[-\frac{1}{2}(\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \Omega^{1/2}_{\mathcal{F}\mathcal{F}} \lambda_t)' \Omega^{-1}_{\mathcal{F}\mathcal{F}} (\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \Omega^{1/2}_{\mathcal{F}\mathcal{F}} \lambda_t)].$$

We expand the above equation as

$$\begin{aligned}
&\exp[-\frac{1}{2}\{\epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1} \Omega^{-1}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + 2\lambda_t' \Omega^{-1/2}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} + \lambda_t' \lambda_t\}] \\
=&\exp[-\frac{1}{2} \epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1} \Omega^{-1}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t+1} - \frac{1}{2}\lambda_t' \lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1}].
\end{aligned}$$

Therefore, if we substitute the last expression into the original equation,

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\exp[-\frac{1}{2}\epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1}\Omega^{-1}_{\mathcal{F}\mathcal{F}}\epsilon^{\mathbb{P}}_{\mathcal{F},t+1}] = \exp[-\frac{1}{2}\epsilon^{\mathbb{P}\prime}_{\mathcal{F},t+1}\Omega^{-1}_{\mathcal{F}\mathcal{F}}\epsilon^{\mathbb{P}}_{\mathcal{F},t+1} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1}].$$

We simplify the above equation with respect to $d\mathbb{Q}/d\mathbb{P}$, and it leads to

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp[-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1}].$$

Therefore, the model-consistent SDF $\mathcal{M}^{\mathbb{P}}_{t+1}$ should be

$$\begin{aligned}\mathcal{M}^{\mathbb{P}}_{t+1} &\equiv \exp[-r_t]\frac{d\mathbb{Q}}{d\mathbb{P}},\\ &= \exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1}]. \end{aligned} \tag{25}$$

3.1 Bond price under $\mathbb{P}$ measure

By equation (25),

$$\begin{aligned}P_{\tau,t} &= \mathbb{E}^{\mathbb{P}}_t[\mathcal{M}^{\mathbb{P}}_{t+1}P_{\tau-1,t+1}],\\ &= \mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1}]P_{\tau-1,t+1}].\end{aligned}$$

Also, by equation (3),

$$\begin{aligned}&\exp[-a_\tau - b_\tau' X_t]\\ =&\mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1} - a_{\tau-1} - b_{\tau-1}'X_{t+1}]],\\ 1 =&\mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1} - a_{\tau-1} - b_{\tau-1}'X_{t+1} + a_\tau + b_\tau' X_t]],\\ 0 =& \log \mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1} - a_{\tau-1} - b_{\tau-1}'X_{t+1} + a_\tau + b_\tau' X_t]],\\ 0 =& \log \mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1} - a_{\tau-1} - b_{\tau-1}'\mathcal{T}_{0,\mathcal{P}} - b_{\tau-1}'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_{t+1} - c)\\ &+ a_\tau + b_\tau'\mathcal{T}_{0,\mathcal{P}} + b_\tau'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c)]],\\ 0 =& \log \mathbb{E}^{\mathbb{P}}_t[\exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t' v^{\mathbb{P}}_{\mathcal{F},t+1} - a_{\tau-1} - b_{\tau-1}'\mathcal{T}_{0,\mathcal{P}} + a_\tau + b_\tau'\mathcal{T}_{0,\mathcal{P}} + b_\tau'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c)\\ &- b_{\tau-1}'\mathcal{T}_{1,\mathcal{P}}\{K^{\mathbb{P}}_{\mathcal{P}} + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{P},t+1}\}]].\end{aligned}$$

We focus on the RHS. By equations (27) and (7), the RHS is

$$
\begin{aligned}
&\log \mathbb{E}_t^{\mathbb{P}}[\exp[-\iota'\mathcal{T}_{0,\mathcal{P}} - \iota'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) - \frac{1}{2}\lambda'_{t,(1:d_{\mathbb{Q}})}\lambda_{t,(1:d_{\mathbb{Q}})} - \frac{1}{2}\lambda'_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})}\lambda_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})} \\
&- \lambda'_{t,(1:d_{\mathbb{Q}})}v^{\mathbb{P}}_{\mathcal{P},t+1} - \lambda'_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})}v^{\mathbb{P}}_{M,t+1} - a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + b'_\tau\mathcal{T}_{0,\mathcal{P}} + b'_\tau\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{P}}_{\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega^{1/2}_{\mathcal{P}\mathcal{P}}v^{\mathbb{P}}_{\mathcal{P},t+1})]], \\
=& -\iota'\mathcal{T}_{0,\mathcal{P}} - \iota'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) - \frac{1}{2}\lambda'_{t,(1:d_{\mathbb{Q}})}\lambda_{t,(1:d_{\mathbb{Q}})} - \frac{1}{2}\lambda'_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})}\lambda_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})} \\
&- a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + b'_\tau\mathcal{T}_{0,\mathcal{P}} + b'_\tau\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{P}}_{\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l} \\
&+ \frac{1}{2}\lambda'_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})}\lambda_{t,(d_{\mathbb{Q}}+1:d_{\mathbb{P}})} + \frac{1}{2}(\lambda'_{t,(1:d_{\mathbb{Q}})} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega^{1/2}_{\mathcal{P}\mathcal{P}})(\lambda'_{t,(1:d_{\mathbb{Q}})} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega^{1/2}_{\mathcal{P}\mathcal{P}})', \\
=& -\iota'\mathcal{T}_{0,\mathcal{P}} - \iota'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) - a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + b'_\tau\mathcal{T}_{0,\mathcal{P}} + b'_\tau\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{P}}_{\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l} \\
&+ b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega^{1/2}_{\mathcal{P}\mathcal{P}}\lambda_{t,(1:d_{\mathbb{Q}})} + \frac{1}{2}b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}.
\end{aligned}
$$

By equation (18), the RHS is

$$
\begin{aligned}
&-\iota'\mathcal{T}_{0,\mathcal{P}} - \iota'\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) - a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + b'_\tau\mathcal{T}_{0,\mathcal{P}} + b'_\tau\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{P}}_{\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l} + \frac{1}{2}b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} \\
&+ b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(K^{\mathbb{P}}_{\mathcal{P}} - K^{\mathbb{Q}}_{\mathcal{P}} + (G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1} - G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}})(\mathcal{P}_t - c) + \sum_{l=2}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l}), \\
=& -a_{\tau-1} + a_\tau + (b'_\tau - \iota' - b'_{\tau-1})\mathcal{T}_{0,\mathcal{P}} + \frac{1}{2}b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{Q}}_{\mathcal{P}} \\
&+ [(b'_\tau - \iota')\mathcal{T}_{1,\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}}](\mathcal{P}_t - c).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
a_\tau &= a_{\tau-1} - (b'_\tau - \iota' - b'_{\tau-1})\mathcal{T}_{0,\mathcal{P}} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{Q}}_{\mathcal{P}} - \frac{1}{2}b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}, \\
b_\tau &= \iota + (\mathcal{T}_{1,\mathcal{P}}G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}}\mathcal{T}_{1,X})'b_{\tau-1}.
\end{aligned}
$$

In section 1.2,

$$
a_\tau = a_{\tau-1} + b'_{\tau-1}K^{\mathbb{Q}}_X - \frac{1}{2}b'_{\tau-1}\Omega_{XX}b_{\tau-1},
$$

$$b_\tau = \iota + G^{\mathbb{Q}\prime}_{XX} b_{\tau-1}. \tag{26}$$

By equations (13), (14) and (15), the above equations are

$$\begin{aligned}
b_\tau &= \iota + (\mathcal{T}_{1,\mathcal{P}} G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} \mathcal{T}_{1,X})' b_{\tau-1}. \\
a_\tau &= a_{\tau-1} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}} K^{\mathbb{Q}}_{\mathcal{P}} + b'_{\tau-1}(I_{d_{\mathbb{Q}}} - G^{\mathbb{Q}}_{XX})\mathcal{T}_{0,\mathcal{P}} - \frac{1}{2} b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}} b_{\tau-1}, \\
&= a_{\tau-1} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}} K^{\mathbb{Q}}_{\mathcal{P}} - (b'_\tau - \iota' - b'_{\tau-1})\mathcal{T}_{0,\mathcal{P}} - \frac{1}{2} b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}} b_{\tau-1}.
\end{aligned}$$

For the last equality, we use (26). In conclusion, the results are the same between $\mathbb{P}$ and $\mathbb{Q}$ measures.

4 Risk premiums

4.1 Expected Return-Beta Representation

Market prices of risks λ_t is a price for

$$v^{\mathbb{P}}_{\mathcal{F},t} \equiv \Omega^{-1/2}_{\mathcal{F}\mathcal{F}} \epsilon^{\mathbb{P}}_{\mathcal{F},t} \backsim \mathcal{N}(O_{d_{\mathbb{P}}\times 1}, I_{d_{\mathbb{P}}}). \tag{27}$$

To show that, we first define the holding-period return,

$$\begin{aligned}
hr_{\tau,t+1} &\equiv \ln P_{\tau-1,t+1} - \ln P_{\tau,t}, \\
&= -a_{\tau-1} - b'_{\tau-1} X_{t+1} + a_\tau + b'_\tau X_t \\
&= -a_{\tau-1} - \underbrace{\begin{bmatrix} b'_{\tau-1} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{bmatrix}}_{b'_{F,\tau-1}} F_{t+1} + a_\tau + b'_{F,\tau} F_t, \\
&= -a_{\tau-1} - b'_{F,\tau-1}(\mathcal{T}_{0,\mathcal{F}} + \mathcal{T}_{1,\mathcal{F}}\mathcal{F}_{t+1}) + a_\tau + b'_{F,\tau}(\mathcal{T}_{0,\mathcal{F}} + \mathcal{T}_{1,\mathcal{F}}\mathcal{F}_t), \\
&= a_\tau - a_{\tau-1} + (b'_{F,\tau} - b'_{F,\tau-1})\mathcal{T}_{0,\mathcal{F}} - b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}\mathcal{F}_{t+1} + b'_{F,\tau}\mathcal{T}_{1,\mathcal{F}}\mathcal{F}_t.
\end{aligned} \tag{28}$$

By $\mathbb{P}$-dynamics $\mathcal{F}_{t+1} = K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^p G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t+1-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t+1}$,

$$\begin{aligned}
hr_{\tau,t+1} =& a_\tau - a_{\tau-1} + (b'_{F,\tau} - b'_{F,\tau-1})\mathcal{T}_{0,\mathcal{F}} + b'_{F,\tau}\mathcal{T}_{1,\mathcal{F}}\mathcal{F}_t \\
&- b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}} K^{\mathbb{P}}_{\mathcal{F}} - b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}} \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t+1-l} - b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}\Omega^{1/2}_{\mathcal{F}\mathcal{F}} v^{\mathbb{P}}_{\mathcal{F},t+1}.
\end{aligned} \tag{29}$$

In the case of the beta,

$$\begin{aligned}
Cov^{\mathbb{P}}_t[hr_{\tau,t+1}, v^{\mathbb{P}}_{\mathcal{F},t+1}] &= \mathbb{E}^{\mathbb{P}}_t[hr_{\tau,t+1} v^{\mathbb{P}\prime}_{\mathcal{F},t+1}] - \mathbb{E}^{\mathbb{P}}_t[hr_{\tau,t+1}]\mathbb{E}^{\mathbb{P}}_t[v^{\mathbb{P}\prime}_{\mathcal{F},t+1}], \\
&= \mathbb{E}^{\mathbb{P}}_t[-b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}\Omega^{1/2}_{\mathcal{F}\mathcal{F}} v^{\mathbb{P}}_{\mathcal{F},t+1} v^{\mathbb{P}\prime}_{\mathcal{F},t+1}],
\end{aligned}$$

$$
\begin{aligned}
&= -b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}\Omega^{1/2}_{\mathcal{F}\mathcal{F}}, \qquad (30)\\
&= -\left[b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}} \quad O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})}\right]\Omega^{1/2}_{\mathcal{F}\mathcal{F}},
\end{aligned}
$$

Therefore, the expected excess return is

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}_t[hr_{\tau,t+1}] - \mathbb{E}^{\mathbb{Q}}_t[hr_{\tau,t+1}] &= -b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}(\mathbb{E}^{\mathbb{P}}_t[\mathcal{F}_{t+1}] - \mathbb{E}^{\mathbb{Q}}_t[\mathcal{F}_{t+1}]) \text{ by equation (28)},\\
&= -b'_{F,\tau-1}\mathcal{T}_{1,\mathcal{F}}\Omega^{1/2}_{\mathcal{F}\mathcal{F}}\Omega^{-1/2}_{\mathcal{F}\mathcal{F}}(\mathbb{E}^{\mathbb{P}}_t[\mathcal{F}_{t+1}] - \mathbb{E}^{\mathbb{Q}}_t[\mathcal{F}_{t+1}]),\\
&= Cov^{\mathbb{P}}_t[hr_{\tau,t+1}, v^{\mathbb{P}}_{\mathcal{F},t+1}]\lambda_t \text{ by equations (30) and (16)},\\
&= \sum_{i=1}^{d_{\mathbb{P}}} \underbrace{\frac{Cov^{\mathbb{P}}_t[hr_{\tau,t+1}, v^{\mathbb{P}}_{\mathcal{F},t+1,(i)}]}{Var^{\mathbb{P}}_t[v^{\mathbb{P}}_{\mathcal{F},t+1,(i)}]}}_{beta}\lambda_{t,(i)}.
\end{aligned}
$$

Our beta representation shows that λ_t is a price for $v^{\mathbb{P}}_{\mathcal{F},t+1}$.

4.2 Sharpe Ratio

We define the Sharpe ratio as

$$
SR_{\tau,t} := \frac{\mathbb{E}^{\mathbb{P}}_t[xr_{\tau,t+1}] + \frac{1}{2}Var^{\mathbb{P}}_t[xr_{\tau,t+1}]}{Std^{\mathbb{P}}_t[xr_{\tau,t+1}]},
$$

where $xr_{\tau,t+1} := \ln P_{\tau-1,t+1} - \ln P_{\tau,t} - r_t$. By equation (4),

$$
\begin{aligned}
1 &= \mathbb{E}^{\mathbb{P}}_t[\mathcal{M}^{\mathbb{P}}_{t+1}\frac{P_{\tau-1,t+1}}{P_{\tau,t}}],\\
1 &= \mathbb{E}^{\mathbb{P}}_t[\exp[\ln\mathcal{M}^{\mathbb{P}}_{t+1} + (xr_{\tau,t+1} + r_t)]],\\
0 &= \mathbb{E}^{\mathbb{P}}_t\ln[\mathcal{M}^{\mathbb{P}}_{t+1}] + \mathbb{E}^{\mathbb{P}}_t[xr_{\tau,t+1}] + r_t\\
&+ \frac{1}{2}Var^{\mathbb{P}}_t[\ln\mathcal{M}^{\mathbb{P}}_{t+1}] + \frac{1}{2}Var^{\mathbb{P}}_t[xr_{\tau,t+1}] + Cov^{\mathbb{P}}_t[\ln\mathcal{M}^{\mathbb{P}}_{t+1}, xr_{\tau,t+1}]). \qquad (31)
\end{aligned}
$$

When $\tau = 1$,

$$
0 = \mathbb{E}^{\mathbb{P}}_t\ln[\mathcal{M}^{\mathbb{P}}_{t+1}] + r_t + \frac{1}{2}Var^{\mathbb{P}}_t[\ln\mathcal{M}^{\mathbb{P}}_{t+1}].
$$

Therefore, we can re-write equation ((31)) as

$$
0 = \mathbb{E}^{\mathbb{P}}_t[xr_{\tau,t+1}] + \frac{1}{2}Var^{\mathbb{P}}_t[xr_{\tau,t+1}] + Cov^{\mathbb{P}}_t[\ln\mathcal{M}^{\mathbb{P}}_{t+1}, xr_{\tau,t+1}]),
$$

$$
\mathbb{E}^{\mathbb{P}}_t[xr_{\tau,t+1}] + \frac{1}{2}Var^{\mathbb{P}}_t[xr_{\tau,t+1}] = -Cov^{\mathbb{P}}_t[\ln\mathcal{M}^{\mathbb{P}}_{t+1}, xr_{\tau,t+1}]).
$$

It simplifies the definition of the Sharpe ratio to

$$\begin{aligned}SR_{\tau,t} &= \frac{-Cov_t^{\mathbb{P}}[\ln \mathcal{M}_{t+1}^{\mathbb{P}}, xr_{\tau,t+1}])}{Std_t^{\mathbb{P}}[xr_{\tau,t+1}]}, \\ &= -Corr_t^{\mathbb{P}}[\ln \mathcal{M}_{t+1}^{\mathbb{P}}, xr_{\tau,t+1}])Std_t^{\mathbb{P}}[\ln \mathcal{M}_{t+1}^{\mathbb{P}}], \\ &\le Std_t^{\mathbb{P}}[\ln \mathcal{M}_{t+1}^{\mathbb{P}}].\end{aligned}$$

We define $Std_t^{\mathbb{P}}[\ln \mathcal{M}_{t+1}^{\mathbb{P}}]$ as the maximum Sharpe ratio.

4.3 Term premiums

The term premium is

$$TP_{\tau,t} = \frac{1}{\tau}\sum_{i=1}^{\tau-1} \mathbb{E}_t^{\mathbb{P}} exr_{\tau-i+1,t+i-1}, \tag{32}$$

$$where\ exr_{\tau,t} \equiv \mathbb{E}_t^{\mathbb{P}}[\ln P_{\tau-1,t+1} - \ln P_{\tau,t} - r_t], \tag{33}$$

By equations (3), (1), and (6),

$$\begin{aligned}exr_{\tau,t} =& \mathbb{E}_t^{\mathbb{P}}[-a_{\tau-1} - b'_{\tau-1}X_{t+1} + a_\tau + (b'_\tau - \iota')X_t], \\ =& \mathbb{E}_t^{\mathbb{P}}[-a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_{t+1} - c) + a_\tau + (b'_\tau - \iota')\mathcal{T}_{0,\mathcal{P}} + (b'_\tau - \iota')\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c)].\end{aligned}$$

By $\mathbb{P}$-dynamics $\mathcal{P}_{t+1} - c = K_{\mathcal{P}}^{\mathbb{P}} + \sum_{l=1}^{p} G_{\mathcal{P}\mathcal{P},l}^{\mathbb{P}}(\mathcal{P}_{t+1-l} - c) + \sum_{l=1}^{p} G_{\mathcal{P}M,l}^{\mathbb{P}} M_{t+1-l} + \epsilon_{\mathcal{P},t+1}^{\mathbb{P}}$,

$$\begin{aligned}exr_{\tau,t} =& \mathbb{E}_t^{\mathbb{P}}[-a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + (b'_\tau - \iota')\mathcal{T}_{0,\mathcal{P}} + (b'_\tau - \iota')\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\ &- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(K_{\mathcal{P}}^{\mathbb{P}} + \sum_{l=1}^{p} G_{\mathcal{P}\mathcal{P},l}^{\mathbb{P}}(\mathcal{P}_{t+1-l} - c) + \sum_{l=1}^{p} G_{\mathcal{P}M,l}^{\mathbb{P}} M_{t+1-l} + \epsilon_{\mathcal{P},t+1}^{\mathbb{P}})], \\ =& - a_{\tau-1} - b'_{\tau-1}\mathcal{T}_{0,\mathcal{P}} + a_\tau + (b'_\tau - \iota')\mathcal{T}_{0,\mathcal{P}} + (b'_\tau - \iota')\mathcal{T}_{1,\mathcal{P}}(\mathcal{P}_t - c) \\ &- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K_{\mathcal{P}}^{\mathbb{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G_{\mathcal{P}\mathcal{P},l}^{\mathbb{P}}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G_{\mathcal{P}M,l}^{\mathbb{P}} M_{t+1-l}, \\ =& (a_\tau - a_{\tau-1}) + (b'_\tau - \iota' - b'_{\tau-1})\mathcal{T}_{0,\mathcal{P}} + (b'_\tau\mathcal{T}_{1,\mathcal{P}} - \iota'\mathcal{T}_{1,\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}G_{\mathcal{P}\mathcal{P},1}^{\mathbb{P}})(\mathcal{P}_t - c) \\ &- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K_{\mathcal{P}}^{\mathbb{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=2}^{p} G_{\mathcal{P}\mathcal{P},l}^{\mathbb{P}}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G_{\mathcal{P}M,l}^{\mathbb{P}} M_{t+1-l}.\end{aligned}$$

By equations (39), (40), (14), and (15),

$$\begin{aligned}a_\tau - a_{\tau-1} &= b'_{\tau-1}K_X^{\mathbb{Q}} - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}, \\ &= b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K_{\mathcal{P}}^{\mathbb{Q}} + b'_{\tau-1}(I_{d_{\mathbb{Q}}} - G_{XX}^{\mathbb{Q}})\mathcal{T}_{0,\mathcal{P}} - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1},\end{aligned} \tag{34}$$

$$b'_\tau - \iota' - b'_{\tau-1} = b'_{\tau-1}(G^{\mathbb{Q}}_{XX} - I_{d_{\mathbb{Q}}}), \tag{35}$$
$$b'_\tau \mathcal{T}_{1,\mathcal{P}} - \iota' \mathcal{T}_{1,\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1} = b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} - G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1}), \tag{36}$$

and these results simplify $exr_{\tau,t}$ to

$$\begin{aligned}
exr_{\tau,t} =& b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{Q}}_{\mathcal{P}} + b'_{\tau-1}(I_{d_{\mathbb{Q}}} - G^{\mathbb{Q}}_{XX})\mathcal{T}_{0,\mathcal{P}} - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} \\
&+ b'_{\tau-1}(G^{\mathbb{Q}}_{XX} - I_{d_{\mathbb{Q}}})\mathcal{T}_{0,\mathcal{P}} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} - G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1})(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}K^{\mathbb{P}}_{\mathcal{P}} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=2}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l}, \\
=& - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(K^{\mathbb{P}}_{\mathcal{P}} - K^{\mathbb{Q}}_{\mathcal{P}}) - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} + b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(G^{\mathbb{Q}}_{\mathcal{P}\mathcal{P}} - G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},1})(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=2}^{p} G^{\mathbb{P}}_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{P}M,l}M_{t+1-l}.
\end{aligned}$$

Using notations defined in equation (17),

$$\begin{aligned}
exr_{\tau,t} =& - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\lambda_{\mathcal{P}} - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Lambda_{\mathcal{P}\mathcal{P},1}(\mathcal{P}_t - c) \\
&- b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=2}^{p}\Lambda_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\sum_{l=1}^{p}\Lambda_{\mathcal{P}M,l}M_{t+1-l}, \\
=& - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} - b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(\lambda_{\mathcal{P}} + \sum_{l=1}^{p}\Lambda_{\mathcal{P}\mathcal{P},l}(\mathcal{P}_{t+1-l} - c) + \sum_{l=1}^{p}\Lambda_{\mathcal{P}M,l}M_{t+1-l}), \quad (37) \\
=& - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1} - \begin{bmatrix} b'_{\tau-1} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{bmatrix}\mathcal{T}_{1,\mathcal{F}}(\Omega^{1/2}_{\mathcal{F}\mathcal{F}}\lambda_t). \quad (38)
\end{aligned}$$

Therefore,

$$\begin{aligned}
TP_{\tau,t} =& \frac{1}{\tau}\sum_{i=1}^{\tau-1}\mathbb{E}^{\mathbb{P}}_t exr_{\tau-i+1,t+i-1}, \\
=& \frac{1}{\tau}\sum_{i=1}^{\tau-1}\mathbb{E}^{\mathbb{P}}_t[-0.5b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-i} - \begin{bmatrix} b'_{\tau-i} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{bmatrix}\mathcal{T}_{1,\mathcal{F}}\Omega^{1/2}_{\mathcal{F}\mathcal{F}}\lambda_{t+i-1}], \\
=& -\frac{1}{\tau}\sum_{i=1}^{\tau-1}0.5b'_i\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_i - \frac{1}{\tau}\sum_{i=1}^{\tau-1}\begin{bmatrix} b'_{\tau-i} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{bmatrix}\mathcal{T}_{1,\mathcal{F}}\mathbb{E}^{\mathbb{P}}_t[\Omega^{1/2}_{\mathcal{F}\mathcal{F}}\lambda_{t+i-1}].
\end{aligned}$$

and, by equation (17),

$$TP_{\tau,t} = -\frac{1}{\tau}\sum_{i=1}^{\tau-1}(0.5b'_i\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_i + \begin{bmatrix} b'_{\tau-i} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{bmatrix}\mathcal{T}_{1,\mathcal{F}}\lambda)$$

$$
- \frac{1}{\tau}\sum_{i=1}^{\tau-1}\begin{bmatrix} b'_{\tau-i} & O_{1\times(d_{\mathbb{P}}-d_{\mathbb{Q}})}\end{bmatrix} \mathcal{T}_{1,\mathcal{F}}(\sum_{l=1}^{p}\Lambda_l \mathbb{E}_t^{\mathbb{P}}[\mathcal{F}_{t+i-l}]).
$$

We calculate $\mathbb{E}_t^{\mathbb{P}}[\mathcal{F}_{t+i-l}]$ iteratively using $\mathbb{P}$-dynamics $\mathcal{F}_t = K_{\mathcal{F}}^{\mathbb{P}} + \sum_{l=1}^{p} G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}}\mathcal{F}_{t-l} + \epsilon_{\mathcal{F},t}^{\mathbb{P}}$.

Or, by using equation (37),

$$
\begin{aligned}
TP_{\tau,t} =& \frac{1}{\tau}\sum_{i=1}^{\tau-1}\mathbb{E}_t^{\mathbb{P}} exr_{\tau-i+1,t+i-1}, \\
=& -\frac{1}{\tau}\sum_{i=1}^{\tau-1}(0.5b_i'\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}_{1,\mathcal{P}}'b_i + b_i'\mathcal{T}_{1,\mathcal{P}}\lambda_{\mathcal{P}}) \\
& -\frac{1}{\tau}\sum_{i=1}^{\tau-1}b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}(\sum_{l=1}^{p}\Lambda_{\mathcal{P}\mathcal{P},l}\mathbb{E}_t^{\mathbb{P}}[\mathcal{P}_{t+i-l}-c] + \sum_{l=1}^{p}\Lambda_{\mathcal{P}M,l}\mathbb{E}_t^{\mathbb{P}}[M_{t+i-l}]).
\end{aligned}
$$

5 Empirical Methodology

We estimate the model in terms of annual returns. Originally, the measurement equation is

$$
\begin{aligned}
\mathcal{O}_t &= W_{\mathcal{O}}R_t = \mathcal{A}_{\mathcal{P}} + \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c) + \mathcal{N}(O_{(N-d_{\mathbb{Q}})\times 1}, \Sigma_{\mathcal{O}}), \\
\mathcal{A}_{\mathcal{P}} &= W_{\mathcal{O}}[\mathcal{A}_X + \mathcal{B}_X\mathcal{T}_{0,\mathcal{P}}], \\
\mathcal{B}_{\mathcal{P}} &= W_{\mathcal{O}}\mathcal{B}_X\mathcal{T}_{1,\mathcal{P}}, \\
\mathcal{T}_{1,\mathcal{P}} &= (W_{\mathcal{P}}\mathcal{B}_X)^{-1}, \\
\mathcal{T}_{0,\mathcal{P}} &= -\mathcal{T}_{1,\mathcal{P}}(W_{\mathcal{P}}\mathcal{A}_X - c) = -\mathcal{T}_{1,\mathcal{P}}\mathcal{T}_{0,X}, \\
\mathcal{A}_X' &= \begin{bmatrix} a_{\tau_1}/\tau_1 & a_{\tau_2}/\tau_2 & \cdots & a_{\tau_N}/\tau_N \end{bmatrix}, \\
\mathcal{B}_X' &= \begin{bmatrix} b_{\tau_1}/\tau_1 & b_{\tau_2}/\tau_2 & \cdots & b_{\tau_N}/\tau_N \end{bmatrix}, \\
a_\tau &= a_{\tau-1} + b'_{\tau-1}K_X^{\mathbb{Q}} - 0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}_{1,\mathcal{P}}'b_{\tau-1}, \\
b_\tau &= \iota + G_{XX}^{\mathbb{Q}\prime}b_{\tau-1},
\end{aligned}
$$

and the transition equation is

$$
\begin{bmatrix} \mathcal{P}_t - c \\ M_t \end{bmatrix} = \begin{bmatrix} K_{\mathcal{P}}^{\mathbb{P}} \\ K_{M|\mathcal{F}}^{\mathbb{P}} \end{bmatrix} + \sum_{l=1}^{p}\begin{pmatrix} G_{\mathcal{P}\mathcal{P},l}^{\mathbb{P}} & G_{\mathcal{P}M,l}^{\mathbb{P}} \\ G_{M\mathcal{P},l}^{\mathbb{P}} & G_{MM,l}^{\mathbb{P}} \end{pmatrix}\begin{bmatrix} \mathcal{P}_{t-l} - c \\ M_{t-l} \end{bmatrix} + \mathcal{N}(O_{d_{\mathbb{P}}\times 1}, \begin{pmatrix} \Omega_{\mathcal{P}\mathcal{P}} & \Omega_{M\mathcal{P}}' \\ \Omega_{M\mathcal{P}} & \Omega_{MM} \end{pmatrix}).
$$

The term premium is

$$
TP_{\tau,t} = -\frac{1}{\tau}\sum_{i=1}^{\tau-1}(0.5b_i'\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}_{1,\mathcal{P}}'b_i + b_i'\mathcal{T}_{1,\mathcal{P}}\lambda_{\mathcal{P}})
$$

$$-\frac{1}{\tau}\sum_{i=1}^{\tau-1} b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}(\sum_{l=1}^{p}\Lambda_{\mathcal{PP},l}\mathbb{E}_t^{\mathbb{P}}[\mathcal{P}_{t+i-l}-c]+\sum_{l=1}^{p}\Lambda_{\mathcal{P}M,l}\mathbb{E}_t^{\mathbb{P}}[M_{t+i-l}]).$$

If we use annual data so yields are $1200\times R_t$, the state-space is modified as

$$\begin{aligned}
\mathbf{1200}\cdot\mathcal{O}_{\mathbf{t}} &= W_{\mathcal{O}}(\mathbf{1200}\cdot\mathbf{R_t}) = \mathbf{1200}\cdot\mathcal{A}_{\mathcal{P}} + \mathcal{B}_{\mathcal{P}}(\mathbf{1200}\cdot\mathcal{P}_{\mathbf{t}} - \mathbf{1200}\cdot\mathbf{c}) + \mathcal{N}(O_{(N-d_{\mathbb{Q}})\times 1}, \mathbf{1200^2}\cdot\mathbf{\Sigma}_{\mathcal{O}}),\\
\mathbf{1200}\cdot\mathcal{A}_{\mathcal{P}} &= W_{\mathcal{O}}[\mathbf{1200}\cdot\mathcal{A}_{\mathbf{X}} + \mathcal{B}_X(\mathbf{1200}\cdot\mathcal{T}_{\mathbf{0},\mathcal{P}})],\\
\mathcal{B}_{\mathcal{P}} &= W_{\mathcal{O}}\mathcal{B}_X\mathcal{T}_{1,\mathcal{P}},\\
\mathcal{T}_{1,\mathcal{P}} &= (W_{\mathcal{P}}\mathcal{B}_X)^{-1},\\
\mathbf{1200}\cdot\mathcal{T}_{\mathbf{0},\mathcal{P}} &= -\mathcal{T}_{1,\mathcal{P}}[\mathbf{W}_{\mathcal{P}}(\mathbf{1200}\cdot\mathcal{A}_{\mathbf{X}}) - \mathbf{1200}\cdot\mathbf{c}] = -\mathcal{T}_{1,\mathcal{P}}(\mathbf{1200}\cdot\mathcal{T}_{\mathbf{0},\mathbf{X}}),\\
\mathbf{1200}\cdot\mathcal{A}'_{\mathbf{X}} &= \begin{bmatrix}(\mathbf{1200}\cdot\mathbf{a}_{\tau_1})/\tau_1 & (\mathbf{1200}\cdot\mathbf{a}_{\tau_2})/\tau_2 & \cdots & (\mathbf{1200}\cdot\mathbf{a}_{\tau_N})/\tau_N\end{bmatrix},\\
\mathcal{B}'_X &= \begin{bmatrix}b_{\tau_1}/\tau_1 & b_{\tau_2}/\tau_2 & \cdots & b_{\tau_N}/\tau_N\end{bmatrix},\\
\mathbf{1200}\cdot\mathbf{a}_{\tau} &= \mathbf{1200}\cdot\mathbf{a}_{\tau-\mathbf{1}} + b'_{\tau-1}(\begin{bmatrix}\mathbf{1200}\cdot\mathbf{k}_{\infty}^{\mathbb{Q}}\\ O_{(d_{\mathbb{Q}}-1)\times 1}\end{bmatrix}) - \frac{0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}(\mathbf{1200^2}\cdot\mathbf{\Omega}_{\mathcal{PP}})\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}}{1200},\\
b_{\tau} &= \iota + G_{XX}^{\mathbb{Q}\prime}b_{\tau-1},
\end{aligned}$$

and

$$\begin{bmatrix}\mathbf{1200}\cdot\mathcal{P}_{\mathbf{t}}-\mathbf{1200}\cdot\mathbf{c}\\ M_t\end{bmatrix} = \begin{bmatrix}\mathbf{1200}\cdot\mathbf{K}_{\mathcal{P}}^{\mathbb{P}}\\ K_{M|\mathcal{F}}^{\mathbb{P}}\end{bmatrix} + \sum_{l=1}^{p}\begin{pmatrix}G_{\mathcal{PP},l}^{\mathbb{P}} & \mathbf{1200}\cdot\mathbf{G}_{\mathcal{P}\mathbf{M},\mathbf{l}}^{\mathbb{P}}\\ \mathbf{1200^{-1}}\cdot\mathbf{G}_{\mathbf{M}\mathcal{P},\mathbf{l}}^{\mathbb{P}} & G_{MM,l}^{\mathbb{P}}\end{pmatrix}\begin{bmatrix}\mathbf{1200}\cdot\mathcal{P}_{\mathbf{t-1}}-\mathbf{1200}\cdot\mathbf{c}\\ M_{t-l}\end{bmatrix}$$
$$+\mathcal{N}\left(O_{d_{\mathbb{P}}\times 1}, \begin{pmatrix}\mathbf{1200^2}\cdot\mathbf{\Omega}_{\mathcal{PP}} & \mathbf{1200}\cdot\mathbf{\Omega}'_{\mathbf{M}\mathcal{P}}\\ \mathbf{1200}\cdot\mathbf{\Omega}_{\mathbf{M}\mathcal{P}} & \Omega_{MM}\end{pmatrix}\right).$$

The term premium is modified to

$$\begin{aligned}
\mathbf{1200}\cdot\mathbf{TP}_{\tau,\mathbf{t}} = &-\frac{1}{\tau}\sum_{i=1}^{\tau-1}[\frac{0.5b'_i\mathcal{T}_{1,\mathcal{P}}(\mathbf{1200^2}\cdot\mathbf{\Omega}_{\mathcal{PP}})\mathcal{T}'_{1,\mathcal{P}}b_i}{1200} + b'_i\mathcal{T}_{1,\mathcal{P}}(\mathbf{1200}\cdot\lambda_{\mathcal{P}})]\\
&-\frac{1}{\tau}\sum_{i=1}^{\tau-1}b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}(\sum_{l=1}^{p}\Lambda_{\mathcal{PP},l}\mathbb{E}_t^{\mathbb{P}}[\mathbf{1200}\cdot\mathcal{P}_{\mathbf{t+i-l}}-\mathbf{1200}\cdot\mathbf{c}] + \sum_{l=1}^{p}\mathbf{1200}\cdot\mathbf{\Lambda}_{\mathcal{P}\mathbf{M},\mathbf{l}}\mathbb{E}_t^{\mathbb{P}}[M_{t+i-l}]),
\end{aligned}$$

and

$$\mathbf{1200}\cdot\lambda_{\mathcal{P}} = \mathbf{1200}\cdot\mathbf{K}_{\mathcal{P}}^{\mathbb{P}} - \mathcal{T}_{1,X}[\begin{bmatrix}\mathbf{1200}\cdot\mathbf{k}_{\infty}^{\mathbb{Q}}\\ O_{(d_{\mathbb{Q}}-1)\times 1}\end{bmatrix} + (G_{XX}^{\mathbb{Q}} - I_{d_{\mathbb{Q}}})(\mathbf{1200}\cdot\mathcal{T}_{\mathbf{0},\mathcal{P}})].$$

When we redefine the notations for scaled parameters, the modified state-space model is almost same as the original. Therefore, is not irrelevant whether we use annual yields or monthly yields. The only thing we have to be careful is the form of the Jensen's inequality term. The Jensen's inequality term $\mathcal{J}(:\kappa^{\mathbb{Q}},\Omega_{\mathcal{P}\mathcal{P}})$ is

$$\mathcal{J}_\tau(:\kappa^{\mathbb{Q}},\Omega_{\mathcal{P}\mathcal{P}})=\begin{cases}0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}, & \text{if yields are monthly decimal data,}\\ \frac{0.5b'_{\tau-1}\mathcal{T}_{1,\mathcal{P}}\Omega_{\mathcal{P}\mathcal{P}}\mathcal{T}'_{1,\mathcal{P}}b_{\tau-1}}{1200}, & \text{if yields are annual percentage data.}\end{cases}$$

Then, the difference equation is

$$a_\tau = a_{\tau-1}+b'_{\tau-1}K^{\mathbb{Q}}_X-\mathcal{J}_\tau(:\kappa^{\mathbb{Q}},\Omega_{\mathcal{P}\mathcal{P}}), \tag{39}$$

$$b_\tau = \iota + G^{\mathbb{Q}\prime}_{XX}b_{\tau-1}. \tag{40}$$

and term premium is

$$\begin{aligned}TP_{\tau,t}=&-\frac{1}{\tau}\sum_{i=1}^{\tau-1}(\mathcal{J}_{i+1}(:\kappa^{\mathbb{Q}},\Omega_{\mathcal{P}\mathcal{P}})+b'_i\mathcal{T}_{1,\mathcal{P}}\lambda_{\mathcal{P}})\\ &-\frac{1}{\tau}\sum_{i=1}^{\tau-1}\sum_{l=1}^{p}[b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}\Lambda_{\mathcal{P}\mathcal{P},l}\mathbb{E}^{\mathbb{P}}_t[\mathcal{P}_{t+i-l}-c]+b'_{\tau-i}\mathcal{T}_{1,\mathcal{P}}\Lambda_{\mathcal{P}M,l}\mathbb{E}^{\mathbb{P}}_t[M_{t+i-l}])].\end{aligned}$$

The estimated parameters are

$$\kappa^{\mathbb{Q}},k^{\mathbb{Q}}_\infty,K^{\mathbb{P}}_{\mathcal{F}},\{G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\}_{l=1}^p,\Omega_{\mathcal{F}\mathcal{F}},\Sigma_{\mathcal{O}}.$$

5.1 Prior distribution

5.1.1 VAR covariance $\Omega_{\mathcal{F}\mathcal{F}}$

For a prior distribution of $\Omega_{\mathcal{F}\mathcal{F}}$, we impose

$$\Omega_{\mathcal{F}\mathcal{F}}\backsim\mathcal{IW}(\nu_0,diag[\Omega_0]).$$

Instead of using the InverseWishart-InverseWishart update, we translate the Wishart distribution into a series of normal-inverse gamma distributions using the following proposition.

Proposition 2. *Consider the following normal-inverse-gamma priors on the diagonal elements of Σ and the lower triangular elements of A:*

$$\sigma_i^2\backsim\mathcal{IG}(\frac{\nu_0+i-n}{2},\frac{s_i^2}{2}),i=1,\cdots,n,$$
$$(A_{i,j}|\sigma_i^2)\backsim\mathcal{N}(0,\frac{\sigma_i^2}{s_j^2}),1\le j<i\le n,i=2,\cdots,n.$$

Then $\tilde{\Sigma}^{-1}=A'\Sigma^{-1}A$ has the Wishart distribution $\tilde{\Sigma}^{-1}\backsim\mathcal{W}(v_0,S^{-1})$, where $S=diag(s_1^2,s_2^2\cdots,s_n^2)$. It follows that $\tilde{\Sigma}\backsim\mathcal{IW}(\nu_0,S)$.

To this end, we orthogonalize the system

$$\mathcal{F}_t = K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t-l} + \epsilon^{\mathbb{P}}_{\mathcal{F},t},$$

so

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \mathcal{C}_{21} & 1 & 0 & \cdots & 0 \\ \mathcal{C}_{31} & \mathcal{C}_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_{d_{\mathbb{P}}1} & \mathcal{C}_{d_{\mathbb{P}}2} & \mathcal{C}_{d_{\mathbb{P}}3} & \cdots & 1 \end{pmatrix}}_{\mathcal{C}} \mathcal{F}_t = \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t-l} + diag(\underbrace{\begin{bmatrix} \sigma_{\mathcal{F}\mathcal{F},1} \\ \vdots \\ \sigma_{\mathcal{F}\mathcal{F},d_{\mathbb{P}}} \end{bmatrix}}_{\tilde{\Omega}^{1/2}_{\mathcal{F}\mathcal{F}}})v^{\mathbb{P}}_{\mathcal{F},t}. \quad (41)$$

$\mathcal{C}$ is implicitly defined as $\Omega^{-1}_{\mathcal{F}\mathcal{F}} = \mathcal{C}' diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F}}]^{-1}\mathcal{C}$ (or $\Omega_{\mathcal{F}\mathcal{F}} = \mathcal{C}^{-1} diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F}}]\mathcal{C}^{-1\prime}$). By the proposition, the InverseWishart prior is the same as

$$\sigma^2_{\mathcal{F}\mathcal{F},i} \backsim \mathcal{IG}(\frac{\nu_0 + i - d_{\mathbb{P}}}{2}, \frac{\Omega_{0,(i)}}{2}), i = 1, \cdots, d_{\mathbb{P}},$$
$$\mathcal{C}_{ij}|\sigma^2_{\mathcal{F}\mathcal{F},i} \backsim \mathcal{N}(0, \frac{\sigma^2_{\mathcal{F}\mathcal{F},i}}{\Omega_{0,(j)}}), 1 \le j < i \le d_{\mathbb{P}}, i = 2, \cdots, d_{\mathbb{P}}.$$

From the recursive structure that $\Omega_{\mathcal{F}\mathcal{F}}$ has, we can derive that

$$\Omega_{\mathcal{P}\mathcal{P}} = (\mathcal{C}^{-1})_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})} diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F},(1:d_{\mathbb{Q}})}]((\mathcal{C}^{-1})_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})'.$$

Because $\mathcal{C}$ is a lower triangular matrix, when we partition

$$\mathcal{C} = \begin{pmatrix} \mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \\ \mathcal{C}_{(d_{\mathbb{Q}}+1:d_{\mathbb{P}},1:d_{\mathbb{Q}})} & \mathcal{C}_{(d_{\mathbb{Q}}+1:d_{\mathbb{P}},d_{\mathbb{Q}}+1:d_{\mathbb{P}})} \end{pmatrix},$$
$$\mathcal{C}^{-1} = \begin{pmatrix} (\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})^{-1} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \\ -(\mathcal{C}_{(d_{\mathbb{Q}}+1:d_{\mathbb{P}},d_{\mathbb{Q}}+1:d_{\mathbb{P}})})^{-1}\mathcal{C}_{(d_{\mathbb{Q}}+1:d_{\mathbb{P}},1:d_{\mathbb{Q}})}(\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})^{-1} & (\mathcal{C}_{(d_{\mathbb{Q}}+1:d_{\mathbb{P}},d_{\mathbb{Q}}+1:d_{\mathbb{P}})})^{-1} \end{pmatrix},$$
$$(\mathcal{C}^{-1})_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})} = (\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})^{-1}.$$
$$\therefore \Omega_{\mathcal{P}\mathcal{P}} = (\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})^{-1} diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F},(1:d_{\mathbb{Q}})}](\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})^{-1\prime}.$$

In terms of a precision,

$$\Omega^{-1}_{\mathcal{P}\mathcal{P}} = (\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})})' diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F},(1:d_{\mathbb{Q}})}]^{-1}\mathcal{C}_{(1:d_{\mathbb{Q}},1:d_{\mathbb{Q}})}.$$

Therefore, sampling $\sigma^2_{\mathcal{F}\mathcal{F},1}$, and $\{C_{(i,1:i-1)}, \sigma^2_{\mathcal{F}\mathcal{F},i}\}_{i=2}^{d_{\mathbb{Q}}}$ is equal to sampling $\Omega_{\mathcal{P}\mathcal{P}}$. For information,

$$\Omega_{\mathcal{P}\mathcal{P}} \backsim \mathcal{IW}(\nu_0 - (d_{\mathbb{P}} - d_{\mathbb{Q}}), diag[\Omega_{0,(1:d_{\mathbb{Q}})}]).$$

5.1.2 VAR conditional mean $K^{\mathbb{P}}_{\mathcal{F}}$ and $\{G^{\mathbb{P}}_{\mathcal{FF},l}\}_{l=1}^{p}$

First, we set prior means of $\{G^{\mathbb{P}}_{\mathcal{FF},l}\}_{l=1}^{p}$. We set prior mean of $G^{\mathbb{P}}_{\mathcal{PP},1}$ to the prior mean of $G^{\mathbb{Q}}_{\mathcal{PP}}$. For other diagonal elements in $G^{\mathbb{P}}_{\mathcal{PP},1}$, we set to zero and 0.9 for growth and level variables, respectively. The other elements in $\{G^{\mathbb{P}}_{\mathcal{FF},l}\}_{l=1}^{p}$ have zero prior means.

For the efficiency of the sampler, we estimate the orthogonalized equation (41). Since

$$(\mathcal{C}G^{\mathbb{P}}_{\mathcal{FF},l})_{(i,j)} = G^{\mathbb{P}}_{\mathcal{FF},l,(i,j)} + \sum_{k=1}^{i-1} \mathcal{C}_{ik}(G^{\mathbb{P}}_{\mathcal{FF},l})_{(k,j)},$$

$$\begin{aligned}
\mathbb{E}[(\mathcal{C}G^{\mathbb{P}}_{\mathcal{FF},l})_{(i,j)}] &= \mathbb{E}[\mathbb{E}[(\mathcal{C}G^{\mathbb{P}}_{\mathcal{FF},l})_{(i,j)}|G^{\mathbb{P}}_{\mathcal{FF},l},\tilde{\Omega}_{\mathcal{FF}}]],\\
&= \mathbb{E}[\mathbb{E}[G^{\mathbb{P}}_{\mathcal{FF},l,(i,j)} + \sum_{k=1}^{i-1} \mathcal{C}_{ik}(G^{\mathbb{P}}_{\mathcal{FF},l})_{(k,j)}|G^{\mathbb{P}}_{\mathcal{FF},l},\tilde{\Omega}_{\mathcal{FF}}]],\\
&= \mathbb{E}[G^{\mathbb{P}}_{\mathcal{FF},l,(i,j)}].
\end{aligned}$$

Therefore, we can use prior means of $\{G^{\mathbb{P}}_{\mathcal{FF},l}\}_{l=1}^{p}$ as prior means of $\{\mathcal{C}G^{\mathbb{P}}_{\mathcal{FF},l}\}_{l=1}^{p}$. For the intercept term, it is

$$(\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}})_{(i)} = K^{\mathbb{P}}_{\mathcal{F},(i)} + \sum_{k=1}^{i-1} \mathcal{C}_{ik}(K^{\mathbb{P}}_{\mathcal{F}})_{(k)},$$

and the corresponding prior mean is

$$\begin{aligned}
\mathbb{E}[(\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}})_{(i)}] &= \mathbb{E}[\mathbb{E}[K^{\mathbb{P}}_{\mathcal{F},(i)} + \sum_{k=1}^{i-1} \mathcal{C}_{ik}(K^{\mathbb{P}}_{\mathcal{F}})_{(k)}|K^{\mathbb{P}}_{\mathcal{F}},\tilde{\Omega}_{ff}]],\\
&= \mathbb{E}[K^{\mathbb{P}}_{\mathcal{F},(i)}].
\end{aligned}$$

Therefore, we can use a reduced form prior for $K^{\mathbb{P}}_{\mathcal{F}}$ as a prior for $\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}}$. Here, we set a mean prior of $K^{\mathbb{P}}_{\mathcal{P}}$ to be same to that of $K^{\mathbb{Q}}_{\mathcal{P}}$. We can calculate a prior mean of $K^{\mathbb{Q}}_{\mathcal{P}}$ with a Monte Carlo simulation. The prior mean of the remaining elements in $K^{\mathbb{P}}_{\mathcal{F}}$ is set to zero.

A remaining thing is to set prior variances. Our prior variances are

$$\frac{Var[(\mathcal{C}G^{\mathbb{P}}_{\mathcal{FF},l})_{(i,j)}]}{\sigma^2_{\mathcal{FF},i}} = \underbrace{\psi_{l,ij}}_{local} \cdot \underbrace{\frac{q_{slope}}{l^{q_{lag}}}\frac{(\nu_0 - d_{\mathbb{P}} - 1)}{\Omega_{0,(j)}}}_{Minnesota},$$

$$\text{where } q_{slope} = \begin{cases} q_{11} & \text{if } i=j \text{ and } i \le d_{\mathbb{Q}},\\ q_{12} & \text{if } i=j \text{ and } i > d_{\mathbb{Q}},\\ q_{21} & \text{if } i\neq j \text{ and } i \le d_{\mathbb{Q}},\\ q_{22} & \text{otherwise,}\end{cases}$$

$$q_{lag} = \begin{cases} q_{31} & \text{if } i \leq d_{\mathbb{Q}}, \\ q_{32} & \text{otherwise,} \end{cases}$$

and $\psi_{l,ij} \backsim \mathcal{G}(\eta_\psi, \eta_\psi)$ or $\psi_{l,ij} = 1$ if we want to shut down a local shrinkage. The benchmark value for q_{lag} is two. If we introduce the local prior,

$$\eta_\psi \backsim \mathcal{G}(1, 1)$$

that reminds us that $\eta_\psi = 1$ corresponds to the Bayesian Lasso. In the case of the intercept,

$$\frac{Var[(\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}})_{(i)}]}{\sigma^2_{\mathcal{F}\mathcal{F},i}} = \underbrace{\psi_{0,i}}_{local} \cdot \underbrace{q_{intercept}}_{Minnesota},$$

$$\text{where } q_{intercept} = \begin{cases} q_{41} & \text{if } i \leq d_{\mathbb{Q}}, \\ q_{42} & \text{otherwise.} \end{cases}$$

The intercept and slope matrices are estimated with $\mathcal{C}$. Specifically, for

$$\phi_i = \begin{bmatrix} \begin{bmatrix} (\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}})_{(i)} & (\mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},1})_{(i,:)} & \cdots & (\mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},p})_{(i,:)} \end{bmatrix} & \begin{bmatrix} \mathcal{C}_{i1} & \cdots & \mathcal{C}_{i(i-1)} \end{bmatrix} \end{bmatrix}',$$

we express the ith row of orthogonalized system (41) as

$$\underbrace{\begin{bmatrix} \mathcal{F}_{1,(i)} \\ \mathcal{F}_{2,(i)} \\ \vdots \\ \mathcal{F}_{T,(i)} \end{bmatrix}}_{y_{\phi_i}} = X_{\phi_i}\phi_i + \mathcal{N}(O_{T\times 1}, \sigma^2_{\mathcal{F}\mathcal{F},i} I_T).$$

By estimating ϕ_i and $\sigma^2_{\mathcal{F}\mathcal{F},i}$ $\forall i$, we can derive a posterior sample of $\Omega_{\mathcal{F}\mathcal{F}}$, $K^{\mathbb{P}}_{\mathcal{F}}$ and $\{G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\}_{l=1}^{p}$. The corresponding prior distribution is

$$\phi_i \backsim \mathcal{N}(m_i, \sigma^2_{\mathcal{F}\mathcal{F},i} \cdot diag[V_i]).$$

5.1.3 prices of risks $k^{\mathbb{Q}}_{\infty}$

We observe $\mathbb{P}$-dynamics only once at a time, but we can observe $\mathbb{Q}$-dynamics several times contemporaneously through forward rates in the bond market. Therefore, $\mathbb{Q}$ dynamics is estimated stably without a sensitivity on prior settings. We impose

$$k^{\mathbb{Q}}_{\infty} \backsim \mathcal{N}(0, \sigma^2_{k^{\mathbb{Q}}_{\infty}}),$$
$$\sigma^2_{k^{\mathbb{Q}}_{\infty}} = 0.01^2.$$

5.1.4 Decay parameter $\kappa^{\mathbb{Q}}$

The DNS decay parameter is typically chosen to maximize the curvature factor loading at a specific maturity. The curvature factor loading is

$$\frac{1-\exp(-\kappa^{\mathbb{Q}}\tau)}{\kappa^{\mathbb{Q}}\tau}-\exp(-\kappa^{\mathbb{Q}}\tau).$$

We calculate $\kappa^{\mathbb{Q}}$ that maximize the curvature factor loading at each maturity $\tau\in\{24,30,\cdots,60\}$. Suppose the calculated decay parameters are $\{\kappa^{\mathbb{Q}}_{24},\kappa^{\mathbb{Q}}_{30}\cdots,\kappa^{\mathbb{Q}}_{60}\}$. We impose a uniform distribution on grid $\{\kappa^{\mathbb{Q}}_{24},\kappa^{\mathbb{Q}}_{30}\cdots,\kappa^{\mathbb{Q}}_{60}\}$ as a prior distribution for $\kappa^{\mathbb{Q}}$. A first derivative of the factor loading w.r.t. τ is

$$\frac{(\kappa^{\mathbb{Q}}\tau+1)\exp(-\kappa^{\mathbb{Q}}\tau)-1}{\kappa^{\mathbb{Q}}\tau^2}+\exp(-\kappa^{\mathbb{Q}}\tau)\kappa^{\mathbb{Q}}.$$

On the grid, the derivative should be zero for each maturity $\tau\in\{24,30,\cdots,60\}$.

5.1.5 Pricing error $\Sigma_{\mathcal{O}}$

We impose a hierarchical prior structure, so that

$$\begin{aligned}\sigma^2_{\mathcal{O},i}&\backsim\mathcal{IG}(2,\gamma_i),\\ \gamma_i&\backsim\mathcal{G}(1,\bar{\gamma}).\end{aligned}$$

To determine $\bar{\gamma}$, we regress the measurement equation (9) using OLS with unrestricted coefficients. And then, we set an inverse of $\bar{\gamma}$ to an average of the residual variances.

5.1.6 Summary

Our prior distribution is

$$\begin{aligned}&p(\theta\equiv\{\{\phi_i,\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{P}}},\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}},\{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}},\eta_\psi,k^{\mathbb{Q}}_\infty,\kappa^{\mathbb{Q}},\Sigma_{\mathcal{O}},\{\gamma_i\}_{i=1}^{N-d_{\mathbb{Q}}}\}|p,\nu_0,\Omega_0,q)\\ =&p(\{\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{P}}}|\nu_0,\Omega_0)p(\{\phi_i\}_{i=1}^{d_{\mathbb{P}}}|p,\{\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{P}}},q,\nu_0,\Omega_0,\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}},\{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}})\\ &\times p(\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}},\{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}}|p,\eta_\psi)p(\eta_\psi)p(k^{\mathbb{Q}}_\infty)p(\kappa^{\mathbb{Q}})p(\Sigma_{\mathcal{O}}|\{\gamma_i\}_{i=1}^{N-d_{\mathbb{Q}}})p(\{\gamma_i\}_{i=1}^{N-d_{\mathbb{Q}}}),\end{aligned}$$

where $q\equiv\begin{bmatrix}q_{slope} & q_{lag} & q_{intercept}\end{bmatrix}'$. The hyper-parameters are p, ν_0, Ω_0, and q. Our MH block contains $\{\phi_i,\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{Q}}}$ and η_ψ.

5.2 Likelihood

Our likelihood function is

$$\begin{aligned}p(\{\mathcal{O}_t,\mathcal{P}_t,M_t\}_{t=1}^{T}|\theta)&=\Pi_{t=1}^{T}p(\mathcal{O}_t,\mathcal{P}_t,M_t|\mathcal{I}_{t-1},\theta),\\ p(\mathcal{O}_t,\mathcal{P}_t,M_t|\mathcal{I}_{t-1},\theta)&=p(\mathcal{O}_t|\mathcal{P}_t,\theta)p(\mathcal{F}_t|\mathcal{I}_{t-1},\theta).\end{aligned}$$

We express the likelihood in terms of orthogonalized system (41), so

$$p(\mathcal{F}_{t,(1)}, \cdots, \mathcal{F}_{t,(d_{\mathbb{P}})}|\mathcal{I}_{t-1}, \theta) = p(\mathcal{F}_{t,(1)}|\mathcal{I}_{t-1}, \theta)\Pi_{i=2}^{d_{\mathbb{P}}} p(\mathcal{F}_{t,(i)}|\mathcal{F}_{t,(i-1)}, \cdots, \mathcal{F}_{t,(1)}, \mathcal{I}_{t-1}, \theta),$$

Therefore,

$$\begin{aligned} p(\{\mathcal{O}_t, \mathcal{P}_t, M_t\}_{t=1}^T|\theta) = &\Pi_{t=1}^T p(\mathcal{O}_t|\mathcal{P}_t, \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}, \Sigma_{\mathcal{O}}) \\ &\times \Pi_{t=1}^T p(\mathcal{F}_{t,(1)}|\mathcal{I}_{t-1}, \phi_1, \sigma^2_{\mathcal{F}\mathcal{F},1}) \\ &\times \Pi_{i=2}^{d_{\mathbb{P}}}[\Pi_{t=1}^T p(\mathcal{F}_{t,(i)}|\mathcal{F}_{t,(i-1)}, \cdots, \mathcal{F}_{t,(1)}, \mathcal{I}_{t-1}, \phi_i, \sigma^2_{\mathcal{F}\mathcal{F},i})]. \end{aligned}$$

The measurement equation is

$$\Pi_{t=1}^T p(\mathcal{O}_t|\mathcal{P}_t, \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}, \Sigma_{\mathcal{O}}) = \Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t|\mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}),$$

and the transition equation is

$$\Pi_{t=1}^T p(\mathcal{F}_{t,(i)}|\mathcal{F}_{t,(i-1)}, \cdots, \mathcal{F}_{t,(1)}, \mathcal{I}_{t-1}, \theta) = \mathcal{N}(y_{\phi_i}|X_{\phi_i}\phi_i, \sigma^2_{\mathcal{F}\mathcal{F},i} I_T).$$

Therefore,

$$\begin{aligned} p(\{\mathcal{O}_t, \mathcal{P}_t, M_t\}_{t=1}^T|\theta) = &\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t|\mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}) \\ &\times \Pi_{i=1}^{d_{\mathbb{P}}} \mathcal{N}(y_{\phi_i}|X_{\phi_i}\phi_i, \sigma^2_{\mathcal{F}\mathcal{F},i} I_T). \end{aligned}$$

5.3 Posterior distribution

We use the Gibbs sampling to sampling our joint posterior distribution.

5.3.1 Long-run level parameter $k_\infty^{\mathbb{Q}}$

The full-conditional posterior distribution is

$$p(k_\infty^{\mathbb{Q}}|\mathcal{I}_T, \theta\backslash\{k_\infty^{\mathbb{Q}}\}) \propto p(k_\infty^{\mathbb{Q}})\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t|\mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}).$$

We first simplify the measurement equation. A close inspection on the difference equations (39) and (40) leads to

$$a_\tau = a_{\tau-1} - \mathcal{J}_\tau(: \kappa^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}) + (\tau - 1)k_\infty^{\mathbb{Q}}.$$

So,

$$\begin{aligned} a_\tau &= 0, \text{for } \tau < 2 \\ a_\tau &= \underbrace{-\sum_{i=2}^{\tau} \mathcal{J}_i(: \kappa^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}})}_{a_{0\tau}} + \underbrace{\sum_{i=2}^{\tau}(i-1)}_{a_{1\tau}} k_\infty^{\mathbb{Q}}, \text{for } \tau \geq 2, \end{aligned}$$

and,

$$\mathcal{A}_X = \underbrace{\begin{bmatrix} a_{0\tau_1}/\tau_1 \\ a_{0\tau_2}/\tau_2 \\ \vdots \\ a_{0\tau_N}/\tau_N \end{bmatrix}}_{\mathcal{A}_{0,k_\infty^{\mathbb{Q}}}} + \underbrace{\begin{bmatrix} \sum_{i=2}^{\tau_1} \frac{(i-1)}{\tau_1} \\ \sum_{i=2}^{\tau_2} \frac{(i-1)}{\tau_2} \\ \vdots \\ \sum_{i=2}^{\tau_N} \frac{(i-1)}{\tau_N} \end{bmatrix}}_{\mathcal{A}_{1,k_\infty^{\mathbb{Q}}}} k_\infty^{\mathbb{Q}}.$$

Based on the result, the measurement equation can be rephrased as

$$\begin{aligned} \mathcal{O}_t &= \mathcal{A}_{\mathcal{P}} + \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c) + \mathcal{N}(O_{N-d_{\mathbb{Q}}}, \Sigma_{\mathcal{O}}), \\ &= W_{\mathcal{O}}(\mathcal{A}_X + \mathcal{B}_X \mathcal{T}_{0,\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c) + \mathcal{N}(O_{N-d_{\mathbb{Q}}}, \Sigma_{\mathcal{O}}), \\ &= W_{\mathcal{O}}(I_N - \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} W_{\mathcal{P}})\mathcal{A}_X + W_{\mathcal{O}} \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} c + \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c) + \mathcal{N}(O_{N-d_{\mathbb{Q}}}, \Sigma_{\mathcal{O}}), \\ &= W_{\mathcal{O}}(I_N - \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} W_{\mathcal{P}})(\mathcal{A}_{0,k_\infty^{\mathbb{Q}}} + \mathcal{A}_{1,k_\infty^{\mathbb{Q}}} k_\infty^{\mathbb{Q}}) + W_{\mathcal{O}} \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} c + \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c) + \mathcal{N}(O_{N-d_{\mathbb{Q}}}, \Sigma_{\mathcal{O}}), \end{aligned}$$

$$\begin{aligned} &\therefore \underbrace{\Sigma_{\mathcal{O}}^{-1/2}[\mathcal{O}_t - W_{\mathcal{O}}(I_N - \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} W_{\mathcal{P}})\mathcal{A}_{0,k_\infty^{\mathbb{Q}}} - W_{\mathcal{O}} \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} c - \mathcal{B}_{\mathcal{P}}(\mathcal{P}_t - c)]}_{y_{k_\infty^{\mathbb{Q}},t}} \\ &= \underbrace{\Sigma_{\mathcal{O}}^{-1/2} W_{\mathcal{O}}(I_N - \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} W_{\mathcal{P}})\mathcal{A}_{1,k_\infty^{\mathbb{Q}}}}_{x_{k_\infty^{\mathbb{Q}}}} k_\infty^{\mathbb{Q}} + \mathcal{N}(O_{(N-d_{\mathbb{Q}})\times 1}, I_{N-d_{\mathbb{Q}}}). \end{aligned}$$

Stacking up $y_{k_\infty^{\mathbb{Q}},t}$ leads to

$$\underbrace{\begin{bmatrix} y_{k_\infty^{\mathbb{Q}},1} \\ y_{k_\infty^{\mathbb{Q}},2} \\ \vdots \\ y_{k_\infty^{\mathbb{Q}},T} \end{bmatrix}}_{Y_{k_\infty^{\mathbb{Q}}}} = \underbrace{\begin{bmatrix} x_{k_\infty^{\mathbb{Q}}} \\ x_{k_\infty^{\mathbb{Q}}} \\ \vdots \\ x_{k_\infty^{\mathbb{Q}}} \end{bmatrix}}_{X_{k_\infty^{\mathbb{Q}}}} k_\infty^{\mathbb{Q}} + \mathcal{N}(O_{T(N-d_{\mathbb{Q}})\times 1}, I_{T(N-d_{\mathbb{Q}})}).$$

This derivation says that we can do the Normal-Normal update for $k_\infty^{\mathbb{Q}}$, so

$$p(k_\infty^{\mathbb{Q}}|\mathcal{I}_T, \theta\backslash\{k_\infty^{\mathbb{Q}}\}) = \mathcal{N}((X'_{k_\infty^{\mathbb{Q}}} X_{k_\infty^{\mathbb{Q}}} + \sigma_{k_\infty^{\mathbb{Q}}}^{-2})^{-1} X'_{k_\infty^{\mathbb{Q}}} Y_{k_\infty^{\mathbb{Q}}}, (X'_{k_\infty^{\mathbb{Q}}} X_{k_\infty^{\mathbb{Q}}} + \sigma_{k_\infty^{\mathbb{Q}}}^{-2})^{-1}).$$

We do not care about the Jacobian term because it does not depend on $k_\infty^{\mathbb{Q}}$.

5.3.2 Decay parameter $\kappa^{\mathbb{Q}}$

The unnormalized kernel of the full-conditional posterior distribution is

$$p^*(\kappa^{\mathbb{Q}} = \kappa_i^{\mathbb{Q}}|\mathcal{I}_T, \theta\backslash\{\kappa^{\mathbb{Q}}\}) = p(\kappa_i^{\mathbb{Q}})\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t|\mathcal{A}_{\mathcal{P}}(: \kappa_i^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{P}\mathcal{P}}) + \mathcal{B}_{\mathcal{P}}(: \kappa_i^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}).$$

Therefore, the posterior is

$$p(\kappa^{\mathbb{Q}} = \kappa_i^{\mathbb{Q}} | \mathcal{I}_T, \theta \backslash \{\kappa^{\mathbb{Q}}\}) = \frac{p^*(\kappa^{\mathbb{Q}} = \kappa_i^{\mathbb{Q}} | \mathcal{I}_T, \theta \backslash \{\kappa^{\mathbb{Q}}\})}{\sum_{j \in \{24,30,\cdots,60\}} p^*(\kappa^{\mathbb{Q}} = \kappa_j^{\mathbb{Q}} | \mathcal{I}_T, \theta \backslash \{\kappa^{\mathbb{Q}}\})},$$

for $i \in \{24, 30, \cdots, 60\}$.

5.3.3 MCMC algorithm for $\{\phi_i, \sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{Q}}}$, and η_ψ

The full-conditional posterior is

$$\begin{aligned}
&\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t | \mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{PP}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}) \times \Pi_{i=1}^{d_{\mathbb{P}}} \mathcal{N}(y_{\phi_i} | X_{\phi_i}\phi_i, \sigma^2_{\mathcal{FF},i} I_T) \\
&\times p(\{\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{P}}} | \nu_0, \Omega_0) p(\{\phi_i\}_{i=1}^{d_{\mathbb{P}}} | p, \{\sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{P}}}, q, \nu_0, \Omega_0, \{\{\psi_{l,ij}\}_{l=1}^p\}_{i,j=1}^{d_{\mathbb{P}}}, \{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}}) \\
&\times \Pi_{l=1}^p \Pi_{i=1}^{d_{\mathbb{P}}} \Pi_{j=1}^{d_{\mathbb{P}}} \mathcal{G}(\psi_{l,ij} | \eta_\psi, \eta_\psi) \times \Pi_{i=1}^{d_{\mathbb{P}}} \mathcal{G}(\psi_{0,i} | \eta_\psi, \eta_\psi) \times \mathcal{G}(\eta_\psi | 1, 1).
\end{aligned}$$

This MCMC sampling is divided into $d_{\mathbb{Q}} + 1$ MCMC blocks, that are $\{\phi_i, \sigma^2_{\mathcal{FF},i}\}$ for each $i \in \{1, \cdots, d_{\mathbb{Q}}\}$ and η_ψ.

Sampling $\{\phi_i, \sigma^2_{\mathcal{FF},i}\}_{i=1}^{d_{\mathbb{Q}}}$ For specific i, the targeted $\{\phi_i, \sigma^2_{\mathcal{FF},i}\}$ has a posterior distribution of

$$\begin{aligned}
&\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t | \mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{PP}}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})(\mathcal{P}_t - c), \Sigma_{\mathcal{O}}) \\
&\times \mathcal{N}(y_{\phi_i} | X_{\phi_i}\phi_i, \sigma^2_{\mathcal{FF},i} I_T) \times \mathcal{N}(\phi_i | m_i, \sigma^2_{\mathcal{FF},i} \cdot diag[V_i]) \times \mathcal{IG}(\sigma^2_{\mathcal{FF},i} | \frac{\nu_0 + i - d_{\mathbb{P}}}{2}, \frac{\Omega_{0,(i)}}{2}).
\end{aligned}$$

Here,

$$\mathcal{N}(y_{\phi_i} | X_{\phi_i}\phi_i, \sigma^2_{\mathcal{FF},i} I_T) \times \mathcal{N}(\phi_i | m_i, \sigma^2_{\mathcal{FF},i} \cdot diag[V_i]) \times \mathcal{IG}(\sigma^2_{\mathcal{FF},i} | \frac{\nu_0 + i - d_{\mathbb{P}}}{2}, \frac{\Omega_{0,(i)}}{2})$$

corresponds to the Normal-InverseGamma-Normal-InverseGamma update. We set it as a proposal distribution. Then, the acceptance probability is

$$min[1, \frac{\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t | \mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{PP},proposal}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})\mathcal{P}_t, \Sigma_{\mathcal{O}})}{\Pi_{t=1}^T \mathcal{N}(\mathcal{O}_t | \mathcal{A}_{\mathcal{P}}(: \kappa^{\mathbb{Q}}, k_\infty^{\mathbb{Q}}, \Omega_{\mathcal{PP},current}) + \mathcal{B}_{\mathcal{P}}(: \kappa^{\mathbb{Q}})\mathcal{P}_t, \Sigma_{\mathcal{O}})}].$$

Sampling η_ψ The log target distribution is

$$\begin{aligned}
&\log p(\eta_\psi | \mathcal{I}_T, \theta \backslash \{\eta_\psi\}) \\
=& \sum_{l=1}^p \sum_{i=1}^{d_{\mathbb{P}}} \sum_{j=1}^{d_{\mathbb{P}}} \log \mathcal{G}(\psi_{l,ij} | \eta_\psi, \eta_\psi) + \sum_{i=1}^{d_{\mathbb{P}}} \log \mathcal{G}(\psi_{0,i} | \eta_\psi, \eta_\psi) + \log \mathcal{G}(\eta_\psi | 1, 1) + \text{constant},
\end{aligned}$$

$$
\begin{aligned}
=&\sum_{l=1}^{p}\sum_{i=1}^{d_{\mathbb{P}}}\sum_{j=1}^{d_{\mathbb{P}}}\log[\frac{\eta_\psi^{\eta_\psi}}{\Gamma(\eta_\psi)}\psi_{l,ij}^{\eta_\psi-1}\exp[-\eta_\psi\psi_{l,ij}]]\\
&+\sum_{i=1}^{d_{\mathbb{P}}}\log[\frac{\eta_\psi^{\eta_\psi}}{\Gamma(\eta_\psi)}\psi_{0,i}^{\eta_\psi-1}\exp[-\eta_\psi\psi_{0,i}]]+\log[\frac{1}{\Gamma(1)}\exp[-\eta_\psi]]+\text{constant}\\
=&\sum_{l=1}^{p}\sum_{i=1}^{d_{\mathbb{P}}}\sum_{j=1}^{d_{\mathbb{P}}}[\eta_\psi\log\eta_\psi-\log\Gamma(\eta_\psi)+(\eta_\psi-1)\log\psi_{l,ij}-\eta_\psi\psi_{l,ij}]\\
&+\sum_{i=1}^{d_{\mathbb{P}}}[\eta_\psi\log\eta_\psi-\log\Gamma(\eta_\psi)+(\eta_\psi-1)\log\psi_{0,i}-\eta_\psi\psi_{0,i}]-\log\Gamma(1)-\eta_\psi+\text{constant},\\
=&(pd_{\mathbb{P}}^2+d_{\mathbb{P}})[\eta_\psi\log\eta_\psi-\log\Gamma(\eta_\psi)]+\sum_{\forall\psi}[(\eta_\psi-1)\log\psi-\eta_\psi\psi]-\log\Gamma(1)-\eta_\psi+\text{constant}.
\end{aligned}
$$

We derive an approximated target distribution and then use it as a proposal distribution. The first derivative is

$$
\frac{\partial\log p(\eta_\psi|\mathcal{I}_T,\theta\backslash\{\eta_\psi\})}{\partial\eta_\psi}=(pd_{\mathbb{P}}^2+d_{\mathbb{P}})[\log\eta_\psi+1-\frac{d\log\Gamma(\eta_\psi)}{d\eta_\psi}]+\sum_{\forall\psi}\log\psi-\sum_{\forall\psi}\psi-1,
$$

where $\frac{d\log\Gamma(\eta_\psi)}{d\eta_\psi}$ is a digamma function. The second derivative is

$$
\frac{\partial^2\log p(\eta_\psi|\mathcal{I}_T,\theta\backslash\{\eta_\psi\})}{(\partial\eta_\psi)^2}=(pd_{\mathbb{P}}^2+d_{\mathbb{P}})[\frac{1}{\eta_\psi}-\frac{d^2\log\Gamma(\eta_\psi)}{(d\eta_\psi)^2}],
$$

where $\frac{d^2\log\Gamma(\eta_\psi)}{(d\eta_\psi)^2}$ is a trigamma function. We use the first derivative to obtain the conditional posterior mode $\hat{\eta}_\psi$. And then, the hessian is evaluated to derive the variance of the proposal distribution. The proposal is a student-t distribution of

$$
t_{15}(\hat{\eta}_\psi,(-\frac{\partial^2\log p(\eta_\psi|\mathcal{I}_T,\theta\backslash\{\eta_\psi\})}{(\partial\eta_\psi)^2})^{-1}|_{\eta_\psi=\hat{\eta}_\psi}).
$$

5.3.4 Other transition equation parameters $\{\phi_i,\sigma^2_{\mathcal{FF},i}\}_{i=d_{\mathbb{Q}}+1}^{d_{\mathbb{P}}}$

The full-conditional distribution is

$$
\begin{aligned}
&\Pi_{i=d_{\mathbb{Q}}+1}^{d_{\mathbb{P}}}\mathcal{N}(y_{\phi_i}|X_{\phi_i}\phi_i,\sigma^2_{\mathcal{FF},i}I_T)\times p(\{\sigma^2_{\mathcal{FF},i}\}_{i=d_{\mathbb{Q}}+1}^{d_{\mathbb{P}}}|\nu_0,\Omega_0)\\
&\times p(\{\phi_i\}_{i=d_{\mathbb{Q}}+1}^{d_{\mathbb{P}}}|p,\{\sigma^2_{\mathcal{FF},i}\}_{i=d_{\mathbb{Q}}+1}^{d_{\mathbb{P}}},q,\nu_0,\Omega_0,\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}},\{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}})
\end{aligned}
$$

For $i\in\{d_{\mathbb{Q}}+1,\cdots,d_{\mathbb{P}}\}$, the kernel is

$$
\mathcal{N}(y_{\phi_i}|X_{\phi_i}\phi_i,\sigma^2_{\mathcal{FF},i}I_T)\times\mathcal{N}(\phi_i|m_i,\sigma^2_{\mathcal{FF},i}diag[V_i])\times\mathcal{IG}(\sigma^2_{\mathcal{FF},i}|\frac{\nu_0+i-d_{\mathbb{P}}}{2},\frac{\Omega_{0,(i)}}{2}),
$$

so it is the Normal-InverseGamma-Normal-InverseGamma update.

In the case of

$$\begin{aligned} y|\beta,\sigma^2 &= X\beta + \mathcal{N}(O, \sigma^2 I_T), \\ \beta|\sigma^2 &\backsim \mathcal{N}(\beta_0, \sigma^2 B_0), \\ \sigma^2 &\backsim \mathcal{IG}(\alpha_0, \delta_0), \end{aligned}$$

the posterior distribution is

$$\begin{aligned} \beta|\sigma^2, y &\backsim \mathcal{N}(\beta_1, \sigma^2 B_1), \\ \sigma^2|y &\backsim \mathcal{IG}(\alpha_0 + \frac{T}{2}, \delta_1), \end{aligned}$$

where $B_1 = (B_0^{-1} + X'X)^{-1}$, $\beta_1 = B_1(B_0^{-1}\beta_0 + X'y)$, and $\delta_1 = \delta_0 + (y'y + \beta_0' B_0^{-1}\beta_0 - \beta_1' B_1^{-1}\beta_1)/2$.

5.3.5 Sparsity parameter $\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}}$, and $\{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}}$

The kernel is

$$\begin{aligned} & p(\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}}, \{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}} | \mathcal{I}_T, \theta \backslash \{\{\{\psi_{l,ij}\}_{l=1}^{p}\}_{i,j=1}^{d_{\mathbb{P}}}, \{\psi_{0,i}\}_{i=1}^{d_{\mathbb{P}}}\}) \\ \propto & \Pi_{l=1}^{p} \Pi_{i,j=1}^{d_{\mathbb{P}}} [\mathcal{N}(\phi_{i,(1+(l-1)d_{\mathbb{P}}+j)} | m_{i,(1+(l-1)d_{\mathbb{P}}+j)}, \psi_{l,ij} \sigma^2_{\mathcal{FF},i} \frac{q_{slope}}{l^{q_{lag}}} \frac{(\nu_0 - d_{\mathbb{P}} - 1)}{\Omega_{0,(j)}}) \\ & \times \mathcal{G}(\psi_{l,ij} | \eta_\psi, \eta_\psi)] \times \Pi_{i=1}^{d_{\mathbb{P}}} [\mathcal{N}(\phi_{i,(1)} | m_{i,(1)}, \psi_{0,i} \sigma^2_{\mathcal{FF},i} q_{intercept}) \times \mathcal{G}(\psi_{0,i} | \eta_\psi, \eta_\psi)]. \end{aligned}$$

Each component in the brackets are updated to distributions of the Generalized Inverse Gaussian (GIG). In general notation, it is

$$\begin{aligned} \phi &\backsim \mathcal{N}(m, \psi V), \\ \psi &\backsim \mathcal{G}(\eta, \eta), \end{aligned}$$

the corresponding kernel is

$$p(\psi|\phi) \propto \psi^{\eta - \frac{1}{2} - 1} \exp[-\frac{1}{2}(2\eta\psi + \psi^{-1}\frac{(\phi - m)^2}{V})],$$

that is a kernel of

$$p(\psi|\phi) \propto \mathcal{GIG}(\eta - \frac{1}{2}, 2\eta, \frac{(\phi - m)^2}{V}).$$

5.3.6 Pricing error $\Sigma_{\mathcal{O}}$

The full-conditional distribution is

$$\Pi_{i=1}^{N-d_{\mathbb{Q}}} [\{\Pi_{t=1}^{T} \mathcal{N}(\mathcal{O}_{t,(i)} | \mathcal{A}_{\mathcal{P},(i)} + \mathcal{B}_{\mathcal{P},(i,:)}(\mathcal{P}_t - c), \sigma^2_{\mathcal{O},i})\} \mathcal{IG}(\sigma^2_{\mathcal{O},i} | 2, \gamma_i)]$$

$$=\Pi_{i=1}^{N-d_{\mathbb{Q}}}[\mathcal{N}(\underbrace{\begin{bmatrix}\mathcal{O}_{1,(i)}\\ \mathcal{O}_{2,(i)}\\ \vdots \\ \mathcal{O}_{T,(i)}\end{bmatrix}}_{y_{\sigma^2_{\mathcal{O},i}}} | \underbrace{\begin{bmatrix}\mathcal{A}_{\mathcal{P},(i)}+\mathcal{B}_{\mathcal{P},(i,:)}(\mathcal{P}_1-c)\\ \mathcal{A}_{\mathcal{P},(i)}+\mathcal{B}_{\mathcal{P},(i,:)}(\mathcal{P}_2-c)\\ \vdots \\ \mathcal{A}_{\mathcal{P},(i)}+\mathcal{B}_{\mathcal{P},(i,:)}(\mathcal{P}_T-c)\end{bmatrix}}_{m_{\sigma^2_{\mathcal{O},i}}}, \sigma^2_{\mathcal{O},i} I_T)\mathcal{IG}(\sigma^2_{\mathcal{O},i}|2,\gamma_i)].$$

Therefore, we can update the components in the brackets using the InverseGamma-InverseGamma update. For each i,

$$p(\sigma^2_{\mathcal{O},i}|\mathcal{I}_T,\theta\backslash\{\Sigma_{\mathcal{O}}\}) = \mathcal{IG}(\sigma^2_{\mathcal{O},i}|2+\frac{T}{2},\gamma_i+\frac{(y_{\sigma^2_{\mathcal{O},i}}-m_{\sigma^2_{\mathcal{O},i}})'(y_{\sigma^2_{\mathcal{O},i}}-m_{\sigma^2_{\mathcal{O},i}})}{2}).$$

5.3.7 Population measurement error γ

The corresponding kernel is

$$\begin{aligned}&\Pi_{i=1}^{N-d_{\mathbb{Q}}} p(\sigma^2_{\mathcal{O},i}|\gamma_i)p(\gamma_i)\\ =&\Pi_{i=1}^{N-d_{\mathbb{Q}}}\mathcal{IG}(\sigma^2_{\mathcal{O},i}|2,\gamma_i)\mathcal{G}(\gamma_i|1,\bar{\gamma}).\end{aligned}$$

Therefore, there is a posterior independence. For each i,

$$\begin{aligned}&\mathcal{IG}(\sigma^2_{\mathcal{O},i}|2,\gamma_i)\mathcal{G}(\gamma_i|1,\bar{\gamma})\\ =&\frac{\gamma_i^2}{\Gamma(2)}(\sigma^2_{\mathcal{O},i})^{-(2+1)}\exp[-\sigma^{-2}_{\mathcal{O},i}\gamma_i]\frac{\bar{\gamma}}{\Gamma(1)}\exp[-\bar{\gamma}\gamma_i],\\ \propto&\gamma_i^{3-1}\exp[-(\sigma^{-2}_{\mathcal{O},i}+\bar{\gamma})\gamma_i].\end{aligned}$$

Therefore,

$$p(\gamma_i|\mathcal{I}_T,\theta\backslash\{\gamma_i\}_{i=1}^{N-d_{\mathbb{Q}}}) = \mathcal{G}(3,\frac{1}{\sigma^2_{\mathcal{O},i}}+\bar{\gamma}).$$

5.4 Empirical Bayes and the marginal likelihood

Hyperparameters, p, ν_0, Ω_0, and q, should be calibrated. We calibrate them in the perspective of the empirical Bayes. We maximize the marginal likelihood for the $\mathbb{P}$-transition equation by calibrating p, ν_0, Ω_0, and q, given $\forall\{\psi_{l,ij},\psi_{0,i}\}=1$. The log marginal likelihood is

$$\begin{aligned}&\sum_{i=1}^{d_{\mathbb{P}}}\log\mathcal{N}(y_{\phi_i}|p,\nu_0,\Omega_0,q,\psi_{l,ij},\psi_{0,i})\\ =&-\frac{Td_{\mathbb{P}}}{2}\log[2\pi]+\sum_{i=1}^{d_{\mathbb{P}}}\{-\frac{1}{2}(\sum_{j=1}^{\#\phi_i}\log[V_{i,(j)}]+\log\det[K_{\phi_i}])\end{aligned}$$

$$+\log\Gamma(\nu_i+\frac{T}{2})+\nu_i\log S_i-\log\Gamma(\nu_i)-(\nu_i+\frac{T}{2})\log[\hat{S}_i]\}.$$

Here, $\nu_i=(\nu_0+i-d_{\mathbb{P}})/2$, $S_i=\Omega_{0,(i)}/2$, $K_{\phi_i}=diag[V_i]^{-1}+X'_{\phi_i}X_{\phi_i}$, $\hat{\phi}_i=K_{\phi_i}^{-1}(diag[V_i]^{-1}m_i+X'_{\phi_i}y_{\phi_i})$, $\hat{S}_i=S_i+(y'_{\phi_i}y_{\phi_i}+m'_i diag[V_i]^{-1}m_i-\hat{\phi}'_iK_{\phi_i}\hat{\phi}_i)/2$.

When we maximize the marginal likelihood, we constrain the maximum Sharpe ratio. To do that, we use the below proposition.

Proposition 3. *In our specification, the second-order Taylor series approximates the conditional expectation for the maximum Sharpe ratio as*

$$\begin{aligned}&\mathbb{E}[\sqrt{\lambda'_t\lambda_t}|\{\mathcal{F}_t\}_{t=1}^T,\Omega_{\mathcal{F}\mathcal{F}},k_\infty^{\mathbb{Q}},\kappa^{\mathbb{Q}}]\\ \approx&\frac{\sqrt{\pi}}{2}\left[\frac{(tr[B_{\lambda_t}]+\beta'_{\lambda_t}\beta_{\lambda_t})^{0.5}}{\Gamma(1.5)}+\frac{(tr[B_{\lambda_t}]+\beta'_{\lambda_t}\beta_{\lambda_t})^{-1.5}(2tr[B_{\lambda_t}^2]+4\beta'_{\lambda_t}B_{\lambda_t}\beta_{\lambda_t})}{2\Gamma(-0.5)}\right].\end{aligned}$$

If we assume $\Omega_{\mathcal{F}\mathcal{F}}$ as a diagonal, the expressions for β_{λ_t} and B_{λ_t} get simplified. Also, under diagonal $\Omega_{\mathcal{F}\mathcal{F}}$,

$$\begin{aligned}&\mathbb{E}[\sqrt{(K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}})'\Omega_{\mathcal{F}\mathcal{F}}^{-1}(K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}})}|\{\mathcal{F}_t\}_{t=1}^T,\Omega_{\mathcal{F}\mathcal{F}},k_\infty^{\mathbb{Q}},\kappa^{\mathbb{Q}}]\\ \approx&\frac{\sqrt{\pi}}{2}\left[\frac{(tr[B_{const}]+\beta'_{const}\beta_{const})^{0.5}}{\Gamma(1.5)}+\frac{(tr[B_{const}]+\beta'_{const}\beta_{const})^{-1.5}(2tr[B_{const}^2]+4\beta'_{const}B_{const}\beta_{const})}{2\Gamma(-0.5)}\right].\end{aligned}$$

Proof. See Appendix B. □

5.5 Sparsity on the covariance matrix

We post-processing posterior samples of $\Omega_{\mathcal{F}\mathcal{F}}$ using the Lasso penalty. Especially, for i'th sample $\Omega_{\mathcal{F}\mathcal{F}}^{(i)}$,

$$\hat{\Omega}_{\mathcal{F}\mathcal{F}}^{-1(i)}=\arg\min_X\{\mathrm{tr}[X\Omega_{\mathcal{F}\mathcal{F}}^{(i)}]-\log\det[X]+\omega_{(i)}\sum_{j,k}\frac{|X_{(j,k)}|}{|(\Omega_{\mathcal{F}\mathcal{F}}^{-1(i)})_{(j,k)}|}\}$$

is the corresponding lasso estimate. To choose $\omega_{(i)}$, we minimize the EBIC.

5.6 Predictions and Scenario Analysis

The state space is

$$\begin{bmatrix}R_t\\M_t\end{bmatrix}=\begin{bmatrix}\mathcal{A}_X+\mathcal{B}_X\mathcal{T}_{0,\mathcal{P}}\\O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1}\end{bmatrix}+\begin{pmatrix}\mathcal{B}_X\mathcal{T}_{1,\mathcal{P}}&O_{N\times(d_{\mathbb{P}}-d_{\mathbb{Q}})}\\O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times d_{\mathbb{Q}}}&I_{(d_{\mathbb{P}}-d_{\mathbb{Q}})}\end{pmatrix}\mathcal{F}_t+\begin{bmatrix}e_{R,t}\\O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1}\end{bmatrix},$$

$$\mathcal{F}_t=K_{\mathcal{F}}^{\mathbb{P}}+\sum_{l=1}^{p}G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}}\mathcal{F}_{t-l}+\epsilon_{\mathcal{F},t}^{\mathbb{P}}.$$

Since

$$\begin{bmatrix} O_{d_{\mathbb{Q}}\times 1} \\ e_{\mathcal{O},t} \end{bmatrix} := \begin{pmatrix} W_{\mathcal{P}} \\ W_{\mathcal{O}} \end{pmatrix} e_{R,t} \backsim \mathcal{N}(O_{N\times 1}, \begin{pmatrix} O_{d_{\mathbb{Q}}\times d_{\mathbb{Q}}} & O_{d_{\mathbb{Q}}\times (N-d_{\mathbb{Q}})} \\ O_{(N-d_{\mathbb{Q}})\times d_{\mathbb{Q}}} & \Sigma_{\mathcal{O}} \end{pmatrix}),$$

then

$$e_{R,t} = \begin{pmatrix} W_{\mathcal{O}} \\ W_{\mathcal{P}} \end{pmatrix}^{-1} \begin{bmatrix} e_{\mathcal{O},t} \\ O_{d_{\mathbb{Q}}\times 1} \end{bmatrix}.$$

Let's think a series of scenarios $\{\mathcal{S}_h, s_h\}_{h=T+1}^{T+d_h}$ that

$$\underset{(d_{s_h}\times(N+d_{\mathbb{P}}-d_{\mathbb{Q}}))}{\mathcal{S}_h} \begin{bmatrix} R_h \\ M_h \end{bmatrix} = \underset{(d_{s_h}\times 1)}{s_h},$$

where h is a forecasting horizon. Given the set $\{\mathcal{S}_h, s_h\}_{h=T+1}^{T+d_h}$,

$$s_h = \mathcal{S}_h \begin{bmatrix} \mathcal{A}_X + \mathcal{B}_X \mathcal{T}_{0,\mathcal{P}} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1} \end{bmatrix} + \mathcal{S}_h \begin{pmatrix} \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} & O_{N\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times d_{\mathbb{Q}}} & I_{(d_{\mathbb{P}}-d_{\mathbb{Q}})} \end{pmatrix} \mathcal{F}_h + \mathcal{S}_{h,(:,1:N)} \left(\begin{pmatrix} W_{\mathcal{O}} \\ W_{\mathcal{P}} \end{pmatrix}^{-1} \right)_{(:,1:N-d_{\mathbb{Q}})} e_{\mathcal{O},t},$$

$$\mathcal{F}_h = K_{\mathcal{F}}^{\mathbb{P}} + \sum_{l=1}^{p} G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}} \mathcal{F}_{h-l} + \epsilon_{\mathcal{F},h}^{\mathbb{P}}.$$

The state-space form is

$$s_h = \mathcal{S}_h \begin{bmatrix} \mathcal{A}_X + \mathcal{B}_X \mathcal{T}_{0,\mathcal{P}} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times 1} \end{bmatrix}$$

$$+ \mathcal{S}_h \begin{pmatrix} \mathcal{B}_X \mathcal{T}_{1,\mathcal{P}} & O_{N\times(d_{\mathbb{P}}-d_{\mathbb{Q}})} & \left(\begin{pmatrix} W_{\mathcal{O}} \\ W_{\mathcal{P}} \end{pmatrix}^{-1} \right)_{(:,1:N-d_{\mathbb{Q}})} & O_{N\times(d_{\mathbb{P}}p-d_{\mathbb{P}})} \\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times d_{\mathbb{Q}}} & I_{(d_{\mathbb{P}}-d_{\mathbb{Q}})} & O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times(N-d_{\mathbb{Q}})} & O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times(d_{\mathbb{P}}p-d_{\mathbb{P}})} \end{pmatrix} \begin{bmatrix} \mathcal{F}_h \\ e_{\mathcal{O},h} \\ \mathcal{F}_{h-1} \\ \mathcal{F}_{h-2} \\ \vdots \\ \mathcal{F}_{h-p+1} \end{bmatrix},$$

$$\begin{bmatrix} \mathcal{F}_h \\ e_{\mathcal{O},h} \\ \mathcal{F}_{h-1} \\ \mathcal{F}_{h-2} \\ \vdots \\ \mathcal{F}_{h-p+1} \end{bmatrix} = \begin{pmatrix} G_{\mathcal{F}\mathcal{F},1}^{\mathbb{P}} & O_{d_{\mathbb{P}}\times(N-d_{\mathbb{Q}})} & G_{\mathcal{F}\mathcal{F},2}^{\mathbb{P}} & \cdots & G_{\mathcal{F}\mathcal{F},p-1}^{\mathbb{P}} & G_{\mathcal{F}\mathcal{F},p}^{\mathbb{P}} \\ O_{(N-d_{\mathbb{Q}})\times d_{\mathbb{P}}} & O_{(N-d_{\mathbb{Q}})\times(N-d_{\mathbb{Q}})} & O_{(N-d_{\mathbb{Q}})\times d_{\mathbb{P}}} & \cdots & O_{(N-d_{\mathbb{Q}})\times d_{\mathbb{P}}} & O_{(N-d_{\mathbb{Q}})\times d_{\mathbb{P}}} \\ I_{d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times(N-d_{\mathbb{Q}})} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & \cdots & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \\ O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times(N-d_{\mathbb{Q}})} & I_{d_{\mathbb{P}}} & \cdots & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times(N-d_{\mathbb{Q}})} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & \cdots & I_{d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \end{pmatrix} \begin{bmatrix} \mathcal{F}_{h-1} \\ e_{\mathcal{O},h-1} \\ \mathcal{F}_{h-2} \\ \mathcal{F}_{h-3} \\ \vdots \\ \mathcal{F}_{h-p} \end{bmatrix}$$

$$
+\begin{bmatrix} K^{\mathbb{P}}_{\mathcal{F}} \\ O_{(N-d_{\mathbb{Q}})\times 1} \\ O_{d_{\mathbb{P}}\times 1} \\ O_{d_{\mathbb{P}}\times 1} \\ \vdots \\ O_{d_{\mathbb{P}}\times 1} \end{bmatrix} + \begin{bmatrix} \mathcal{N}(O_{d_{\mathbb{P}}\times 1}, \Omega_{\mathcal{F}\mathcal{F}}) \\ \mathcal{N}(O_{(N-d_{\mathbb{Q}})\times 1}, \Sigma_{\mathcal{O}}) \\ O_{d_{\mathbb{P}}\times 1} \\ O_{d_{\mathbb{P}}\times 1} \\ \vdots \\ O_{d_{\mathbb{P}}\times 1} \end{bmatrix}.
$$

Using the above state space and the Kalman filter, we can derive a distribution of $\{\mathcal{P}_h - c, M_h, e_{R,t}, R_h, TP_{\tau,h}\}_{h=T+1}^{\infty}$. From $h = T + d_h + 1$, the posterior distribution is sampled by only using the predictive part in the Kalman filter.

A Proof of Proposition 1

The restricted JSZ form is

$$
\begin{aligned}
r_t &= \iota' X_t, \\
X_t &= \underbrace{\begin{bmatrix} k^{\mathbb{Q}}_{\infty} \\ 0 \\ 0 \end{bmatrix}}_{K^{\mathbb{Q}}_X} + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp[-\kappa^{\mathbb{Q}}] & 1 \\ 0 & 0 & \exp[-\kappa^{\mathbb{Q}}] \end{pmatrix}}_{G^{\mathbb{Q}}_{XX}} X_{t-1} + \epsilon^{\mathbb{Q}}_{X,t}, \\
X_t &= K^{\mathbb{P}}_X + \sum_{l=1}^{p} G^{\mathbb{P}}_{XX,l} X_{t-l} + \sum_{l=1}^{p} G^{\mathbb{P}}_{XM,l} M_{t-l} + \epsilon^{\mathbb{P}}_{X,t}, \\
M_t &= K^{\mathbb{P}}_M + \sum_{l=1}^{p} G^{\mathbb{P}}_{MX,l} X_{t-l} + \sum_{l=1}^{p} G^{\mathbb{P}}_{MM,l} M_{t-l} + \epsilon^{\mathbb{P}}_{M,t}.
\end{aligned}
$$

Rotating X_t using

$$
X^*_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\kappa^{\mathbb{Q}}}{1-\exp[-\kappa^{\mathbb{Q}}]} & \frac{1+2\kappa^{\mathbb{Q}} - \exp[\kappa^{\mathbb{Q}}] - \kappa^{\mathbb{Q}}\exp[-\kappa^{\mathbb{Q}}]}{(1-\exp[-\kappa^{\mathbb{Q}}])^2} \\ 0 & 0 & \frac{\exp[\kappa^{\mathbb{Q}}]}{1-\exp[-\kappa^{\mathbb{Q}}]} \end{pmatrix} X_t
$$

yields

$$
\begin{aligned}
r_t &= \underbrace{\begin{bmatrix} 1 & \frac{1-\exp(-\kappa^{\mathbb{Q}})}{\kappa^{\mathbb{Q}}} & \frac{1-\exp(-\kappa^{\mathbb{Q}})}{\kappa^{\mathbb{Q}}} - \exp(-\kappa^{\mathbb{Q}}) \end{bmatrix}}_{\beta^{*\prime}} X^*_t, \\
X^*_t &= K^{\mathbb{Q}}_X + \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \exp[-\kappa^{\mathbb{Q}}] & \kappa^{\mathbb{Q}}\exp[-\kappa^{\mathbb{Q}}] \\ 0 & 0 & \exp[-\kappa^{\mathbb{Q}}] \end{pmatrix}}_{G^{\mathbb{Q}}_{X^*X^*}} X^*_{t-1} + \epsilon^{\mathbb{Q}}_{X^*,t},
\end{aligned}
$$

In a companion form, the $\mathbb{P}$-dynamics is

$$\begin{bmatrix} \begin{bmatrix} X_t^* \\ M_t \end{bmatrix} \\ \begin{bmatrix} X_{t-1}^* \\ M_{t-1} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} X_{t-p+1}^* \\ M_{t-p+1} \end{bmatrix} \end{bmatrix} = \underbrace{\begin{bmatrix} \begin{bmatrix} K_{X^*}^{\mathbb{P}} \\ K_M^{\mathbb{P}} \end{bmatrix} \\ O_{d_{\mathbb{P}}(p-1)\times 1} \end{bmatrix}}_{K^{\mathbb{P}}} + G^{\mathbb{P}} \begin{bmatrix} \begin{bmatrix} X_{t-1}^* \\ M_{t-1} \end{bmatrix} \\ \begin{bmatrix} X_{t-2}^* \\ M_{t-2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} X_{t-p}^* \\ M_{t-p} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} \epsilon_{X^*,t}^{\mathbb{P}} \\ \epsilon_{M,t}^{\mathbb{P}} \end{bmatrix} \\ O_{d_{\mathbb{P}}(p-1)\times 1} \end{bmatrix},$$

where

$$G^{\mathbb{P}} \equiv \begin{pmatrix} \begin{pmatrix} G_{X^*X^*,1}^{\mathbb{P}} & G_{X^*M,1}^{\mathbb{P}} \\ G_{MX^*,1}^{\mathbb{P}} & G_{MM,1}^{\mathbb{P}} \end{pmatrix} & \begin{pmatrix} G_{X^*X^*,2}^{\mathbb{P}} & G_{X^*M,2}^{\mathbb{P}} \\ G_{MX^*,2}^{\mathbb{P}} & G_{MM,2}^{\mathbb{P}} \end{pmatrix} & \cdots & \begin{pmatrix} G_{X^*X^*,p-1}^{\mathbb{P}} & G_{X^*M,p-1}^{\mathbb{P}} \\ G_{MX^*,p-1}^{\mathbb{P}} & G_{MM,p-1}^{\mathbb{P}} \end{pmatrix} & \begin{pmatrix} G_{X^*X^*,p}^{\mathbb{P}} & G_{X^*M,p}^{\mathbb{P}} \\ G_{MX^*,p}^{\mathbb{P}} & G_{MM,p}^{\mathbb{P}} \end{pmatrix} \\ I_{d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & \cdots & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \\ O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & I_{d_{\mathbb{P}}} & \cdots & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} & \cdots & I_{d_{\mathbb{P}}} & O_{d_{\mathbb{P}}\times d_{\mathbb{P}}} \end{pmatrix}.$$

And then, relocating the factors using

$$\underbrace{\begin{bmatrix} \begin{bmatrix} X_t^{**} \\ M_t^{**} \end{bmatrix} \\ \begin{bmatrix} X_{t-1}^{**} \\ M_{t-1}^{**} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} X_{t-p+1}^{**} \\ M_{t-p+1}^{**} \end{bmatrix} \end{bmatrix}}_{F_t^{**}} = \begin{bmatrix} \begin{bmatrix} X_t^* \\ M_t \end{bmatrix} \\ \begin{bmatrix} X_{t-1}^* \\ M_{t-1} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} X_{t-p+1}^* \\ M_{t-p+1} \end{bmatrix} \end{bmatrix} + \underbrace{(G^{\mathbb{P}} - I)^{-1} K^{\mathbb{P}}}_{z^{\mathbb{P}}}$$

leads to

$$F_t^{**} - (G^{\mathbb{P}} - I)^{-1} K^{\mathbb{P}} = K^{\mathbb{P}} + (G^{\mathbb{P}} - I + I)(F_{t-1}^{**} - (G^{\mathbb{P}} - I)^{-1} K^{\mathbb{P}}) + \underbrace{\begin{bmatrix} \begin{bmatrix} \epsilon_{X^*,t}^{\mathbb{P}} \\ \epsilon_{M,t}^{\mathbb{P}} \end{bmatrix} \\ O_{d_{\mathbb{P}}(p-1)\times 1} \end{bmatrix}}_{\epsilon_t^{\mathbb{P}**}},$$

$$F_t^{**} = G^{\mathbb{P}} F_{t-1}^{**} + \epsilon_t^{\mathbb{P}**}.$$

Therefore, we can erase the intercept in $\mathbb{P}$-dynamics.

Except for the case of a singularity, $z^{\mathbb{P}}$ would be a dense vector. Then, the transformation is

$$X_t^{**} = X_t^* + \underbrace{\begin{bmatrix} z_1^{\mathbb{P}} \\ z_2^{\mathbb{P}} \\ z_3^{\mathbb{P}} \end{bmatrix}}_{z_{(1:3)}^{\mathbb{P}}}.$$

The short-term interest rate is

$$\begin{aligned} r_t &= \beta^{*\prime} X_t^*, \\ &= \underbrace{-\beta^{*\prime} z_{(1:3)}^{\mathbb{P}}}_{\delta^{**}} + \beta^{*\prime} X_t^{**}. \end{aligned}$$

It is an affine form.

Lastly, $\mathbb{Q}$-dynamics is

$$\begin{aligned} X_t^{**} - z_{(1:3)}^{\mathbb{P}} &= K_X^{\mathbb{Q}} + G_{X^*X^*}^{\mathbb{Q}}(X_{t-1}^{**} - z_{(1:3)}^{\mathbb{P}}) + \epsilon_{X^*,t}^{\mathbb{Q}}, \\ X_t^{**} &= \underbrace{(K_X^{\mathbb{Q}} + (I_{d_{\mathbb{Q}}} - G_{X^*X^*}^{\mathbb{Q}}) z_{(1:3)}^{\mathbb{P}})}_{K_{X^{**}}^{\mathbb{Q}}} + G_{X^*X^*}^{\mathbb{Q}} X_{t-1}^{**} + \epsilon_{X^*,t}^{\mathbb{Q}}. \end{aligned}$$

Vector $K_{X^{**}}^{\mathbb{Q}}$ is a dense vector, because

$$\begin{aligned} K_{X^{**}}^{\mathbb{Q}} &= K_X^{\mathbb{Q}} + (I_{d_{\mathbb{Q}}} - G_{X^*X^*}^{\mathbb{Q}}) z_{(1:3)}^{\mathbb{P}}, \\ &= \begin{bmatrix} k_\infty^{\mathbb{Q}} \\ 0 \\ 0 \end{bmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1-\exp(-\kappa^{\mathbb{Q}}) & -\kappa^{\mathbb{Q}}\exp(-\kappa^{\mathbb{Q}}) \\ 0 & 0 & 1-\exp(-\kappa^{\mathbb{Q}}) \end{pmatrix} \begin{bmatrix} z_1^{\mathbb{P}} \\ z_2^{\mathbb{P}} \\ z_3^{\mathbb{P}} \end{bmatrix} \\ &= \begin{bmatrix} k_\infty^{\mathbb{Q}} \\ z_2^{\mathbb{P}}[1-\exp(-\kappa^{\mathbb{Q}})] - z_3^{\mathbb{P}}\kappa^{\mathbb{Q}}\exp(-\kappa^{\mathbb{Q}}) \\ z_3^{\mathbb{P}}[1-\exp(-\kappa^{\mathbb{Q}})] \end{bmatrix}. \end{aligned}$$

Since there are sufficient degree of freedoms in expressions of δ^{**} and $K_{X^{**}}^{\mathbb{Q}}$, these parameters are free-parameters. Also, check that we cannot make $K_{X^{**}}^{\mathbb{Q}}$ to a zero vector, because there is no such $z^{\mathbb{P}}$ that make the first element in $K_{X^{**}}^{\mathbb{Q}}$ zero. In other words, $K_{X^{**}}^{\mathbb{Q}} = O_{d_{\mathbb{Q}}\times 1}$ is an over-identification.

B Proof of Proposition 3

B.1 $\Omega_{\mathcal{F}\mathcal{F}}$ is non-diagonal

Our market prices of risks specification (17) is

$$\lambda_t = \Omega_{\mathcal{F}\mathcal{F}}^{-1/2}[(K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}}) + \sum_{l=1}^{p}(G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F},l}^{\mathbb{Q}})\mathcal{F}_{t+1-l}].$$

We first derive the conditional posterior distribution of $\{\mathcal{C}K^{\mathbb{P}}_{\mathcal{F}}, \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\}$ using a likelihood of equation (41),

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \mathcal{C}_{21} & 1 & 0 & \cdots & 0 \\ \mathcal{C}_{31} & \mathcal{C}_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_{d_{\mathbb{P}}1} & \mathcal{C}_{d_{\mathbb{P}}2} & \mathcal{C}_{d_{\mathbb{P}}3} & \cdots & 1 \end{pmatrix}}_{\mathcal{C}} \mathcal{F}_t = \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} + \sum_{l=1}^{p} \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},l}\mathcal{F}_{t-l} + diag(\underbrace{\begin{bmatrix} \sigma_{\mathcal{F}\mathcal{F},1} \\ \vdots \\ \sigma_{\mathcal{F}\mathcal{F},d_{\mathbb{P}}} \end{bmatrix}}_{\tilde{\Omega}^{1/2}_{\mathcal{F}\mathcal{F}}}) v^{\mathbb{P}}_{\mathcal{F},t}.$$

Since we know $\Omega_{\mathcal{F}\mathcal{F}}$ and $\{\mathcal{F}_t\}_{t=1}^T$, $\{\mathcal{C}\mathcal{F}_t\}_{t=1}^T$ and $\{\sigma_{\mathcal{F}\mathcal{F},i}\}_{i=1}^{d_{\mathbb{P}}}$ are also known. Since the above system has a diagonal covariance matrix, parameters in

$$\begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}_{(i,:)} \equiv \begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},1} & \cdots & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},p} \end{bmatrix}_{(i,:)}$$

are posteriorly independent across i. By Normal-Normal update, we can derive the conditional posterior distribution as

$$p(\begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}_{(i,:)} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k^{\mathbb{Q}}_{\infty}, \kappa^{\mathbb{Q}}) \backsim \mathcal{N}(\beta_{CKCG,i}, B_{CKCG,i}),$$

where

$$B_{CKCG,i} \equiv (diag[V_{i,(1:1+d_{\mathbb{P}}p)}]^{-1} + \sigma^{-2}_{\mathcal{F}\mathcal{F},i} X'_{\phi_i,(1:1+d_{\mathbb{P}}p)} X_{\phi_i,(1:1+d_{\mathbb{P}}p)})^{-1},$$
$$\beta_{CKCG,i} \equiv B_{CKCG,i}(diag[V_{i,(1:1+d_{\mathbb{P}}p)}]^{-1} m_{i,(1:1+d_{\mathbb{P}}p)} + \sigma^{-2}_{\mathcal{F}\mathcal{F},i} X'_{\phi_i,(1:1+d_{\mathbb{P}}p)} \begin{bmatrix} (\mathcal{C}\mathcal{F}_1)_{(i)} \\ \vdots \\ (\mathcal{C}\mathcal{F}_T)_{(i)} \end{bmatrix}).$$

And then, we derive the posterior distribution of the reduced form parameters. To do that, we check the relationship between two types of parameters,

$$\begin{aligned} vec\left[\begin{bmatrix} K^{\mathbb{P}}_{\mathcal{F}} & G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}'\right] &= vec\left[\begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}' \mathcal{C}^{-1\prime}\right], \\ &= (\mathcal{C}^{-1} \otimes I_{1+d_{\mathbb{P}}p}) vec\left[\begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}'\right], \end{aligned}$$

where $G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \equiv \begin{pmatrix} G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},1} & \cdots & G^{\mathbb{P}}_{\mathcal{F}\mathcal{F},p} \end{pmatrix}$. The distribution of $vec\left[\begin{bmatrix} \mathcal{C}K^{\mathbb{P}}_{\mathcal{F}} & \mathcal{C}G^{\mathbb{P}}_{\mathcal{F}\mathcal{F}} \end{bmatrix}'\right]$ conditional on $\{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k^{\mathbb{Q}}_{\infty}, \kappa^{\mathbb{Q}}$ is

$$\mathcal{N}\left(\beta_{CKCG} \equiv \begin{bmatrix} \beta_{CKCG,1} \\ \vdots \\ \beta_{CKCG,d_{\mathbb{P}}} \end{bmatrix}, B_{CKCG} \equiv \begin{pmatrix} B_{CKCG,1} & \cdots & O_{(d_{\mathbb{P}}p+1)\times(d_{\mathbb{P}}p+1)} \\ \vdots & \ddots & \vdots \\ O_{(d_{\mathbb{P}}p+1)\times(d_{\mathbb{P}}p+1)} & \cdots & B_{CKCG,d_{\mathbb{P}}} \end{pmatrix}\right).$$

Therefore, the distribution of $vec\left[\begin{bmatrix}K_{\mathcal{F}}^{\mathbb{P}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}\end{bmatrix}'|\{\mathcal{F}_t\}_{t=1}^T,\Omega_{\mathcal{F}\mathcal{F}},k_\infty^{\mathbb{Q}},\kappa^{\mathbb{Q}}\right]$ is

$$\mathcal{N}\left(\beta_{KG}\equiv(\mathcal{C}^{-1}\otimes I_{1+d_{\mathbb{P}}p})\beta_{CKCG},B_{KG}\equiv(\mathcal{C}^{-1}\otimes I_{1+d_{\mathbb{P}}p})B_{CKCG}(\mathcal{C}^{-1\prime}\otimes I_{1+d_{\mathbb{P}}p})\right).$$

Next, by the definition of the $\mathbb{Q}$-dynamics,

$$\begin{aligned}&vec\left[\begin{bmatrix}K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}-G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}}\end{bmatrix}'\right]\\ =&vec\left[\begin{bmatrix}K_{\mathcal{P}}^{\mathbb{P}}-K_{\mathcal{P}}^{\mathbb{Q}} & G_{\mathcal{P}\mathcal{P},1}^{\mathbb{P}}-G_{\mathcal{P}\mathcal{P}}^{\mathbb{Q}} & G_{\mathcal{P}\mathcal{M},1}^{\mathbb{P}} & \cdots & G_{\mathcal{P}\mathcal{P},p}^{\mathbb{P}} & G_{\mathcal{P}\mathcal{M},p}^{\mathbb{P}}\\ & & O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})\times(d_{\mathbb{P}}p+1)} & & & \end{bmatrix}'\right].\end{aligned}$$

Therefore, it has a degenerated distribution. The degenerated distribution is

$$\begin{aligned}&\begin{bmatrix}\mathcal{N}\left(\beta_{KG,(1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)}-vec\left[\begin{bmatrix}K_{\mathcal{P}}^{\mathbb{Q}} & G_{\mathcal{P}\mathcal{P}}^{\mathbb{Q}} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}p-d_{\mathbb{Q}})}\end{bmatrix}'\right],B_{KG,(1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p,1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)}\right)\\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)\times 1}\end{bmatrix}\\ =&\mathcal{N}\left(\beta_{KGPQ}\equiv\begin{bmatrix}\beta_{KG,(1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)}-vec\left[\begin{bmatrix}K_{\mathcal{P}}^{\mathbb{Q}} & G_{\mathcal{P}\mathcal{P}}^{\mathbb{Q}} & O_{d_{\mathbb{Q}}\times(d_{\mathbb{P}}p-d_{\mathbb{Q}})}\end{bmatrix}'\right]\\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)\times 1}\end{bmatrix},\right.\\ &\left.B_{KGPQ}\equiv\begin{pmatrix}B_{KG,(1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p,1:d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)} & O_{(d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)\times(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)}\\ O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)\times(d_{\mathbb{Q}}+d_{\mathbb{Q}}d_{\mathbb{P}}p)} & O_{(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)\times(d_{\mathbb{P}}-d_{\mathbb{Q}})(d_{\mathbb{P}}p+1)}\end{pmatrix}\right).\end{aligned}$$

Given $\{\mathcal{F}_t\}_{t=1}^T$, since

$$(K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}})+\sum_{l=1}^{p}(G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}}-G_{\mathcal{F}\mathcal{F},l}^{\mathbb{Q}})\mathcal{F}_{t+1-l}=\underbrace{\left(I_{d_{\mathbb{P}}}\otimes\begin{bmatrix}1\\ \mathcal{F}_t\\ \vdots\\ \mathcal{F}_{t+1-p}\end{bmatrix}'\right)}_{\mathcal{F}_{kron,t}}vec\left[\begin{bmatrix}K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}-G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}}\end{bmatrix}'\right],$$

the distribution of $(K_{\mathcal{F}}^{\mathbb{P}}-K_{\mathcal{F}}^{\mathbb{Q}})+\sum_{l=1}^{p}(G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}}-G_{\mathcal{F}\mathcal{F},l}^{\mathbb{Q}})\mathcal{F}_{t+1-l}|\{\mathcal{F}_t\}_{t=1}^T,\Omega_{\mathcal{F}\mathcal{F}},k_\infty^{\mathbb{Q}},\kappa^{\mathbb{Q}}$ is

$$\mathcal{N}(\mathcal{F}_{kron,t}\beta_{KGPQ},\mathcal{F}_{kron,t}B_{KGPQ}\mathcal{F}_{kron,t}').$$

Therefore, the distribution of $\lambda_t|\{\mathcal{F}_t\}_{t=1}^T,\Omega_{\mathcal{F}\mathcal{F}},k_\infty^{\mathbb{Q}},\kappa^{\mathbb{Q}}$ is

$$\mathcal{N}\left(\beta_{\lambda_t}\equiv\Omega_{\mathcal{F}\mathcal{F}}^{-1/2}\mathcal{F}_{kron,t}\beta_{KGPQ},B_{\lambda_t}\equiv(\Omega_{\mathcal{F}\mathcal{F}}^{-1/2}\mathcal{F}_{kron,t})B_{KGPQ}(\Omega_{\mathcal{F}\mathcal{F}}^{-1/2}\mathcal{F}_{kron,t})'\right).$$

We want to derive $\mathbb{E}[\sqrt{\lambda_t'\lambda_t}]$. By the second-order Taylor approximation,

$$\mathbb{E}[\sqrt{\lambda_t'\lambda_t}]=\mathbb{E}[\sum_{n=0}^{\infty}\frac{\sqrt{\pi}}{2}\frac{\mathbb{E}[\lambda_t'\lambda_t]^{0.5-n}(\lambda_t'\lambda_t-\mathbb{E}[\lambda_t'\lambda_t])^n}{\Gamma(1.5-n)n!}],$$

$$\approx \frac{\sqrt{\pi}}{2}\frac{\mathbb{E}[\lambda_t'\lambda_t]^{0.5}}{\Gamma(1.5)} + \frac{\sqrt{\pi}}{2}\frac{\mathbb{E}[\lambda_t'\lambda_t]^{-1.5}Var[\lambda_t'\lambda_t]}{2\Gamma(-0.5)}.$$

By the formula for the quadratic form in statistics,

$$\mathbb{E}[\lambda_t'\lambda_t|\{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}] = tr[B_{\lambda_t}] + \beta_{\lambda_t}'\beta_{\lambda_t},$$
$$Var[\lambda_t'\lambda_t|\{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}] = 2tr[B_{\lambda_t}^2] + 4\beta_{\lambda_t}'B_{\lambda_t}\beta_{\lambda_t}.$$

Therefore, we can calculate

$$\begin{aligned}
&\mathbb{E}[\sqrt{\lambda_t'\lambda_t}|\{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}] \\
\approx&\frac{\sqrt{\pi}}{2}\left[\frac{(tr[B_{\lambda_t}] + \beta_{\lambda_t}'\beta_{\lambda_t})^{0.5}}{\Gamma(1.5)} + \frac{(tr[B_{\lambda_t}] + \beta_{\lambda_t}'\beta_{\lambda_t})^{-1.5}(2tr[B_{\lambda_t}^2] + 4\beta_{\lambda_t}'B_{\lambda_t}\beta_{\lambda_t})}{2\Gamma(-0.5)}\right].
\end{aligned}$$

B.2 $\Omega_{\mathcal{F}\mathcal{F}}$ is a diagonal

If $\Omega_{\mathcal{F}\mathcal{F}}$ is a diagonal,

$$\Omega_{\mathcal{F}\mathcal{F}}^{1/2} = diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F}}^{1/2}].$$

and

$$\lambda_t = diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F}}^{-1/2}][(K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}}) + \sum_{l=1}^{p}(G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F},l}^{\mathbb{Q}})\mathcal{F}_{t+1-l}].$$

Also, since $\mathcal{C}$ is an identity matrix, the priors between $\{K_{\mathcal{F}}^{\mathbb{P}}, G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}\}$ and $\{\mathcal{C}K_{\mathcal{F}}^{\mathbb{P}}, \mathcal{C}G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}\}$ are the same. Therefore, we can use a likelihood of equation (41),

$$\mathcal{F}_t = K_{\mathcal{F}}^{\mathbb{P}} + \sum_{l=1}^{p} G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}}\mathcal{F}_{t-l} + diag(\underbrace{\begin{bmatrix}\sigma_{\mathcal{F}\mathcal{F},1}\\ \vdots \\ \sigma_{\mathcal{F}\mathcal{F},d_{\mathbb{P}}}\end{bmatrix}}_{\tilde{\Omega}_{\mathcal{F}\mathcal{F}}^{1/2}})v_{\mathcal{F},t}^{\mathbb{P}}$$

to use Normal-Normal update. Specifically, we can derive the conditional posterior distribution as

$$p(\begin{bmatrix}K_{\mathcal{F}}^{\mathbb{P}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}}\end{bmatrix}_{(i,:)}|\{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}) \backsim \mathcal{N}(\beta_{CKCG,i}, B_{CKCG,i}),$$

where

$$B_{CKCG,i} \equiv (diag[V_{i,(1:1+d_{\mathbb{P}}p)}]^{-1} + \sigma_{\mathcal{F}\mathcal{F},i}^{-2}X_{\phi_i,(:,1:1+d_{\mathbb{P}}p)}'X_{\phi_i,(:,1:1+d_{\mathbb{P}}p)})^{-1},$$
$$\beta_{CKCG,i} \equiv B_{CKCG,i}(diag[V_{i,(1:1+d_{\mathbb{P}}p)}]^{-1}m_{i,(1:1+d_{\mathbb{P}}p)} + \sigma_{\mathcal{F}\mathcal{F},i}^{-2}X_{\phi_i,(:,1:1+d_{\mathbb{P}}p)}'\begin{bmatrix}\mathcal{F}_{1,(i)}\\ \vdots \\ \mathcal{F}_{T,(i)}\end{bmatrix}).$$

Because of the definition of $\mathbb{Q}$-dynamics,

$$p(\begin{bmatrix} K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}} \end{bmatrix}_{(i,:)} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}})$$
$$\backsim \mathcal{N}(\beta_{KGPQ,i} \equiv \beta_{CKCG,i} - \begin{bmatrix} K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}} \end{bmatrix}_{(i,:)}, B_{KGPQ,i} \equiv B_{CKCG,i}) \text{ for } i \leq d_{\mathbb{Q}},$$

and for $i > d_{\mathbb{Q}}$ the distribution $\begin{bmatrix} K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}} \end{bmatrix}_{(i,:)} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}$ degenerates to zero vector. Then,

$$[(K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}}) + \sum_{l=1}^{p} (G_{\mathcal{F}\mathcal{F},l}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F},l}^{\mathbb{Q}}) \mathcal{F}_{t+1-l}]_{(i,:)}$$
$$= \begin{bmatrix} K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}} & G_{\mathcal{F}\mathcal{F}}^{\mathbb{P}} - G_{\mathcal{F}\mathcal{F}}^{\mathbb{Q}} \end{bmatrix}_{(i,:)} \underbrace{\begin{bmatrix} 1 \\ \mathcal{F}_t \\ \vdots \\ \mathcal{F}_{t+1-p} \end{bmatrix}}_{\mathcal{F}_{const,t}}$$
$$\backsim \mathcal{N}(\beta_{KGPQ,i} \mathcal{F}_{const,t}, \mathcal{F}_{const,t}' B_{KGPQ,i} \mathcal{F}_{const,t}) \text{ for } i \leq d_{\mathbb{Q}},$$

or zero, otherwise. Finally,

$$\lambda_{t,(i)} = \mathcal{N}(\beta_{\lambda_{t,(i)}} \equiv \sigma_{\mathcal{F}\mathcal{F},i}^{-1} \beta_{KGPQ,i} \mathcal{F}_{const,t}, B_{\lambda_{t,(i)}} \equiv \sigma_{\mathcal{F}\mathcal{F},i}^{-2} \mathcal{F}_{const,t}' B_{KGPQ,i} \mathcal{F}_{const,t}) \text{ for } i \leq d_{\mathbb{Q}},$$

or zero, otherwise. Since, the last elements are zeros,

$$\lambda_t' \lambda_t = \lambda_{t,(1:d_{\mathbb{Q}})}' \lambda_{t,(1:d_{\mathbb{Q}})}.$$

Therefore, it is enough to focus on $\lambda_{t,(1:d_{\mathbb{Q}})}$. The distribution of $\lambda_{t,(1:d_{\mathbb{Q}})} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}$ is

$$\mathcal{N}\left(\beta_{\lambda_t} \equiv \begin{bmatrix} \beta_{\lambda_{t,(1)}} \\ \vdots \\ \beta_{\lambda_{t,(d_{\mathbb{Q}})}} \end{bmatrix}, B_{\lambda_t} \equiv diag \begin{bmatrix} B_{\lambda_{t,(1)}} \\ \vdots \\ B_{\lambda_{t,(d_{\mathbb{Q}})}} \end{bmatrix} \right).$$

Using the above result, we can derive

$$\mathbb{E}[\sqrt{\lambda_t' \lambda_t} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_\infty^{\mathbb{Q}}, \kappa^{\mathbb{Q}}]$$
$$\approx \frac{\sqrt{\pi}}{2} \left[\frac{(tr[B_{\lambda_t}] + \beta_{\lambda_t}' \beta_{\lambda_t})^{0.5}}{\Gamma(1.5)} + \frac{(tr[B_{\lambda_t}] + \beta_{\lambda_t}' \beta_{\lambda_t})^{-1.5} (2tr[B_{\lambda_t}^2] + 4\beta_{\lambda_t}' B_{\lambda_t} \beta_{\lambda_t})}{2\Gamma(-0.5)} \right].$$

For deriving the time invariant part

$$\sqrt{\lambda_{t,const}' \lambda_{t,const}},$$

$$\text{where } \lambda_{t,const} := diag[\tilde{\Omega}_{\mathcal{F}\mathcal{F}}^{-1/2}](K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}}),$$

first check that

$$\begin{aligned} & p(\left[K_{\mathcal{F}}^{\mathbb{P}} - K_{\mathcal{F}}^{\mathbb{Q}}\right]_{(i,:)} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_{\infty}^{\mathbb{Q}}, \kappa^{\mathbb{Q}}) \\ \backsim & \mathcal{N}(\beta_{KGPQ,i,(1)}, B_{KGPQ,i,(1,1)}) \text{ for } i \leq d_{\mathbb{Q}}, \end{aligned}$$

and zero, otherwise. Therefore,

$$\begin{aligned} & p(\lambda_{t,const,(i)} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_{\infty}^{\mathbb{Q}}, \kappa^{\mathbb{Q}}) \\ \backsim & \mathcal{N}(\sigma_{\mathcal{F}\mathcal{F},i}^{-1} \beta_{KGPQ,i,(1)}, \sigma_{\mathcal{F}\mathcal{F},i}^{-2} B_{KGPQ,i,(1,1)}) \text{ for } i \leq d_{\mathbb{Q}}, \end{aligned}$$

and zero, otherwise. Similarly, since the last elements are zeros,

$$\lambda_{t,const}' \lambda_{t,const} = \lambda_{t,const,(1:d_{\mathbb{Q}})}' \lambda_{t,const,(1:d_{\mathbb{Q}})}.$$

With the same logic, since the distribution of $\lambda_{t,const,(1:d_{\mathbb{Q}})} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_{\infty}^{\mathbb{Q}}, \kappa^{\mathbb{Q}}$ is

$$\mathcal{N}\left(\beta_{const} \equiv \begin{bmatrix} \sigma_{\mathcal{F}\mathcal{F},1}^{-1} \beta_{KGPQ,1,(1)} \\ \vdots \\ \sigma_{\mathcal{F}\mathcal{F},d_{\mathbb{Q}}}^{-1} \beta_{KGPQ,d_{\mathbb{Q}},(1)} \end{bmatrix}, B_{const} \equiv diag \begin{bmatrix} \sigma_{\mathcal{F}\mathcal{F},1}^{-2} B_{KGPQ,1,(1,1)} \\ \vdots \\ \sigma_{\mathcal{F}\mathcal{F},d_{\mathbb{Q}}}^{-2} B_{KGPQ,d_{\mathbb{Q}},(1,1)} \end{bmatrix} \right),$$

the approximation is

$$\begin{aligned} & \mathbb{E}[\sqrt{\lambda_{t,const}' \lambda_{t,const}} | \{\mathcal{F}_t\}_{t=1}^T, \Omega_{\mathcal{F}\mathcal{F}}, k_{\infty}^{\mathbb{Q}}, \kappa^{\mathbb{Q}}] \\ \approx & \frac{\sqrt{\pi}}{2} \left[\frac{(tr[B_{const}] + \beta_{const}' \beta_{\lambda})^{0.5}}{\Gamma(1.5)} + \frac{(tr[B_{const}] + \beta_{const}' \beta_{const})^{-1.5} (2tr[B_{const}^2] + 4\beta_{const}' B_{const} \beta_{const})}{2\Gamma(-0.5)} \right]. \end{aligned}$$