Review of Derivatives Research (2023) 26:171–206
https://doi.org/10.1007/s11147-023-09196-4

Pricing swaptions and zero-coupon futures options under the discrete-time arbitrage-free Nelson–Siegel model

Frédéric Godin[1,2] · **Ramin Eghbalzadeh**[1] · **Patrice Gaillardetz**[1,2]

Accepted: 31 August 2023 / Published online: 4 October 2023

Abstract
The paper outlines pricing procedures for several interest rate derivatives under the discrete-time arbitrage-free Nelson–Siegel (DTAFNS) model of Eghbalzadeh et al. (The discrete-time arbitrage-free Nelson–Siegel model: a closed-form solution and applications to mixed funds representation, 2022). Derivatives considered include swaptions, zero-coupon futures, and options on such futures. Formulas for expected excess returns are also provided for options on futures. Whereas swaption pricing relies on Monte-Carlo simulation, closed-form formulas are obtained for all other derivatives.

Keywords Interest rate derivatives · Swaptions · Options on futures · Option premium · Option excess returns · Discrete-time arbitrage-free Nelson–Siegel model

JEL Classification E43 · G13

1 Introduction

Interest rate risk management is of paramount importance for financial institutions such as banks and insurance companies. Several financial derivatives can be used for this purpose, e.g. interest rate swaps, zero-coupon futures, and options on such contracts.

✉ Frédéric Godin
frederic.godin@concordia.ca

Ramin Eghbalzadeh
r.nirvan@gmail.com

Patrice Gaillardetz
patrice.gaillardetz@concordia.ca

1 Department of Mathematics and Statistics, Concordia University, Montreal, Canada

2 Quantact Laboratory, Centre de Recherches Mathématiques, Montreal, Canada

Developing pricing procedures for these derivatives is therefore essential. Indeed, the calibration of interest rate dynamics models relies on these pricing procedures. Furthermore, a strand of literature is interested in studying the risk premium embedded in option prices, see for example Coval and Shumway (2001), Bakshi et al. (2023), Bakshi et al. (2022). In the context of interest rate options, Bakshi et al. (2023) recall the puzzling stylized fact of negative excess returns for both out-of-the-money call and put options on treasury futures, and propose a pricing kernel model explaining such feature. The ability to calculate option risk premia implied by interest rate models is therefore useful to better analyse whether a given model is consistent with observed properties of option prices.

The main objective of this study is to provide procedures and formulas to obtain prices for several interest derivatives, namely swaps, swaptions, zero-coupon futures and European options on such futures. Expected excess return formulas for options on zero-coupon futures are also provided. The formulas presented are based on the discrete-time arbitrage-free Nelson–Siegel (DTAFNS) model of Eghbalzadeh et al. (2022), which is a discrete-time version of the original arbitrage-free Nelson–Siegel model developed in Christensen et al. (2011). Such model has numerous advantages. Firstly, being within the family of affine term structure (ATS) models (see Duffie and Kan, 1996) where spot rates are linear combinations of risk factors, it is highly tractable. Secondly, it provides a clear interpretation for factors driving term structure movements: they respectively drive the yield curve's level, its slope and its curvature. Finally it ensures absence of arbitrage.

Several other works also study the pricing of swaptions and other interest rate options in the context of multi-factor or ATS models. We list a few here. The pioneering works of Black et al. (1990) and Black and Karasinski (1991) show how to price zero-coupon options based respectively on a one-factor binomial tree model and a log-normal diffusion model. Munk (1999) demonstrates that the price of a European option on a coupon bond (e.g. a swaption) is roughly equal to some multiple of the price of a European option on a zero-coupon bond with maturity equal to the coupon bond's stochastic duration. Collin-Dufresne and Goldstein (2002) propose to apply an Edgeworth expansion to approximate the density of the coupon bond price and obtain the price of a swaption. Singleton and Umantsev (2002) rely on Fourier inversion methods to calculate swaption prices in the ATS framework. Schrager and Pelsser (2006) propose to approximate the swap rate volatility under the swap measure, which is a low-variance martingale, by its time-zero value. Such strategy leads to a closed-form formula for the swaption price.

The paper is structured as follows. Section 2 provides a description of the DTAFNS term structure model. Section 3 presents pricing procedures for swaptions, whereas Sect. 4 provides formulas for prices and expected excess returns of options on zero-coupon futures. Section 5 briefly discusses calibration methods for the DTAFNS model relying on option prices. Section 6 concludes.

2 The DTAFNS model

This section discusses interest rate dynamics in the DTAFNS model of Eghbalzadeh et al. (2022). Dynamics are provided under three probability measures: the physical measure, the risk-neutral measure and the forward measure. All three measures are required for the computation of prices and associated risk premia of derivatives considered in this study.

2.1 Risk-neutral dynamics in the DTAFNS model

This section provides a description of risk-neutral dynamics in the DTAFNS interest rate term structure model. Consider a discrete-time setting with monthly time points $t = 0, \dots, T$ and time elapse Δ year between each point. The filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^{T}$ characterizes the information flow in the market. The DTAFNS model assumes that the term structure of interest rates is determined by three factors: the long-term level of interest rates, the slope of the yield curve and its curvature. The time-t short rate applying over period $[t, t+1)$ is

$$r_t = X_t^{(1)} + X_t^{(2)}, \tag{2.1}$$

with $\{X_t\}_{t=0}^{T}$ denoting the term structure factor process, where time-t factors are the triplet $X_t = [X_t^{(1)}, X_t^{(2)}, X_t^{(3)}]^\top$.

Under the risk-neutral measure $\mathbb{Q}$, factors exhibit the following auto-regressive dynamics:

$$\begin{pmatrix} X_{t+1}^{(1)} - X_t^{(1)} \\ X_{t+1}^{(2)} - X_t^{(2)} \\ X_{t+1}^{(3)} - X_t^{(3)} \end{pmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}}_{\kappa^{\mathbb{Q}}} \underbrace{\begin{bmatrix} \theta_1^{\mathbb{Q}} - X_t^{(1)} \\ \theta_2^{\mathbb{Q}} - X_t^{(2)} \\ \theta_3^{\mathbb{Q}} - X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{Q}} - X_t} + \underbrace{\begin{pmatrix} \Sigma_{1,1} & 0 & 0 \\ 0 & \Sigma_{2,2} & 0 \\ 0 & 0 & \Sigma_{3,3} \end{pmatrix}}_{\Sigma} \begin{pmatrix} Z_{t+1,1}^{\mathbb{Q}} \\ Z_{t+1,2}^{\mathbb{Q}} \\ Z_{t+1,3}^{\mathbb{Q}} \end{pmatrix}, \tag{2.2}$$

where scalar $\lambda \in (0, 1)$ and matrices $\theta^{\mathbb{Q}}$, $\kappa^{\mathbb{Q}}$ and Σ, with $\Sigma_{i,i} > 0$, represent model parameters, and $\{Z_{t,i}^{\mathbb{Q}}\}_{t=1}^{n}$, $i = 1, 2, 3$ are $\mathcal{F}$-adapted standard Gaussian white noises with contemporaneous correlation $\mathrm{Corr}[Z_{t_1,i}^{\mathbb{Q}}, Z_{t_2,j}^{\mathbb{Q}}] = \mathbb{1}_{\{t_1=t_2\}}\rho_{i,j}$ represented by correlation matrix $\rho = \left[\rho_{i,j}\right]_{i,j=1}^{3}$. We set $\theta_1^{\mathbb{Q}} = 0$ since such parameter is unused.

As shown in Eghbalzadeh et al. (2022), the time-t price of a risk-free zero-coupon bond paying one dollar on maturity $\mathcal{T} > t$ is, under the such model,

$$P(t, \mathcal{T}) = A_\tau \exp\left[-\Delta \mathcal{B}_\tau^\top X_t\right], \tag{2.3}$$

where $\tau = \mathcal{T} - t$, $\mathcal{B}_\tau = \left[\mathcal{B}_\tau^{(1)},\ \mathcal{B}_\tau^{(2)},\ \mathcal{B}_\tau^{(3)}\right]^\top$ and

$$\mathcal{B}_\tau^{(1)} = \tau, \quad \mathcal{B}_\tau^{(2)} = \frac{1 - (1-\lambda)^\tau}{\lambda}, \quad \mathcal{B}_\tau^{(3)} = \frac{1 - (1-\lambda)^{\tau-1}}{\lambda} - (\tau - 1)(1-\lambda)^{\tau-1}, \tag{2.4}$$

$$\log A_\tau = -\Delta\theta_2^{\mathbb{Q}}\left(\mathcal{B}_\tau^{(1)} - \mathcal{B}_\tau^{(2)}\right) + \Delta\theta_3^{\mathbb{Q}}\mathcal{B}_\tau^{(3)} + \frac{1}{2}\Delta^2 v_\tau, \tag{2.5}$$

with

$$\begin{aligned}
v_\tau &= \left(\sum_{i=1}^{3}\sum_{j=1}^{3} v_\tau^{(i,j)}\right),\\
v_\tau^{(1,1)} &= \Sigma_{1,1}^2 \frac{\tau(\tau-1)(2\tau-1)}{6},\\
v_\tau^{(2,2)} &= \frac{\Sigma_{2,2}^2}{\lambda^2}\left(\tau - 2\left[\frac{1-(1-\lambda)^\tau}{\lambda}\right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2}\right),\\
v_\tau^{(3,3)} &= \mathbb{1}_{\{\tau>1\}}\frac{\Sigma_{3,3}^2}{\lambda^2}\Big[\tau - 2 + \zeta_0\left((1-\lambda)^2, \tau-1\right) + \lambda^2\zeta_2\left((1-\lambda)^2, \tau-1\right)\\
&\quad - 2\zeta_0((1-\lambda), \tau-1) - 2\lambda\zeta_1((1-\lambda), \tau-1) + 2\lambda\zeta_1\left((1-\lambda)^2, \tau-1\right)\Big],\\
v_\tau^{(1,2)} &= v_t^{(2,1)} = \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda), \tau)\right),\\
v_\tau^{(1,3)} &= v_t^{(3,1)} = \mathbb{1}_{\{\tau>1\}}\rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\\
&\quad \Big[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0((1-\lambda), \tau-1) - (\lambda+1)\zeta_1((1-\lambda), \tau-1)\\
&\quad - \lambda\zeta_2((1-\lambda), \tau-1)\Big],\\
v_\tau^{(2,3)} &= v_t^{(3,2)} = \mathbb{1}_{\{\tau>1\}}\rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\\
&\quad \left(\frac{\tau - 2 - (2-\lambda)\zeta_0((1-\lambda), \tau-1) + (1-\lambda)\zeta_0\left((1-\lambda)^2, \tau-1\right)}{\lambda^2}\right.\\
&\quad \left. + \frac{-\zeta_1((1-\lambda), \tau-1) + (1-\lambda)\zeta_1\left((1-\lambda)^2, \tau-1\right)}{\lambda}\right),
\end{aligned} \tag{2.6}$$

and

$$\zeta_0(r,\tau) \equiv \sum_{u=1}^{\tau-1} r^u = \frac{r - r^\tau}{1-r}, \tag{2.7}$$

$$\zeta_1(r,\tau) \equiv \sum_{u=1}^{\tau-1} u r^u = \frac{r - \tau r^\tau + (\tau-1)r^{\tau+1}}{(1-r)^2}, \tag{2.8}$$

$$\zeta_2(r,\tau) \equiv \sum_{u=1}^{\tau-1} u^2 r^u = \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1) r^{\tau+1} - \tau^2 r^\tau + r^2 + r}{(1-r)^3}. \tag{2.9}$$

Remark 2.1 For the rest of the paper, the convention $\mathcal{B}_0^{(1)} = \mathcal{B}_0^{(2)} = \mathcal{B}_0^{(3)} = 0$, $A_0 = 1$ and $v_0 = 0$ is used, which makes (2.3) hold for $\tau = 0$.

2.2 Forward measure dynamics in the DTAFNS model

Prices of financial derivatives are expressed as expected discounted payoffs under the risk-neutral measure. However, the interaction between the stochastic discount factor and the derivatives payoff is often non-trivial and complexifies the pricing. A common technique to ease the calculation of prices in the context of stochastic interest rates is the change of numéraire (see for instance Geman et al., 1995). This approach relies on the construction of a new probability measure, called the forward measure, under which the price of a zero-coupon is used a numéraire for discounting. This allows directly discounting with zero-coupon bond prices, thus circumventing the difficulty associated with representing the potentially complex dependence between the payoff and the stochastic discount factor.

The probability measure using the risk-free zero-coupon bond maturing at time $\mathcal{T}$ as a numéraire is known as the $\mathcal{T}$-forward measure and is denoted by $\mathbb{Q}^{\mathcal{T}}$. The Radon–Nikodym derivative allowing to pass from the risk-neutral to the $\mathcal{T}$-forward measure, which is provided by Jamshidian (1996) or Brigo and Mercurio (2007), is

$$\frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}} = \frac{B(0)P(\mathcal{T},\mathcal{T})}{P(0,\mathcal{T})B(\mathcal{T})} = \frac{D(0,\mathcal{T})}{P(0,\mathcal{T})}, \tag{2.10}$$

where $B(t) = \exp(\Delta \sum_{s=0}^{t-1} r_s)$ is the year-t bank account numéraire under the risk-neutral measure and $D(t_1,t_2) = B(t_1)/B(t_2)$ is the stochastic discount factor for any $0 \le t_1 \le t_2 \le T$. Note that the Radon–Nikodym derivative allowing to go from the forward measure to the risk-neutral measure is $\frac{d\mathbb{Q}}{d\mathbb{Q}^{\mathcal{T}}} = \left(\frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}}\right)^{-1}$.

Let $\mathbb{E}^{\mathcal{T}}[\cdot]$ represent the expectation under the $\mathcal{T}$-forward measure. Asset prices discounted by the zero-coupon price maturing at $\mathcal{T}$ are martingales under the forward measure (Geman, 1989). As a consequence, as discussed in Brigo and Mercurio (2007), the time-t price H_t of an asset providing a payoff $H_{\mathcal{T}}$ at time $\mathcal{T}$ is

$$H_t = P(t,\mathcal{T})\mathbb{E}^{\mathcal{T}}\left[H_{\mathcal{T}} | \mathcal{F}_t\right]. \tag{2.11}$$

The following proposition defines so-called *forward measure innovations* and outlines their dynamics.

Proposition 2.1 *For any $\mathcal{T} \in \{1, \dots, T\}$, $t < \mathcal{T}$ and $\tau = \mathcal{T} - t$, conditional on $\mathcal{F}_t$, the forward measure innovation defined as $Z^{\mathcal{T}}_{t+1} = Z^{\mathbb{Q}}_{t+1} + \Delta\rho\Sigma\mathcal{B}_{\tau-1}$ follows the multivariate Gaussian distribution with mean vector zero and covariance matrix ρ under the $\mathcal{T}$-forward measure.*

Proof See "Appendix". □

Corollary 2.1 *Since the conditional distribution of $Z^{\mathcal{T}}_{t+1}$ with respect $\mathcal{F}_t$ under the $\mathcal{T}$-forward measure does not depend on $Z^{\mathcal{T}}_1, \dots, Z^{\mathcal{T}}_t$, and since the latter variables characterize the information contained in $\mathcal{F}_t$, elements of the sequence $\{Z^{\mathcal{T}}_j\}^{\mathcal{T}}_{j=1}$ are independent.*

Based on the above results, we now provide an expression for the dynamics of term structure factors $\{X_t\}^T_{t=0}$ analogous to (2.2), but using instead the $\mathcal{T}$-forward measure innovations. Define

$$\theta^{\mathcal{T}} = \theta^{\mathbb{Q}}, \quad \kappa^{\mathcal{T}} = \kappa^{\mathbb{Q}}, \quad \eta^{\mathcal{T}}_t = \Delta\Sigma\rho\Sigma\mathcal{B}_{\tau-1} = \Delta\Sigma\rho\Sigma\mathcal{B}_{\mathcal{T}-t-1}. \tag{2.12}$$

A direct consequence of the application of Proposition 2.1 into (2.2) is that

$$X_{t+1} = X_t - \eta^{\mathcal{T}}_t + \kappa^{\mathcal{T}}(\theta^{\mathcal{T}} - X_t) + \Sigma Z^{\mathcal{T}}_{t+1}. \tag{2.13}$$

Representation (2.13), along with some additional lemmas provided in "Appendix", allow obtaining the t-conditional distribution of $X_{\mathcal{T}}$ under the $\mathcal{T}$-forward measure.

Proposition 2.2 *Under the $\mathcal{T}$-forward measure, conditionally on $\mathcal{F}_t$ and for $t + n \leq \mathcal{T}$, factors X_{t+n} follow the multivariate Gaussian distribution with mean vector $\mathcal{M}_{t,n} = \left[\mathcal{M}^{(i)}_{t,n}\right]^3_{i=1}$ and covariance matrix $\mathcal{V}_n = \left[\mathcal{V}^{(i,j)}_n\right]^3_{i,j=1}$ where*

$$\mathcal{M}_{t,n}^{(1)} = X_t^{(1)} - \sum_{l=0}^{n-1} \eta_{t+l}^{\mathcal{T},(1)},$$

$$\begin{aligned}\mathcal{M}_{t,n}^{(2)} &= X_t^{(2)}(1-\lambda)^n + (\theta_2^{\mathcal{T}} - \theta_3^{\mathcal{T}})(1-(1-\lambda)^n) - \sum_{l=0}^{n-1} \eta_{t+l}^{\mathcal{T},(2)}(1-\lambda)^{n-1-l} \\ &\quad + \lambda\Bigg(nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\Bigg(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1}\Bigg) \\ &\quad - \sum_{l=0}^{n-1}(n-l-1)\eta_{t+l}^{\mathcal{T},(3)}(1-\lambda)^{n-l-2}\Bigg),\end{aligned}$$

$$\mathcal{M}_{t,n}^{(3)} = X_t^{(3)}(1-\lambda)^n + \theta_3^{\mathcal{T}}(1-(1-\lambda)^n) - \sum_{l=0}^{n-1} \eta_{t+l}^{\mathcal{T},(3)}(1-\lambda)^{n-1-l},$$

$$\begin{aligned}\mathcal{V}_n^{(1,1)} &= n\Sigma_{1,1}^2, \\ \mathcal{V}_n^{(2,2)} &= \Sigma_{2,2}^2\big(1+\zeta_0((1-\lambda)^2,n)\big) + \lambda^2\Sigma_{3,3}^2(1-\lambda)^{-2}\zeta_2((1-\lambda)^2,n) \\ &\quad + 2\Sigma_{2,2}\lambda\Sigma_{3,3}\rho_{2,3}(1-\lambda)^{-1}\zeta_1\big((1-\lambda)^2,n\big), \\ \mathcal{V}_n^{(3,3)} &= \Sigma_{3,3}^2\big(1+\zeta_0((1-\lambda)^2,n)\big), \\ \mathcal{V}_n^{(1,2)} &= \mathcal{V}_n^{(2,1)} = \Sigma_{1,1}\Sigma_{2,2}\rho_{1,2}\big(1+\zeta_0(1-\lambda,n)\big) + \lambda\Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\frac{\zeta_1(1-\lambda,n)}{1-\lambda}, \\ \mathcal{V}_n^{(1,3)} &= \mathcal{V}_n^{(3,1)} = \Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\big(1+\zeta_0(1-\lambda,n)\big), \\ \mathcal{V}_n^{(2,3)} &= \mathcal{V}_n^{(3,2)} = \Sigma_{2,2}\Sigma_{3,3}\rho_{2,3}\big(1+\zeta_0((1-\lambda)^2,n)\big) + \lambda\Sigma_{3,3}^2\frac{\zeta_1((1-\lambda)^2,n)}{1-\lambda}.\end{aligned}$$

Proof See "Appendix". □

The following quantities appearing in Proposition 2.2 can be further simplified.

Lemma 2.1 *Considering the case* $n = \mathcal{T} - t$,

$$\begin{aligned}\sum_{l=0}^{\mathcal{T}-t-1} \eta_{t+l}^{\mathcal{T},(1)} &= \Delta\Sigma_{1,1}\Bigg[\Sigma_{1,1}\frac{(\mathcal{T}-t-1)(\mathcal{T}-t)}{2} \\ &\quad + \frac{\Sigma_{2,2}\rho_{1,2}}{\lambda}\big(\mathcal{T}-t-1-\zeta_0(1-\lambda,\mathcal{T}-t)\big) \\ &\quad + \Sigma_{3,3}\rho_{1,3}\Bigg(\frac{\mathcal{T}-t-1-\big(1+\zeta_0(1-\lambda,\mathcal{T}-t-1)\big)}{\lambda} \\ &\quad - \zeta_1(1-\lambda,\mathcal{T}-t-1)\Bigg)\Bigg].\end{aligned}$$

Moreover, for $i = 2, 3$,

$$\sum_{l=0}^{\mathcal{T}-t-1} \eta_{t+l}^{\mathcal{T},(i)}(1-\lambda)^{\mathcal{T}-t-1-l} = \Delta\Sigma_{i,i}\Bigg[\Sigma_{1,1}\rho_{i,1}\zeta_1(1-\lambda,\mathcal{T}-t)$$
$$+\frac{\Sigma_{2,2}\rho_{i,2}}{\lambda}\big(\zeta_0(1-\lambda,\mathcal{T}-t)-\zeta_0\big((1-\lambda)^2,\mathcal{T}-t\big)\big)$$
$$+\Sigma_{3,3}\rho_{i,3}\Bigg(\frac{\zeta_0(1-\lambda,\mathcal{T}-t)-(1-\lambda)^{-1}\zeta_0\big((1-\lambda)^2,\mathcal{T}-t\big)}{\lambda}$$
$$-(1+\lambda)\zeta_1\big((1-\lambda)^2,\mathcal{T}-t-1\big)\Bigg)\Bigg].$$

Lastly,

$$\sum_{l=0}^{\mathcal{T}-t-1}(\mathcal{T}-t-l-1)\eta_{t+l}^{\mathcal{T},(3)}(1-\lambda)^{\mathcal{T}-t-l-2}$$
$$=\frac{\Delta\Sigma_{3,3}}{1-\lambda}\Bigg(\Sigma_{1,1}\rho_{3,1}\zeta_2(1-\lambda,\mathcal{T}-t)$$
$$+\frac{\Sigma_{2,2}\rho_{2,1}}{\lambda}\big[\zeta_1(1-\lambda,\mathcal{T}-t)-\zeta_1\big((1-\lambda)^2,\mathcal{T}-t\big)\big]$$
$$+\Sigma_{3,3}\rho_{3,1}\Bigg[\frac{\zeta_1(1-\lambda,\mathcal{T}-t)}{\lambda}-\frac{1}{\lambda}\zeta_1\big((1-\lambda)^2,\mathcal{T}-t\big)$$
$$-(1-\lambda)^{-1}\zeta_2\big((1-\lambda)^2,\mathcal{T}-t\big)\Bigg]\Bigg).$$

Proof See "Appendix". □

2.3 Physical measure dynamics in the DTAFNS model

To determine option risk premia and expected excess returns, dynamics of interest rates under the physical measure $\mathbb{P}$ must be specified. The $\mathbb{P}$-dynamics considered in Eghbalzadeh et al. (2022) are used here since they are shown in that paper to exhibit natural compatibility with the $\mathbb{Q}$-dynamics model; the form of the pricing kernel allowing to pass from such $\mathbb{P}$-measure to the $\mathbb{Q}$-measure outlined in Sect. 2.1 is provided in that paper.

Under the risk-neutral measure $\mathbb{P}$, factors are assumed to have the following autoregressive dynamics:

$$\begin{pmatrix} X_{t+1}^{(1)}-X_t^{(1)} \\ X_{t+1}^{(2)}-X_t^{(2)} \\ X_{t+1}^{(3)}-X_t^{(3)} \end{pmatrix} = \underbrace{\begin{bmatrix} \kappa_{1,1}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{2,2}^{\mathbb{P}} & -\lambda \\ 0 & 0 & \kappa_{3,3}^{\mathbb{P}} \end{bmatrix}}_{\kappa^{\mathbb{P}}} \underbrace{\begin{bmatrix} \theta_1^{\mathbb{P}}-X_t^{(1)} \\ \theta_2^{\mathbb{P}}-X_t^{(2)} \\ \theta_3^{\mathbb{P}}-X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{P}}-X_t} + \underbrace{\begin{pmatrix} \Sigma_{1,1} & 0 & 0 \\ 0 & \Sigma_{2,2} & 0 \\ 0 & 0 & \Sigma_{3,3} \end{pmatrix}}_{\Sigma} \begin{pmatrix} Z_{t+1,1}^{\mathbb{P}} \\ Z_{t+1,2}^{\mathbb{P}} \\ Z_{t+1,3}^{\mathbb{P}} \end{pmatrix}, \tag{2.14}$$

where $\kappa^{\mathbb{P}}_{1,1} \in [0,1)$, $\kappa^{\mathbb{P}}_{i,i} \in (0,1)$, $i = 2,3$, and $\{(Z^{\mathbb{P}}_{t,1}, Z^{\mathbb{P}}_{t,2}, Z^{\mathbb{P}}_{t,3})\}_{t=1}^{T}$ is again a 3-dimensional Gaussian standard white noise with $\mathrm{Corr}[Z^{\mathbb{P}}_{t_1,i}, Z^{\mathbb{P}}_{t_2,j}] = \mathbb{1}_{\{t_1=t_2\}}\rho_{i,j}$. The $\mathbb{P}$ -dynamics are slightly more flexible than the $\mathbb{Q}$-dynamics since the components in the diagonal of $\kappa^{\mathbb{P}}$ are not required to be either 0 or λ. A possibility would be to impose $\kappa^{\mathbb{P}}_{1,1} = 0$ to replicate the non-stationary dynamics of factor $X^{(1)}$ under $\mathbb{Q}$. Nevertheless Eghbalzadeh et al. (2022) argue that not imposing such restriction provides a better fit for their dataset.

Under the above $\mathbb{P}$-dynamics, Eghbalzadeh et al. (2022) provide the relationship between the long-term mean parameters $\theta^{\mathbb{P}}$ and $\theta^{\mathbb{Q}}$ for both measures:

$$\theta^{\mathbb{Q}}_2 = \lambda^{-1}(\kappa^{\mathbb{P}}_{2,2}\theta^{\mathbb{P}}_2 + \kappa^{\mathbb{P}}_{3,3}\theta^{\mathbb{P}}_3 - \lambda\theta^{\mathbb{P}}_3), \quad \theta^{\mathbb{Q}}_3 = \lambda^{-1}\kappa^{\mathbb{P}}_{3,3}\theta^{\mathbb{P}}_3$$

and

$$\theta^{\mathbb{P}}_3 = \lambda\theta^{\mathbb{Q}}_3/\kappa^{\mathbb{P}}_{3,3}, \quad \theta^{\mathbb{P}}_2 = \frac{\lambda}{\kappa^{\mathbb{P}}_{2,2}}\left[\theta^{\mathbb{Q}}_2 - \frac{\theta^{\mathbb{Q}}_3}{\kappa^{\mathbb{P}}_{3,3}}(\kappa^{\mathbb{P}}_{3,3} - \lambda)\right],$$

whereas the $\kappa^{\mathbb{P}}_{i,i}$, $i = 1,2,3$ are allowed to vary freely.

Remark 2.2 A straightforward possible extension of the model could consist in still using (2.14) for $\mathbb{P}$-dynamics, but instead considering a physical volatility matrix $\Sigma^{\mathbb{P}}$ whose components are different (most likely lower) than those of the risk-neutral volatility Σ. This could help producing negative excess returns for out-of-the-money options (either calls or put) on zero-coupon futures, which are observed empirically as documented in Bakshi et al. (2023).

The following proposition, whose proof is in "Appendix", is analogous to Proposition 2.2 and provides transition distributions for factors X under the physical measure $\mathbb{P}$. Such proposition is used subsequently in the derivation of option expected excess returns.

Proposition 2.3 *Assume $\kappa^{\mathbb{P}}_{2,2} \neq \kappa^{\mathbb{P}}_{3,3}$ and define $\omega = \frac{1-\kappa^{\mathbb{P}}_{3,3}}{1-\kappa^{\mathbb{P}}_{2,2}}$. Under measure $\mathbb{P}$, conditionally on $\mathcal{F}_t$ and for $t+n \leq T$, factors X_{t+n} follow the multivariate Gaussian distribution with mean vector $\mathcal{M}^{\mathbb{P}}_{t,n} = \left[\mathcal{M}^{\mathbb{P},(i)}_{t,n}\right]_{i=1}^{3}$ and covariance matrix $\mathcal{V}^{\mathbb{P}}_n = \left[\mathcal{V}^{\mathbb{P},(i,j)}_n\right]_{i,j=1}^{3}$ where*

$$\begin{aligned}\mathcal{M}^{\mathbb{P},(i)}_{t,n} &= X^{(i)}_t(1-\kappa^{\mathbb{P}}_{i,i})^n + \theta^{\mathbb{P}}_i\left(1-(1-\kappa^{\mathbb{P}}_{i,i})^n\right)\\ &\quad + \mathbb{1}_{\{i=2\}}\lambda(X^{(3)}_t - \theta^{\mathbb{P}}_3)\frac{1-\omega^n}{1-\omega}(1-\kappa^{\mathbb{P}}_{2,2})^{n-1}\end{aligned}$$

and

$$
\begin{aligned}
\mathcal{V}_n^{\mathbb{P},(1,1)} &= \begin{cases} \Sigma_{1,1}^2\Big(1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})^2,n)\Big) & \text{if } \kappa_{1,1}^{\mathbb{P}} \in (0,1), \\ n\Sigma_{1,1}^2 \quad \text{if } \kappa_{1,1}^{\mathbb{P}} = 0, \end{cases} \\
\mathcal{V}_n^{\mathbb{P},(2,2)} &= \Sigma_{2,2}^2\Big(1+\zeta_0((1-\kappa_{2,2}^{\mathbb{P}})^2,n)\Big) \\
&\quad + \lambda^2\Sigma_{3,3}^2\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{2n-2}}{(1-\omega)^2}\Big[\zeta_0\Big((1-\kappa_{2,2}^{\mathbb{P}})^{-2},n\Big) - 2\omega^n\zeta_0\Big(\omega(1-\kappa_{3,3}^{\mathbb{P}})^{-2},n\Big) \\
&\quad +\omega^{2n}\zeta_0\Big((1-\kappa_{3,3}^{\mathbb{P}})^{-2},n\Big)\Big] \\
&\quad + 2\frac{\rho_{2,3}\lambda\Sigma_{2,2}\Sigma_{3,3}}{1-\omega}(1-\kappa_{2,2}^{\mathbb{P}})^{(2n-1)}\left[\zeta_0\left(\frac{\omega}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})},n\right)\right. \\
&\quad \left.-\omega^n\zeta_0\left(\frac{1}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})},n\right)\right] \\
\mathcal{V}_n^{\mathbb{P},(3,3)} &= \Sigma_{3,3}^2\Big(1+\zeta_0((1-\kappa_{3,3}^{\mathbb{P}})^2,n)\Big) \\
\mathcal{V}_n^{\mathbb{P},(1,2)} &= \Sigma_{1,1}\Sigma_{2,2}\rho_{1,2}\Big[1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{2,2}^{\mathbb{P}}),n)\Big] \\
&\quad + \lambda\Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\frac{(1-\kappa_{1,1}^{\mathbb{P}})^n(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{1-\omega}\times \\
&\quad \left(\zeta_0\left(\frac{\omega}{(1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})},n\right) - \omega^n\zeta_0\left(\frac{1}{(1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})},n\right)\right), \\
\mathcal{V}_n^{\mathbb{P},(1,3)} &= \Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\Big[1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}}),n)\Big], \\
\mathcal{V}_n^{\mathbb{P},(2,3)} &= \Sigma_{2,2}\Sigma_{3,3}\rho_{2,3}\Big[1+\zeta_0((1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}}),n)\Big] \\
&\quad + \lambda\Sigma_{3,3}^2\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}(1-\kappa_{3,3}^{\mathbb{P}})^n}{1-\omega}\left(\zeta_0\left(\frac{1}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})},n\right)\right. \\
&\quad \left.-\omega^n\zeta_0\Big((1-\kappa_{3,3}^{\mathbb{P}})^{-2},n\Big)\right).
\end{aligned}
$$

Remark 2.3 Closed-form formulas for conditional moments of X_{t+n} given X_t under $\mathbb{P}$ could also be derived in the case $\kappa_{2,2}^{\mathbb{P}} = \kappa_{3,3}^{\mathbb{P}}$. However, since such case is very unlikely to occur in practice, it is omitted.

3 European swaption pricing

This section describes two pricing procedures for European swaptions and outlines their respective advantages. The first relies on a risk-neutral simulation, whereas the second uses the forward measure to perform the simulation.

3.1 Risk-neutral pricing of European swaptions

Swaptions are classified into three types: European, Bermudan, and American, which differ in their possible exercise dates. Whereas American and Bermudan swaptions allow the exercise of the option on multiple dates, the European swaption has a single possible exercise date. We shall focus on European swaptions in this study. The European swaption considered, which is a payer swaption, is a financial option that gives the holder the right to enter, at time T_α, into a swap with payment dates $T_{\alpha+1}, \dots, T_\beta$ on which the holder pays the strike rate as the fixed rate, and receives the prevailing floating rate on each payment date.[1] Typically, the floating rate is tied to an interbank offered rate, such as LIBOR in the United Kingdom or the CDOR in Canada.

As shown in Brigo and Mercurio (2007), for $t < T_\alpha$, the time-t price of a European payer swaption with maturity T_α, strike K, nominal value N and payment dates $\{T_i\}_{i=\alpha+1}^{\beta}$ is

$$PS\left[t;\{T_i\}_{i=\alpha}^{\beta};K;N\right] = \mathbb{E}^{\mathbb{Q}}\left[D(t,T_\alpha)\left(N\left(S_{\alpha,\beta}(T_\alpha)-K\right)^+ \sum_{i=\alpha+1}^{\beta} \delta_i P(T_\alpha,T_i)\right)\middle|\mathcal{F}_t\right] \tag{3.1}$$

where $\delta_i = T_i - T_{i-1}$. The time-t forward swap rate $S_{\alpha,\beta}(t)$ is

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} \delta_i P(t,T_i)}. \tag{3.2}$$

The swap rate $S_{\alpha,\beta}(t)$ corresponds to a value of the fixed rate which would make the time-t value of the swap nil. The rationale underlying (3.1) is that a market participant could, while exercising the option, enter without fee into a receiver swap with the swap rate as the fixed rate. Combining both positions would lead to a net payment being the difference between the swap rate and the strike rate at each payment date.

A straightforward approach to obtain the swaption price via (3.1) is to conduct a Monte-Carlo simulation of the term structure factors under the risk-neutral measure and to average discounted cash flows, thereby approximating the expectation in (3.1). Algorithm 1 summarizes this process. The risk-neutral approach for swaption pricing has the advantage of requiring a single simulation to price multiple swaptions at once, which could be desirable in a calibration exercise. The drawback of using such an approach is that it requires simulating the entire path of the term structure factors, which might not be needed if a single swaption needs to be priced, as explained in the following section.

[1] For a payment date $T_{\xi+1}$, $\xi = \alpha, \dots, \beta - 1$, the floating rate is determined at the reset date T_ξ.

Algorithm 1 Calculating a European payer swaption price under the risk-neutral measure

Input: t, $X_t^{(1)}$, $X_t^{(2)}$, $X_t^{(3)}$, K, N, $\{T_i\}_{i=\alpha}^{\beta}$, Δ, $\kappa^{\mathbb{Q}}$, $\theta^{\mathbb{Q}}$, ρ **and** Σ
for $m \in \{1, \ldots, M\}$ **do** ▷ M is the number of simulated paths
  Simulate the mth path of the state variables, X, denoted by $\{ {}_mX_{t'}^{(i)}\}_{t'=t+1}^{T_\alpha}, i = 1, 2, 3$ under the risk neural measure $\mathbb{Q}$
  for $t' \in \{t+1, \ldots, T_\alpha\}$ **do**
    Calculate the interest rate, $r({}_mX_{t'})$, through (2.1)
  Calculate $D(t, T_\alpha)$
  for $t'' \in \{T_{\alpha+1}, \ldots, T_\beta\}$ **do**
    Calculate ${}_mP(T_\alpha, t'')$ through (2.3)
  Calculate ${}_mS_{\alpha,\beta}(T_\alpha)$ through (3.2)
  Calculate the payoff for m^{th} path

$$ {}_mPS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right] = D(t, T_\alpha) N \left({}_mS_{\alpha,\beta}(T_\alpha) - K\right)^+ \sum_{i=\alpha+1}^{\beta} \delta_i \, {}_mP(T_\alpha, T_i) $$

$$ PS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right] \approx \frac{\sum_{m=1}^{M} {}_mPS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right]}{M} $$

3.2 Pricing swaptions under the forward measure

Calculating European swaption prices using Algorithm 1 requires simulating the entire path of risk-free rate factors, which might be numerically cumbersome in some situations. By applying a change of numéraire, we can obtain a pricing approach which is more time-efficient. Detailing such an approach in the context of the DTAFNS model is the objective of this subsection.

Considering the zero-coupon bond maturing at $\mathcal{T} = T_\alpha$ as the new numéraire makes the computation of the swaption price much more convenient. In such case, the payer swaption price may therefore be rewritten based on (2.11), (3.1) and (3.2) as

$$ \begin{aligned} PS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right] &= P(t, T_\alpha) \mathbb{E}^{T_\alpha}\left[\left(N\left(S_{\alpha,\beta}(T_\alpha) - K\right)^+ \sum_{i=\alpha+1}^{\beta} \delta_i P(T_\alpha, T_i)\right) \middle| \mathcal{F}_t\right] \\ &= P(t, T_\alpha) \mathbb{E}^{T_\alpha}\left[\left(N\left(1 - P(T_\alpha, T_\beta) - K \sum_{i=\alpha+1}^{\beta} \delta_i P(T_\alpha, T_i)\right)^+\right) \middle| \mathcal{F}_t\right]. \end{aligned} \tag{3.3} $$

Equation (3.3) involves the t-conditional expectation of a function of time-T_α zero-coupon bond prices, which are fully characterized by term structure factors $X_{T_\alpha} = \left[X_{T_\alpha}^{(1)}, X_{T_\alpha}^{(2)}, X_{T_\alpha}^{(3)}\right]^\top$. As a result, Proposition 2.2 can be used to calculate (3.3).

Algorithm 2 highlights the procedure to price swaptions using such an approach. When pricing a single swaption, such $\mathcal{T}$-forward measure simulation is much

quicker than the Algorithm 1 based on the risk-neutral measure, which requires computing expectations over entire paths of the term structure factors.

Algorithm 2 Calculating a European payer swaptions price under the forward measure

Input: t, $X_t^{(1)}$, $X_t^{(2)}$, $X_t^{(3)}$, K, N, $\{T_i\}_{i=\alpha}^{\beta}$, Δ, $\kappa^{\mathbb{Q}}$, $\theta^{\mathbb{Q}}$, ρ, **and** Σ
Calculate $P(t, T_\alpha)$ *with* (2.3) *and* $\kappa^{\mathcal{T}}$, $\theta^{\mathcal{T}}$ *using* (2.12)
for $m \in \{1, \ldots, M\}$ **do** ▷ M is the number of simulated paths
  Simulate ${}_mX_{T_\alpha}^{(1)}$, ${}_mX_{T_\alpha}^{(2)}$, ${}_mX_{T_\alpha}^{(3)}$ *from multivariate Gaussian distribution with mean vector* $\mathcal{M}_{t,T_\alpha - t}$ *and matrix variance-covariance* $\mathcal{V}_{T_\alpha - t}$ *defined in Proposition 2.2*
  for $i \in \{\alpha + 1, \ldots, \beta\}$ **do**
    Calculate the zero coupon bond price, ${}_mP(T_\alpha, T_i)$ *through* (2.3)
  Calculate the payoff for m^{th} *path*

$$ {}_mPS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right] = P(t, T_\alpha) N \left(1 - {}_mP(T_\alpha, T_\beta) - K \sum_{i=\alpha+1}^{\beta} \delta_i \, {}_mP(T_\alpha, T_i)\right)^+ $$

$$ PS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right] \approx \frac{\sum_{m=1}^{M} {}_mPS\left[t; \{T_i\}_{i=\alpha}^{\beta}; K; N\right]}{M} $$

4 Zero-coupon futures and options on futures

This section discusses calculation steps for prices and expected excess returns associated with European options on risk-free zero-coupon futures.

4.1 Futures price

Consider a futures contract with maturity $\mathcal{T}_2$ on a zero-coupon bond maturing on $\mathcal{T}_3$. Its time-$\mathcal{T}_1$ price $F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$ is given by

$$ F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} = \mathbb{E}^{\mathbb{Q}}\left[P(\mathcal{T}_2, \mathcal{T}_3) \middle| \mathcal{F}_{\mathcal{T}_1}\right], $$

see for instance Björk (2009). Such expression can be calculated in closed-form, as indicated by the following theorem whose proof is found in "Appendix".

Theorem 4.1 *The time-$\mathcal{T}_1$ price of a zero-coupon bond futures whose maturity is $\mathcal{T}_2$ and whose underlying risk-free zero-coupon bond matures at $\mathcal{T}_3$ is*

$$ F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} = \tilde{A}_{\tau_2,\tau_3} \exp\left[-\Delta \sum_{i=1}^{3} \tilde{\mathcal{B}}_{\tau_3}^{(i)} X_{\mathcal{T}_1}^{(i)}\right] \tag{4.1} $$

with $\tau_2 = \mathcal{T}_2 - \mathcal{T}_1$, $\tau_3 = \mathcal{T}_3 - \mathcal{T}_2$, $\mathcal{V}_{\tau_2}$ being defined in Proposition 2.2 and

$$\begin{aligned}\tilde{A}_{\tau_2,\tau_3} = A_{\tau_3} \exp\Bigg[&\frac{\Delta^2}{2}\mathcal{B}_{\tau_3}^\top \mathcal{V}_{\tau_2}\mathcal{B}_{\tau_3} - \Delta\mathcal{B}_{\tau_3}^{(2)}(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}})(1-(1-\lambda)^{\tau_2}) \\ &- \Delta\mathcal{B}_{\tau_3}^{(2)}\lambda\theta_3^{\mathbb{Q}}\left(\frac{\zeta_0(1-\lambda,\tau_2+1)}{1-\lambda} - \tau_2(1-\lambda)^{\tau_2-1}\right) \\ &- \Delta\mathcal{B}_{\tau_3}^{(3)}\theta_3^{\mathbb{Q}}(1-(1-\lambda)^{\tau_2})\Bigg],\end{aligned}$$

$$\tilde{\mathcal{B}}_n^{(1)} = \mathcal{B}_n^{(1)}, \quad \tilde{\mathcal{B}}_n^{(2)} = \mathcal{B}_n^{(2)}(1-\lambda)^n, \quad \tilde{\mathcal{B}}_n^{(3)} = \mathcal{B}_n^{(3)}(1-\lambda)^n + \mathcal{B}_n^{(2)}\lambda n(1-\lambda)^{n-1}.$$

4.2 Price for options on futures

The closed-form solution for futures prices leads to a Black-Scholes-type formula for the price of options on the zero-coupon futures. Indeed, denote by $\mathrm{Call}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K)$ the time-t price of a European option with strike K maturing at $\mathcal{T}_1$ on a futures maturing at $\mathcal{T}_2$ whose underlying asset is a zero-coupon maturing at $\mathcal{T}_3$.

Recall the following result, see for instance Lemma A.1 from Godin (2019) for a proof. Suppose Y is a Gaussian random variable with mean μ and standard deviation σ. Then $\mathbb{E}\left[e^Y \mathbb{1}_{\{Y>y\}}\right] = e^{\mu+\sigma^2/2}\bar{\Phi}\left(\frac{y-\mu-\sigma^2}{\sigma}\right)$ where $\bar{\Phi}$ is the survival function (i.e. one minus the CDF) of the Gaussian distribution.

Lemma 4.1 *Suppose Y is a Gaussian random variable with mean μ and standard deviation σ. Then,*

$$\begin{aligned}\mathbb{E}[\max(0, e^Y - K)] =& \mathbb{E}[e^Y \mathbb{1}_{\{Y>\log(K)\}}] - K\mathbb{E}[\mathbb{1}_{\{Y>\log(K)\}}] \\ =& e^{\mu+\sigma^2/2}\bar{\Phi}\left(\frac{\log(K)-\mu-\sigma^2}{\sigma}\right) - K\bar{\Phi}\left(\frac{\log(K)-\mu}{\sigma}\right).\end{aligned}$$

The following result is obtained by combining Proposition 2.2 with (4.1).

Lemma 4.2 *The forward measure $\mathbb{Q}^{\mathcal{T}_1}$ distribution of time-$\mathcal{T}_1$ the log-futures price $\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$ conditional on time-t information is Gaussian with mean $\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$ and variance $\varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$, where*

$$\begin{aligned}\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} =& \log\tilde{A}_{\tau_2,\tau_3} - \Delta\sum_{i=1}^{3}\tilde{\mathcal{B}}_{\tau_3}^{(i)}\mathcal{M}_{t,\tau_1}^{(i)}, \\ \varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} =& \Delta\tilde{\mathcal{B}}_{\tau_3}^\top \mathcal{V}_{\tau_1}\tilde{\mathcal{B}}_{\tau_3},\end{aligned}$$

with $\tau_1 = \mathcal{T}_1 - t$ and $\tilde{\mathcal{B}}_\tau = \left[\tilde{\mathcal{B}}_\tau^{(1)},\ \tilde{\mathcal{B}}_\tau^{(2)},\ \tilde{\mathcal{B}}_\tau^{(3)}\right]^\top$.

Using Lemma 4.1, the time-t call option price is

$$\begin{aligned}\mathrm{Call}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) =& P(t,\mathcal{T}_1)\mathbb{E}^{\mathcal{T}_1}[\max(0, F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - K)|\mathcal{F}_t] \\ =& P(t,\mathcal{T}_1)\left[e^{\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}+\varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}/2}\bar{\Phi}\left(\frac{\log(K) - \nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - \varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{\varsigma_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}\right)\right. \\ & \left. - K\bar{\Phi}\left(\frac{\log(K) - \nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{\varsigma_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}\right)\right].\end{aligned}$$

Furthermore, denoting by $\mathrm{Put}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K)$ the corresponding European put option, the put-call parity leads to

$$\mathrm{Put}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) = \mathrm{Call}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) - (\mathbb{E}^{\mathcal{T}_1}[F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}|\mathcal{F}_t] - K)P(t,\mathcal{T}_1)$$

where Lemma 4.2 leads to

$$\mathbb{E}^{\mathcal{T}_1}[F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}|\mathcal{F}_t] = \exp\left(\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + \varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}/2\right).$$

4.3 Price for quadratic options on futures

A quadratic option on futures with time-$\mathcal{T}_1$ payoff $\left(\frac{F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} - 1\right)^2$ can also be considered. Such option bears resemblance to a straddle option since it is more likely to produce higher payoffs in higher volatility environments for interest rates. Using Lemma 4.2, its time-t price is given by

$$\begin{aligned}\mathrm{Quad}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} =& P(t,\mathcal{T}_1)\mathbb{E}^{\mathcal{T}_1}\left[\left(\frac{F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} - 1\right)^2\middle|\mathcal{F}_t\right] \\ =& P(t,\mathcal{T}_1)\mathbb{E}^{\mathcal{T}_1}\left[\frac{\exp(2\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})}{F^2_{t,\mathcal{T}_2,\mathcal{T}_3}} - 2\frac{\exp(\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} + 1\middle|\mathcal{F}_t\right] \\ =& P(t,\mathcal{T}_1)\left[\frac{\exp\left(2\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + 2\varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}\right)}{F^2_{t,\mathcal{T}_2,\mathcal{T}_3}}\right. \\ & \left. -2\frac{\exp\left(\nu_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + \varsigma^2_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}/2\right)}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} + 1\middle|\mathcal{F}_t\right].\end{aligned}$$

4.4 Option expected excess returns

Consider a European-type derivative whose time-t price is Price_t and whose time $\mathcal{T}_1$ payoff is $\text{Payoff}_{\mathcal{T}_1}$. Its (periodic) expected excess return (EER) could be calculated in two ways:

$$\begin{aligned}\text{EER}^{\text{Approach 1}}_{t,\mathcal{T}_1} &= \frac{1}{\mathcal{T}_1 - t}\log\frac{\mathbb{E}^{\mathbb{P}}\left[\text{Payoff}_{\mathcal{T}_1}|\mathcal{F}_t\right]}{\text{Price}_{t,\mathcal{T}_1}} - s(t,\mathcal{T}_1),\\ \text{EER}^{\text{Approach 2}}_{t,\mathcal{T}_1} &= \frac{1}{\mathcal{T}_1 - t}\log\frac{\mathbb{E}^{\mathbb{P}}\left[D(t,\mathcal{T}_1)\text{Payoff}_{\mathcal{T}_1}|\mathcal{F}_t\right]}{\text{Price}_{t,\mathcal{T}_1}},\end{aligned} \tag{4.2}$$

where the risk-free spot rate is obtained through $s(t,\mathcal{T}_1) = -\frac{1}{\mathcal{T}_1 - t}\log P(t,\mathcal{T}_1)$. The first formulation relies on a future value perspective, whereas the second sees the expected excess return through the lens of a present value. The second approach has the conceptual advantage of producing an exactly null premium if $\mathbb{P} = \mathbb{Q}$. Nevertheless, it is more cumbersome to compute and as such we consider (4.2) in this work. Bakshi et al. (2023) also use a formulation similar to (4.2) in their work.

The expected excess return for the European call, the European put and the quadratic options presented above are therefore respectively

$$\text{EER}^{\text{Call}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) = \frac{1}{\mathcal{T}_1 - t}\log\frac{\mathbb{E}^{\mathbb{P}}\left[\max(0;F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - K)|\mathcal{F}_t\right]}{\text{Call}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K)} - s(t,\mathcal{T}_1), \tag{4.3}$$

$$\text{EER}^{\text{Put}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) = \frac{1}{\mathcal{T}_1 - t}\log\frac{\mathbb{E}^{\mathbb{P}}\left[\max(0;K - F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})|\mathcal{F}_t\right]}{\text{Put}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K)} - s(t,\mathcal{T}_1), \tag{4.4}$$

$$\text{EER}^{\text{Quad}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K) = \frac{1}{\mathcal{T}_1 - t}\log\frac{\mathbb{E}^{\mathbb{P}}\left[\left(\frac{F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} - 1\right)^2\middle|\mathcal{F}_t\right]}{\text{Quad}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}(K)} - s(t,\mathcal{T}_1). \tag{4.5}$$

The following result is obtained by combining Proposition 2.3 with (4.1).

Lemma 4.3 *Assuming $\kappa^{\mathbb{P}}_{2,2} \neq \kappa^{\mathbb{P}}_{3,3}$, the $\mathbb{P}$-distribution of time-$\mathcal{T}_1$ the log-futures price $\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$ conditional on time-t information is Gaussian with mean $\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}$ and variance $(\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2$, where*

$$\begin{aligned}\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} &= \log\tilde{A}_{\tau_2,\tau_3} - \Delta\sum_{i=1}^{3}\tilde{\mathcal{B}}^{(i)}_{\tau_3}\mathcal{M}^{\mathbb{P},(i)}_{t,\tau_1},\\ (\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2 &= \Delta\tilde{\mathcal{B}}^{\top}_{\tau_3}\mathcal{V}^{\mathbb{P}}_{\tau_1}\tilde{\mathcal{B}}_{\tau_3},\end{aligned}$$

with $\tau_1 = \mathcal{T}_1 - t$ and $\tilde{\mathcal{B}}_\tau = \left[\tilde{\mathcal{B}}^{(1)}_\tau,\ \tilde{\mathcal{B}}^{(2)}_\tau,\ \tilde{\mathcal{B}}^{(3)}_\tau\right]^{\top}$.

Again, using Lemma 4.1, the time-t European call and put option expected payoffs are respectively

$$\mathbb{E}^{\mathbb{P}}[\max(0, F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - K)|\mathcal{F}_t] = \Bigg[e^{\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + (\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2/2} \times \bar{\Phi}\left(\frac{\log(K) - \nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - (\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2}{\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}} \right) - K\bar{\Phi}\left(\frac{\log(K) - \nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}} \right) \Bigg],$$

$$\mathbb{E}^{\mathbb{P}}[\max(0, K - F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})|\mathcal{F}_t] = \mathbb{E}^{\mathbb{P}}[\max(0, F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} - K)|\mathcal{F}_t] - (\mathbb{E}^{\mathbb{P}}[F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}|\mathcal{F}_t] - K)$$

where, from (4.2),

$$\mathbb{E}^{\mathbb{P}}[F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}|\mathcal{F}_t] = \exp\left(\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + (\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2/2 \right).$$

Moreover, the quadratic option's expected payoff under $\mathbb{P}$ is

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\left(\frac{F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3}}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} - 1 \right)^2 \middle| \mathcal{F}_t \right] &= \mathbb{E}^{\mathbb{P}}\left[\frac{\exp(2\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})}{F^2_{t,\mathcal{T}_2,\mathcal{T}_3}} - 2\frac{\exp(\log F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} + 1 \middle| \mathcal{F}_t \right] \\ &= \Bigg[\frac{\exp\left(2\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + 2(\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2 \right)}{F^2_{t,\mathcal{T}_2,\mathcal{T}_3}} \\ &\quad -2\frac{\exp\left(\nu^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} + (\varsigma^{\mathbb{P}}_{t,\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3})^2/2 \right)}{F_{t,\mathcal{T}_2,\mathcal{T}_3}} + 1 \Bigg| \mathcal{F}_t \Bigg]\end{aligned}$$

Substituting the above formulas in (4.3)–(4.5) provides values for the option's expected excess return.

5 Methods for the calibration of the DTANFS model to option prices

While a full-blown calibration of the model to interest rate derivatives prices is left out-of-scope, we briefly highlight potential approaches for such a purpose. Assume a set of d derivatives prices $Y_t = [Y_t^{(1)}, \dots, Y_t^{(d)}]$ is available on any period t, each of which are associated with a set of deterministically chosen strike prices $K_t = [K_t^{(1)}, \dots, K_t^{(d)}]$.[2] Observed derivatives prices are assumed to be a noisy version

[2] To produce more homogeneous errors N_t, it might be desirable to fix the moneyness of options rather than their strike price.

of their true prices, and thus we can consider the following system of equations to depict the dynamics of derivatives prices:

$$Y_t = G(X_t, K_t) + N_t, \quad X_{t+1} = \mathbf{a} + bX_t + \Sigma Z_{t+1}, \quad t = 0, \dots, T, \tag{5.1}$$

where G is the non-linear function mapping risk factors and strikes into option prices, and the process $N = \{N_t\}_{t=0}^T$ is assumed to be a Gaussian d-dimensional white noise. Furthermore, $\mathbf{a} = \kappa^{\mathbb{P}}\theta^{\mathbb{P}}$ and $b = I - \kappa^{\mathbb{P}}$ with I being the 3×3 identity matrix.

Since (5.1) involves a non-linear transformation G of the Gaussian latent factors, the conventional Kalman filter cannot be applied. Nevertheless, the unscented Kalman filter (UKF) developed in Julier and Uhlmann (1997) can be used, as the non-linear system (5.1) is a special case of their equations (1)-(2). The UKF is a generalization of the Kalman filter allowing to tackle non-linearities through a deterministic sampling method leading to a better approximation of filtered moments of the observable quantities. An alternative approach consists in using particle filters, which instead apply stochastic sampling of latent quantities. See for instance Del Moral (1997), Creal (2012) or Remillard (2013) for more information about particle filters. Both the UKF and particle filters have been applied to term structure models estimation in the literature, for instance by Christoffersen et al. (2014).

6 Conclusion

This paper describes how to calculate prices of swaptions and European options (either conventional or quadratic) on zero-coupon futures under the DTAFNS model. Expressions for the option expected excess return associated with the European options are also provided. Whereas Monte-Carlo simulation is used for swaptions, closed-form solutions are provided for the zero-coupon futures options. All pricing expressions are obtained after deriving exact formulas for transition distributions of risk factors underlying the term structure dynamics under the following three respective measures: the physical measure, the risk-neutral measure and the forward measure. A potential future work could consist of studying option risk premia produced by the DTAFNS model and determining whether or not they are consistent with empirical stylized facts outlined for instance in Bakshi et al. (2023). Tackling American options could also be an interesting subsequent work.

Appendix: Proofs

Before proving Proposition 2.1, several lemmas are presented.

Lemma 6.1 *For $i = 1, 2, 3$ and any integer $\tau > 1$,*

$$\begin{aligned}\mathcal{B}_{\tau}^{(i)} - 1 &= \mathcal{B}_{\tau-1}^{(i)}\left(1 - \kappa_{i,i}^{\mathbb{Q}}\right) - \mathbb{1}_{\{i=3\}}(1-\lambda)^{\tau-1} \\ &= \mathcal{B}_{\tau-1}^{(i)}(1 - \lambda\mathbb{1}_{\{i>1\}}) - \mathbb{1}_{\{i=3\}}(1-\lambda)^{\tau-1}.\end{aligned}$$

Proof of Lemma 6.1 See Lemma A.2 of Eghbalzadeh et al. (2022).

Lemma 6.2 *The following recursive relationship between the time-t and time-$t+1$ zero-coupon bond prices presented in (2.3) holds for any integer $\tau > 0$:*

$$P(t+1, t+\tau) = P(t, t+\tau)e^{\Delta r_t} \exp\left[\log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}_{\tau-1}^{\top}\left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \Sigma Z_{t+1}^{\mathbb{Q}}\right)\right].$$

Proof of Lemma 6.2 The case $\tau = 1$ is trivial. For $\tau > 1$, using (2.1) and (2.3) for the first equality, and then Lemma 6.1 for the third one,

$$\begin{aligned}\log\left(\frac{P(t+1,t+\tau)}{P(t,t+\tau)}\right) - \Delta r_t &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\sum_{i=1}^{3}\Big(\mathcal{B}_{\tau-1}^{(i)}X_{t+1}^{(i)} \\ &\quad -\left(\mathcal{B}_{\tau}^{(i)} - 1\right)X_t^{(i)}\Big) + \Delta\sum_{i=1}^{3}X_t^{(i)} - \Delta(X_t^{(1)} + X_t^{(2)}) \\ &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\sum_{i=1}^{2}\mathcal{B}_{\tau-1}^{(i)}\left(X_{t+1}^{(i)} - \left(\frac{\mathcal{B}_{\tau}^{(i)} - 1}{\mathcal{B}_{\tau-1}^{(i)}}\right)X_t^{(i)}\right) + \Delta X_t^{(3)} \\ &\quad - \Delta\left(\mathcal{B}_{\tau-1}^{(3)}X_{t+1}^{(3)} - \left(\mathcal{B}_{\tau}^{(3)} - 1\right)X_t^{(3)}\right) \\ &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\sum_{i=1}^{2}\mathcal{B}_{\tau-1}^{(i)}\left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}}\right)X_t^{(i)}\right) + \Delta X_t^{(3)} \\ &\quad - \left(\Delta\mathcal{B}_{\tau-1}^{(3)}X_{t+1}^{(3)} - \left(\Delta\mathcal{B}_{\tau-1}^{(3)}(1-\lambda) - (1-\lambda)^{\tau-1}\right)X_t^{(3)}\right) \\ &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\sum_{i=1}^{3}\mathcal{B}_{\tau-1}^{(i)}\left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}}\right)X_t^{(i)}\right) \\ &\quad + \Delta\left(1 - (1-\lambda)^{\tau-1}\right)X_t^{(3)} \\ &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}_{\tau-1}^{\top}\left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \Sigma Z_{t+1}^{\mathbb{Q}}\right) + \Delta\left(1 - (1-\lambda)^{\tau-1} - \mathcal{B}_{\tau-1}^{(2)}\lambda\right)X_t^{(3)} \\ &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}_{\tau-1}^{\top}\left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \Sigma Z_{t+1}^{\mathbb{Q}}\right).\end{aligned}$$

Therefore,

$$P(t+1, t+\tau) = P(t, t+\tau)e^{\Delta r_t} \exp\left[\log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}_{\tau-1}^{\top}\left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \Sigma Z_{t+1}^{\mathbb{Q}}\right)\right].$$

□

Lemma 6.3 *Consider any integer $\tau > 0$ and any real number r. Using the convention $0^0 = 0$, for functions $\zeta_0(r,\tau)$, $\zeta_1(r,\tau)$ and $\zeta_2(r,\tau)$ defined in (2.7)–(2.9),*

$$\zeta_0(r,\tau-1) - \zeta_0(r,\tau) = -r^{\tau-1}, \tag{6.1}$$

$$\zeta_1(r,\tau-1) - \zeta_1(r,\tau) = -r^{\tau-1}(\tau-1), \tag{6.2}$$

$$\zeta_2(r,\tau-1) - \zeta_2(r,\tau) = -r^{\tau-1}(\tau-1)^2. \tag{6.3}$$

Proof of Lemma 6.3 This result is a direct consequence of the sum representations of $\zeta_0(r,\tau)$, $\zeta_1(r,\tau)$ and $\zeta_2(r,\tau)$ provided in (2.7)–(2.9). □

Lemma 6.4 *The following recursive connection holds for the quantity υ_τ defined in (2.6) for any integer $\tau > 0$:*

$$\upsilon_\tau = \mathcal{B}_{\tau-1}^\top \Sigma\rho(\mathcal{B}_{\tau-1}^\top \Sigma)^\top + \upsilon_{\tau-1}. \tag{6.4}$$

Proof of Lemma 6.4 The case $\tau = 1$ is trivial as it leads to $0 = 0$. For $\tau > 1$, first,

$$\mathcal{B}_{\tau-1}^\top \Sigma\rho(\mathcal{B}_{\tau-1}^\top \Sigma)^\top = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathcal{B}_{\tau-1}^{(i)}\mathcal{B}_{\tau-1}^{(j)}\Sigma_{i,i}\Sigma_{j,j}\rho_{i,j}.$$

Moreover, based on (2.6),

$$\upsilon_\tau - \upsilon_{\tau-1} = \sum_{i=1}^{3}\sum_{j=1}^{3} a_\tau^{(i,j)}\Sigma_{i,i}\Sigma_{j,j}\rho_{i,j},$$

with $a_\tau^{(i,j)} = (\upsilon_\tau^{(i,j)} - \upsilon_{\tau-1}^{(i,j)})/(\Sigma_{i,i}\Sigma_{j,j}\rho_{i,j})$.

To complete the proof, we now show that $a_\tau^{(i,j)} = \mathcal{B}_{\tau-1}^{(i)}\mathcal{B}_{\tau-1}^{(j)}$ for any $i,j = 1,2,3$.

First, for $i = j = 1$,

$$\begin{aligned} a_\tau^{(1,1)} &= \frac{\tau(\tau-1)(2\tau-1)}{6} - \frac{(\tau-1)(\tau-2)(2\tau-3)}{6}. \\ &= \frac{(\tau-1)(2\tau^2-\tau-2\tau^2+7\tau-6)}{6} \\ &= (\tau-1)^2 \\ &= \mathcal{B}_{\tau-1}^{(1)}\mathcal{B}_{\tau-1}^{(1)}. \end{aligned}$$

Secondly, for $i = 1$ and $j = 2$, using (6.2),

$$\begin{aligned} a_\tau^{(1,2)} &= \frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1(1-\lambda,\tau) - \frac{(\tau-1)(\tau-2)}{2} + \zeta_1(1-\lambda,\tau-1)\right) \\ &= \frac{(\tau-1)-(1-\lambda)^{\tau-1}(\tau-1)}{\lambda} \\ &= \mathcal{B}_{\tau-1}^{(1)}\mathcal{B}_{\tau-1}^{(2)}. \end{aligned}$$

Thirdly, consider $i = 1$ and $j = 3$. For $\tau = 2$, $a_2^{(1,3)} = \mathcal{B}_1^{(1)}\mathcal{B}_1^{(3)} = 0$ since $v_2^{(1,3)} = v_1^{(1,3)} = 0$. For $\tau > 2$, using (6.1), (6.2), and (6.3),

$$\begin{aligned} a_\tau^{(1,3)} &= \frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0(1-\lambda,\tau-1) - (1+\lambda)\zeta_1(1-\lambda,\tau-1) \\ &\quad - \lambda\zeta_2(1-\lambda,\tau-1) - \frac{(\tau-1)(\tau-2)}{2} + 1 + \zeta_0(1-\lambda,\tau-2) \\ &\quad + (1+\lambda)\zeta_1(1-\lambda,\tau-2) + \lambda\zeta_2(1-\lambda,\tau-2)\Bigg] \\ &= \frac{1}{\lambda}\left[\tau - 1 - (1-\lambda)^{\tau-2} - (1-\lambda)^{\tau-2}(\tau-2)(1+\lambda) - \lambda(1-\lambda)^{\tau-2}(\tau-2)^2\right] \\ &= \frac{1}{\lambda}\Bigg[\tau - 1 - (1-\lambda)^{\tau-2} - (1-\lambda)^{\tau-2}(\tau-1-1)(1+\lambda) \\ &\quad - \lambda(1-\lambda)^{\tau-2}(\tau-2)(\tau-1-1)\Bigg] \\ &= \frac{1}{\lambda}\Bigg[\tau - 1 - (1-\lambda)^{\tau-2}(\tau-1)(1+\lambda) \\ &\quad + \lambda(1-\lambda)^{\tau-2} - \lambda(1-\lambda)^{\tau-2}(\tau-2)(\tau-1) + \lambda(1-\lambda)^{\tau-2}(\tau-1-1)\Bigg] \\ &= \frac{\tau-1}{\lambda}\left[1 - (1-\lambda)^{\tau-2} - \lambda(1-\lambda)^{\tau-2}(\tau-2)\right] \\ &= \mathcal{B}_{\tau-1}^{(1)}\mathcal{B}_{\tau-1}^{(3)}. \end{aligned}$$

Fourthly, for $i = j = 2$,

$$\begin{aligned} a_\tau^{(2,2)} &= \frac{1}{\lambda^2}\Bigg(\tau - 2\left[\frac{1-(1-\lambda)^\tau}{\lambda}\right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2} - \tau + 1 + 2\left[\frac{1-(1-\lambda)^{\tau-1}}{\lambda}\right] \\ &\quad - \frac{1-(1-\lambda)^{2(\tau-1)}}{1-(1-\lambda)^2}\Bigg) \\ &= \frac{1}{\lambda^2}\left(1 - 2(1-\lambda)^{\tau-1} + (1-\lambda)^{2(\tau-1)}\right) \\ &= \mathcal{B}_{\tau-1}^{(2)}\mathcal{B}_{\tau-1}^{(2)}. \end{aligned}$$

Fifthly, consider $i = 2$ and $j = 3$. For $\tau = 2$, $a_2^{(2,3)} = \mathcal{B}_1^{(2)}\mathcal{B}_1^{(3)} = 0$ since $v_2^{(2,3)} = v_1^{(2,3)} = 0$. For $\tau > 2$, using (6.1) and (6.2),

$$
\begin{aligned}
a_\tau^{(2,3)} &= \frac{\tau-2-(2-\lambda)\zeta_0(1-\lambda,\tau-1)+(1-\lambda)\zeta_0\big((1-\lambda)^2,\tau-1\big)}{\lambda^2} \\
&\quad + \frac{-\zeta_1(1-\lambda,\tau-1)+(1-\lambda)\zeta_1\big((1-\lambda)^2,\tau-1\big)}{\lambda} \\
&\quad - \frac{\tau-3-(2-\lambda)\zeta_0(1-\lambda,\tau-2)+(1-\lambda)\zeta_0\big((1-\lambda)^2,\tau-2\big)}{\lambda^2} \\
&\quad - \frac{-\zeta_1(1-\lambda,\tau-2)+(1-\lambda)\zeta_1\big((1-\lambda)^2,\tau-2\big)}{\lambda} \\
&= \frac{1-(2-\lambda)(1-\lambda)^{\tau-2}+(1-\lambda)(1-\lambda)^{2(\tau-2)}}{\lambda^2} \\
&\quad + \frac{-(1-\lambda)^{\tau-2}(\tau-2)+(1-\lambda)(1-\lambda)^{2(\tau-2)}(\tau-2)}{\lambda} \\
&= \frac{\big(1-(1-\lambda)^{\tau-2}\big)^2+\lambda(1-\lambda)^{\tau-2}\big(1-(1-\lambda)^{\tau-2}\big)}{\lambda^2} \\
&\quad + \frac{(1-\lambda)^{\tau-2}(\tau-2)\big((1-\lambda)^{\tau-1}-1\big)}{\lambda} \\
&= \frac{\big(1-(1-\lambda)^{\tau-2}\big)\big(1-(1-\lambda)^{\tau-2}(1-\lambda)\big)}{\lambda^2} + \frac{(1-\lambda)^{\tau-2}(\tau-2)\big((1-\lambda)^{\tau-1}-1\big)}{\lambda} \\
&= \frac{1-(1-\lambda)^{\tau-1}}{\lambda}\left(\frac{1-(1-\lambda)^{\tau-2}}{\lambda}-(1-\lambda)^{\tau-2}(\tau-2)\right) \\
&= \mathcal{B}_{\tau-1}^{(2)}\mathcal{B}_{\tau-1}^{(3)}.
\end{aligned}
$$

Lastly, consider $i=j=3$. For $\tau=2$, $a_2^{(3,3)}=\mathcal{B}_1^{(3)}\mathcal{B}_1^{(3)}=0$ since $\upsilon_2^{(3,3)}=\upsilon_1^{(3,3)}=0$. For $\tau>2$, using (6.1), (6.2), and (6.3),

$$
\begin{aligned}
a_\tau^{(3,3)} &= \frac{1}{\lambda^2}\Big[\tau-2+\zeta_0\big((1-\lambda)^2,\tau-1\big)+\lambda^2\zeta_2\big((1-\lambda)^2,\tau-1\big) \\
&\quad -2\zeta_0(1-\lambda,\tau-1)-2\lambda\zeta_1(1-\lambda,\tau-1)+2\lambda\zeta_1\big((1-\lambda)^2,\tau-1\big) \\
&\quad -\tau+3-\zeta_0\big((1-\lambda)^2,\tau-2\big)-\lambda^2\zeta_2\big((1-\lambda)^2,\tau-2\big) \\
&\quad +2\zeta_0(1-\lambda,\tau-2)+2\lambda\zeta_1(1-\lambda,\tau-2)-2\lambda\zeta_1\big((1-\lambda)^2,\tau-2\big)\Big] \\
&= \frac{1}{\lambda^2}\Big[1+(1-\lambda)^{2(\tau-2)}+\lambda^2(1-\lambda)^{2(\tau-2)}(\tau-2)^2-2(1-\lambda)^{\tau-2} \\
&\quad -2\lambda(1-\lambda)^{\tau-2}(\tau-2)+2\lambda(1-\lambda)^{2(\tau-2)}(\tau-2)\Big] \\
&= \frac{1}{\lambda^2}\Big[\big(1-(1-\lambda)^{\tau-2}\big)^2+\lambda^2(1-\lambda)^{2(\tau-2)}(\tau-2)^2 \\
&\quad -2\lambda(1-\lambda)^{\tau-2}(\tau-2)\big(1-(1-\lambda)^{\tau-2}\big)\Big] \\
&= \frac{1}{\lambda^2}\Big[\big(1-(1-\lambda)^{\tau-2}-\lambda(1-\lambda)^{\tau-2}(\tau-2)\big)^2\Big] \\
&= \left(\frac{1-(1-\lambda)^{\tau-2}}{\lambda}-(1-\lambda)^{\tau-2}(\tau-2)\right)^2 \\
&= \mathcal{B}_{\tau-1}^{(3)}\mathcal{B}_{\tau-1}^{(3)}.
\end{aligned}
$$

□

Lemma 6.5 *For $i = 1, 2, 3$ and any integer $\tau > 0$, the following recursive relationships between $\mathcal{B}_\tau^{(i)}$ and $\mathcal{B}_{\tau-1}^{(i)}$ hold*:

$$\begin{aligned}\mathcal{B}_\tau^{(1)} =& \mathcal{B}_{\tau-1}^{(1)} + 1,\\ \mathcal{B}_\tau^{(2)} =& \mathcal{B}_{\tau-1}^{(2)} + (1-\lambda)^{\tau-1},\\ \mathcal{B}_\tau^{(3)} =& \mathcal{B}_{\tau-1}^{(3)} + (1-\lambda)^{\tau-2}\lambda(\tau-1).\end{aligned}$$

Proof of Lemma 6.5 The case $\tau = 1$ is trivial. For $\tau > 1$, based on (2.4), $\mathcal{B}_\tau^{(1)} = \tau - 1 + 1 = \mathcal{B}_{\tau-1}^{(1)} + 1$. Furthermore,

$$\begin{aligned}\mathcal{B}_\tau^{(2)} &= \frac{1-(1-\lambda)^\tau}{\lambda} - 1 + 1 = (1-\lambda)\mathcal{B}_{\tau-1}^{(2)} + 1\\ &= \mathcal{B}_{\tau-1}^{(2)} - 1 + (1-\lambda)^{\tau-1} + 1 = \mathcal{B}_{\tau-1}^{(2)} + (1-\lambda)^{\tau-1}.\end{aligned}$$

Lastly,

$$\begin{aligned}\mathcal{B}_\tau^{(3)} &= \frac{1-(1-\lambda)^{\tau-1}}{\lambda} - 1 + 1 - (\tau-2+1)(1-\lambda)^{\tau-1}\\ &= (1-\lambda)\left(\frac{1-(1-\lambda)^{\tau-2}}{\lambda} + \frac{1}{1-\lambda} - (\tau-2)(1-\lambda)^{\tau-2} - (1-\lambda)^{\tau-2}\right)\\ &= \mathcal{B}_{\tau-1}^{(3)} + \frac{1}{1-\lambda} - (1-\lambda)^{\tau-2}\\ &\quad - 1 + (1-\lambda)^{\tau-2} - \frac{\lambda}{1-\lambda} + \lambda(\tau-2)(1-\lambda)^{\tau-2} + \lambda(1-\lambda)^{\tau-2}\\ &= \mathcal{B}_{\tau-1}^{(3)} + (1-\lambda)^{\tau-2}\lambda(\tau-1).\end{aligned}$$

□

Lemma 6.6 *The following recursive relationship holds for quantity A_τ defined in* (2.3) *and any integer $\tau > 0$*:

$$A_\tau = A_{\tau-1}\exp\left(\frac{1}{2}\Delta^2\mathcal{B}_{\tau-1}^\top\Sigma\rho(\mathcal{B}_{\tau-1}^\top\Sigma)^\top - \Delta\mathcal{B}_{\tau-1}^\top\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}}\right).$$

Proof of Lemma 6.6 For $\tau = 1$, the proof is trivial. For $\tau > 1$, using Lemma 6.4, Lemma 6.5 and (2.4) to substitute into (2.5),

$$
\begin{aligned}
\log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}_{\tau-1}^\top \kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} &= -\Delta\theta_2^{\mathbb{Q}}\left(\mathcal{B}_{\tau-1}^{(1)} - \mathcal{B}_\tau^{(1)} + \mathcal{B}_\tau^{(2)} - \mathcal{B}_{\tau-1}^{(2)}\right) \\
&\quad + \Delta\theta_3^{\mathbb{Q}}\left(\mathcal{B}_{\tau-1}^{(3)} - \mathcal{B}_\tau^{(3)}\right) - \Delta\mathcal{B}_{\tau-1}^\top \kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} - \frac{1}{2}\Delta^2\left(\upsilon_\tau - \upsilon_{\tau-1}\right) \\
&= -\Delta\theta_2^{\mathbb{Q}}\left(-1 + (1-\lambda)^{\tau-1}\right) - \Delta\theta_3^{\mathbb{Q}}(1-\lambda)^{\tau-2}\lambda(\tau-1) \\
&\quad - \Delta\mathcal{B}_{\tau-1}^\top \kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} - \frac{1}{2}\Delta^2\left(\upsilon_\tau - \upsilon_{\tau-1}\right) \\
&= \Delta\theta_2^{\mathbb{Q}}\lambda\mathcal{B}_{\tau-1}^{(2)} - \Delta\theta_3^{\mathbb{Q}}\lambda(1-\lambda)^{\tau-2}(\tau-1) - \Delta\mathcal{B}_{\tau-1}^{(2)}\lambda(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}) - \Delta\mathcal{B}_{\tau-1}^{(3)}\lambda\theta_3^{\mathbb{Q}} \\
&\quad - \frac{1}{2}\Delta^2\left(\upsilon_\tau - \upsilon_{\tau-1}\right) \\
&= -\Delta\theta_3^{\mathbb{Q}}\lambda(1-\lambda)^{\tau-2}(\tau-1) + \Delta\lambda\theta_3^{\mathbb{Q}}(\mathcal{B}_{\tau-1}^{(2)} - \mathcal{B}_{\tau-1}^{(3)}) - \frac{1}{2}\Delta^2\mathcal{B}_{\tau-1}^\top\Sigma\rho(\mathcal{B}_{\tau-1}^\top\Sigma)^\top \\
&= -\frac{1}{2}\Delta^2\mathcal{B}_{\tau-1}^\top\Sigma\rho(\mathcal{B}_{\tau-1}^\top\Sigma)^\top.
\end{aligned}
$$

□

Proof of Proposition 2.1 The proof relies on calculating the moments generating function of innovations under the $\mathcal{T}$-forward measure. Consider the row vector $\Gamma = [\Gamma_1, \Gamma_2, \Gamma_3]$ with $\Gamma_i \in \mathbb{R}$ for $i = 1, 2, 3$. Then

$$
\begin{aligned}
\mathbb{E}^{\mathcal{T}}\left[\exp(\Gamma Z_{t+1}^{\mathcal{T}})\middle|\mathcal{F}_t\right] &= \mathbb{E}^{\mathcal{T}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}} + \Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1})\middle|\mathcal{F}_t\right] \\
&= \frac{\mathbb{E}^{\mathbb{Q}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}} + \Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1})\frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}}\middle|\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[\frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}}\middle|\mathcal{F}_t\right]} \\
&= \exp\left[\Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1}\right]\frac{\mathbb{E}^{\mathbb{Q}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}})\frac{P(t+\tau,t+\tau)B(0)}{P(0,t+\tau)B(t+\tau)}\middle|\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[\frac{P(t+\tau,t+\tau)B(0)}{P(0,t+\tau)B(t+\tau)}\middle|\mathcal{F}_t\right]} \\
&= \exp\left[\Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1}\right]\frac{1}{P(0,t+\tau)}\frac{\mathbb{E}^{\mathbb{Q}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}})\frac{P(t+\tau,t+\tau)}{B(t+\tau)}\middle|\mathcal{F}_t\right]}{\frac{P(t,t+\tau)}{P(0,t+\tau)B(t)}} \\
&= \exp\left[\Gamma\rho\Delta\Sigma\mathcal{B}_{\tau-1}\right]\frac{B(t)}{P(t,t+\tau)}\mathbb{E}^{\mathbb{Q}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}})\mathbb{E}^{\mathbb{Q}}\left[\frac{P(t+\tau,t+\tau)}{B(t+\tau)}\middle|\mathcal{F}_{t+1}\right]\middle|\mathcal{F}_t\right] \\
&= \exp\left[\Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1}\right]\mathbb{E}^{\mathbb{Q}}\left[\exp(\Gamma Z_{t+1}^{\mathbb{Q}})\frac{B(t)}{P(t,t+\tau)}\frac{P(t+1,t+\tau)}{B(t+1)}\middle|\mathcal{F}_t\right],
\end{aligned}
$$

where the fourth and fifth equalities rely on the fact that $\frac{P(\cdot,\mathcal{T})}{B(\cdot)}$ is a $\mathbb{Q}$-martingale.

Define

$$Y(Z^{\mathbb{Q}}_{t+1}) \equiv \Gamma Z^{\mathbb{Q}}_{t+1} + \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}^\top_{\tau-1}\left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \Sigma Z^{\mathbb{Q}}_{t+1}\right).$$

Using Lemma 6.2 therefore leads to

$$\mathbb{E}^{\mathcal{T}}\left[\exp\left(\Gamma Z^{\mathcal{T}}_{t+1}\right)\middle|\mathcal{F}_t\right] = \exp\left[\Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1}\right]\mathbb{E}^{\mathbb{Q}}\left[\exp\left(Y(Z^{\mathbb{Q}}_{t+1})\right)\middle|\mathcal{F}_t\right],$$

where, given $\mathcal{F}_t$, $Y(Z^{\mathbb{Q}}_{t+1})$ follows the Gaussian distribution with conditional mean and variance

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[Y(Z^{\mathbb{Q}}_{t+1})\middle|\mathcal{F}_t\right] &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta\mathcal{B}^\top_{\tau-1}\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}},\\ \mathrm{Var}^{\mathbb{Q}}\left[Y(Z^{\mathbb{Q}}_{t+1})\middle|\mathcal{F}_t\right] &= \Gamma\rho\Gamma^\top + \Delta\mathcal{B}^\top_{\tau-1}\Sigma\rho(\Delta\mathcal{B}^\top_{\tau-1}\Sigma)^\top - 2\Gamma\Delta\rho\Sigma\mathcal{B}_{\tau-1}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{E}^{\mathcal{T}}\left[\exp(\Gamma Z^{\mathcal{T}}_{t+1})\middle|\mathcal{F}_t\right] = \frac{A_{\tau-1}}{A_\tau}\exp\left(-\Delta\mathcal{B}^\top_{\tau-1}\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} + \frac{1}{2}\Delta\mathcal{B}^\top_{\tau-1}\Sigma\rho(\Delta\mathcal{B}^\top_{\tau-1}\Sigma)^\top\right)\times\\ \exp\left(\frac{1}{2}\Gamma\rho\Gamma^\top\right).\end{aligned}$$

Therefore, using Lemma 6.6 leads to

$$\mathbb{E}^{\mathcal{T}}\left[\exp(\Gamma Z^{\mathcal{T}}_{t+1})\middle|\mathcal{F}_t\right] = \exp\left(\frac{1}{2}\Gamma\rho\Gamma^\top\right).$$

□

Lemma 6.7 *Assume* (2.13) *holds. For* $t = 0, \dots, \mathcal{T}$ *and* $n = 0, \dots, \mathcal{T} - t$,

$$\begin{aligned}X^{(i)}_{t+n} &= X^{(i)}_t(1-\kappa^{\mathcal{T}}_{i,i})^n + \kappa^{\mathcal{T}}_{i,i}\theta^{\mathcal{T}}_i\sum_{l=1}^{n}(1-\kappa^{\mathcal{T}}_{i,i})^{(n-l)} + \Sigma_{i,i}\sum_{l=1}^{n}Z^{\mathcal{T}}_{t+l,i}(1-\kappa^{\mathcal{T}}_{i,i})^{(n-l)}\\ &+ \sum_{l=1}^{n}\sum_{j\neq i}^{3}\kappa^{\mathcal{T}}_{i,j}(\theta^{\mathcal{T}}_j - X^{(j)}_{t+l-1})(1-\kappa^{\mathcal{T}}_{i,i})^{(n-l)} - \sum_{l=0}^{n-1}\eta^{\mathcal{T},(i)}_{t+l}(1-\kappa^{\mathcal{T}}_{i,i})^{n-1-l}.\end{aligned} \tag{6.5}$$

Proof of Lemma 6.7 This proof is analogous to that of Lemma A.1 from Eghbalzadeh et al. (2022), which is based on induction. We apply the convention $\sum_{l=0}^{-1} x_l = \sum_{l=1}^{0} x_l \equiv 0$. The case $n = 0$ is therefore trivial. Then, assume (6.5) holds for some $n \leq \mathcal{T} - t - 1$. Using (2.13), for any $i = 1, 2, 3$,

$$
\begin{aligned}
X_{t+n+1}^{(i)} &= X_{t+n}^{(i)} + \sum_{j=1}^{3} \kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+n}^{(j)}) + \Sigma_{i,i} Z_{t+n+1,i}^{\mathcal{T}} - \eta_{t+n}^{\mathcal{T},(i)} \\
&= X_{t+n}^{(i)}(1-\kappa_{i,i}^{\mathcal{T}}) + \kappa_{i,i}^{\mathcal{T}}\theta_i^{\mathcal{T}} + \sum_{j\neq i}^{3} \kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+n}^{(j)}) + \Sigma_{i,i} Z_{t+n+1,i}^{\mathcal{T}} - \eta_{t+n}^{\mathcal{T},(i)}.
\end{aligned}
$$

Applying (6.5) in the latter equality yields

$$
\begin{aligned}
X_{t+n+1}^{(i)} &= X_t^{(i)}(1-\kappa_{i,i}^{\mathcal{T}})^{n+1} + \kappa_{i,i}^{\mathcal{T}}\theta_i^{\mathcal{T}} \sum_{l=1}^{n}(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} + \Sigma_{i,i}\sum_{l=1}^{n} Z_{t+l,i}^{\mathcal{T}}(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} \\
&\quad + \sum_{l=1}^{n}\sum_{j\neq i}^{3} \kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+l-1}^{(j)})(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} + \kappa_{i,i}^{\mathcal{T}}\theta_i^{\mathcal{T}} - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(i)}(1-\kappa_{i,i}^{\mathcal{T}})^{n-l} \\
&\quad + \sum_{j\neq i}^{3} \kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+n}^{(j)}) + \Sigma_{i,i} Z_{t+n+1,i}^{\mathcal{T}} - \eta_{t+n}^{\mathcal{T},(i)} \\
&= X_t^{(i)}(1-\kappa_{i,i}^{\mathcal{T}})^{n+1} + \kappa_{i,i}^{\mathcal{T}}\theta_i^{\mathcal{T}} \sum_{l=1}^{n+1}(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} + \Sigma_{i,i}\sum_{l=1}^{n+1} Z_{t+l,i}^{\mathcal{T}}(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} \\
&\quad + \sum_{l=1}^{n+1}\sum_{j\neq i}^{3} \kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+l-1}^{(j)})(1-\kappa_{i,i}^{\mathcal{T}})^{(n+1-l)} - \sum_{l=0}^{n}\eta_{t+l}^{\mathcal{T},(i)}(1-\kappa_{i,i}^{\mathcal{T}})^{n-l},
\end{aligned}
$$

thereby finishing the induction. □

Lemma 6.8 *For $n = 0, \ldots, \mathcal{T} - t$, the factors $X_{t+n}^{(i)}$ can be expressed in terms of X_t and innovations $\{Z_{t+l}^{\mathcal{T}}\}_{l=1}^{\mathcal{T}-t}$ as follows*:

$$
\begin{aligned}
X_{t+n}^{(1)} &= X_t^{(1)} + \Sigma_{1,1}\sum_{l=1}^{n} Z_{t+l,1}^{\mathcal{T}} - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(1)}, \\
X_{t+n}^{(2)} &= X_t^{(2)}(1-\lambda)^n + (\theta_2^{\mathcal{T}} - \theta_3^{\mathcal{T}})(1-(1-\lambda)^n) + \Sigma_{2,2}\sum_{l=1}^{n} Z_{t+l,2}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad + \lambda\Bigg(nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\Bigg(\frac{(1-(1-\lambda)^n)}{\lambda} - n(1-\lambda)^{n-1}\Bigg) \\
&\quad + \Sigma_{3,3}\sum_{k=1}^{n-1}(n-k)(1-\lambda)^{n-k-1} Z_{t+k,3}^{\mathcal{T}} - \sum_{k=0}^{n-1}(n-k-1)\eta_{t+k}^{\mathcal{T},(3)}(1-\lambda)^{n-k-2}\Bigg) \\
&\quad - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(2)}(1-\lambda)^{n-1-l}, \\
X_{t+n}^{(3)} &= X_t^{(3)}(1-\lambda)^n + \theta_3^{\mathcal{T}}(1-(1-\lambda)^n) + \Sigma_{3,3}\sum_{l=1}^{n} Z_{t+l,3}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(3)}(1-\lambda)^{n-1-l}.
\end{aligned}
$$

Proof of Lemma 6.8 From (2.12), $\kappa^{\mathcal{T}}_{1,1} = \kappa^{\mathcal{T}}_{1,2} = \kappa^{\mathcal{T}}_{1,3} = \kappa^{\mathcal{T}}_{2,1} = \kappa^{\mathcal{T}}_{3,1} = \kappa^{\mathcal{T}}_{3,2} = 0$, $\kappa^{\mathcal{T}}_{2,2} = \kappa^{\mathcal{T}}_{3,3} = \lambda$ and $\kappa^{\mathcal{T}}_{2,3} = -\lambda$. When placed into (6.5), this leads to

$$\begin{aligned}
X^{(1)}_{t+n} =& X^{(1)}_t + \Sigma_{1,1} \sum_{l=1}^{n} Z^{\mathcal{T}}_{t+l,1} - \sum_{l=0}^{n-1} \eta^{\mathcal{T},(1)}_{t+l}, \\
X^{(2)}_{t+n} =& X^{(2)}_t (1-\lambda)^n + \lambda \theta^{\mathcal{T}}_2 \sum_{l=1}^{n} (1-\lambda)^{(n-l)} + \Sigma_{2,2} \sum_{l=1}^{n} Z^{\mathcal{T}}_{t+l,2} (1-\lambda)^{(n-l)} \\
& - \lambda \sum_{l=1}^{n} (\theta^{\mathcal{T}}_3 - X^{(3)}_{t+l-1})(1-\lambda)^{(n-l)} - \sum_{l=0}^{n-1} \eta^{\mathcal{T},(2)}_{t+l} (1-\lambda)^{n-1-l},
\end{aligned} \tag{6.6}$$

$$\begin{aligned}
X^{(3)}_{t+n} =& X^{(3)}_t (1-\lambda)^n + \lambda \theta^{\mathcal{T}}_3 \sum_{l=1}^{n} (1-\lambda)^{(n-l)} + \Sigma_{3,3} \sum_{l=1}^{n} Z^{\mathcal{T}}_{t+l,3} (1-\lambda)^{(n-l)} \\
& - \sum_{l=0}^{n-1} \eta^{\mathcal{T},(3)}_{t+l} (1-\lambda)^{n-1-l}.
\end{aligned} \tag{6.7}$$

Furthermore,

$$\begin{aligned}
\sum_{l=1}^{n} (1-\lambda)^{n-l} X^{(3)}_{t+l-1} &= \sum_{l=0}^{n-1} (1-\lambda)^{n-l-1} X^{(3)}_{t+l} \\
&= \sum_{l=0}^{n-1} (1-\lambda)^{n-l-1} \Bigg[X^{(3)}_t (1-\lambda)^l + \lambda \theta^{\mathcal{T}}_3 \sum_{k=1}^{l} (1-\lambda)^{(l-k)} \\
&\quad + \Sigma_{3,3} \sum_{k=1}^{l} Z^{\mathcal{T}}_{t+k,3} (1-\lambda)^{(l-k)} - \sum_{k=0}^{l-1} (1-\lambda)^{l-1-k} \eta^{\mathcal{T},(3)}_{t+k} \Bigg] \\
&= n X^{(3)}_t (1-\lambda)^{n-1} + \lambda \theta^{\mathcal{T}}_3 \sum_{l=0}^{n-1} \frac{(1-\lambda)^{n-l-1} - (1-\lambda)^{n-1}}{\lambda} \\
&\quad + \Sigma_{3,3} \sum_{l=0}^{n-1} \sum_{k=1}^{n-1} \mathbb{1}_{\{k \le l\}} Z^{\mathcal{T}}_{t+k,3} (1-\lambda)^{n-k-1} - \sum_{l=0}^{n-1} \sum_{k=0}^{n-1} \mathbb{1}_{\{k \le l-1\}} (1-\lambda)^{n-k-2} \eta^{\mathcal{T},(3)}_{t+k} \\
&= n X^{(3)}_t (1-\lambda)^{n-1} + \theta^{\mathcal{T}}_3 \left(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1} \right) \\
&\quad + \Sigma_{3,3} \sum_{k=1}^{n-1} Z^{\mathcal{T}}_{t+k,3} \sum_{l=k}^{n-1} (1-\lambda)^{n-k-1} - \sum_{k=0}^{n-1} \eta^{\mathcal{T},(3)}_{t+k} \sum_{l=k+1}^{n-1} (1-\lambda)^{n-k-2} \\
&= n X^{(3)}_t (1-\lambda)^{n-1} + \theta^{\mathcal{T}}_3 \left(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1} \right) \\
&\quad + \Sigma_{3,3} \sum_{k=1}^{n-1} (n-k)(1-\lambda)^{n-k-1} Z^{\mathcal{T}}_{t+k,3} - \sum_{k=0}^{n-1} (n-k-1) \eta^{\mathcal{T},(3)}_{t+k} (1-\lambda)^{n-k-2}.
\end{aligned} \tag{6.8}$$

Using (6.8) in (6.6),

$$
\begin{aligned}
X_{t+n}^{(2)} &= X_t^{(2)}(1-\lambda)^n + \theta_2^{\mathcal{T}}(1-(1-\lambda)^n) + \Sigma_{2,2}\sum_{l=1}^{n} Z_{t+l,2}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad - \theta_3^{\mathcal{T}}(1-(1-\lambda)^n) \\
&\quad + \lambda\Bigg(nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\Bigg(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1}\Bigg) \\
&\quad + \Sigma_{3,3}\sum_{k=1}^{n-1}(n-k)(1-\lambda)^{n-k-1}Z_{t+k,3}^{\mathcal{T}} - \sum_{k=0}^{n-1}(n-k-1)\eta_{t+k}^{\mathcal{T},(3)}(1-\lambda)^{n-k-2}\Bigg) \\
&\quad - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(2)}(1-\lambda)^{n-1-l}.
\end{aligned}
$$

Moreover, using (2.7) in (6.7) leads to

$$
\begin{aligned}
X_{t+n}^{(3)} &= X_t^{(3)}(1-\lambda)^n + \theta_3^{\mathcal{T}}(1-(1-\lambda)^n) + \Sigma_{3,3}\sum_{l=1}^{n} Z_{t+l,3}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad - \sum_{l=0}^{n-1}\eta_{t+l}^{\mathcal{T},(3)}(1-\lambda)^{n-1-l}.
\end{aligned}
$$

□

Proof of Proposition 2.2 The joint normality of X_{t+n} given $\mathcal{F}_t$ is a direct consequence of Lemma 6.8, which expresses X_{t+n} as a linear combination of the $\mathcal{F}_t$-measurable elements of X_t and of jointly Gaussian innovations $Z_{t+1}^{\mathcal{T}}, \dots, Z_{t+n}^{\mathcal{T}}$ that are independent of $\mathcal{F}_t$. The composition of $\mathcal{M}_{t,n}$ is also a direct consequence of Lemma 6.8, and of the null expectation of innovations $Z_{t+1}^{\mathcal{T}}, \dots, Z_{t+n}^{\mathcal{T}}$ given $\mathcal{F}_t$.

Components of $\mathcal{V}_n$ are also obtained through Lemma 6.8 and (2.7)–(2.9):

$$
\begin{aligned}
\mathcal{V}_n^{(1,1)} &= \mathrm{Var}^{\mathcal{T}}(X_{t+n}^{(1)}|\mathcal{F}_t) = \Sigma_{1,1}^2\sum_{l=1}^{n}\mathrm{Var}^{\mathcal{T}}(Z_{t+l,1}^{\mathcal{T}}) = n\Sigma_{1,1}^2, \\
\mathcal{V}_n^{(2,2)} &= \mathrm{Var}^{\mathcal{T}}(X_{t+n}^{(2)}|\mathcal{F}_t) = \Sigma_{2,2}^2\sum_{l=1}^{n}(1-\lambda)^{2(n-l)}\mathrm{Var}^{\mathcal{T}}(Z_{t+l,2}^{\mathcal{T}}) \\
&\quad + \lambda^2\Sigma_{3,3}^2\sum_{l=1}^{n-1}(n-l)^2(1-\lambda)^{2(n-l-1)}\mathrm{Var}^{\mathcal{T}}(Z_{t+l,3}^{\mathcal{T}}) \\
&\quad + 2\Sigma_{2,2}\lambda\Sigma_{3,3}\sum_{l=1}^{n-1}(n-l)(1-\lambda)^{2(n-l)-1}\mathrm{Cov}^{\mathcal{T}}(Z_{t+l,2}^{\mathcal{T}}, Z_{t+l,3}^{\mathcal{T}}) \\
&= \Sigma_{2,2}^2\big(1+\zeta_0((1-\lambda)^2,n)\big) + \lambda^2\Sigma_{3,3}^2(1-\lambda)^{-2}\zeta_2((1-\lambda)^2,n) \\
&\quad + 2\Sigma_{2,2}\lambda\Sigma_{3,3}\rho_{2,3}(1-\lambda)^{-1}\zeta_1\big((1-\lambda)^2,n\big), \\
\mathcal{V}_n^{(3,3)} &= \mathrm{Var}^{\mathcal{T}}(X_{t+n}^{(3)}|\mathcal{F}_t) = \Sigma_{3,3}^2\sum_{l=1}^{n}(1-\lambda)^{2(n-l)}\mathrm{Var}^{\mathcal{T}}(Z_{t+l,3}^{\mathcal{T}}) = \Sigma_{3,3}^2\big(1+\zeta_0((1-\lambda)^2,n)\big)
\end{aligned}
$$

and

$$\begin{aligned}
\mathcal{V}_n^{(1,2)} = \mathcal{V}_n^{(2,1)} &= \mathrm{Cov}^{\mathcal{T}}(X_{t+n}^{(1)}, X_{t+n}^{(2)}|\mathcal{F}_t) = \Sigma_{1,1}\Sigma_{2,2}\sum_{l=1}^{n}(1-\lambda)^{n-l}\mathrm{Cov}^{\mathcal{T}}(Z_{t+l,1}^{\mathcal{T}}, Z_{t+l,2}^{\mathcal{T}}) \\
&\quad + \lambda\Sigma_{1,1}\Sigma_{3,3}\sum_{l=1}^{n-1}(n-l)(1-\lambda)^{n-l-1}\mathrm{Cov}^{\mathcal{T}}(Z_{t+l,1}^{\mathcal{T}}, Z_{t+l,3}^{\mathcal{T}}) \\
&= \Sigma_{1,1}\Sigma_{2,2}\rho_{1,2}\big(1+\zeta_0(1-\lambda,n)\big) + \lambda\Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\frac{\zeta_1(1-\lambda,n)}{1-\lambda}, \\
\mathcal{V}_n^{(1,3)} = \mathcal{V}_n^{(3,1)} &= \mathrm{Cov}^{\mathcal{T}}(X_{t+n}^{(1)}, X_{t+n}^{(3)}|\mathcal{F}_t) = \Sigma_{1,1}\Sigma_{3,3}\sum_{l=1}^{n}(1-\lambda)^{n-l}\mathrm{Cov}^{\mathcal{T}}(Z_{t+l,1}^{\mathcal{T}}, Z_{t+l,3}^{\mathcal{T}}) \\
&= \Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\big(1+\zeta_0(1-\lambda,n)\big), \\
\mathcal{V}_n^{(2,3)} = \mathcal{V}_n^{(3,2)} &= \mathrm{Cov}^{\mathcal{T}}(X_{t+n}^{(2)}, X_{t+n}^{(3)}|\mathcal{F}_t) \\
&= \Sigma_{2,2}\Sigma_{3,3}\sum_{l=1}^{n}(1-\lambda)^{2(n-l)}\mathrm{Cov}^{\mathcal{T}}(Z_{t+l,2}^{\mathcal{T}}, Z_{t+l,3}^{\mathcal{T}}) \\
&\quad + \lambda\Sigma_{3,3}^2\sum_{l=1}^{n-1}(n-l)(1-\lambda)^{2(n-l)-1}\mathrm{Var}^{\mathcal{T}}(Z_{t+l,3}^{\mathcal{T}}) \\
&= \Sigma_{2,2}\Sigma_{3,3}\rho_{2,3}\big(1+\zeta_0((1-\lambda)^2,n)\big) + \lambda\Sigma_{3,3}^2\frac{\zeta_1((1-\lambda)^2,n)}{1-\lambda}.
\end{aligned}$$

□

Proof of Lemma 2.1 From, (2.12), for $i = 1,2,3$, $\eta_t^{\mathcal{T},(i)} = \Delta\Sigma_{i,i}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{i,j}\mathcal{B}_{\mathcal{T}-t-1}^{(j)}$. Therefore,

$$\begin{aligned}
\sum_{l=0}^{\mathcal{T}-t-1}\eta_{t+l}^{\mathcal{T},(1)} &= \Delta\Sigma_{1,1}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{1,j}\sum_{l=0}^{\mathcal{T}-t-1}\mathcal{B}_{\mathcal{T}-t-l-1}^{(j)} \\
&= \Delta\Sigma_{1,1}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{1,j}\sum_{l=0}^{\mathcal{T}-t-1}\mathcal{B}_{l}^{(j)} \\
&= \Delta\Sigma_{1,1}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{1,j}\sum_{l=1}^{\mathcal{T}-t-1}\mathcal{B}_{l}^{(j)} \\
&= \Delta\Sigma_{1,1}\Bigg[\Sigma_{1,1}\frac{(\mathcal{T}-t-1)(\mathcal{T}-t)}{2} + \frac{\Sigma_{2,2}\rho_{1,2}}{\lambda}\big(\mathcal{T}-t-1-\zeta_0(1-\lambda,\mathcal{T}-t)\big) \\
&\quad + \Sigma_{3,3}\rho_{1,3}\Bigg(\frac{\mathcal{T}-t-1-\big(1+\zeta_0(1-\lambda,\mathcal{T}-t-1)\big)}{\lambda} \\
&\quad - \zeta_1(1-\lambda,\mathcal{T}-t-1)\Bigg)\Bigg].
\end{aligned}$$

Moreover, for $i = 2,3$,

$$
\begin{aligned}
&\sum_{l=0}^{\mathcal{T}-t-1} \eta_{t+l}^{\mathcal{T},(i)}(1-\lambda)^{\mathcal{T}-t-1-l} \\
&\quad = \Delta\Sigma_{i,i} \sum_{l=0}^{\mathcal{T}-t-1} (1-\lambda)^{\mathcal{T}-t-1-l}\Big(\Sigma_{1,1}\rho_{i,1}\mathcal{B}_{\mathcal{T}-t-l-1}^{(1)} + \Sigma_{2,2}\rho_{i,2}\mathcal{B}_{\mathcal{T}-t-l-1}^{(2)} \\
&\qquad + \Sigma_{3,3}\rho_{i,3}\mathcal{B}_{\mathcal{T}-t-l-1}^{(3)}\Big) \\
&\quad = \Delta\Sigma_{i,i} \sum_{l=1}^{\mathcal{T}-t-1} (1-\lambda)^{l}\Big(\Sigma_{1,1}\rho_{i,1}\mathcal{B}_{l}^{(1)} + \Sigma_{2,2}\rho_{i,2}\mathcal{B}_{l}^{(2)} + \Sigma_{3,3}\rho_{i,3}\mathcal{B}_{l}^{(3)}\Big) \\
&\quad = \Delta\Sigma_{i,i}\Bigg[\Sigma_{1,1}\rho_{i,1}\zeta_1(1-\lambda,\mathcal{T}-t) + \frac{\Sigma_{2,2}\rho_{i,2}}{\lambda}\big(\zeta_0(1-\lambda,\mathcal{T}-t) - \zeta_0\big((1-\lambda)^2,\mathcal{T}-t\big)\big) \\
&\qquad + \Sigma_{3,3}\rho_{i,3}\Bigg(\frac{\zeta_0(1-\lambda,\mathcal{T}-t) - (1-\lambda)^{-1}\zeta_0\big((1-\lambda)^2,\mathcal{T}-t\big)}{\lambda} \\
&\qquad - (1+\lambda)\zeta_1\big((1-\lambda)^2,\mathcal{T}-t-1\big)\Bigg)\Bigg].
\end{aligned}
$$

Lastly,

$$
\begin{aligned}
&\sum_{k=0}^{\mathcal{T}-t-1} (\mathcal{T}-t-k-1)\eta_{t+k}^{\mathcal{T},(3)}(1-\lambda)^{\mathcal{T}-t-k-2} \\
&\quad = \sum_{k=0}^{\mathcal{T}-t-1} (\mathcal{T}-t-k-1)\Delta\Sigma_{3,3}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{3,j}\mathcal{B}_{\mathcal{T}-t-k-1}^{(j)}(1-\lambda)^{\mathcal{T}-t-k-2} \\
&\quad = \Delta\Sigma_{3,3}\sum_{k=0}^{\mathcal{T}-t-1} k\sum_{j=1}^{3}\Sigma_{j,j}\rho_{3,j}\mathcal{B}_{k}^{(j)}(1-\lambda)^{k-1} \\
&\quad = \Delta\Sigma_{3,3}\sum_{k=1}^{\mathcal{T}-t-1}\sum_{j=1}^{3}\Sigma_{j,j}\rho_{3,j}k\mathcal{B}_{k}^{(j)}(1-\lambda)^{k-1} \\
&\quad = \frac{\Delta\Sigma_{3,3}}{1-\lambda}\Bigg(\Sigma_{1,1}\rho_{3,1}\zeta_2(1-\lambda,\mathcal{T}-t) + \frac{\Sigma_{2,2}\rho_{2,1}}{\lambda}\big[\zeta_1(1-\lambda,\mathcal{T}-t) \\
&\qquad -\zeta_1\big((1-\lambda)^2,\mathcal{T}-t\big)\big] \\
&\qquad + \Sigma_{3,3}\rho_{3,1}\Bigg[\frac{\zeta_1(1-\lambda,\mathcal{T}-t)}{\lambda} - \frac{1}{\lambda}\zeta_1\big((1-\lambda)^2,\mathcal{T}-t\big) \\
&\qquad -(1-\lambda)^{-1}\zeta_2\big((1-\lambda)^2,\mathcal{T}-t\big)\Bigg]\Bigg).
\end{aligned}
$$

Lemma 6.9 *Under the risk-neutral measure $\mathbb{Q}$, conditionally on $\mathcal{F}_t$, factors X_{t+n} follow the multivariate Gaussian distribution with mean vector $\tilde{\mathcal{M}}_{t,n} = \left[\tilde{\mathcal{M}}_{t,n}^{(i)}\right]_{i=1}^{3}$ and covariance matrix $\mathcal{V}_n = \left[\mathcal{V}_n^{(i,j)}\right]_{i,j=1}^{3}$, where*

$$
\begin{aligned}
\tilde{\mathcal{M}}^{(1)}_{t,n} &= X^{(1)}_t, \\
\tilde{\mathcal{M}}^{(2)}_{t,n} &= X^{(2)}_t(1-\lambda)^n + (\theta^{\mathbb{Q}}_2 - \theta^{\mathbb{Q}}_3)(1-(1-\lambda)^n) \\
&\quad + \lambda\left(nX^{(3)}_t(1-\lambda)^{n-1} + \theta^{\mathbb{Q}}_3\left(\frac{\zeta_0(1-\lambda, n+1)}{1-\lambda} - n(1-\lambda)^{n-1}\right)\right), \\
\tilde{\mathcal{M}}^{(3)}_{t,n} &= X^{(3)}_t(1-\lambda)^n + \theta^{\mathbb{Q}}_3(1-(1-\lambda)^n).
\end{aligned}
$$

Proof of Lemma 6.9 The proof is analogous to that of Proposition 2.2, replacing forward measure superscripts $\mathcal{T}$ by $\mathbb{Q}$ and setting the process η to zero. □

Proof of Proposition 2.3 From (2.14), For $t = 0, \dots, T$ and $n = 0, \dots, T-t$,

$$
\begin{aligned}
X^{(i)}_{t+n} &= X^{(i)}_t(1-\kappa^{\mathbb{P}}_{i,i})^n + \kappa^{\mathbb{P}}_{i,i}\theta^{\mathbb{P}}_i \sum_{l=1}^{n}(1-\kappa^{\mathbb{P}}_{i,i})^{(n-l)} + \Sigma_{i,i}\sum_{l=1}^{n} Z^{\mathbb{P}}_{t+l,i}(1-\kappa^{\mathbb{P}}_{i,i})^{(n-l)} \\
&\quad + \sum_{l=1}^{n}\sum_{j\neq i}^{3}\kappa^{\mathbb{P}}_{i,j}(\theta^{\mathbb{P}}_j - X^{(j)}_{t+l-1})(1-\kappa^{\mathbb{P}}_{i,i})^{(n-l)}.
\end{aligned} \tag{6.9}
$$

Indeed, steps for the proof of (6.9) are identical to these for the proof of Lemma 6.7, setting the process η to zero. This leads to

$$
X^{(i)}_{t+n} = X^{(i)}_t(1-\kappa^{\mathbb{P}}_{i,i})^n + \theta^{\mathbb{P}}_i\left(1-(1-\kappa^{\mathbb{P}}_{i,i})^n\right) + \Sigma_{i,i}\sum_{l=1}^{n} Z^{\mathbb{P}}_{t+l,i}(1-\kappa^{\mathbb{P}}_{i,i})^{(n-l)}, \quad i = 1, 3. \tag{6.10}
$$

and

$$
\begin{aligned}
X^{(2)}_{t+n} =& X^{(2)}_t(1-\kappa^{\mathbb{P}}_{2,2})^n + \kappa^{\mathbb{P}}_{2,2}\theta^{\mathbb{P}}_2 \sum_{l=1}^{n}(1-\kappa^{\mathbb{P}}_{2,2})^{(n-l)} + \Sigma_{2,2}\sum_{l=1}^{n} Z^{\mathbb{P}}_{t+l,2}(1-\kappa^{\mathbb{P}}_{2,2})^{(n-l)} \\
&- \lambda\sum_{l=1}^{n}(\theta^{\mathbb{P}}_3 - X^{(3)}_{t+l-1})(1-\kappa^{\mathbb{P}}_{2,2})^{(n-l)}.
\end{aligned} \tag{6.11}
$$

Furthermore, assuming $\kappa^{\mathbb{P}}_{2,2} \neq \kappa^{\mathbb{P}}_{3,3}$ and denoting $\omega = \frac{1-\kappa^{\mathbb{P}}_{3,3}}{1-\kappa^{\mathbb{P}}_{2,2}}$,

$$
\begin{aligned}
-\sum_{l=1}^{n}(1-\kappa_{2,2}^{\mathbb{P}})^{n-l}\left(\theta_3^{\mathbb{P}}-X_{t+l-1}^{(3)}\right) &= -\sum_{l=0}^{n-1}(1-\kappa_{2,2}^{\mathbb{P}})^{n-l-1}\left(\theta_3^{\mathbb{P}}-X_{t+l}^{(3)}\right)\\
&= \sum_{l=0}^{n-1}(1-\kappa_{2,2}^{\mathbb{P}})^{n-l-1}\Bigg[X_t^{(3)}(1-\kappa_{3,3}^{\mathbb{P}})^l-\theta_3^{\mathbb{P}}(1-\kappa_{3,3}^{\mathbb{P}})^l\\
&\quad+\Sigma_{3,3}\sum_{k=1}^{l}Z_{t+k,3}^{\mathbb{P}}(1-\kappa_{3,3}^{\mathbb{P}})^{(l-k)}\Bigg]\\
&= (X_t^{(3)}-\theta_3^{\mathbb{P}})\frac{1-\omega^n}{1-\omega}(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}\\
&\quad+\Sigma_{3,3}\sum_{l=0}^{n-1}\sum_{k=1}^{n-1}\mathbb{1}_{\{k\le l\}}Z_{t+k,3}^{\mathbb{P}}\omega^l\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^k}\\
&= (X_t^{(3)}-\theta_3^{\mathbb{P}})\frac{1-\omega^n}{1-\omega}(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}\\
&\quad+\Sigma_{3,3}\sum_{k=1}^{n-1}Z_{t+k,3}^{\mathbb{P}}\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^k}\sum_{l=k}^{n-1}\omega^l\\
&= (X_t^{(3)}-\theta_3^{\mathbb{P}})\frac{1-\omega^n}{1-\omega}(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}\\
&\quad+\Sigma_{3,3}\sum_{k=1}^{n-1}Z_{t+k,3}^{\mathbb{P}}\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^k}\frac{\omega^k-\omega^n}{1-\omega}.
\end{aligned}
\tag{6.12}
$$

Substituting (6.12) into (6.11) leads to

$$
\begin{aligned}
X_{t+n}^{(2)} &= X_t^{(2)}(1-\kappa_{2,2}^{\mathbb{P}})^n+\theta_2^{\mathbb{P}}\left(1-(1-\kappa_{2,2}^{\mathbb{P}})^n\right)+\Sigma_{2,2}\sum_{l=1}^{n}Z_{t+l,2}^{\mathbb{P}}(1-\kappa_{2,2}^{\mathbb{P}})^{(n-l)}\\
&\quad+\lambda\left((X_t^{(3)}-\theta_3^{\mathbb{P}})\frac{1-\omega^n}{1-\omega}(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}+\Sigma_{3,3}\sum_{l=1}^{n-1}Z_{t+l,3}^{\mathbb{P}}\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^l}\frac{\omega^l-\omega^n}{1-\omega}\right).
\end{aligned}
\tag{6.13}
$$

Combining (6.10) and (6.13) directly yields expressions for $\mathcal{M}_{t,n}^{\mathbb{P},(i)}$. Additionally,

$$
\begin{aligned}
\mathcal{V}_n^{\mathbb{P},(1,1)} &= \mathrm{Var}^{\mathbb{P}}(X_{t+n}^{(1)}|\mathcal{F}_t) = \Sigma_{1,1}^2 \sum_{l=1}^{n} (1-\kappa_{1,1}^{\mathbb{P}})^{2(n-l)} \mathrm{Var}^{\mathbb{P}}(Z_{t+l,1}^{\mathbb{P}}) \\
&= \begin{cases} \Sigma_{1,1}^2 \Big(1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})^2, n)\Big) & \text{if } \kappa_{1,1}^{\mathbb{P}} \in (0,1), \\ n\Sigma_{1,1}^2 \quad \text{if } \kappa_{1,1}^{\mathbb{P}} = 0, \end{cases} \\
\mathcal{V}_n^{\mathbb{P},(2,2)} &= \mathrm{Var}^{\mathbb{P}}(X_{t+n}^{(2)}|\mathcal{F}_t) = \Sigma_{2,2}^2 \sum_{l=1}^{n} (1-\kappa_{2,2}^{\mathbb{P}})^{2(n-l)} \mathrm{Var}^{\mathbb{P}}(Z_{t+l,2}^{\mathbb{P}}) \\
&\quad + \lambda^2 \Sigma_{3,3}^2 \sum_{l=1}^{n-1} \left(\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^l} \frac{\omega^l - \omega^n}{1-\omega} \right)^2 \mathrm{Var}^{\mathbb{P}}(Z_{t+l,3}^{\mathbb{P}}) \\
&\quad + 2\Sigma_{2,2}\lambda\Sigma_{3,3} \sum_{l=1}^{n-1} (1-\kappa_{2,2}^{\mathbb{P}})^{(2n-l-1)} \left(\frac{1}{1-\omega} \left(\frac{\omega}{1-\kappa_{3,3}^{\mathbb{P}}} \right)^l \right. \\
&\quad \left. - \frac{\omega^n}{1-\omega} \left(\frac{1}{1-\kappa_{3,3}^{\mathbb{P}}} \right)^l \right) \mathrm{Cov}^{\mathbb{P}}(Z_{t+l,2}^{\mathbb{P}}, Z_{t+l,3}^{\mathbb{P}}) \\
&= \Sigma_{2,2}^2 \Big(1+\zeta_0((1-\kappa_{2,2}^{\mathbb{P}})^2, n)\Big) + \lambda^2 \Sigma_{3,3}^2 \frac{(1-\kappa_{2,2}^{\mathbb{P}})^{2n-2}}{(1-\omega)^2} \sum_{l=1}^{n-1} \left(\frac{\omega^{2l} - 2\omega^{l+n} + \omega^{2n}}{(1-\kappa_{3,3}^{\mathbb{P}})^{2l}} \right) \\
&\quad + 2\frac{\rho_{2,3}\lambda\Sigma_{2,2}\Sigma_{3,3}}{1-\omega} (1-\kappa_{2,2}^{\mathbb{P}})^{(2n-1)} \sum_{l=1}^{n-1} \left(\left(\frac{\omega}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})} \right)^l \right. \\
&\quad \left. - \omega^n \left(\frac{1}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})} \right)^l \right) \\
&= \Sigma_{2,2}^2 \Big(1+\zeta_0((1-\kappa_{2,2}^{\mathbb{P}})^2, n)\Big) \\
&\quad + \lambda^2 \Sigma_{3,3}^2 \frac{(1-\kappa_{2,2}^{\mathbb{P}})^{2n-2}}{(1-\omega)^2} \Big[\zeta_0\Big((1-\kappa_{2,2}^{\mathbb{P}})^{-2}, n\Big) - 2\omega^n \zeta_0\Big(\omega(1-\kappa_{3,3}^{\mathbb{P}})^{-2}, n\Big) \\
&\quad + \omega^{2n} \zeta_0\Big((1-\kappa_{3,3}^{\mathbb{P}})^{-2}, n\Big) \Big] \\
&\quad + 2\frac{\rho_{2,3}\lambda\Sigma_{2,2}\Sigma_{3,3}}{1-\omega} (1-\kappa_{2,2}^{\mathbb{P}})^{(2n-1)} \left[\zeta_0\left(\frac{\omega}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})}, n \right) \right. \\
&\quad \left. - \omega^n \zeta_0\left(\frac{1}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})}, n \right) \right], \\
\mathcal{V}_n^{\mathbb{P},(3,3)} &= \mathrm{Var}^{\mathbb{P}}(X_{t+n}^{(3)}|\mathcal{F}_t) = \Sigma_{3,3}^2 \sum_{l=1}^{n} (1-\kappa_{3,3}^{\mathbb{P}})^{2(n-l)} \mathrm{Var}^{\mathbb{P}}(Z_{t+l,3}^{\mathbb{P}}) \\
&= \Sigma_{3,3}^2 \Big(1+\zeta_0((1-\kappa_{3,3}^{\mathbb{P}})^2, n)\Big)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{V}_n^{\mathbb{P},(1,2)} &= \mathcal{V}_n^{\mathbb{P},(2,1)} = \mathrm{Cov}^{\mathbb{P}}(X_{t+n}^{(1)}, X_{t+n}^{(2)}|\mathcal{F}_t) \\
&= \Sigma_{1,1}\Sigma_{2,2}\sum_{l=1}^{n}(1-\kappa_{1,1}^{\mathbb{P}})^{n-l}(1-\kappa_{2,2}^{\mathbb{P}})^{n-l}\mathrm{Cov}^{\mathbb{P}}(Z_{t+l,1}^{\mathbb{P}}, Z_{t+l,2}^{\mathbb{P}}) \\
&\quad + \lambda\Sigma_{1,1}\Sigma_{3,3}\sum_{l=1}^{n-1}(1-\kappa_{1,1}^{\mathbb{P}})^{n-l}\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^{l}}\frac{\omega^l-\omega^n}{1-\omega}\mathrm{Cov}^{\mathbb{P}}(Z_{t+l,1}^{\mathbb{P}}, Z_{t+l,3}^{\mathbb{P}}) \\
&= \Sigma_{1,1}\Sigma_{2,2}\rho_{1,2}\left[1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{2,2}^{\mathbb{P}}), n)\right] \\
&\quad + \lambda\Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\frac{(1-\kappa_{1,1}^{\mathbb{P}})^{n}(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{1-\omega}\times \\
&\quad \left(\zeta_0\left(\frac{\omega}{(1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})}, n\right) - \omega^n\zeta_0\left(\frac{1}{(1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})}, n\right)\right), \\
\mathcal{V}_n^{\mathbb{P},(1,3)} &= \mathcal{V}_n^{\mathbb{P},(3,1)} = \mathrm{Cov}^{\mathbb{P}}(X_{t+n}^{(1)}, X_{t+n}^{(3)}|\mathcal{F}_t) \\
&= \Sigma_{1,1}\Sigma_{3,3}\sum_{l=1}^{n}(1-\kappa_{1,1}^{\mathbb{P}})^{(n-l)}(1-\kappa_{3,3}^{\mathbb{P}})^{(n-l)}\mathrm{Cov}^{\mathbb{P}}(Z_{t+l,1}^{\mathbb{P}}, Z_{t+l,3}^{\mathbb{P}}) \\
&= \Sigma_{1,1}\Sigma_{3,3}\rho_{1,3}\left[1+\zeta_0((1-\kappa_{1,1}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}}), n)\right], \\
\mathcal{V}_n^{\mathbb{P},(2,3)} &= \mathcal{V}_n^{\mathbb{P},(3,2)} = \mathrm{Cov}^{\mathbb{P}}(X_{t+n}^{(2)}, X_{t+n}^{(3)}|\mathcal{F}_t) \\
&= \Sigma_{2,2}\Sigma_{3,3}\sum_{l=1}^{n}(1-\kappa_{2,2}^{\mathbb{P}})^{n-l}(1-\kappa_{3,3}^{\mathbb{P}})^{n-l}\mathrm{Cov}^{\mathbb{P}}(Z_{t+l,2}^{\mathbb{P}}, Z_{t+l,3}^{\mathbb{P}}) \\
&\quad + \lambda\Sigma_{3,3}^2\sum_{l=1}^{n-1}\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}}{(1-\kappa_{3,3}^{\mathbb{P}})^{l}}\frac{\omega^l-\omega^n}{1-\omega}(1-\kappa_{3,3}^{\mathbb{P}})^{n-l}\mathrm{Var}^{\mathbb{P}}(Z_{t+l,3}^{\mathbb{P}}) \\
&= \Sigma_{2,2}\Sigma_{3,3}\rho_{2,3}\left[1+\zeta_0((1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}}), n)\right] \\
&\quad + \lambda\Sigma_{3,3}^2\frac{(1-\kappa_{2,2}^{\mathbb{P}})^{n-1}(1-\kappa_{3,3}^{\mathbb{P}})^{n}}{1-\omega}\left(\zeta_0\left(\frac{1}{(1-\kappa_{2,2}^{\mathbb{P}})(1-\kappa_{3,3}^{\mathbb{P}})}, n\right)\right. \\
&\quad \left. - \omega^n\zeta_0\left((1-\kappa_{3,3}^{\mathbb{P}})^{-2}, n\right)\right).
\end{aligned}
$$

□

Proof of Theorem 4.1 Using Lemma (6.9), the futures prices can be obtained from the moment generating function of the normal distribution:

$$
\begin{aligned}
F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} =& \mathbb{E}^{\mathbb{Q}}\left[P(\mathcal{T}_2,\mathcal{T}_3)\middle|\mathcal{F}_{\mathcal{T}_1}\right],\\
=& \mathbb{E}^{\mathbb{Q}}\left[A_{\tau_3}\exp\left[-\Delta\mathcal{B}_{\tau_3}^{\top}X_{\mathcal{T}_2}\right]\middle|\mathcal{F}_{\mathcal{T}_1}\right]\\
=& A_{\tau_3}\exp\left[-\Delta\sum_{i=1}^{3}\mathcal{B}_{\tau_3}^{(i)}\tilde{\mathcal{M}}_{\mathcal{T}_1,\tau_2}^{(i)}+\frac{\Delta^2}{2}\mathcal{B}_{\tau_3}^{\top}\mathcal{V}_{\tau_2}\mathcal{B}_{\tau_3}\right].
\end{aligned}
$$

This implies

$$
F_{\mathcal{T}_1,\mathcal{T}_2,\mathcal{T}_3} = \tilde{A}_{\tau_2,\tau_3}\exp\left[-\Delta\sum_{i=1}^{3}\tilde{\mathcal{B}}_{\tau_3}^{(i)}X_{\mathcal{T}_1}^{(i)}\right]
$$

where

$$
\begin{aligned}
\tilde{A}_{\tau_2,\tau_3} = A_{\tau_3}\exp\Bigg[&\frac{\Delta^2}{2}\mathcal{B}_{\tau_3}^{\top}\mathcal{V}_{\tau_2}\mathcal{B}_{\tau_3}-\Delta\mathcal{B}_{\tau_3}^{(2)}(\theta_2^{\mathbb{Q}}-\theta_3^{\mathbb{Q}})(1-(1-\lambda)^{\tau_2})\\
&-\Delta\mathcal{B}_{\tau_3}^{(2)}\lambda\theta_3^{\mathbb{Q}}\left(\frac{\zeta_0(1-\lambda,\tau_2+1)}{1-\lambda}-\tau_2(1-\lambda)^{\tau_2-1}\right)\Bigg)\\
&-\Delta\mathcal{B}_{\tau_3}^{(3)}\theta_3^{\mathbb{Q}}(1-(1-\lambda)^{\tau_2})\Bigg],
\end{aligned}
$$

$$
\tilde{\mathcal{B}}_n^{(1)} = \mathcal{B}_n^{(1)},\quad \tilde{\mathcal{B}}_n^{(2)} = \mathcal{B}_n^{(2)}(1-\lambda)^n,\quad \tilde{\mathcal{B}}_n^{(3)} = \mathcal{B}_n^{(3)}(1-\lambda)^n+\mathcal{B}_n^{(2)}\lambda n(1-\lambda)^{n-1}.
$$

□

Acknowledgements We thank the Natural Sciences and Engineering Research Council of Canada (Godin: RGPIN-2017-06837, Gaillardetz: RGPIN-2020-06821) for their financial support.

Declarations

Conflict of interest Authors have no conflict of interests to report in relation to this work.

References

Bakshi, G., Crosby, J., Gao, X., & Hansen, J. W. (2023). Treasury option returns and models with unspanned risks. *Journal of Financial Economics*(**forthcoming**).

Bakshi, G., Crosby, J., & Gao, X. (2022). Dark matter in (volatility and) equity option risk premiums. *Operations Research, 70*(6), 3108–3124.

Björk, T. (2009). *Arbitrage theory in continuous time*. Oxford University Press.

Black, F., Derman, E., & Toy, W. (1990). A one-factor model of interest rates and its application to treasury bond options. *Financial Analysts Journal, 46*(1), 33–39.

Black, F., & Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal, 47*(4), 52–59.

Brigo, D., & Mercurio, F. (2007). *Interest rate models-theory and practice: With smile, inflation and credit*. Springer.

Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics, 164*(1), 4–20.
Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman filtering in affine term structure models. *Management Science, 60*(9), 2248–2268.
Collin-Dufresne, P., & Goldstein, R. S. (2002). Pricing swaptions within an affine framework. *The Journal of Derivatives, 10*(1), 9–26.
Coval, J. D., & Shumway, T. (2001). Expected option returns. *The Journal of Finance, 56*(3), 983–1009.
Creal, D. (2012). A survey of sequential Monte Carlo methods for economics and finance. *Econometric Reviews, 31*(3), 245–296.
Del Moral, P. (1997). Nonlinear filtering: Interacting particle resolution. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 325*(6), 653–658.
Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, 6*(4), 379–406.
Eghbalzadeh, R., Godin, F., & Gaillardetz, P. (2022). The discrete-time arbitrage-free Nelson–Siegel model: a closed-form solution and applications to mixed funds representation.
Geman, H. (1989). The importance of the forward neutral probability in a stochastic approach of interest rates. Technical report, Working paper, ESSEC.
Geman, H., El Karoui, N., & Rochet, J.-C. (1995). Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability, 32*(2), 443–458.
Godin, F. (2019). A closed-form solution for the global quadratic hedging of options under geometric Gaussian random walks. *The Journal of Derivatives, 26*(3), 97–107.
Jamshidian, F. (1996). Bond, futures and option evaluation in the quadratic interest rate model. *Applied Mathematical Finance, 3*(2), 93–115.
Julier, S. J. & Uhlmann, J. K. (1997). New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI* (Vol. 3068, pp. 182–193). SPIE.
Munk, C. (1999). Stochastic duration and fast coupon bond option pricing in multi-factor models. *Review of Derivatives Research, 3*(2), 157–181.
Remillard, B. (2013). *Statistical methods for financial engineering*. CRC Press.
Schrager, D. F., & Pelsser, A. A. (2006). Pricing swaptions and coupon bond options in affine term structure models. *Mathematical Finance, 16*(4), 673–694.
Singleton, K. J., & Umantsev, L. (2002). Pricing coupon-bond options and swaptions in affine term structure models. *Mathematical Finance, 12*(4), 427–446.