Annals of Actuarial Science (2024), 1–32
doi:10.1017/S1748499524000010

ORIGINAL RESEARCH PAPER

The discrete-time arbitrage-free Nelson-Siegel model: a closed-form solution and applications to mixed funds representation

Ramin Eghbalzadeh[1], Frédéric Godin[1,2] and Patrice Gaillardetz[1,2]

[1]Department of Mathematics and Statistics, Concordia University, Montréal, Canada; and [2]Quantact Laboratory, Centre de Recherches Mathématiques, Montréal, Canada
Corresponding author: Frédéric Godin; Email: frederic.godin@concordia.ca

(Received 23 June 2023; revised 30 November 2023; accepted 22 December 2023)

Abstract

A closed-form solution for zero-coupon bonds is obtained for a version of the discrete-time arbitrage-free Nelson-Siegel model. An estimation procedure relying on a Kalman filter is provided. The model is shown to produce adequate fit when applied to historical Canadian spot rate data and to improve distributional predictive performance over benchmarks. An adaptation of the mixed fund return model from Augustyniak *et al.* ((2021). *ASTIN Bulletin: The Journal of the IAA*, *51*(1), 131–159.) is also provided to include the discrete-time arbitrage-free Nelson-Siegel model as one of its building blocks.

Keywords: Interest rate models; term structure; arbitrage-free Nelson-Siegel; Gaussian factor models; Kalman filter; mixed funds

JEL classification: E43; G12

1. Introduction

Dynamics of interest rates have implications in different areas such as asset and derivatives pricing, portfolio management, and risk measurement. In the context of actuarial science, interest rate models are central to the evaluation of long-term financial contracts such as standard insurance and annuities, as well as equity-linked products.

The seminal work of Nelson & Siegel (1987) depicts the term structure as a linear combination of interpretable factors representing the level, the slope, and a hump in the yield curve. Although this model was initially developed in a static setting, Diebold & Li (2006) extend the Nelson & Siegel (1987) model into a dynamic framework. However, a drawback associated with the dynamic version of the model is its incompatibility with no-arbitrage theory, an issue that is discussed in Filipović (1999). The work of Christensen *et al.* (2011) reconciles both the no-arbitrage strand of literature and the Nelson & Siegel (1987) model by developing the so-called arbitrage-free Nelson-Siegel (AFNS) model. Such model is an affine term structure (ATS) model (see Duffie & Kan, 1996) where the specification of the diffusion followed by term structure factors is carefully chosen such that factor loadings exactly coincide with these of the Nelson-Siegel model. The discrepancy between spot rates of the Nelson-Siegel model and of the AFNS model only stems from a corrective term whose inclusion is necessary to preclude theoretical arbitrage opportunities.

The AFNS model is expressed in a continuous-time settings. It is however of interest to develop a discrete-time counterpart to the model. Indeed, many of existing interest rate models already have such discrete-time versions, as described for instance in Wüthrich & Merz (2013). Discrete-time models often possess the advantage of being more simple to manipulate, which is favorable in applications. For instance, when Monte-Carlo simulations are used to generate discount factors, discretization must often be applied. Whenever such approximation is used, numerical testing is required to ensure that discretization errors exhibit sufficiently low bias and variance. Discrete-time models allow circumventing all such analyses by providing exact simulations. Furthermore, multiple time series and econometric models involving regressions (e.g. these including lagged components) are naturally formulated in discrete-time, and compatibility with such frameworks is often a key consideration when developing interest rate models. Finally, dynamic optimization schemes such as reinforcement learning are more extensively developed in a discrete-time context and exhibit much more tractability in that case. The integration of stochastic interest rates to applications of such methods to finance, such as the deep hedging scheme of Buehler *et al.* (2019), will be simplified if the model is already specified in discrete-time.

A discrete-time version of the AFNS model has been developed in Hong *et al.* (2019). This paper further expands on the work of Hong *et al.* (2019) and revisits the discrete-time arbitrage-free Nelson-Siegel (DTAFNS) model, providing the following three main contributions. The first is to obtain a closed-form expression for risk-free spot rates. The closed-form solution is obtained after slightly modifying the interest rate dynamics from Hong *et al.* (2019) so as to obtain additional tractability (i.e. simpler formulas), while still retaining asymptotic equivalence with the AFNS model. As highlighted in Section 2.5 of Hong *et al.* (2019), ATS models often provide enough tractability to study the decomposition of yields into expected rates and risk premia analytically. The closed-form expression we obtain can therefore further facilitate such analysis. The second contribution consists in presenting evidence of superior distributional out-of-sample predictive ability of the DTAFNS model over benchmarks when applied on historical Canadian term structure data. Finally, as an actuarial application, our third contribution is to develop a modified version of the mixed fund return model of Augustyniak *et al.* (2021) by substituting their three-factor Gaussian model with the DTAFNS model. The latter is justified by the higher performance of the DTAFNS model outlined in this paper. Such model is useful when pricing life insurance contracts embedding guarantees related to financial market performance such as variable annuities, universal life or participating contracts.

The remainder of this paper is organized as follows. Section 2 specifies the proposed interest rate model and provides closed-form expressions for risk-free zero-coupon bond prices and spot rates. In Section 3, the model calibration is illustrated and numerical experiments assessing predictive performance are provided. Section 4 adapts the mixed fund model of Augustyniak *et al.* (2021) to include the DTAFNS model as a building block for the term structure model. Section 5 concludes.

2. Interest rate term structure model

This section presents the mathematical construction of the discrete-time version of the arbitrage-free Nelson-Siegel term structure model, referred to subsequently as the DTAFNS model. First, the specification of the traditional Nelson-Siegel model and its dynamic continuous-time arbitrage-free extension are recalled. Then, the dynamics of the short rate in the DTAFNS model is provided and closed-form formulas for spot rates are derived. As explained in more details subsequently, the DTAFNS model presented in this section has a slighlty different specification from that of Hong *et al.* (2019).

2.1. The Nelson-Siegel term structure representation

Nelson & Siegel (1987) propose a parametric representation of the yield curve of the form

$$y(t,T) = X^{(1)} + X^{(2)}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + X^{(3)}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right), \tag{2.1}$$

where $y(t, T)$ is the time-t spot rate with tenor $\tau = T - t$, and $X^{(1)}, X^{(2)}, X^{(3)}$ and $\lambda > 0$ are model parameters. The model leads to a convenient interpretation of the parameters, with $X^{(1)}$, $X^{(2)}$, and $X^{(3)}$, respectively, characterizing the level, slope, and curvature of the yield curve, while λ is a dilation parameter. Although this model has first been considered in a static setting, Diebold & Li (2006) embed it in a dynamic setting where time-varying factors $X_t^{(1)}, X_t^{(2)}, X_t^{(3)}$ are considered instead of $X^{(1)}, X^{(2)}, X^{(3)}$. Several auto-regressive models are used for the dynamics of such parameters, thereby making it possible to forecast the term structure.

A pitfall of the dynamic Nelson-Siegel model of Diebold & Li (2006) is its incompatibility with no-arbitrage theory as discussed Filipović (1999). Christensen *et al.* (2011) therefore adapt the Nelson-Siegel framework and develop the AFNS model, an arbitrage-free term structure model whose spot rate formulas match exactly these of the Nelson-Siegel model up to a correction term. It is obtained by specifying the short rate risk-neutral dynamics as the following affine Gaussian Itô diffusion:

$$r_t \equiv X_t^{(1)} + X_t^{(2)}, \tag{2.2}$$

$$\begin{bmatrix} dX_t^{(1)} \\ dX_t^{(2)} \\ dX_t^{(3)} \end{bmatrix} \equiv \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}}_{\kappa^{\mathbb{Q}}} \underbrace{\begin{bmatrix} \theta_1^{\mathbb{Q}} - X_t^{(1)} \\ \theta_2^{\mathbb{Q}} - X_t^{(2)} \\ \theta_3^{\mathbb{Q}} - X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{Q}} - \mathbf{X}_t} dt + \Sigma \begin{pmatrix} dW_{t,1}^{\mathbb{Q}} \\ dW_{t,2}^{\mathbb{Q}} \\ dW_{t,3}^{\mathbb{Q}} \end{pmatrix} \tag{2.3}$$

where $\{W_{t,1}^{\mathbb{Q}}\}_{t\geq 0}$, $\{W_{t,2}^{\mathbb{Q}}\}_{t\geq 0}$ and $\{W_{t,3}^{\mathbb{Q}}\}_{t\geq 0}$ are independent Brownian motions under the risk-neutral probability measure $\mathbb{Q}$, Σ is a 3×3 positive semi-definite matrix and $\theta_1^{\mathbb{Q}}, \theta_2^{\mathbb{Q}}, \theta_3^{\mathbb{Q}}, \lambda$ are real numbers. As shown in Christensen *et al.* (2011), this leads to the following formulas for spot rates:

$$y(t, T) = -\frac{1}{\tau} \sum_{j=1}^{3} \tilde{B}^{(j)}(t, T) X_t^{(j)} - \frac{\tilde{C}(t, T)}{\tau}, \tag{2.4}$$

$$\tilde{C}(t, T) = \sum_{j=2}^{3} \left(\kappa^{\mathbb{Q}} \theta^{\mathbb{Q}}\right)_j \int_t^T \tilde{B}^{(j)}(s, T) ds + \frac{1}{2} \sum_{j=1}^{3} \int_t^T \left(\Sigma^\top \tilde{B}(s, T) \tilde{B}(s, T)^\top \Sigma\right)_{j,j} ds \tag{2.5}$$

with

$$\begin{aligned} \tilde{B}(s, T) &\equiv \left[\tilde{B}^{(1)}(s, T),\ \tilde{B}^{(2)}(s, T),\ \tilde{B}^{(3)}(s, T)\right]^\top \\ &\equiv \left[-\tau,\ -\left(\frac{1 - e^{-\lambda\tau}}{\lambda}\right),\ \left(\tau e^{-\lambda\tau} - \frac{1 - e^{-\lambda\tau}}{\lambda}\right)\right]^\top \end{aligned} \tag{2.6}$$

and $()_j$ or $()_{j,j}$ denoting respectively the j^{th} element of a vector or the element at row j and column j of a matrix, respectively.

Striking similarities can be observed when comparing (2.1) with (2.4) and (2.6). Indeed, identifying $X_t^{(1)}, X_t^{(2)}, X_t^{(3)}$ with $X^{(1)}, X^{(2)}, X^{(3)}$, discrepancies between the spot rates from both models are only caused by the corrective term $-\frac{\tilde{C}(t,T)}{\tau}$ appearing in (2.4). This leads Christensen *et al.* (2011) to describe the arbitrage-free Nelson-Siegel model (2.2)–(2.3) as the "*closest match to the Nelson-Siegel yield function*" within some class of exponential affine term structure models.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

2.2. Discrete-time arbitrage-free Nelson-Siegel model

The purpose of this section is to define a discrete-time model analogous to model (2.2)–(2.3). In what follows, an arbitrage-free market with discrete time steps $t = 0, 1, ..., T$, each separated by a time elapse Δ, is considered. Market dynamics are defined on probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$ where $\mathbb{P}$ is the real-world probability measure and $\mathcal{F} := \{\mathcal{F}_t\}_{t=0}^{T}$ is a filtration characterizing the information flow in the market.

In the discrete-time context, the time-t short risk-free rate r_t is the $\mathcal{F}_t$-measurable rate effective for the period $[t, t+1)$. The short rate dynamics considered are

$$r_t = X_t^{(1)} + X_t^{(2)}, \tag{2.7}$$

$$X_{t+1} = X_t + \kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - X_t\right) + \Sigma Z_{t+1}^{\mathbb{P}}, \tag{2.8}$$

where the stochastic factor vectors are $X_t = \left[X_t^{(1)}, X_t^{(2)}, X_t^{(3)}\right]^\top$, $t = 0, \ldots, T$, $\kappa^{\mathbb{P}}$ and Σ are constant parameter matrices of dimension 3×3, and $\theta^{\mathbb{P}}$ is a constant column vector of dimension 3. The process $Z^{\mathbb{P}} = \left\{Z_t^{\mathbb{P}}\right\}_{t=1}^{T}$ with $Z_t^{\mathbb{P}} = \left[Z_{t,1}^{\mathbb{P}}, Z_{t,2}^{\mathbb{P}}, Z_{t,3}^{\mathbb{P}}\right]^\top$ is a multivariate standard Gaussian white noise under $\mathbb{P}$ with contemporaneous correlation parameters $\rho_{i,j}^{(Z)} \equiv \mathrm{corr}\left(Z_{t,i}^{\mathbb{P}}, Z_{t,j}^{\mathbb{P}}\right)$, $t = 1, \ldots, T$ and $i, j = 1, 2, 3$. The initial value of the factor process is fixed at $X_0 = x_0 \in \mathbb{R}^3$. Σ is assumed to be a diagonal matrix with strictly positive elements. Equations (2.7)–(2.8) are meant to mimic (2.2)–(2.3), except that the model is specified under the physical measure $\mathbb{P}$ instead of under the risk-neutral measure $\mathbb{Q}$. A more general class of Gaussian discrete-time affine structure models is studied in Wüthrich & Merz (2013), but without specializing to the DTAFNS model which is the object of this paper.

To obtain the specification of spot rates under the DTAFNS model, risk-neutral dynamics need to be specified. Indeed, spot rates are given by

$$y(t, T) = -\frac{\log P(t, T)}{(T-t)\Delta}$$

where $P(t,T)$ is the time-t of a risk-free zero coupon of unit face value with maturity on time point T. In discrete-time models of the short rate, the relationship between $P(t,T)$ and the risk-free rate is

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\Delta\sum_{j=t}^{T-1} r_j\right)\middle|\mathcal{F}_t\right].$$

To obtain risk-neutral dynamics, following the lines of Augustyniak *et al.* (2021) for instance, a discrete-time version of the Girsanov theorem can be invoked to apply a translation to the drift of short rate factors X_t without altering their volatilities, while keeping a multivariate Gaussian distribution for innovations.[1] It follows from this theorem that for a given constant market price of risk matrix $\gamma \in \mathbb{R}^{3\times 3}$, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the process $Z^{\mathbb{Q}} = \left\{Z_t^{\mathbb{Q}}\right\}_{t=1}^{T}$ defined through $Z_{t+1}^{\mathbb{Q}} = Z_{t+1}^{\mathbb{P}} - \gamma X_t$ is also an $\mathcal{F}$-adapted standard Gaussian white noise under $\mathbb{Q}$, still with contemporaneous correlation matrix $\rho^{(Z)} = \left[\rho_{i,j}^{(Z)}\right]_{i,j=1}^{3}$. Substituting $Z_{t+1}^{\mathbb{P}} = Z_{t+1}^{\mathbb{Q}} + \gamma X_t$ into (2.8) leads to

$$X_{t+1} = X_t + \kappa^{\mathbb{P}}\theta^{\mathbb{P}} - \left(\kappa^{\mathbb{P}} - \Sigma\gamma\right)X_t + \Sigma Z_{t+1}^{\mathbb{Q}}.$$

[1]Because of the discrete-time nature of the model, alternative changes of measures could have been contemplated. Nevertheless, the Girsanov-type family of changes of measure is considered to keep the model as similar as possible to the continuous-time version of the model.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Assuming there is a solution $(\kappa^{\mathbb{Q}}, \theta^{\mathbb{Q}})$ with $\kappa^{\mathbb{Q}} \in \mathbb{R}^{3\times 3}$ and $\theta^{\mathbb{Q}} \in \mathbb{R}^{3\times 1}$ to the system of equations

$$\begin{cases} \kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} = \kappa^{\mathbb{P}}\theta^{\mathbb{P}}, \\ \kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} - \Sigma\gamma, \end{cases} \tag{2.9}$$

then

$$X_{t+1} = X_t + \kappa^{\mathbb{Q}}(\theta^{\mathbb{Q}} - X_t) + \Sigma Z^{\mathbb{Q}}_{t+1} \tag{2.10}$$

and thus, the factor process $\{X_t\}_{t=0}^{T}$ preserves its auto-regressive structure under the risk-neutral measure. Note that as a future research avenue, other pricing kernels can be considered. For instance, these associated with essentially affine models defined in Duffee (2002) are shown in that work to enhance the forecasting ability of the model by generalizing the physical dynamics of term structure factors while retaining the same risk-neutral dynamics.

The DTAFNS model relies on specific choices of parameters mimicking the AFNS model. Zero-coupon bond prices under such specific choices are hereby derived. The proofs for the next results are provided in Appendix A.

Lemma 2.1. *For some integer $\tau > 0$ and a real number $r \neq 1$,*

$$\zeta_0(r,\tau) \equiv \sum_{u=1}^{\tau-1} r^u = \frac{r - r^{\tau}}{1-r}, \tag{2.11}$$

$$\zeta_1(r,\tau) \equiv \sum_{u=1}^{\tau-1} u r^u = \frac{r - \tau r^{\tau} + (\tau-1) r^{\tau+1}}{(1-r)^2}, \tag{2.12}$$

$$\zeta_2(r,\tau) \equiv \sum_{u=1}^{\tau-1} u^2 r^u = \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1) r^{\tau+1} - \tau^2 r^{\tau} + r^2 + r}{(1-r)^3}. \tag{2.13}$$

Proposition 2.1. *Consider $\lambda \in (0, 1)$ and assume*

$$\kappa^{\mathbb{Q}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix}. \tag{2.14}$$

In other words, suppose that the state variables X_t have the following risk-neutral dynamics:

$$\begin{pmatrix} X^{(1)}_{t+1} - X^{(1)}_t \\ X^{(2)}_{t+1} - X^{(2)}_t \\ X^{(3)}_{t+1} - X^{(3)}_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{bmatrix} \theta^{\mathbb{Q}}_1 - X^{(1)}_t \\ \theta^{\mathbb{Q}}_2 - X^{(2)}_t \\ \theta^{\mathbb{Q}}_3 - X^{(3)}_t \end{bmatrix} + \begin{pmatrix} \Sigma_{1,1} & 0 & 0 \\ 0 & \Sigma_{2,2} & 0 \\ 0 & 0 & \Sigma_{3,3} \end{pmatrix} \begin{pmatrix} Z^{\mathbb{Q}}_{t+1,1} \\ Z^{\mathbb{Q}}_{t+1,2} \\ Z^{\mathbb{Q}}_{t+1,3} \end{pmatrix}.$$

Then, the time-t arbitrage-free price of a zero-coupon maturing at T is, for $t = 0, \ldots, T-1$,

$$P(t,T) = A_{\tau} \exp\left[-\Delta B_{\tau}^{\top} X_t\right], \tag{2.15}$$

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

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

$$\log A_{\tau} = -\Delta\theta^{\mathbb{Q}}_2 \left(B^{(1)}_{\tau} - B^{(2)}_{\tau}\right) + \Delta\theta^{\mathbb{Q}}_3 B^{(3)}_{\tau} + \frac{1}{2}\Delta^2 \upsilon_{\tau}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

with

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

$$\upsilon_\tau^{(1,1)} = \Sigma_{1,1}^2 \frac{\tau(\tau-1)(2\tau-1)}{6},$$

$$\upsilon_\tau^{(2,2)} = \frac{\Sigma_{2,2}^2}{\lambda^2}\left(\tau - 2\left[\frac{1-(1-\lambda)^\tau}{\lambda}\right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2}\right),$$

$$\upsilon_\tau^{(3,3)} = \mathbb{1}_{\{\tau>1\}}\frac{\Sigma_{3,3}^2}{\lambda^2}\Bigg[\tau - 2 + \zeta_0\big((1-\lambda)^2, \tau-1\big) + \lambda^2\zeta_2\big((1-\lambda)^2, \tau-1\big)$$
$$- 2\zeta_0((1-\lambda), \tau-1) - 2\lambda\zeta_1((1-\lambda), \tau-1) + 2\lambda\zeta_1\big((1-\lambda)^2, \tau-1\big)\Bigg],$$

$$\upsilon_\tau^{(1,2)} = \upsilon_\tau^{(2,1)} = \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda),\tau)\right),$$

$$\upsilon_\tau^{(1,3)} = \upsilon_\tau^{(3,1)} = \mathbb{1}_{\{\tau>1\}}\rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0((1-\lambda), \tau-1)$$
$$- (\lambda+1)\zeta_1((1-\lambda), \tau-1) - \lambda\zeta_2((1-\lambda), \tau-1)\Bigg],$$

$$\upsilon_\tau^{(2,3)} = \upsilon_\tau^{(3,2)} = \mathbb{1}_{\{\tau>1\}}\rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\Bigg(\frac{\tau - 2 - (2-\lambda)\zeta_0((1-\lambda), \tau-1) + (1-\lambda)\zeta_0\big((1-\lambda)^2, \tau-1\big)}{\lambda^2}$$
$$+ \frac{-\zeta_1((1-\lambda), \tau-1) + (1-\lambda)\zeta_1\big((1-\lambda)^2, \tau-1\big)}{\lambda}\Bigg).$$

This entails that spot rates have the form

$$y(t,T) = X_t^{(1)} + \frac{(1-(1-\lambda)^\tau)}{\tau\lambda}X_t^{(2)} + \left[\frac{(1-(1-\lambda)^{\tau-1})}{\tau\lambda} - \frac{\tau-1}{\tau}(1-\lambda)^{\tau-1}\right]X_t^{(3)} - \frac{1}{\Delta\tau}\log A_\tau. \tag{2.16}$$

Remark 2.1. In Proposition 2.1, parameter $\theta_1^{\mathbb{Q}}$ does not appear in the zero-coupon price formula (2.15) and thus can be set to $\theta_1^{\mathbb{Q}} = 0$.

To obtain risk-neutral dynamics of the form proposed in Proposition 2.1, the following physical dynamics is considered:

$$\kappa^{\mathbb{P}} = \begin{pmatrix} \kappa_{1,1}^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_{2,2}^{\mathbb{P}} & -\lambda \\ 0 & 0 & \kappa_{3,3}^{\mathbb{P}} \end{pmatrix}, \quad \theta^{\mathbb{P}} = \begin{pmatrix} 0 \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix}$$

which allows meeting the second condition in (2.9) with

$$\gamma = \begin{pmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{pmatrix}, \quad \gamma_i = \left(\kappa_{i,i}^{\mathbb{P}} - \mathbb{1}_{\{i>1\}}\lambda\right)/\Sigma_{i,i},\ i = 1,2,3. \tag{2.17}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Furthermore, substituting (2.17) in the first equation of (2.9) leads to

$$\theta_2^{\mathbb{Q}} = \lambda^{-1}\left(\kappa_{2,2}^{\mathbb{P}}\theta_2^{\mathbb{P}} + \kappa_{3,3}^{\mathbb{P}}\theta_3^{\mathbb{P}} - \lambda\theta_3^{\mathbb{P}}\right), \quad \theta_3^{\mathbb{Q}} = \lambda^{-1}\kappa_{3,3}^{\mathbb{P}}\theta_3^{\mathbb{P}}.$$

Conversely, physical parameters can be retrieved from risk-neutral parameters through

$$\kappa_{i,i}^{\mathbb{P}} = \mathbb{1}_{\{i>1\}}\lambda + \Sigma_{i,i}\gamma_i,\ i = 1, 2, 3, \quad \theta_3^{\mathbb{P}} = \lambda\theta_3^{\mathbb{Q}}/\kappa_{3,3}^{\mathbb{P}}, \quad \theta_2^{\mathbb{P}} = \frac{\lambda}{\kappa_{2,2}^{\mathbb{P}}}\left[\theta_2^{\mathbb{Q}} - \frac{\theta_3^{\mathbb{Q}}}{\kappa_{3,3}^{\mathbb{P}}}\left(\kappa_{3,3}^{\mathbb{P}} - \lambda\right)\right]. \tag{2.18}$$

Remark 2.2. Since $\kappa_{1,1}^{\mathbb{Q}} = 0$, an alternative specification enforcing $\kappa_{1,1}^{\mathbb{P}} = 0$ has been tested, but it was ultimately not retained due to lower performance. See Appendix E of the Online Appendix for more details.

Remark 2.3. The DTANFS model of this paper is specified slightly differently than in Hong et al. (2019). In the latter work, the matrix $\kappa^{\mathbb{Q}}$ is chosen such that factor loadings B_τ exactly match these of the Nelson-Siegel model (2.1). Conversely, this paper emphasizes an auto-regressive representation of factors, i.e. Equation (2.10), which is the discrete-time counterpart of the diffusive dynamics of the AFNS model (2.3). Nevertheless, both specifications of the model are quite similar conceptually. In Hong et al. (2019), only a recursive representation of A_τ is provided, which makes the derivation of the closed-form solution for A_τ a useful contribution of the present work.

3. Model estimation

This section presents a method to estimate the DTAFNS model using historical spot curves data. The model is estimated on Canadian interest rate data. The fit performance is analyzed and benchmarked against the discrete-time Gaussian three-factor model of Augustyniak *et al.* (2021) and a version of the Dynamic Nelson-Siegel model of Diebold & Li (2006).

3.1. Estimation framework

Several methods have been considered in the literature to estimate factor models of interest rates, such as conventional maximum likelihood estimation (Chen & Scott, 1993), maximum likelihood with Kalman filters (Duan & Simonato, 1999; Lemke, 2006; Park, 2014; Augustyniak *et al.*, 2021), the method of simulated moments (Dai & Singleton, 2000), or Bayesian approaches (Hong et al., 2019). In this paper, the Kalman filter approach is applied. Observed spot rates (observable data) are considered to be a noisy version of the true spot rates (the signal). Setting up the Kalman filter requires to first derive the recursive Gaussian linear relationship between the observable quantities and factors driving the signal, which is hereby provided.

Assume that on each time t a set of M annualized continuously compounded spot rates with fixed times-to-maturity $n_1, ..., n_M$ are observed, which are denoted by the column vector

$$\hat{\mathsf{y}}(t) \equiv (\hat{y}(t, t+n_1), ..., \hat{y}(t, t+n_M))^\top.$$

Define also the corresponding model-implied spot rate vector

$$\mathsf{y}(t) \equiv \left(y(t, t+n_1), ..., y(t, t+n_M)\right)^\top$$

where (2.16) leads to

$$y(t, t+n) = \frac{-1}{n\Delta}\log A_n + \frac{1}{n}B_n^\top X_t. \tag{3.1}$$

The relationship between observed and model-implied spot rate is assumed to be

$$\hat{\mathsf{y}}(t) = \mathsf{y}(t) + \eta_{\mathrm{t}}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

where $\{\eta_t\}_{t=1}^{T}$ is a multivariate Gaussian white noise with diagonal variance-covariance matrix H. Furthermore, for simplicity, all values on the diagonal of matrix H are assumed to be identical and equal to some parameter h. This leads to

$$\hat{\mathsf{y}}(t) = \mathsf{a} + \mathsf{B}X_t + \eta_t, \tag{3.2}$$

where

$$\mathsf{a} \equiv \left(\frac{-1}{n_1\Delta}\log A_{n_1}, \ldots, \frac{-1}{n_M\Delta}\log A_{n_M}\right)^{\top}, \tag{3.3}$$

$$\beta_n \equiv \left(\frac{B_n^{(1)}}{n}, \frac{B_n^{(2)}}{n}, \frac{B_n^{(3)}}{n}\right)^{\top},$$

$$\mathsf{B} \equiv \left(\beta_{n_1}, \ldots, \beta_{n_M}\right)^{\top}. \tag{3.4}$$

Furthermore, the transition equation describing the dynamics of the latent factors X_t is obtained from (2.8):

$$X_{t+1} = \mathbf{b} + \mathbf{D}X_t + \xi_{t+1}, \tag{3.5}$$

where

$$\mathbf{b} \equiv \kappa^{\mathbb{P}}\theta^{\mathbb{P}}, \quad \mathbf{D} \equiv I - \kappa^{\mathbb{P}}, \quad \xi_{t+1} \equiv \Sigma Z_{t+1}^{\mathbb{P}}. \tag{3.6}$$

The sequence $\{\xi_t\}_{t=1}^{T}$ is therefore a multivariate Gaussian white noise with covariance matrix

$$\mathbf{Q} = \Sigma\rho\Sigma. \tag{3.7}$$

The representation (3.2)–(3.5) of state space variables allows estimating the DTAFNS model using a Kalman filter. The general Kalman filter algorithm is presented in Appendix D of the Online Appendix. The Kalman filter allows deriving the log-likelihood of observed spot rate curves for a candidate set of parameters. Algorithm 1 indicates steps to calculate the log-likelihood function in the context of the DTAFNS model. In this algorithm, p denotes joint density functions. In the present work, the implementation of the Kalman filter is performed with the `R` package `FKF`. Moreover, to find maximum likelihood estimates of the parameters, consistently with Augustyniak *et al.* (2021), an optimization of the likelihood is conducted with the `R` package `Rsolnp` (Ghalanos & Theussl, 2015) applying the nonlinear augmented Lagrange multiplier optimization method of Ye (1987). The optimization is performed under the following constraints:

$$\Sigma_{i,i} \geq 0, \quad h \geq 0, \quad -1 \leq \rho_{i,j} \leq 1,\ i, j = 1, 2, 3, \quad \lambda \in (0, 1).$$

Additional constraints could be applied to enforce the positive semi-definiteness of matrices Σ and ρ, but optimization results obtained with our dataset satisfy such constraints without explicitly specifying them to the optimizer.

The smoothed state density is the density of time-t factors conditional on all observations from the sample, namely $p(X_t|\hat{y}(1), \ldots, \hat{y}(T))$. Such distribution is multivariate Gaussian with mean $\bar{x}_{t|T}$ and covariance matrix $P_{t|T}$. Such mean and variance are obtained with the Kalman smoother algorithm (Shumway & Stoffer, 2017) which is also shown in Algorithm 1.

3.2. Model estimation and analysis

This section presents the DTAFNS model estimation using historical Canadian spot rate data.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Algorithm 1 Kalman filter algorithm for the calculation of likelihood function and smoothed state densities

Input: Initial values $\bar{x}_{1|0} \in \mathbb{R}^3$, $P_{1|0} \in \mathbb{R}^{3\times 3}$, parameters h, λ, $\theta^{\mathbb{Q}}$, ρ, γ, Σ and spot rate dataset $\hat{\mathsf{y}}(t)$, $t = 0, \ldots, T$.
Outputs: Log-likelihood $L(\hat{y}) = \log\big(p\big(\hat{\mathsf{y}}(T), \hat{\mathsf{y}}(T-1), ..., \hat{\mathsf{y}}(1)\big)\big)$ and smoothed density moments $\bar{x}_{t|T}, P_{t|T}$, $t = 1, \ldots, T$.

Calculating the log-likelihood:
Calculate a, $\mathbf{b}$, $\mathbf{D}$, $\mathbf{Q}$ *and* B *with* (2.18), (3.3), (3.3), (3.6), (3.7) *and Proposition* 2.1.
for $t \in \{1, ..., T\}$ ***do***
Calculate $p(\hat{\mathsf{y}}(t)|\hat{\mathsf{y}}(1), ..., \hat{\mathsf{y}}(t-1))$:
$\bar{y}(t) \leftarrow \mathsf{a} + \mathsf{B}\bar{x}_{t|t-1}$
$\bar{\Sigma}_t \leftarrow \mathsf{B}P_{t|t-1}\mathsf{B}^\top + H$
$p(\hat{\mathsf{y}}(t)|\hat{\mathsf{y}}(1), ..., \hat{\mathsf{y}}(t-1)) \sim N(\bar{y}(t), \bar{\Sigma}_t)$

Calculate $\bar{x}_{t|t}$ *and* $P_{t|t}$:
$\bar{x}_{t|t} \leftarrow \bar{x}_{t|t-1} + P_{t|t-1}\mathsf{B}^\top\bar{\Sigma}_t^{-1}(\hat{\mathsf{y}}(t) - \bar{\mathsf{y}}(t))$
$P_{t|t} \leftarrow P_{t|t-1} - P_{t|t-1}\mathsf{B}^\top\bar{\Sigma}_t^{-1}\mathsf{B}P_{t|t-1}$
$\bar{x}_{t+1|t} \leftarrow \mathbf{b} + \mathbf{D}\bar{x}_{t|t}$
$P_{t+1|t} \leftarrow \mathbf{D}P_{t|t}\mathbf{D}^\top + \mathbf{Q}$
$L(\hat{y}) \leftarrow \sum_{t=1}^{T} \log p\big(\hat{\mathsf{y}}(t)|\hat{\mathsf{y}}(1), ..., \hat{\mathsf{y}}(t-1)\big)$

Calculating the smoothed state density:
for $t \in \{T-1, \ldots, 1\}$ **do**
$\mathbf{J}_t \leftarrow P_{t|t}\mathbf{D}^\top P_{t+1|t}^{-1}$,
$\bar{x}_{t|T} \leftarrow \bar{x}_{t|t} + \mathbf{J}_t(\bar{x}_{t+1|T} - \bar{x}_{t+1|t})$,
$P_{t|T} \leftarrow P_{t|t} + \mathbf{J}_t(P_{t+1|T} - P_{t+1|t})\mathbf{J}_t^\top$.

3.2.1. The dataset

To estimate the model, end-of-month Canadian spot rate curves from January 1986 to January 2022 (434 months) are considered. Such yield curves are provided publicly on the website of the Bank of Canada[2] and are constructed through an exponential smoothing methodology described in Bolder *et al.* (2004). 33 spot rate tenors are considered, which include short-end maturities of 3, 6, and 9 months and all integer times-to-maturity from 1 to 30 years.[3]

3.2.2. Estimated parameters and performance

Table 1 gives estimates of DTAFNS model parameters obtained with Algorithm 1 which applies the Kalman filter. Factor 1 innovations are negatively correlated with these of the two other factors, which are positively correlated between themselves. The negative correlation between the level factor $X^{(1)}$ and the slope and curvature factors $X^{(2)}$ and $X^{(3)}$ tends to reduce the frequency of very large or low values of the short rate r_t. Furthermore, it implies that when the long-term rates increase through higher values of $X^{(1)}$, the slope given by $-X^{(2)}$ also tends to go up. Thus, movements in the long-end of the curve are not necessarily matched by movements of a similar

[2]Source: https://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/.
[3]For all years up to 1990 inclusively, times-to-maturity between 26 years and 30 years are missing in the dataset and are therefore not considered.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table 1. Maximum likelihood estimates of the DTAFNS model parameters

i	$\kappa_{i,i}^{\mathbb{P}}$	γ_i	$\Sigma_{i,i}$	$\theta_i^{\mathbb{P}}$	$\theta_i^{\mathbb{Q}}$	$\bar{x}_{i,1\|0}$	λ	h
1	0.0075	2.7923	0.0027	0	0	0.0491		
2	0.0288	1.2016	0.0045	0.0301	0.0633	0.0391	0.0233	3.76×10^{-6}
3	0.0354	1.7167	0.0070	0.0505	0.0766	0.0291		

$$\rho = \begin{bmatrix} 1 & -0.6303 & -0.4097 \\ -0.6303 & 1 & 0.2993 \\ -0.4097 & 0.2993 & 1 \end{bmatrix}, P_{1|0} = \begin{bmatrix} 4.45\times10^{-6} & 0 & 0 \\ 0 & 4.45\times10^{-6} & 0 \\ 0 & 0 & 4.45\times10^{-6} \end{bmatrix}$$

Notes: Maximum likelihood parameter estimates for the DTAFNS model presented in Section 2.2 obtained with the Kalman filter (see Algorithm 1). The data sample includes Canadian end-of-month yield curves from January 1986 to January 2022. $\bar{x}_{i,1|0}$ refers to element i of the vector $\bar{x}_{1|0}$.

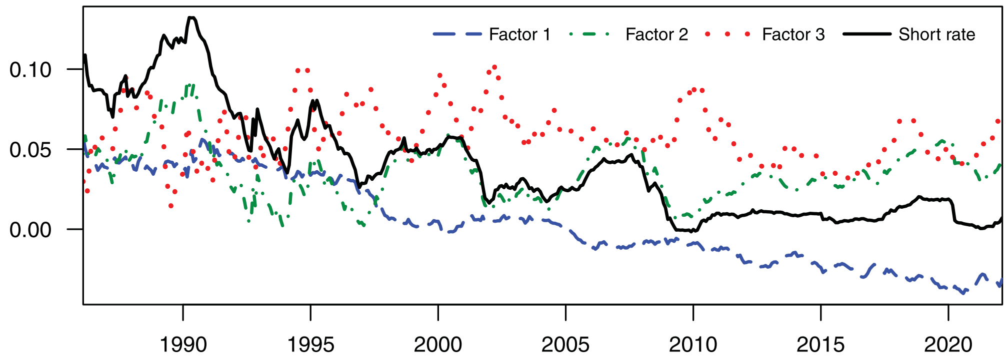

Figure 1. Model-implied factors and short rate time series.
Notes: Time series of DTAFNS model-implied factors which correspond to the smoothed state inferences $E^{\mathbb{P}}\left[x_t^{(i)}|\hat{y}(1), \hat{y}(2), ..., \hat{y}(T)\right], i = 1, 2, 3$ provided by Algorithm 1, and implied short rates obtained by summing smoothed values of the first two factors. The model is estimated on the end-of-month Canadian spot rate curves extending from January 1986 to January 2022.

magnitude and direction in the short-end. The hump in the curve reflected by $X^{(3)}$ also tends to decrease during periods of increasing of long-term spot rates. The speed of reversion $\kappa_{i,i}^{\mathbb{P}}$ is much lower for the first factor ($i = 1$) than for the other two, which is not surprising due to the overall rates level having had a clear lasting downward trend in the historical sample, representing close-to-non-stationary behavior. Finally, the estimation implies a long-term average of short rates of $\theta_2^{\mathbb{P}} = 0.0301$ under physical dynamics.

Fig. 1 shows the smoothed value (the smoothed distribution expected value) of latent factors provided by the Kalman smoother procedure described in Algorithm 1. The black curve is the short rate implied by the model which is obtained by summing the smoothed values of the first two factors, consistently with (2.7). The decreasing trend of the short rate throughout the data sample is recovered by the downward trend of Factor 1. Conversely, Factor 2 and Factor 3, driving the slope and humps in the curve respectively, exhibit more stationary fluctuations.

To assess the ability of the DTAFNS model to match observed short- and long-term spot rates of the term structure, time series of observed spot rates of three-month and 10-year tenors are compared to the model-implied spot rates for the same tenors. Such time series are provided in Fig. 2. The model can be seen to adequately match short- and long-term rates throughout the entire sample, as differences between the model-implied and observed spot rates are very small.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

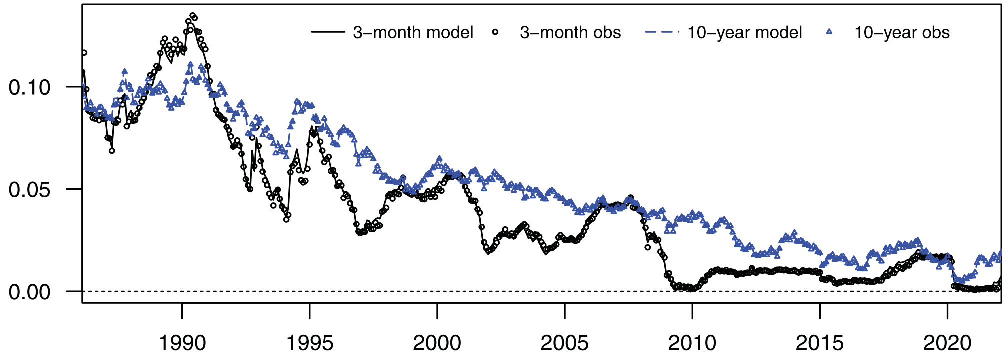

Figure 2. Model-implied and observed spot rate time series for 3-month and 10-year tenors.
Notes: Time series of observed 3-month (short-term) and 10-year (long-term) maturity spot rates (dotted curves), with corresponding spot rates implied by the fitted DTAFNS model. The dataset considered is the end-of-month Canadian spot rate curves extending from January 1986 to January 2022.

3.3. Performance assessment and benchmarking

This section analyzes the adequacy of the fit and the predictive performance of the DTAFNS model through benchmarking.

3.3.1. Benchmarks

The discrete-time Gaussian three-factor model of Augustyniak *et al.* (2021), subsequently referred to as the DG3 model, and a version of the dynamic Nelson-Siegel (DNS) model introduced by Diebold & Li (2006) are benchmarks considered in this study for comparison of performance results. See Appendix B for a formal specification of these two models. The main differences between DG3 and the DTAFNS model are that (i) the latent factors in the former only regress on themselves instead of allowing cross-factor auto-regressions, (ii) the DTAFNS has a short rate equal to the sum of two of the factors instead of the three as in the benchmark, and (iii) the DTAFNS has a fixed structure for the matrix $\kappa^{\mathbb{Q}}$ driving the auto-regression intensity, unlike the Augustyniak *et al.* (2021) model allowing optimizing such quantities. Regarding the second benchmark, unlike in the DTAFNS model, the DNS model does not include a deterministic adjustment term in spot rates to rule out arbitrage opportunities. Additionally, a nested version of the DTAFNS model which sets all correlation parameters to zero, i.e. $\rho_{i,j} = 0$, $i, j = 1, 2, 3$, is also considered for benchmarking purposes. This benchmark, which is denoted by the acronym DTAFNS-U, is used to evaluate the impact of the innovation correlations on performance. Table 2 shows parameter estimates for the DTAFNS-U model, which are slightly different but qualitatively similar to these of the DTAFNS model presented in Table 1.

3.3.2. Goodness-of-fit of the model and forecasting ability

To assess the model's ability to adequately fit the observed data, we compare the log-likelihood of the DTAFNS, the DTAFNS-U, the DG3 and the DNS models. Two types of log-likelihoods are computed: an in-sample version obtained by fitting the models to the entire dataset, and an out-of-sample version which is obtained by considering an expanding window sequential approach. For the latter, in the first iteration, data from January 1986 up to the end of 2016 is considered as a first training set, and data from 2017 is considered as a test set. In any following iteration, the training sets are expanded by one year and the test year moves to the year following that of the previous iteration. The last iteration has a test set including data for both 2021 and 2022 since the dataset only includes data for January in 2022. The aggregated out-of-sample log-likelihood is obtained

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table 2. Maximum likelihood estimates of the DTAFNS-U model parameters

i	$\kappa^{\mathbb{P}}_{i,i}$	γ_i	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_i$	$\theta^{\mathbb{Q}}_i$	$\bar{x}_{i,1\|0}$	λ	h
1	0.0075	2.7250	0.0021	0	0	0.0502		
2	0.0288	1.0161	0.0038	0.0343	0.0653	0.0403	0.0227	3.81×10^{-6}
3	0.0354	1.8645	0.0059	0.0523	0.0775	0.0303		

$$\rho = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_{1|0} = \begin{bmatrix} 4\times10^{-6} & 0 & 0 \\ 0 & 4\times10^{-6} & 0 \\ 0 & 0 & 4\times10^{-6} \end{bmatrix}$$

Notes: Maximum likelihood parameter estimates for the DTAFNS-U model obtained with the Kalman filter. The DTAFNS-U is identical to the DTAFNS model except that the term structure innovations are uncorrelated, i.e. the matrix ρ is set to the identity matrix. The data sample includes Canadian end-of-month yield curves from January 1986 to January 2022. $\bar{x}_{i,1|0}$ refers to element i of the vector $\bar{x}_{1|0}$.

Table 3. Log-likelihood of the DTAFNS model and its benchmarks

	Out-of-sample						
Model	2017	2018	2019	2020	2021–2022	Aggregated	In-sample
DTAFNS	2,034	2,001	2,035	2,015	2,183	10,268	65,702
DTAFNS-U	2,003	1,976	2,015	2,001	2,335	10,330	65,480
DG3	2,012	1,978	2,024	2,027	2,200	10,241	66,069
DNS	1,985	1,982	1,993	2,001	2,163	10,124	64,610

Notes: Comparison of the log-likelihood of the DTAFNS model of Section 2.2 and of the DG3 and DNS benchmarks described in Appendix B. The data sample is the Canadian end-of-month yield curve. The in-sample dataset starts in January 1986 and ends in January 2022. The out-of-sample estimation procedure uses an expanding window approach described in Section 3.3.2. The aggregated out-of-sample log-likelihood is ultimately obtained by summing the log-likelihood for all test years.

by summing these associated with all test sets. Table 3 provides the results. The best-performing model in terms of aggregate log-likelihood over the full sample is the DTAFNS-U. However, when looking at year-by-year performance, the DTAFNS outperforms the DTAFNS-U for the three first years, while only underperforming the latter for the last year. Moreover, the DTAFNS beats both the DG3 and DNS benchmarks out-of-sample for all years except for the DG3 in years 2020 and 2021–2022. While the in-sample log-likelihood of the DG3 is higher than that of the three other competing models, such better fit does not translate to better (out-of-sample) predictive ability. These tests imply that either the DTAFNS or the DTAFNS-U should be favored over the DG3 and DNS benchmarks.

To further analyze the predictive ability of the competing interest rate models, Appendix C provides descriptive statistics on out-of-sample point forecasts of spot rates for various tenors. The difference between the performance of the DTAFNS model and that of benchmarks seems overall marginal, with slightly higher (lower) performance exhibited by the DTAFNS for long (medium) tenors. Therefore, our analyses indicate that point forecasting performance of the various models is somewhat similar, but the distributional predictive ability of the DTAFNS model is slightly superior.

To visualize the ability of the DTAFNS model and benchmarks to replicate observed yield curves, Fig. 3 presents the realized spot curves and model-implied counterparts for four selected dates. These four days are the same that are considered in Augustyniak *et al.* (2021) to display fitting performance, namely December 29, 2006, December 31, 2008, June 30, 2016 and October 31, 2018. Such dates exhibit several spot curve shapes: flat, upward sloping or humped. The figure shows that the DTAFNS and all three benchmarks are able to reproduce observed yield curves reasonably well.

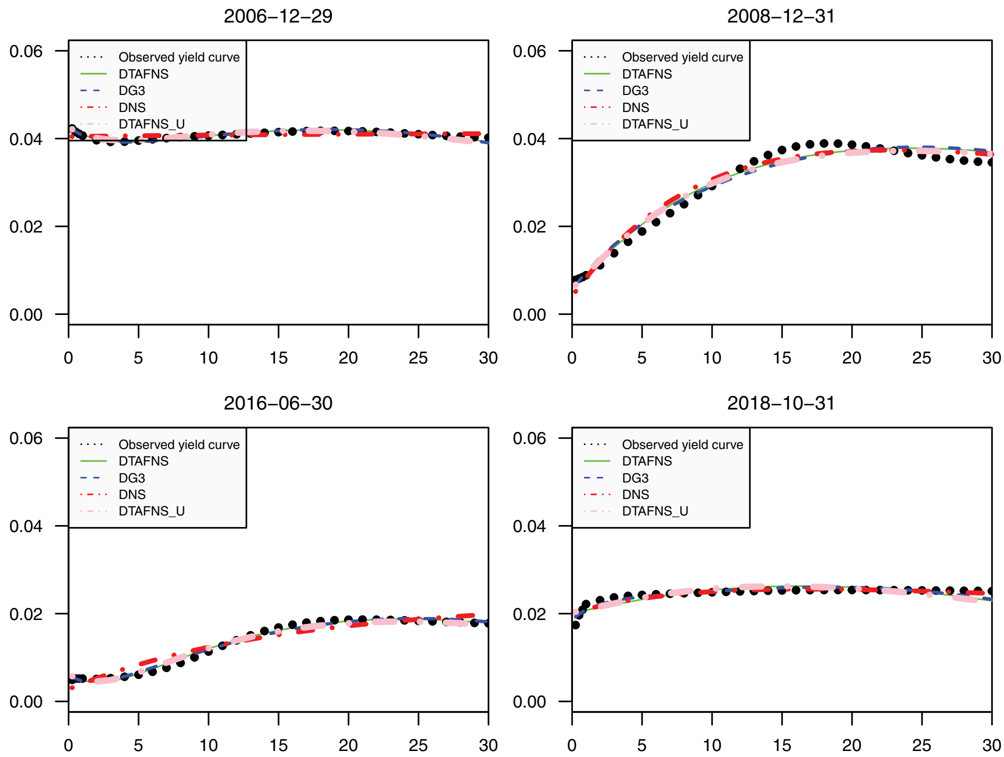

Figure 3. Model-implied and observed yield curves.
Notes: Realized and model-implied spot rate curves on the four following dates: December 29, 2006, December 31, 2008, June 30, 2016 and October 31, 2018. Dotted black line: observed spot rates. Full green line: DTAFNS model-implied curve. Dashed blue line: DG3 benchmark implied curve. Red dotted-dashed line: DNS benchmark implied curve. Pink dotted-dashed line: DTAFNS-U benchmark implied curve. The observed data are end-of-month Canadian spot rates provided by the Bank of Canada.

3.3.3. Frequency of negative values

A drawback associated with the use of the Gaussian distribution within interest rate models is the possibility of producing negative short rates. Although negative rates have been observed in some European markets, such phenomenon almost did not occur in North America. To investigate the propensity of the DTAFNS model and its benchmarks to generate negative interest rates, 200,000 five-year monthly paths are simulated under the physical measure with the estimated parameters drawn from Tables 1, 2, B.1 and B.2, respectively, and starting values for factors being their smoothed values on January 2022. Note that the sample period used for the development of the model excludes interest rate hikes that occurred after January 2022. The January 2022 starting point for the simulation therefore explains the large proportion of simulated negative rates.

Table 4 reports (i) the proportion of simulated observations below the thresholds 0, −0.01, −0.02, or −0, 03, and (ii) the proportion of paths with at least one observation below such thresholds. The DTAFNS model significantly reduces such proportions when compared with the DTAFNS-U, the DG3, and the DNS models, which is desirable.

4. An application to mixed fund modeling

One of the main building blocks of the mixed fund (a fund containing both fixed income and equity) model of Augustyniak *et al.* (2021), which is used for variable annuity pricing, is the underlying factor-based DG3 interest rate model. Since the DTAFNS model is shown herein to

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table 4. Probability of observing negative short rates with the DTAFNS model or its benchmarks

Model	Proportion of months				Proportion of paths with at least one month			
	< 0	< -0.01	< -0.02	< -0.03	< 0	< -0.01	< -0.02	< -0.03
DTAFNS	0.177	0.057	0.015	0.003	0.644	0.299	0.106	0.029
DTAFNS-U	0.227	0.097	0.037	0.012	0.714	0.419	0.206	0.085
DG3	0.425	0.176	0.058	0.015	0.942	0.629	0.303	0.109
DNS	0.240	0.082	0.022	0.005	0.730	0.379	0.142	0.040

Notes: Within 200,000 five-year monthly simulated paths of the DTAFNS model and the DG3 and DNS benchmarks described in Appendix B, the proportion of (i) simulated months with short rates being smaller than, respectively, 0, −0.01, −0.02 and −0.03, and (ii) simulated paths with at least one month below such thresholds are reported. Model parameters are drawn from Tables 1, 2, B.1, and B.2, respectively, with starting values of the factors being their smoothed values on January 2022.

improve predictive performance over the DG3 benchmark, it is relevant to adapt the Augustyniak *et al.* (2021) mixed fund model by replacing the DG3 model with the DTAFNS model.

4.1. Rolling bond fund returns

The dynamics of the mixed fund model of Augustyniak *et al.* (2021) involves a regression on movements of interest rate factors, among other quantities. The inclusion of such regressors is justified by an analysis of returns of a rolling bond fund, which are shown to depend on such drivers. Therefore, to adapt the Augustyniak *et al.* (2021), the rolling bond fund dynamics is derived under the DTAFNS dynamics.

Consider a fund containing a single zero-coupon of time-to-maturity τ that is rolled-over on each period. Denoting its time-$t+1$ log-return by $R^{(\tau)}_{t+1}$, (2.15) leads to

$$
\begin{aligned}
R^{(\tau)}_{t+1} &= \log\left(\frac{P(t+1,t+\tau)}{P(t,t+\tau)}\right) \\
&= \log A_{\tau-1} - \log A_\tau - \Delta B^\top_{\tau-1} X_{t+1} + \Delta B^\top_\tau X_t \\
&= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} \left(B^{(i)}_{\tau-1} X^{(i)}_{t+1} - B^{(i)}_\tau X^{(i)}_t\right).
\end{aligned} \tag{4.1}
$$

Moreover, recalling that $r_t = X^{(1)}_t + X^{(2)}_t$ and using (2.14), (2.18), Lemma A.2 and (4.1), the excess return can be expressed as

$$
\begin{aligned}
R^{(\tau)}_{t+1} - \Delta r_t &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} \left(B^{(i)}_{\tau-1} X^{(i)}_{t+1} - (B^{(i)}_\tau - 1) X^{(i)}_t\right) \\
&\quad + \Delta \sum_{i=1}^{3} X^{(i)}_t - \Delta\left(X^{(1)}_t + X^{(2)}_t\right) \\
&= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} \left(B^{(i)}_{\tau-1} X^{(i)}_{t+1} - \left(B^{(i)}_{\tau-1}\left(1-\kappa^{\mathbb{Q}}_{i,i}\right)\right.\right. \\
&\quad \left.\left. - \mathbb{1}_{\{i=3\}} (1-\lambda)^{\tau-1}\right) X^{(i)}_t\right) + \Delta X^{(3)}_t \\
&= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} B^{(i)}_{\tau-1}\left(X^{(i)}_{t+1} - \left(1-\kappa^{\mathbb{Q}}_{i,i}\right) X^{(i)}_t\right) + \Delta\left(1-(1-\lambda)^{\tau-1}\right) X^{(3)}_t.
\end{aligned} \tag{4.2}
$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Mimicking Augustyniak *et al.* (2021) by considering a portfolio consisting of M positions with weights $\omega_1, \ldots, \omega_M$ (where $\sum_{j=1}^{M} \omega_j = 1$) that are rolled-over on bonds with fixed time-to-maturities $\tau_1, \ldots, \tau_M$, the log-return of the rolling bond fund between times t and $t+1$, denoted by $R^{(B)}_{t+1}$, is approximately characterized by

$$\begin{aligned}
R^{(B)}_{t+1} - \Delta r_t &\approx \omega_1 R^{(\tau_1)}_{t+1} + \cdots + \omega_M R^{(\tau_M)}_{t+1} - \Delta r_t \\
&= \omega_1 \left(R^{(\tau_1)}_{t+1} - \Delta r_t \right) + \cdots + \omega_M \left(R^{(\tau_M)}_{t+1} - \Delta r_t \right) \\
&= \sum_{j=1}^{M} \omega_j \Bigg(\log A_{\tau_j - 1} - \log A_{\tau_j} - \Delta \sum_{i=1}^{3} B^{(i)}_{\tau_j - 1} \left(X^{(i)}_{t+1} - \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right) X^{(i)}_t \right) \\
&\quad + \Delta \left(1 - (1-\lambda)^{\tau_j - 1}\right) X^{(3)}_t \Bigg) \\
&= \sum_{j=1}^{M} \omega_j \left(\log A_{\tau_j - 1} - \log A_{\tau_j} \right) - \Delta \sum_{i=1}^{3} \sum_{j=1}^{M} \omega_j B^{(i)}_{\tau_j - 1} \left(X^{(i)}_{t+1} - \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right) X^{(i)}_t \right) \\
&\quad + \Delta \sum_{j=1}^{M} \omega_j \left(1 - (1-\lambda)^{\tau_j - 1}\right) X^{(3)}_t .
\end{aligned} \tag{4.3}$$

Relationship (4.3) will serve as the basis to model the fixed income part of the mixed fund in the proposed model. The key difference with the Augustyniak *et al.* (2021) model is the presence of the last term proportional to $X^{(3)}_t$ in (4.3), which is not present in their work.

4.2. Risky assets returns

The adaptation of the mixed fund model proposed in this paper keeps the same model for excess equity returns as in Augustyniak *et al.* (2021). Returns of q equity assets are represented with the exponential generalized auto-regressive conditional heteroskedastic (EGARCH) model of Nelson (1991):

$$R^{(S)}_{t+1,j} - \Delta r_t = \lambda^{(S)}_j \sqrt{h^{(S)}_{t,j}} - \frac{1}{2} h^{(S)}_{t,j} + \sqrt{h^{(S)}_{t,j}} Z^{(S)}_{t+1,j}, \tag{4.4}$$

$$\log h^{(S)}_{t,j} = \omega^{(S)}_j + \alpha^{(S)}_j Z^{(S)}_{t,j} + \gamma^{(S)}_j \left(|Z^{(S)}_{t,j}| - \frac{2}{\sqrt{2\pi}} \right) + \beta^{(S)}_j \log h^{(S)}_{t-1,j}, \tag{4.5}$$

where $\left(\omega^{(S)}_j, \alpha^{(S)}_j, \gamma^{(S)}_j, \beta^{(S)}_j\right)$ are parameters associated with the conditional volatility. The process $\left\{h^{(S)}_{t,j}\right\}_{t=1}^{T-1}$ is $\mathcal{F}$-adapted. Parameters $\lambda^{(S)}_j$ represent equity risk premia. For $j = 1, \ldots, q$, $Z^{(S)}_j \equiv \left\{Z^{(S)}_{t,j}\right\}_{t=1}^{T}$ are standard Gaussian white noises under $\mathbb{P}$, independent of Z_t, where the correlation between $Z^{(S)}_{t,i}$ and $Z^{(S)}_{t,j}$ is denoted $\Gamma_{i,j}$.

As in Augustyniak *et al.* (2021), the following risk-neutral dynamics for the q equity assets is assumed:

$$R^{(S)}_{t+1,j} - \Delta r_t = -\frac{1}{2} h^{(S)}_{t,j} + \sqrt{h^{(S)}_{t,j}} Z^{\mathbb{Q}(S)}_{t+1,j}, \tag{4.6}$$

$$\log h^{(S)}_{t,j} = \omega^{(S)}_j + \alpha^{(S)}_j \left(Z^{\mathbb{Q}(S)}_{t,j} - \lambda^{(S)}_j \right) + \gamma^{(S)}_j \left(|Z^{\mathbb{Q}(S)}_{t,j} - \lambda^{(S)}_j| - \frac{2}{\sqrt{2\pi}} \right) + \beta^{(S)}_j \log h^{(S)}_{t-1,j}, \tag{4.7}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

where, under the risk-neutral measure $\mathbb{Q}$, $\left\{Z_{t,j}^{\mathbb{Q}(S)}\right\}_{t=1}^{T} = \left\{Z_{t,j}^{(S)} + \lambda_j^{(S)}\right\}_{t=1}^{T}$ are also standard Gaussian white noises for each $j = 1, \ldots, q$, still with a contemporaneous correlation matrix $\Gamma = [\Gamma_{i,j}]_{i,j=1}^{q}$. The independence between the equity innovations $\left\{Z_{t,j}^{\mathbb{Q}(S)}\right\}_{t=1}^{T}$ and interest rate innovations $\left\{Z_t^{\mathbb{Q}}\right\}_{t=1}^{T}$ is assumed to also hold under $\mathbb{Q}$.

4.3. Mixed fund dynamics

To motivate the structure of the proposed mixed fund model, consider a portfolio invested in the rolling bond fund with weight W and in equity indices with respective weights $\tilde{\psi}_1^{(S)}, \ldots, \tilde{\psi}_q^{(S)}$, where $W + \sum_{j=1}^{q} \tilde{\psi}_j^{(S)} = 1$. Its excess return would be represented as

$$\begin{aligned} \tilde{R}_{t+1}^{(F)} - \Delta r_t &\approx W\left(R_{t+1}^{(B)} - \Delta r_t\right) + \sum_{j=1}^{q} \tilde{\psi}_j^{(S)}\left(R_{t+1,j}^{(S)} - \Delta r_t\right) \\ &= \tilde{\psi}_0 + \sum_{i=1}^{3} \tilde{\psi}_i\left(X_{t+1}^{(i)} - \left(1-\kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)}\right) + \tilde{\psi}_3' X_t^{(3)} + \sum_{j=1}^{q} \tilde{\psi}_j^{(S)}\left(R_{t+1,j}^{(S)} - \Delta r_t\right) \end{aligned} \tag{4.8}$$

where

$$\tilde{\psi}_0 = W \sum_{j=1}^{M} \omega_j \left(\log A_{\tau_j - 1} - \log A_{\tau_j}\right), \tag{4.9}$$

$$\tilde{\psi}_i = -W\Delta \sum_{j=1}^{M} \omega_j B_{\tau_j - 1}^{(i)}, \quad i = 1, 2, 3, \tag{4.10}$$

$$\tilde{\psi}_3' = W\Delta \sum_{j=1}^{M} \omega_j \left(1 - (1-\lambda)^{\tau_j - 1}\right). \tag{4.11}$$

Equation (4.8) serves as the conceptual basis for the proposed mixed fund model. However, instead of relying on (4.9)–(4.11), model parameters can be directly estimated from the data in an econometric fashion. Indeed, allocations to fixed income and equity assets from the mixed funds are not exactly identical to those implied by previous assumptions – that is an investment in a rolling horizon bond fund and in equity indices – and basis risk needs to be taken into account. Moreover, as in Augustyniak *et al.* (2021), basis risk is further represented through a noise component also following EGARCH dynamics. Mixed fund excess returns are thus assumed to be of the form

$$R_{t+1}^{(F)} - \Delta r_t = \psi_0 + \sum_{i=1}^{3} \psi_i\left(X_{t+1}^{(i)} - \left(1-\kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)}\right) + \psi_3' X_t^{(3)} + \sum_{j=1}^{q} \psi_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t\right) + \sqrt{h_t^{(F)}} Z_{t+1}^{(F)}, \tag{4.12}$$

$$\log h_t^{(F)} = \omega^{(F)} + \alpha^{(F)} Z_t^{(F)} + \gamma^{(F)}\left(|Z_t^{(F)}| - \frac{2}{\sqrt{2\pi}}\right) + \beta^{(F)} \log h_{t-1}^{(F)}, \tag{4.13}$$

where the basis risk innovation process $Z^{(F)} = \{Z_t^{(F)}\}_{t=1}^{T}$ is a standard Gaussian white noise under $\mathbb{P}$, independent of interest rate innovations $Z^{\mathbb{P}}$ and equity innovations $Z_j^{(S)}, j = 1, \ldots, q$. Basis risk volatility parameters $\left(\omega^{(F)}, \alpha^{(F)}, \gamma^{(F)}, \beta^{(F)}\right)$ and linear coefficients $\left(\psi_0, \psi_1, \psi_2, \psi_3, \psi_3', \psi_1^{(S)}, \ldots, \psi_q^{(S)}\right)$ are model parameters to be estimated.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

The risk-neutral dynamics can also be obtained along the lines of Augustyniak *et al.* (2021) by translating basis risk innovations $\{Z_t^{(F)}\}_{t=1}^T$ by some adapted process $\{\lambda_t^{(F)}\}_{t=0}^{T-1}$: under $\mathbb{Q}$, innovations $\{Z_t^{\mathbb{Q}(F)}\}_{t=1}^T \equiv \{Z_t^{(F)} + \lambda_{t-1}^{(F)}\}_{t=1}^T$ form a Gaussian white noise independent of risk-neutral term structure and equity innovations $\{Z_t^{\mathbb{Q}}\}_{t=1}^T$ and $\{Z_{t,j}^{\mathbb{Q}(S)}\}_{t=1}^T$, respectively. The determination of the process $\{\lambda_t^{(F)}\}_{t=0}^{T-1}$ stems from the martingale property that must be satisfied by the mixed fund value process. Its value, which characterizes the risk-neutral dynamics of the mixed fund, is provided in the following proposition whose proof is in Appendix A.

Proposition 4.1. *The risk-neutral dynamics of mixed fund excess returns are given by*

$$R_{t+1}^{(F)} - \Delta r_t = -\frac{1}{2}\left(\sigma_t^{(F)}\right)^2 + \sigma_t^{(F)} \epsilon_{t+1}^{\mathbb{Q}(F)} \tag{4.14}$$

where

$$\left(\sigma_t^{(F)}\right)^2 \equiv \sum_{i=1}^{3}\sum_{l=1}^{3} \psi_i \psi_l \Sigma_{i,i}\Sigma_{l,l}\rho_{i,l} + \sum_{j=1}^{q}\sum_{k=1}^{q} \psi_j^{(S)}\psi_k^{(S)}\Gamma_{j,k}\sqrt{h_{t,j}^{(S)} h_{t,k}^{(S)}} + h_t^{(F)}, \tag{4.15}$$

$$\phi_t \equiv \psi_0 + \sum_{i=1}^{3} \psi_i\left(\kappa_{i,i}^{\mathbb{Q}}\theta_i^{\mathbb{Q}} - \lambda\theta_3^{\mathbb{Q}}\mathbb{1}_{\{i=2\}}\right) + (\psi_2\lambda + \psi_3')X_t^{(3)} - \sum_{j=1}^{q}\psi_j^{(S)}\left(\frac{1}{2}h_{t,j}^{(S)}\right),$$

$$\epsilon_{t+1}^{\mathbb{Q}(F)} \equiv \frac{\sum_{i=1}^{p}\psi_i\Sigma_{i,i}Z_{t+1}^{\mathbb{Q}(i)} + \sum_{j=1}^{q}\psi_j^{(S)}\sqrt{h_{t,j}^{(S)}}Z_{t+1,j}^{\mathbb{Q}(S)} + \sqrt{h_t^{(F)}}Z_{t+1}^{\mathbb{Q}(F)}}{\sigma_t^{(F)}},$$

$$\log h_t^{(F)} = \omega^{(F)} + \alpha^{(F)}\left(Z_t^{\mathbb{Q}(F)} - \lambda_{t-1}^{(F)}\right) + \gamma^{(F)}\left(|Z_t^{\mathbb{Q}(F)} - \lambda_{t-1}^{(F)}| - \frac{2}{\sqrt{2\pi}}\right) + \beta^{(F)}\log h_{t-1}^{(F)},$$

$$\lambda_t^{(F)} \equiv \frac{1}{\sqrt{h_t^{(F)}}}\left[\phi_t + \frac{1}{2}\left(\sigma_t^{(F)}\right)^2\right].$$

and the process $\left\{\epsilon_{t+1}^{\mathbb{Q}(F)}\right\}_{t=1}^T$ *is a sequence of independent standardized Gaussian variables under* $\mathbb{Q}$.

4.4. Estimation of the mixed fund model

The mixed fund model developed above is estimated on real data. The equity model (4.4) and (4.5) is estimated for two equity assets ($q = 2$), namely the *S&P/TSX Composite* and *S&P 500* stock indices. For the mixed fund model (4.12)–(4.13), we consider the *Assumption/Louisbourg Balanced Fund A*.[4] This is a mixed bond and equity fund composed approximately of 39% of Canadian fixed income, 36% of Canadian equity, 15% of US equity, and 10% of other products. Monthly price return data (NAV returns for the mixed fund) in Canadian currency from February 1986 (from February 1996 for the mixed fund) to January 2022 are considered, a span that matches the term structure data used in earlier sections. Note that for the mixed fund, 14 out of the 326 monthly returns are missing values.

Tables 5 and 6 respectively show maximum likelihood estimated parameter and corresponding standard errors for the equity model (4.4)–(4.5) and the mixed fund model (4.12)–(4.13).[5]

[4]https://assumption.lipperweb.com/assumplife/list#FundDetailhttps://assumption.lipperweb.com/assumplife/list#FundDetail.

[5]Monthly periods with a missing value for the mixed fund return are discarded in the likelihood calculation.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table 5. Bivariate EGARCH equity model parameter estimates

Stock index	j	λ_j^S	$\omega_j^{(S)}$	$\alpha_j^{(S)}$	$\gamma_j^{(S)}$	$\beta_j^{(S)}$	$\Gamma_{1,2}$	$\sqrt{12h_{0,j}^{(S)}}$
S&P/TSX	1	0.08443	−2.38375	−0.16171	0.38711	0.62836		14.69%
		(0.04574)	(0.53917)	(0.06079)	(0.09043)	(0.08330)	0.75281	
S&P 500	2	0.12605	−1.92871	−0.14922	0.32486	0.69715	(0.02126)	14.98%
		(0.04390)	(0.61691)	(0.05631)	(0.08286)	(0.09706)		

Notes: Maximum likelihood estimates and standard errors (in parentheses) of the equity model (4.4)–(4.5). Indices $j = 1$ and $j = 2$ are respectively the S&P/TSX Composite and the S&P 500. The estimation is performed on monthly price return time series extending from February 1986 to January 2022 (432 returns by index).

Table 6. Mixed fund model parameter estimates

Parameter	ψ_0	ψ_1	ψ_2	ψ_3	ψ_3'	$\psi_1^{(S)}$
Estimate	0.00404	−1.72446	−0.76950	−0.44537	−0.05524	0.40076
	(0.00123)	(0.40433)	(0.24252)	(0.11151)	(0.01935)	(0.01464)
Parameter	$\psi_2^{(S)}$	$\omega^{(F)}$	$\alpha^{(F)}$	$\gamma^{(F)}$	$\beta^{(F)}$	$\sqrt{12h_0^{(F)}}$
Estimate	0.13458	−1.03122	−0.09390	0.06653	0.89453	2.61%
	(0.01390)	(0.01238)	(0.03527)	(0.02091)	(0.00028)	

Notes: Maximum likelihood estimates and standard errors (in parentheses) of the mixed fund model (4.12)–(4.13). The estimation is performed on a monthly NAV return time series for the Assumption/Louisbourg Balanced Fund A extending from February 1996 to January 2022 (312 returns).

Note that initial variances $h_{0,j}^{(S)}$, $j = 1, 2$ and $h_0^{(F)}$ are also considered as parameters to optimize during the estimation. Equity indices are highly correlated with $\Gamma_{1,2} = 0.753$. Furthermore, the persistence of the volatility for both indices is high and similar, with $\beta_1^{(S)} = 0.628$ for the S&P TSX Composite and $\beta_2^{(S)} = 0.697$ the S&P 500. Furthermore, the presence of a leverage effect is implied by negative values of $\alpha_j^{(S)}$, $j = 1, 2$. For the mixed fund model, all parameter estimates are significant at the 99% confidence level. Negative values for estimates of coefficient ψ_i, $i = 1, 2$ reflect negative correlation between variations of the short rate and returns of the mixed fund. This is expected due to the fund being partially invested in fixed income, where bond prices are inversely related to interest rates. Interestingly, the estimate of coefficient ψ_3', a parameter not present in the Augustyniak *et al.* (2021) model, is also significant and negative. It implies that mixed fund returns are negatively related to the humpiness of the term structure (and not only to changes in humpiness).

The performance of our mixed fund model (4.12)–(4.13) is compared to that of Augustyniak *et al.* (2021), which basically imposes the constraint $\psi_3' = 0$ and uses the DG3 instead of the DTAFNS model for the computation of smoothed values for term structure factors X. Our model has an in-sample log-likelihood of 1038.719, whereas that of the Augustyniak *et al.* (2021) benchmark is 1035.714. Such reported figures are conditional log-likelihood for the mixed fund returns specification given smoothed values of the term structure factors for each respective models and equity index returns. Since the use of our mixed fund model does not lead to a deterioration in fitting performance, and since such model is compatible with the DTAFNS model which has arguably more favorable properties than the DG3 counterpart, the use of our mixed fund model shall be preferred over that of Augustyniak *et al.* (2021).

5. Conclusion

This paper develops a discrete-time version of the arbitrage-free Nelson-Siegel model for dynamic term structure modeling which is slightly different than but similar to that in Hong *et al.* (2019).

A closed-form solution for the price of risk-free zero-coupon bonds is obtained, which makes it possible to devise a convenient Kalman filter-based joint estimation procedure for physical and risk-neutral dynamics of factors driving the term structure. The estimation of the model on historical data for the Canadian spot curve reveal that the model accurately represents term structure movements and possesses superior distributional predictive power in comparison to some version of the dynamic Nelson-Siegel model and to a discrete-time G3 model previously proposed by Augustyniak *et al.* (2021), at least for the considered dataset. The discrete-time arbitrage-free Nelson-Siegel model also has two other advantages over the latter benchmark: it tends to produce less frequent negative short rates and it provides better interpretability for the three factors underlying the model. Finally, the mixed fund model of Augustyniak *et al.* (2021) characterizing the dynamics of a mutual fund invested in both equity and fixed income is adapted to integrate the discrete-time arbitrage-free Nelson-Siegel model as its component to depict interest rates dynamics.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S1748499524000010

Acknowledgements. We thank the anonymous referee for their valuable feedback which helped improving the paper.

Data availability statement. The R code allowing for the replication of our results is made available in the following GitHub repository: https://github.com/FredericGodinQC/DTAFNS_model.

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

Competing interests. We have no conflict of interest or competing interests to declare in relation to this project.

References

Augustyniak, M., Godin, F. & Hamel, E. (2021). A mixed bond and equity fund model for the valuation of variable annuities. *ASTIN Bulletin: The Journal of the IAA*, **51**(1), 131–159.

Bolder, D., Metzler, A. & Johnson, G. (2004). An empirical analysis of the Canadian term structure of zero-coupon interest rates. Staff Working Paper 2004-48, Bank of Canada.

Buehler, H., Gonon, L., Teichmann, J. & Wood, B. (2019). Deep hedging. *Quantitative Finance*, **19**(8), 1271–1291.

Chen, R.-R. & Scott, L. (1993). Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *The Journal of Fixed Income*, **3**(3), 14–31.

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.

Dai, Q. & Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, **55**(5), 1943–1978.

Diebold, F. X. and **Li, C.** (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, **130**(2), 337–364.

Duan, J.-C. & Simonato, J.-G. (1999). Estimating and testing exponential-affine term structure models by Kalman filter. *Review of Quantitative Finance and Accounting*, **13**(2), 111–135.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, **57**(1), 405–443.

Duffie, D. & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, **6**(4), 379–406.

Filipović, D. (1999). A note on the Nelson–Siegel family. *Mathematical Finance*, **9**(4), 349–359.

Ghalanos, A. & Theussl, S. (2015). *Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method*. R package version 1.16.

Hong, Z., Niu, L. & Zeng, G. (2019). US and Chinese yield curve responses to RMB exchange rate policy shocks: An analysis with the arbitrage-free Nelson-Siegel term structure model. *China Finance Review International*, **9**(3), 360–385.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Transactions of the ASME–Journal of Basic Engineering, (Series D)(82), 35–45.

Lemke, W. (2006). *Term structure modeling and estimation in a state space framework* (Vol. **565**). Springer Science & Business Media.

Nelson, C. R. & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business*, **60**(4), 473–489.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, **59**(2), 347–370.
Park, H. (2014). Estimation of affine term structure models under the Milstein approximation. *Applied Economics Letters*, **21**(9), 651–656.
Rémillard, B. (2013). *Statistical methods for financial engineering*. CRC Press.
Shumway, R. H. & **Stoffer, D. S.** (2017). Time series analysis and its applications. With R examples (Vol. 4). Springer Texts in Statistics.
Wüthrich, M. V. & **Merz, M.** (2013). *Financial modeling, actuarial valuation and solvency in insurance*. Springer.
Ye, Y. (1987). *Interior algorithms for linear, quadratic, and linearly constrained non-linear programming*. PhD Thesis, Department of ESS, Stanford University.

Appendix A. Proofs

In this section, proofs to several results mentioned in the body of the manuscript are provided.

To prove Proposition 2.1, the following lemma is needed.

Lemma A.1. *Assume* (2.10) *holds. For* $t = 0, \ldots, T$ *and* $n = 0, \ldots, T - t$,

$$X^{(i)}_{t+n} = X^{(i)}_t \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^n + \kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i \sum_{l=1}^{n} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n-l)} + \Sigma_{i,i} \sum_{l=1}^{n} Z^{\mathbb{Q}}_{t+l,i} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n-l)}$$
$$+ \sum_{l=1}^{n} \sum_{j\neq i}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+l-1}\right) \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n-l)}. \tag{A.1}$$

Proof of Lemma A.1: The proof relies on induction. The case $n = 0$ is trivial when using the convention $\sum_{l=1}^{0} x_l \equiv 0$ for any sequence of real numbers $\{x_l\}$.

Then, assume (A.1) holds for some $n < T - t - 1$. Using (2.10), for any $i = 1, 2, 3$,

$$X^{(i)}_{t+n+1} = X^{(i)}_{t+n} + \sum_{j=1}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+n}\right) + \Sigma_{i,i} Z^{\mathbb{Q}}_{t+n+1,i}$$
$$= X^{(i)}_{t+n} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right) + \kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i + \sum_{j\neq i}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+n}\right) + \Sigma_{i,i} Z^{\mathbb{Q}}_{t+n+1,i}.$$

Applying (A.1) in the latter equality yields

$$X^{(i)}_{t+n+1} = X^{(i)}_t \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{n+1} + \kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i \sum_{l=1}^{n} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)} + \Sigma_{i,i} \sum_{l=1}^{n} Z^{\mathbb{Q}}_{t+l,i} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)}$$
$$+ \sum_{l=1}^{n} \sum_{j\neq i}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+l-1}\right) \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)} + \kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i + \sum_{j\neq 1}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+n}\right)$$
$$+ \Sigma_{i,i} Z^{\mathbb{Q}}_{t+n+1,i}$$
$$= X^{(i)}_t \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{n+1} + \kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i \sum_{l=1}^{n+1} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)} + \Sigma_{i,i} \sum_{l=1}^{n+1} Z^{\mathbb{Q}}_{t+l,i} \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)}$$
$$+ \sum_{l=1}^{n+1} \sum_{j\neq i}^{3} \kappa^{\mathbb{Q}}_{i,j} \left(\theta^{\mathbb{Q}}_j - X^{(j)}_{t+l-1}\right) \left(1 - \kappa^{\mathbb{Q}}_{i,i}\right)^{(n+1-l)},$$

hence completing the induction. □

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Remark A.1. The result of Lemma A.1 is also valid under the $\mathbb{P}$-measure, i.e. when replacing all $\mathbb{Q}$'s with $\mathbb{P}$'s in the statement of the lemma and the proof.

Proof of Lemma 2.1: Equation (2.11) is trivial. To show (2.12),

$$\begin{aligned}
\sum_{u=1}^{\tau-1} ur^u &= \sum_{u=1}^{\tau-1} r\frac{d}{dr}r^u \\
&= r\frac{d}{dr}\sum_{u=1}^{\tau-1} r^u \\
&= r\frac{(1-\tau r^{\tau-1})(1-r)+(r-r^\tau)}{(1-r)^2} \\
&= r\frac{1-\tau r^{\tau-1}+\tau r^\tau - r^\tau}{(1-r)^2} \\
&= \frac{r-\tau r^\tau+(\tau-1)r^{\tau+1}}{(1-r)^2}.
\end{aligned}$$

Finally, to show (2.13),

$$\begin{aligned}
\sum_{u=1}^{\tau-1} u^2r^u &= r\sum_{u=1}^{\tau-1} u\frac{d}{dr}r^u \\
&= r\frac{d}{dr}\sum_{u=1}^{\tau-1} ur^u \\
&= r\frac{d}{dr}\frac{r-\tau r^\tau+(\tau-1)r^{\tau+1}}{(1-r)^2} \\
&= r\frac{(1-r)^2[1-\tau^2r^{\tau-1}+(\tau+1)(\tau-1)r^\tau]+2(1-r)[r-\tau r^\tau+(\tau-1)r^{\tau+1}]}{(1-r)^4} \\
&= \frac{[r-r^2-\tau^2r^\tau+\tau^2r^{\tau+1}+(\tau+1)(\tau-1)r^{\tau+1}-(\tau+1)(\tau-1)r^{\tau+2}]+2[r^2-\tau r^{\tau+1}+(\tau-1)r^{\tau+2}]}{(1-r)^3} \\
&= \frac{-(\tau-1)^2r^{\tau+2}+(2\tau^2-2\tau-1)r^{\tau+1}-\tau^2r^\tau+r^2+r}{(1-r)^3}.
\end{aligned}$$

□

Proof of Proposition 2.1: From (2.7), for any $\tau = 1, \ldots, T-t$,

$$\sum_{s=0}^{\tau-1} r_{t+s} = \sum_{s=0}^{\tau-1} X^{(1)}_{t+s} + \sum_{s=0}^{\tau-1} X^{(2)}_{t+s}.$$

Furthermore, since $\kappa^{\mathbb{Q}}_{1,j} = 0$ for $j = 1, \ldots, 3$, Equation (A.1) leads to

$$\begin{aligned}
\sum_{s=0}^{\tau-1} X^{(1)}_{t+s} &= \sum_{s=0}^{\tau-1}\left[X^{(1)}_t + \Sigma_{1,1}\sum_{l=1}^{s} Z^{\mathbb{Q}}_{t+l,1}\right] \\
&= \tau X^{(1)}_t + \Sigma_{1,1}\sum_{s=1}^{\tau-1}\sum_{l=1}^{\tau-1} \mathbb{1}_{\{l\le s\}} Z^{\mathbb{Q}}_{t+l,1}
\end{aligned}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

$$
\begin{aligned}
&= \tau X_t^{(1)} + \Sigma_{1,1} \sum_{l=1}^{\tau-1} \sum_{s=l}^{\tau-1} Z_{t+l,1}^{\mathbb{Q}} \\
&= \tau X_t^{(1)} + \Sigma_{1,1} \sum_{l=1}^{\tau-1} (\tau - l) Z_{t+l,1}^{\mathbb{Q}}.
\end{aligned} \tag{A.2}
$$

Moreover, relying again on Equation (A.1) and recalling that $\kappa_{2,1}^{\mathbb{Q}} = 0$, $\kappa_{2,2}^{\mathbb{Q}} = \lambda$ and $\kappa_{2,3}^{\mathbb{Q}} = -\lambda$,

$$
\begin{aligned}
\sum_{s=0}^{\tau-1} X_{t+s}^{(2)} &= \sum_{s=0}^{\tau-1} \Bigg\{ X_t^{(2)} (1-\lambda)^s + \lambda \theta_2^{\mathbb{Q}} \sum_{l=1}^{s} (1-\lambda)^{s-l} + \Sigma_{2,2} \sum_{l=1}^{s} Z_{t+l,2}^{\mathbb{Q}} (1-\lambda)^{s-l} \\
&\quad - \lambda \sum_{l=1}^{s} (1-\lambda)^{s-l} \theta_3^{\mathbb{Q}} + \lambda \sum_{l=1}^{s} (1-\lambda)^{s-l} X_{t+l-1}^{(3)} \Bigg\} \\
&= X_t^{(2)} \sum_{s=0}^{\tau-1} (1-\lambda)^s + \lambda \theta_2^{\mathbb{Q}} \sum_{s=1}^{\tau-1} \sum_{l=1}^{s} (1-\lambda)^{s-l} + \Sigma_{2,2} \sum_{s=1}^{\tau-1} \sum_{l=1}^{s} (1-\lambda)^{s-l} Z_{t+l,2}^{\mathbb{Q}} \\
&\quad - \lambda \theta_3^{\mathbb{Q}} \sum_{s=1}^{\tau-1} \sum_{l=1}^{s} (1-\lambda)^{s-l} + \lambda \sum_{s=1}^{\tau-1} \sum_{l=1}^{s} (1-\lambda)^{s-l} X_{t+l-1}^{(3)} \\
&= X_t^{(2)} \sum_{s=0}^{\tau-1} (1-\lambda)^s + \lambda \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}} \right) \sum_{s=1}^{\tau-1} \sum_{l=0}^{s-1} (1-\lambda)^l \\
&\quad + \sum_{s=1}^{\tau-1} \sum_{l=1}^{\tau-1} \mathbb{1}_{\{l \le s\}} (1-\lambda)^{s-l} \left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)} \right) \\
&= X_t^{(2)} \sum_{s=0}^{\tau-1} (1-\lambda)^s + \lambda \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}} \right) \sum_{s=1}^{\tau-1} \frac{1-(1-\lambda)^s}{\lambda} \\
&\quad + \sum_{l=1}^{\tau-1} \sum_{s=1}^{\tau-1} \mathbb{1}_{\{l \le s\}} (1-\lambda)^{s-l} \left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)} \right) \\
&= X_t^{(2)} \frac{1-(1-\lambda)^\tau}{\lambda} + \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}} \right) \left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^\tau}{\lambda} \right) \\
&\quad + \sum_{l=1}^{\tau-1} \left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)} \right) \sum_{s=1}^{\tau-1} \mathbb{1}_{\{l \le s\}} (1-\lambda)^{s-l}.
\end{aligned} \tag{A.3}
$$

Note that

$$
\sum_{s=1}^{\tau-1} \mathbb{1}_{\{l \le s\}} (1-\lambda)^{s-l} = \sum_{s=l}^{\tau-1} (1-\lambda)^{s-l} = \sum_{s=0}^{\tau-1-l} (1-\lambda)^s = \frac{1-(1-\lambda)^{\tau-l}}{\lambda}. \tag{A.4}
$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Placing (A.4) in (A.3) yields

$$\sum_{s=0}^{\tau-1} X_{t+s}^{(2)} = X_t^{(2)} \frac{1-(1-\lambda)^{\tau}}{\lambda} + \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}\right)\left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^{\tau}}{\lambda}\right)$$
$$+ \Sigma_{2,2}\sum_{l=1}^{\tau-1} Z_{t+l,2}^{\mathbb{Q}} \frac{1-(1-\lambda)^{\tau-l}}{\lambda} + \sum_{l=1}^{\tau-1}\left[1-(1-\lambda)^{\tau-l}\right] X_{t+l-1}^{(3)}. \tag{A.5}$$

Additionally, combining (A.1) with $\kappa_{3,1}^{\mathbb{Q}} = \kappa_{3,2}^{\mathbb{Q}} = 0$ and $\kappa_{3,3}^{\mathbb{Q}} = \lambda$ leads to

$$\sum_{l=1}^{\tau-1}\left[1-(1-\lambda)^{\tau-l}\right] X_{t+l-1}^{(3)}$$
$$= \sum_{l=0}^{\tau-2}\left[1-(1-\lambda)^{\tau-l-1}\right] X_{t+l}^{(3)}$$
$$= \sum_{l=0}^{\tau-2}\left[1-(1-\lambda)^{\tau-l-1}\right]\left[X_t^{(3)}(1-\lambda)^{l} + \lambda\theta_3^{\mathbb{Q}}\sum_{k=1}^{l}(1-\lambda)^{(l-k)} + \Sigma_{3,3}\sum_{k=1}^{l} Z_{t+k,3}^{\mathbb{Q}}(1-\lambda)^{(l-k)}\right]$$
$$= X_t^{(3)}\left(\sum_{l=0}^{\tau-2}(1-\lambda)^{l} - \sum_{l=0}^{\tau-2}(1-\lambda)^{\tau-1}\right) + \lambda\theta_3^{\mathbb{Q}}\sum_{l=0}^{\tau-2}\frac{1-(1-\lambda)^{l}-(1-\lambda)^{\tau-l-1}+(1-\lambda)^{\tau-1}}{\lambda}$$
$$+ \Sigma_{3,3}\sum_{l=0}^{\tau-2}\sum_{k=1}^{\tau-2}\mathbb{1}_{\{k\le l\}} Z_{t+k,3}^{\mathbb{Q}}\left[(1-\lambda)^{(l-k)} - (1-\lambda)^{\tau-k-1}\right]$$
$$= X_t^{(3)}\left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}\right)$$
$$+ \theta_3^{\mathbb{Q}}\sum_{l=0}^{\tau-2}\left(1-(1-\lambda)^{l}-(1-\lambda)^{\tau-l-1}+(1-\lambda)^{\tau-1}\right)$$
$$+ \Sigma_{3,3}\sum_{k=1}^{\tau-2} Z_{t+k,3}^{\mathbb{Q}}\sum_{l=k}^{\tau-2}\left[(1-\lambda)^{(l-k)} - (1-\lambda)^{\tau-k-1}\right]$$
$$= X_t^{(3)}\left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}\right) + \theta_3^{\mathbb{Q}}\left((\tau-1)\left(1+(1-\lambda)^{\tau-1}\right)\right.$$
$$\left. - \frac{2-(1-\lambda)^{\tau-1}-\lambda-(1-\lambda)^{\tau}}{\lambda}\right)$$
$$+ \Sigma_{3,3}\sum_{k=1}^{\tau-2} Z_{t+k,3}^{\mathbb{Q}}\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]. \tag{A.6}$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Placing (A.6) in (A.5) gives

$$\begin{aligned}\sum_{s=0}^{\tau-1} X_{t+s}^{(2)} =& X_t^{(2)} \frac{1-(1-\lambda)^{\tau}}{\lambda} + X_t^{(3)} \left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1} \right) \\ & + \theta_2^{\mathbb{Q}} \left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^{\tau}}{\lambda} \right) + \theta_3^{\mathbb{Q}} \left((\tau-1)(1-\lambda)^{\tau-1} - \frac{1-(1-\lambda)^{\tau-1}}{\lambda} \right) \\ & + \Sigma_{2,2} \sum_{l=1}^{\tau-1} Z_{t+l,2}^{\mathbb{Q}} \frac{1-(1-\lambda)^{\tau-l}}{\lambda} \\ & + \Sigma_{3,3} \sum_{k=1}^{\tau-2} Z_{t+k,3}^{\mathbb{Q}} \left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k) \right]. \end{aligned} \tag{A.7}$$

Define $R_{t,T} \equiv \Delta \sum_{j=t}^{T-1} r_j = \Delta \sum_{j=0}^{T-1} X_{t+j}^{(1)} + X_{t+j}^{(2)}$. Define also

$$B_\tau^{(1)} = \tau, \quad B_\tau^{(2)} = \frac{1-(1-\lambda)^{\tau}}{\lambda}, \quad B_\tau^{(3)} = \frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}.$$

Based on (A.2) and (A.7), given $\mathcal{F}_t$, $-R_{t,T}$ has a Gaussian distribution with mean

$$\begin{aligned} & -\Delta \sum_{j=1}^{3} B_\tau^{(j)} X_t^{(j)} - \Delta\theta_2^{\mathbb{Q}} \left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^{\tau}}{\lambda} \right) \\ & - \Delta\theta_3^{\mathbb{Q}} \left((\tau-1)(1-\lambda)^{\tau-1} - \frac{1-(1-\lambda)^{\tau-1}}{\lambda} \right) \end{aligned}$$

and variance $\Delta^2 \upsilon_\tau$ where

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

with

$$\upsilon_\tau^{(1,1)} = \mathrm{Var}^{\mathbb{Q}} \left(\Sigma_{1,1} \sum_{l=1}^{\tau-1} (\tau-l) Z_{t+l,1}^{\mathbb{Q}} \right) = \Sigma_{1,1}^2 \sum_{l=1}^{\tau-1} (\tau-l)^2 = \Sigma_{1,1}^2 \sum_{l=1}^{\tau-1} l^2 = \Sigma_{1,1}^2 \frac{\tau(\tau-1)(2\tau-1)}{6},$$

$$\begin{aligned}\upsilon_\tau^{(2,2)} &= \mathrm{Var}^{\mathbb{Q}} \left(\Sigma_{2,2} \sum_{l=1}^{\tau-1} Z_{t+l,2}^{\mathbb{Q}} \frac{1-(1-\lambda)^{\tau-l}}{\lambda} \right) \\ &= \frac{\Sigma_{2,2}^2}{\lambda^2} \sum_{l=1}^{\tau-1} \left(1-(1-\lambda)^{l} \right)^2 \\ &= \frac{\Sigma_{2,2}^2}{\lambda^2} \left(\tau - 1 - 2\left[\frac{1-(1-\lambda)^{\tau}}{\lambda} - 1 \right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2} - 1 \right) \\ &= \frac{\Sigma_{2,2}^2}{\lambda^2} \left(\tau - 2\left[\frac{1-(1-\lambda)^{\tau}}{\lambda} \right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2} \right), \end{aligned}$$

$$\upsilon_\tau^{(3,3)} = \mathrm{Var}^{\mathbb{Q}}\left(\Sigma_{3,3}\sum_{k=1}^{\tau-2} Z^{\mathbb{Q}}_{t+k,3}\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]\right)$$

$$= \Sigma_{3,3}^2\sum_{k=1}^{\tau-2}\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]^2$$

$$= \frac{\Sigma_{3,3}^2}{\lambda^2}\sum_{k=1}^{\tau-2}\left[1-(1-\lambda)^k - \lambda(1-\lambda)^k k\right]^2$$

$$= \frac{\Sigma_{3,3}^2}{\lambda^2}\sum_{k=1}^{\tau-2}\left[1+(1-\lambda)^{2k} + \lambda^2(1-\lambda)^{2k}k^2 - 2(1-\lambda)^k - 2k\lambda(1-\lambda)^k + 2\lambda k(1-\lambda)^{2k}\right]$$

$$= \mathbb{1}_{\{\tau>1\}}\frac{\Sigma_{3,3}^2}{\lambda^2}\Bigg[\tau-2+\zeta_0\big((1-\lambda)^2,\tau-1\big) + \lambda^2\zeta_2\big((1-\lambda)^2,\tau-1\big)$$

$$- 2\zeta_0((1-\lambda),\tau-1) - 2\lambda\zeta_1((1-\lambda),\tau-1) + 2\lambda\zeta_1\big((1-\lambda)^2,\tau-1\big)\Bigg]$$

and

$$\upsilon_\tau^{(1,2)} = \upsilon_\tau^{(2,1)} = \mathrm{Cov}^{\mathbb{Q}}\left(\Sigma_{1,1}\sum_{l=1}^{\tau-1}(\tau-l)Z^{\mathbb{Q}}_{t+l,1},\ \Sigma_{2,2}\sum_{l=1}^{\tau-1}Z^{\mathbb{Q}}_{t+l,2}\frac{1-(1-\lambda)^{\tau-l}}{\lambda}\right)$$

$$= \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\sum_{l=1}^{\tau-1} l\frac{1-(1-\lambda)^l}{\lambda}$$

$$= \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda),\tau)\right),$$

$$\upsilon_\tau^{(1,3)} = \upsilon_\tau^{(3,1)} = \mathrm{Cov}^{\mathbb{Q}}\Bigg(\Sigma_{1,1}\sum_{k=1}^{\tau-1}(\tau-k)Z^{\mathbb{Q}}_{t+k,1},\ \Sigma_{3,3}\sum_{k=1}^{\tau-2}Z^{\mathbb{Q}}_{t+k,3}\Bigg[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda}$$

$$-(1-\lambda)^{\tau-k-1}(\tau-1-k)\Bigg]\Bigg)$$

$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\sum_{k=1}^{\tau-2}(\tau-k)\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]$$

$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\sum_{k=1}^{\tau-2}(k+1)\left[\frac{1-(1-\lambda)^k}{\lambda} - (1-\lambda)^k k\right]$$

$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\sum_{k=1}^{\tau-2}(k+1)\left[1-(1+\lambda k)(1-\lambda)^k\right]$$

$$= \mathbb{1}_{\{\tau>1\}}\rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\left(\frac{(\tau-1)\tau}{2} - 1 - \sum_{k=1}^{\tau-2}(1+(\lambda+1)k+\lambda k^2)(1-\lambda)^k\right)$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

$$= \mathbb{1}_{\{\tau>1\}}\rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\left[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0((1-\lambda),\tau-1) - (\lambda+1)\zeta_1((1-\lambda),\tau-1)\right.$$
$$\left. - \lambda\zeta_2((1-\lambda),\tau-1)\right],$$

$$\upsilon_\tau^{(2,3)} = \upsilon_\tau^{(3,2)} = \mathrm{Cov}^{\mathbb{Q}}\left(\Sigma_{2,2}\sum_{l=1}^{\tau-1} Z^{\mathbb{Q}}_{t+l,2}\frac{1-(1-\lambda)^{\tau-l}}{\lambda}, \Sigma_{3,3}\sum_{k=1}^{\tau-2} Z^{\mathbb{Q}}_{t+k,3}\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda}\right.\right.$$
$$\left.\left. - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]\right)$$
$$= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\sum_{k=1}^{\tau-2}\left(\frac{1-(1-\lambda)^{\tau-k}}{\lambda}\right)\left[\frac{1-(1-\lambda)^{\tau-k-1}}{\lambda}\right.$$
$$\left. -(1-\lambda)^{\tau-k-1}(\tau-k-1)\right]$$
$$= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\sum_{k=1}^{\tau-2}\left(\frac{1-(1-\lambda)^{k+1}}{\lambda}\right)\left[\frac{1-(1-\lambda)^{k}}{\lambda} - (1-\lambda)^k k\right]$$
$$= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\sum_{k=1}^{\tau-2}\left(\frac{1-(2-\lambda)(1-\lambda)^k + (1-\lambda)(1-\lambda)^{2k}}{\lambda^2}\right.$$
$$\left. + \frac{-(1-\lambda)^k k + (1-\lambda)(1-\lambda)^{2k}k}{\lambda}\right)$$
$$= \mathbb{1}_{\{\tau>1\}}\rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\left(\frac{\tau-2-(2-\lambda)\zeta_0((1-\lambda),\tau-1) + (1-\lambda)\zeta_0\big((1-\lambda)^2,\tau-1\big)}{\lambda^2}\right.$$
$$\left. + \frac{-\zeta_1((1-\lambda),\tau-1) + (1-\lambda)\zeta_1\big((1-\lambda)^2,\tau-1\big)}{\lambda}\right).$$

□

Note that in the final expression of $\upsilon_\tau^{(j,3)}$, $j = 1,2,3$, the indicator $\mathbb{1}_{\{\tau>1\}}$ could have been replaced with $\mathbb{1}_{\{\tau>2\}}$ since $\upsilon_2^{(j,3)} = 0$.

Lemma A.2. *Consider assumptions of Proposition 2.1. Set* $B_0^{(1)} = B_0^{(2)} = B_0^{(3)} = 0$. *Then, for* $i = 1,2,3$ *and any integer* $\tau \geq 1$,

$$B_\tau^{(i)} - 1 = B_{\tau-1}^{(i)}\left(1-\kappa^{\mathbb{Q}}_{i,i}\right) - \mathbb{1}_{\{i=3\}}(1-\lambda)^{\tau-1} = B_{\tau-1}^{(i)}\left(1-\lambda\mathbb{1}_{\{i>1\}}\right) - \mathbb{1}_{\{i=3\}}(1-\lambda)^{\tau-1}.$$

Proof of Lemma A.2: Based on Proposition 2.1, the case $\tau = 1$ is trivial. Now consider $\tau > 1$.

$$B_\tau^{(1)} - 1 = \tau - 1 = B_{\tau-1}^{(1)}.$$

Furthermore,

$$\frac{B_\tau^{(2)}-1}{B_{\tau-1}^{(2)}} = \frac{\frac{1-(1-\lambda)^\tau}{\lambda} - 1}{\frac{1-(1-\lambda)^{\tau-1}}{\lambda}} = \frac{(1-\lambda)\frac{1-(1-\lambda)^{\tau-1}}{\lambda}}{\frac{1-(1-\lambda)^{\tau-1}}{\lambda}} = 1-\lambda.$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Finally,

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

□

Proof of Proposition 4.1: Using (2.10), (2.14), (4.6) and (4.12),

$$\begin{aligned}
R_{t+1}^{(F)} - \Delta r_t &= \psi_0 + \sum_{i=1}^{3} \psi_i \left(X_{t+1}^{(i)} - \left(1-\kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)} \right) + \psi_3' X_t^{(3)} \\
&\quad + \sum_{j=1}^{q} \psi_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t \right) + \sqrt{h_t^{(F)}} Z_{t+1}^{(F)} \\
&= \psi_0 + \sum_{i=1}^{3} \psi_i \left(\Sigma_{i,i} Z_{t+1}^{\mathbb{Q}(i)} + \lambda \theta_i^{\mathbb{Q}} \mathbb{1}_{\{i>1\}} - \lambda \theta_3^{\mathbb{Q}} \mathbb{1}_{\{i=2\}} + \lambda \mathbb{1}_{\{i=2\}} X_t^{(3)} \right) + \psi_3' X_t^{(3)} \\
&\quad + \sum_{j=1}^{q} \psi_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t \right) + \sqrt{h_t^{(F)}} Z_{t+1}^{(F)} \\
&= \underbrace{\psi_0 + \sum_{i=1}^{3} \psi_i \left(\kappa_{i,i}^{\mathbb{Q}} \theta_i^{\mathbb{Q}} - \lambda \theta_3^{\mathbb{Q}} \mathbb{1}_{\{i=2\}} \right) + (\psi_2 \lambda + \psi_3') X_t^{(3)} - \sum_{j=1}^{q} \psi_j^{(S)} \left(\frac{1}{2} h_{t,j}^{(S)} \right)}_{=\phi_t} - \sqrt{h_t^{(F)}} \lambda_t^{(F)} \\
&\quad + \sum_{i=1}^{3} \psi_i \Sigma_{i,i} Z_{t+1}^{\mathbb{Q}(i)} + \sum_{j=1}^{q} \psi_j^{(S)} \sqrt{h_{t,j}^{(S)}} Z_{t+1,j}^{\mathbb{Q}(S)} + \sqrt{h_t^{(F)}} Z_{t+1}^{\mathbb{Q}(F)}.
\end{aligned}$$

Conditionally upon $\mathcal{F}_t$, $R_{t+1}^{(F)} - \Delta r_t$ is therefore Gaussian with variance

$$\mathrm{Var}^{\mathbb{Q}}\left[R_{t+1}^{(F)} - \Delta r_t \middle| \mathcal{F}_t \right] = \left(\sigma_t^{(F)} \right)^2 \tag{A.8}$$

with $\sigma_t^{(F)}$ given by (4.15). Enforcing that the discounted mixed fund value is a $\mathbb{Q}$-martingale requires

$$\begin{aligned}
1 &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(R_{t+1}^{(F)} - \Delta r_t \right) | \mathcal{F}_t \right] \\
&= \exp\left(\phi_t - \sqrt{h_t^{(F)}} \lambda_t^{(F)} + \frac{1}{2} \left(\sigma_t^{(F)} \right)^2 \right)
\end{aligned}$$

which leads to

$$\lambda_t^{(F)} = \frac{1}{\sqrt{h_t^{(F)}}} \left[\phi_t + \frac{1}{2} \left(\sigma_t^{(F)} \right)^2 \right].$$

Finally, due to (A.8), the innovation

$$\epsilon_{t+1}^{\mathbb{Q}(F)} = \frac{\sum_{i=1}^{3} \psi_i \Sigma_{i,i} Z_{t+1}^{\mathbb{Q}(i)} + \sum_{j=1}^{q} \psi_j^{(S)} \sqrt{h_{t,j}^{(S)}} Z_{t+1,j}^{\mathbb{Q}(S)} + \sqrt{h_t^{(F)}} Z_{t+1}^{\mathbb{Q}(F)}}{\sigma_t^{(F)}}$$

is standard Gaussian under $\mathbb{Q}$, and thus $\{\epsilon_{t+1}^{\mathbb{Q}(F)}\}_{t=1}^{T}$ is a standard Gaussian white noise. □

Appendix B. Benchmarks

In this paper, the discrete-time Gaussian three-factor model introduced by Augustyniak *et al.* (2021) and the dynamic Nelson-Siegel model from Diebold & Li (2006) are considered as benchmarks for the DTAFNS model. Such benchmarks are presented below.

B.1. Discrete-time Gaussian three-factor model

Augustyniak *et al.* (2021) assume that the physical measure $\mathbb{P}$ dynamics of the short rate are that of a discrete-time version of the three-factor Gaussian G3 model:

$$\begin{aligned} r_t &= X_t^{(1)} + X_t^{(2)} + X_t^{(3)}, \\ X_{t+1}^{(i)} &= X_t^{(i)} + \kappa_i \left(\mu_i - X_t^{(i)}\right) + \sigma_i Z_{t+1}^{(i)}, \qquad i = 1, 2, 3, \end{aligned} \tag{B.1}$$

where $(\kappa_i, \mu_i, \sigma_i)$ are the model parameters and $\left\{Z_t^{(1)}, Z_t^{(2)}, Z_t^{(3)}\right\}_{t=1}^{T}$ is a Gaussian white noise process with contemporaneous correlation matrix Γ.

The risk-neutral dynamics of the factors is obtained using a discrete-time version of the Girsanov theorem. The processes $Z_i^{\mathbb{Q}} = \left\{Z_{t,i}^{\mathbb{Q}}\right\}_{t=1}^{T}$, $i = 1, 2, 3$ defined through $Z_{t+1,i}^{\mathbb{Q}} = Z_{t+1,i} - \lambda_i X_t^{(i)}$, with $\lambda_i \in \mathbb{R}$, are standard Gaussian white noises under the risk-neutral measure $\mathbb{Q}$, still with contemporaneous correlation matrix Γ. Therefore,

$$X_{t+1}^{(i)} = X_t^{(i)} + \kappa_i^{\mathbb{Q}} \left(\mu_i^{\mathbb{Q}} - X_t\right) + \sigma_i Z_{t+1}^{\mathbb{Q}}, \quad i = 1, 2, 3 \tag{B.2}$$

where

$$\kappa_i^{\mathbb{Q}} = \kappa_i - \sigma_i \lambda_i, \qquad \mu_i^{\mathbb{Q}} = \frac{\kappa_i \mu_i}{\kappa_i - \sigma_i \lambda_i}.$$

Augustyniak *et al.* (2021) show that under (B.2), the time-t price of a risk-free zero-coupon bond paying one dollar at the maturity time T is given by

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left[\exp\left(-\Delta \sum_{j=t}^{T-t-1} r_{t-j} \right) \right] = \exp\left(A_\tau - \Delta \sum_{i=1}^{3} B_\tau^{(i)} X_t^{(i)} \right),$$

with $\tau = T - t$ and

$$A_\tau = \frac{\Delta^2}{2} \mathbb{1}_3^\top \upsilon_\tau \mathbb{1}_3 - \Delta \sum_{i=1}^{3} \mu_i^{\mathbb{Q}} \left(\tau - B_\tau^{(i)}\right),$$

$$B_\tau^{(i)} = \frac{1 - \left(1 - \kappa_i^{\mathbb{Q}}\right)^\tau}{\kappa_i^{\mathbb{Q}}}, \quad i = 1, 2, 3,$$

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table B.1. DG3 model parameter estimates

i	κ_i	μ_i	σ_i	λ_i
1	0.00523	0.01780	0.00538	0.72614
2	0.04409	−0.00323	0.00489	−4.76851
3	0.02063	0.05016	0.00810	0.96129

$$\Gamma = \begin{bmatrix} 1 & 0.146 & -0.785 \\ 0.146 & 1 & -0.569 \\ -0.785 & -0.569 & 1 \end{bmatrix}$$

Notes: Parameter estimates for the discrete-time G3 model of Augustyniak *et al.* (2021) given by Equations (B.1)-(B.2). The maximum likelihood estimation procedure is conducted with a Kalman filter on the Canadian end-of-month yield curve data extending from January 1986 to January 2022.

Table B.2. Dynamic Nelson-Siegel model parameters estimates

i	$\kappa^{\mathbb{P}}_{i,i}$	γ_i	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_i$	$\theta^{\mathbb{Q}}_i$	$\bar{x}_{i,1\mid 0}$	λ	h
1	0.0090	1.5184	0.0060	0	0.2000	0.1374		
2	0.0142	1.1746	0.0062	0.0175	0.1912	−0.0351	0.0069	4.4281×10^{-6}
3	0.0226	1.0405	0.0151	0.0686	0.2238	−0.0531		

$$\rho = \begin{bmatrix} 1 & -0.8397 & -0.9202 \\ 0.8397 & 1 & 0.7106 \\ -0.9202 & 0.7106 & 1 \end{bmatrix}, P_{1\mid 0} = \begin{bmatrix} 4\times 10^{-6} & 0 & 0 \\ 0 & 4\times 10^{-6} & 0 \\ 0 & 0 & 4\times 10^{-6} \end{bmatrix}$$

Notes: Parameter estimates for the dynamic Nelson-Siegel model. The estimation is conducted with the `R` package `Rsolnp`. The data sample consists of Canadian end-of-month yield curve from January 1986 to January 2022.

where $\mathbb{1}_3$ is the three-dimensional column vector containing ones as elements and υ_τ is a 3×3 matrix whose element on row i and column l, $\upsilon^{i,l}_\tau$, is

$$\upsilon^{i,l}_\tau = \frac{\sigma_i \sigma_l}{\kappa^{\mathbb{Q}}_i \kappa^{\mathbb{Q}}_l} \Gamma_{i,l} \left[\tau - B^{(i)}_\tau - B^{(l)}_\tau + \frac{1 - \left(1 - \kappa^{\mathbb{Q}}_i\right)^\tau \left(1 - \kappa^{\mathbb{Q}}_l\right)^\tau}{1 - \left(1 - \kappa^{\mathbb{Q}}_i\right) \left(1 - \kappa^{\mathbb{Q}}_l\right)} \right].$$

The model is estimated with maximum likelihood using a Kalman filter on the Canadian end-of-month yield curve data from January 1986 to January 2022. Resulting parameter estimates are given in Table B.1.

B.2. Dynamic Nelson-Siegel model

To represent the DNS model dynamics, we consider the same model specification than for the DTAFNS model, except that we that we use $\log A_\tau = 0$ and $B_\tau = \left[\tau, \frac{1 - e^{-\lambda\tau}}{\lambda}, \frac{1 - e^{-\lambda\tau}}{\lambda} - \tau e^{-\lambda\tau}\right]^\top$ in the bond pricing formula (2.15).

A global optimization of the log-likelihood is conducted with the `R` package `Rsolnp` on the Canadian end-of-month yield curve data extending from January 1986 to January 2022. Corresponding parameter estimates are provided in Table B.2.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Appendix C. Point prediction performance for spot rates

In this section, we compare the ability of the various considered models to produce accurate point monthly forecasts for spot rates. Let $\hat{y}(t, t+\tau)$ be the realized time-t spot rates with tenor τ, and $\tilde{y}(t, t+\tau)$ be its associated model-implied forecast produced at time $t-1$. Out-of-sample predictions are performed with the expanding window approach detailed in Section 3.3.2, i.e. models are retrained yearly, with yearly out-of-sample test sets covering the period extending from January 2017 to January 2022. Define the following performance metrics:

$$\begin{aligned}
\text{Mean error}^{(\tau)} &= \frac{1}{T^*}\sum_{t=1}^{T^*}\left(\hat{y}(t, t+\tau) - \tilde{y}(t, t+\tau)\right), \\
RMSE^{(\tau)} &= \sqrt{\frac{1}{T^*}\sum_{t=1}^{T^*}\left(\hat{y}(t, t+\tau) - \tilde{y}(t, t+\tau)\right)^2}, \\
MAE^{(\tau)} &= \frac{1}{T^*}\sum_{t=1}^{T^*}\left|\hat{y}(t, t+\tau) - \tilde{y}(t, t+\tau)\right|, \\
rMAE^{(\tau)}_{\mathcal{M}} &= \frac{MAE^{(\tau)}_{DTAFNS}}{MAE^{(\tau)}_{\mathcal{M}}}
\end{aligned} \tag{C.1}$$

where T^* is the total number of forecasting days (e.g. in the union of all test sets for out-of-sample tests) and $MAE^{(\tau)}_{\mathcal{M}}$ is the $MAE^{(\tau)}$ for model $\mathcal{M}$. Tables C.1 and C.2 report the point predictive performance results in-sample and out-of-sample, respectively. Since DTAFNS and DTAFNS-U forecasts are extremely close in the in-sample experiment, the DTAFNS-U benchmark is left out of the out-of-sample test. RMSE and MAE metrics indicate that the point forecasting performance of the DTAFNS model is quite similar to that of the DG3, the DNS and the DTAFNS-U models.

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press

Table C.1. Performance metrics for in-sample spot rate point predictions

	Mean error (in %)				RMSE (in %)				MAE (in %)				rMAE		
Tenor (months)	DTAFNS	DTAFNS-U	DG3	DNS	DTAFNS	DTAFNS-U	DG3	DNS	DTAFNS	DTAFNS-U	DG3	DNS	$\text{rMAE}_{\text{DTAFNS-U}}$	rMAE_{DG3}	rMAE_{DNS}
6	0.044	0.049	0.030	−0.007	0.339	0.340	0.337	0.360	0.206	0.207	0.205	0.235	0.996	1.007	0.878
12	0.013	0.023	0.020	−0.004	0.353	0.354	0.351	0.349	0.236	0.235	0.236	0.238	1.001	0.998	0.991
36	−0.023	−0.001	0.005	0.025	0.339	0.341	0.325	0.374	0.243	0.242	0.230	0.272	1.004	1.056	0.894
60	−0.007	0.021	0.004	0.050	0.298	0.299	0.293	0.331	0.216	0.215	0.211	0.243	1.004	1.021	0.887
84	0.024	0.051	0.016	0.068	0.282	0.286	0.280	0.301	0.208	0.212	0.207	0.224	0.983	1.008	0.931
108	0.057	0.078	0.040	0.082	0.268	0.274	0.267	0.279	0.205	0.210	0.204	0.210	0.976	1.006	0.976
132	0.073	0.085	0.056	0.077	0.268	0.272	0.270	0.274	0.204	0.208	0.205	0.206	0.983	0.998	0.991
156	0.060	0.059	0.048	0.045	0.277	0.276	0.280	0.280	0.206	0.207	0.209	0.207	0.998	0.987	0.996
180	0.021	0.009	0.019	−0.008	0.278	0.275	0.280	0.281	0.204	0.203	0.207	0.205	1.008	0.987	0.998
204	−0.024	−0.049	−0.018	−0.063	0.286	0.287	0.286	0.293	0.203	0.202	0.206	0.209	1.002	0.984	0.972
228	−0.053	−0.088	−0.041	−0.095	0.317	0.324	0.317	0.329	0.212	0.216	0.214	0.220	0.984	0.991	0.966
252	−0.040	−0.083	−0.027	−0.078	0.307	0.316	0.309	0.319	0.208	0.213	0.210	0.215	0.977	0.994	0.968
276	0.034	−0.015	0.042	0.008	0.265	0.266	0.264	0.253	0.189	0.192	0.189	0.184	0.982	0.998	1.028
300	0.133	0.082	0.130	0.128	0.600	0.593	0.591	0.563	0.309	0.312	0.311	0.297	0.993	0.996	1.043
324	0.012	−0.038	0.016	0.039	0.188	0.191	0.198	0.178	0.144	0.147	0.151	0.138	0.982	0.958	1.046
348	0.021	−0.024	0.010	0.087	0.255	0.254	0.260	0.242	0.178	0.180	0.181	0.174	0.992	0.986	1.023
Average	0.020	0.008	0.019	0.021	0.301	0.303	0.301	0.308	0.209	0.210	0.209	0.216	0.972	0.991	0.999

Notes: Performance metrics for the in-sample point forecasting of spot rates by the DTAFNS model and its three benchmarks. Performance metrics are reported for a selection of 16 tenors (in months) among the 33 available in dataset, while the average row at the bottom considers all tenors available.

Table C.2. Performance metrics for out-of-sample spot rate point predictions

	Mean error (in %)			RMSE (in %)			MAE (in %)			rMAE	
Tenor (months)	DTAFNS	DG3	DNS	DTAFNS	DG3	DNS	DTAFNS	DG3	DNS	$rMAE_{DG3}$	$rMAE_{DNS}$
6	0.048	0.075	0.120	0.707	0.712	0.734	0.612	0.609	0.620	1.004	0.986
12	0.155	0.160	0.166	0.731	0.732	0.742	0.629	0.630	0.634	0.999	0.992
36	0.089	0.054	−0.009	0.661	0.651	0.648	0.527	0.517	0.513	1.019	1.027
60	−0.043	−0.072	−0.137	0.626	0.623	0.628	0.472	0.472	0.484	1.000	0.975
84	−0.186	−0.192	−0.237	0.632	0.631	0.640	0.476	0.477	0.489	1.000	0.975
108	−0.296	−0.280	−0.298	0.652	0.644	0.647	0.484	0.477	0.480	1.014	1.009
132	−0.375	−0.346	−0.337	0.671	0.657	0.648	0.491	0.479	0.469	1.026	1.048
156	−0.433	−0.400	−0.366	0.689	0.671	0.648	0.500	0.483	0.463	1.035	1.081
180	−0.472	−0.443	−0.390	0.701	0.683	0.649	0.510	0.493	0.462	1.034	1.103
204	−0.494	−0.473	−0.408	0.704	0.692	0.649	0.517	0.505	0.465	1.024	1.113
228	−0.498	−0.489	−0.422	0.699	0.694	0.649	0.517	0.511	0.469	1.011	1.102
252	−0.488	−0.492	−0.434	0.684	0.688	0.648	0.505	0.508	0.472	0.993	1.069
276	−0.466	−0.481	−0.445	0.662	0.673	0.648	0.485	0.496	0.475	0.979	1.021
300	−0.433	−0.457	−0.456	0.633	0.650	0.649	0.465	0.478	0.477	0.972	0.974
324	−0.389	−0.417	−0.466	0.599	0.618	0.651	0.441	0.455	0.482	0.969	0.913
348	−0.331	−0.361	−0.476	0.559	0.577	0.652	0.415	0.426	0.489	0.974	0.848
Average	−0.286	−0.286	−0.287	0.661	0.660	0.658	0.502	0.501	0.497	1.003	1.011

Notes: Performance metrics for the out-of-sample point forecasting of spot rates by the DTAFNS model and the DNG3 and DNS benchmarks explained in Appendix B. The expanding window approach detailed in Section 3.3.2 is used for testing: models are retrained yearly, with yearly out-of-sample test sets covering the period extending from January 2017 to January 2022. Performance metrics are reported for a selection of 16 tenors (in months) among the 33 available in dataset, while the average row at the bottom considers all tenors available.

Cite this article: Eghbalzadeh R, Godin F and Gaillardetz P (2024). The discrete-time arbitrage-free Nelson-Siegel model: a closed-form solution and applications to mixed funds representation, *Annals of Actuarial Science*, 1–32. https://doi.org/10.1017/S1748499524000010

https://doi.org/10.1017/S1748499524000010 Published online by Cambridge University Press