Discrete-Time Arbitrage-Free Nelson-Siegel Model and its Applications to Participating Life Insurance Contracts and Swaptions Pricing

Ramin Eghbalzadeh

A Thesis

In the Department of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Mathematics) at

Concordia University

Montréal, Québec, Canada

May, 2023

© Ramin Eghbalzadeh, 2023

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Ramin Eghbalzadeh

Entitled: Discrete-Time Arbitrage-Free Nelson-Siegel Model and Its Applications in Participating Life Insurance Contracts and Swaptions Pricing

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Rahul Ravi

_____ Co-supervisor
Dr. Frédéric Godin

_____ Co-supervisor
Dr. Patrice Gaillardetz

_____ Examiner
Dr. Cody Hyndman

_____ Examiner
Dr. Mélina Mailhot

_____ Examiner
Dr. Yang Lu

_____ External Examiner
Dr. Adam Metzler

Approved by _____ ,Graduate Program Director
Dr. Yogendra Chaubey

8 May 2023 _____ ,Dean of Faculty
Dr. Pascale Sicotte

Abstract

Title: Discrete-Time Arbitrage-Free Nelson-Siegel Model and Its Applications in Participating Life Insurance Contracts and Swaptions Pricing

Ramin Eghbalzadeh, Ph.D.

Concordia University, 2023

This dissertation explores the importance of interest rate modeling in finance and actuarial science. It emphasizes the significance of yield curve modeling in pricing financial instruments, such as participating life insurance contracts and swaptions. The dissertation extends the work of previous studies and proposes a slightly different version of the discrete-time arbitrage-free Nelson-Siegel model (DTAFNS), providing a closed-form expression for risk-free spot rates and demonstrating its superior out-of-sample predictive ability. Additionally, the dissertation focuses on stochastic interest rates and mortality dynamics' impact on the pricing, reserving, and risk measurement approaches of participating life insurance contracts, with the introduction of a shadow reserve to improve accuracy. Lastly, the dissertation outlines procedures for pricing swaptions under the DTAFNS model. Overall, this dissertation contributes to the stability of the financial sector and protecting the financial well-being of individuals and institutions.

Acknowledgments

First and foremost, I would like to express my heartfelt gratitude to my wife, Elnaz, who has been my constant source of love and support throughout my Ph.D. journey. Your unwavering belief in me has given me the strength and motivation to keep going, even during the most challenging times. Thank you for being my rock and for always being there for me.

I am also deeply grateful to my supervisors, Professors Frédéric Godin and Patrice Gaillardetz, for their invaluable guidance and support throughout my research. Their expertise and insights have been instrumental in shaping my work, and I am thankful for all the opportunities they have provided me with.

I would also like to extend my gratitude to my newborn daughter, Emma, for reminding me of the importance of finding joy and balance in life. Your arrival has brought a new perspective to my work and has made me even more determined to succeed.

Furthermore, I would like to thank Professors Cody Hyndman, Mélina Mailhot, Yang Lu, and Adam Metzler, for generously agreeing to evaluate my dissertation. Their valuable feedback and constructive criticism will undoubtedly help me improve my work and contribute to the advancement of my field. Thank you for your commitment to excellence and for the contributions you have made to my academic journey.

Finally, I want to acknowledge the unwavering support of my friends and family, who have been with me every step of the way. Your encouragement and love have meant more to me than words can express.

Thank you all for being a part of my journey and helping me reach this incredible milestone.

Contribution of Authors

This thesis is based on three research articles:

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

II : Evaluation of participating endowment life insurance policies in a stochastic environment

III : Swaption pricing under the discrete-time arbitrage-free Nelson-Siegel model

Contents

1 Introduction **1**

2 Discrete-time arbitrage-free Nelson–Siegel model **4**

2.1 Introduction 4
2.2 Interest rate term structure model 8
2.2.1 The Nelson-Siegel term structure representation 8
2.2.2 Discrete-time arbitrage-free Nelson-Siegel model 10
2.3 Model estimation 15
2.3.1 Estimation framework 16
2.3.2 Model estimation and analysis 20
2.3.3 Performance assessment and benchmarking 23
2.4 An application to mixed fund modeling 27
2.4.1 Rolling bond fund returns 27
2.4.2 Risky assets returns 29
2.4.3 Mixed fund dynamics 30
2.4.4 Estimation of the mixed fund model 32
2.5 Conclusion 33

3 Participating Life Insurance Contract **36**

3.1 Introduction 36

3.2 Standard endowment life insurance 39
3.2.1 Actuarial notations for mortality 40
3.2.2 Endowment life insurance 40
3.3 Participating policies 43
3.3.1 Standard variable benefit increase mechanism 44
3.3.2 Compound Reversionary Bonus Methods 46
3.4 Pricing, reserving, and risk evaluation for participating contracts 48
3.4.1 The shadow reserve 48
3.4.2 The fair premium of a participating contract 49
3.4.3 Risk measurement 51
3.5 Financial model 52
3.5.1 Interest rate term structure model 52
3.5.2 Reference account dynamics 55
3.6 Numerical analyses 57
3.6.1 Fair premiums 58
3.6.2 Analysis of benefits 60
3.6.3 Risk exposure 62
3.6.4 Analysis of reserves 65
3.7 Conclusion 69

4 Swaption pricing under the discrete-time arbitrage-free Nelson-Siegel model 71

4.1 Introduction 71
4.2 Preliminaries: the DTAFNS model and swaptions 73
4.2.1 DTAFNS model 74
4.2.2 European swaptions 76
4.3 Pricing swaptions under the forward measure 77
4.3.1 Dynamics of the state variables under the forward measure 79

4.4 Conclusion 82

5 Conclusion **84**

A Appendice **86**

A.1 Discrete-time arbitrage-free Nelson-Siegel model 86

A.1.1 Proofs 86

A.1.2 Benchmarks 95

A.1.3 Kalman filter 98

A.1.4 Alternative parameters selection for DTAFNS 100

A.1.5 Point prediction performance for spot rates 101

A.1.6 Fit diagnostics output for the mixed fund model 102

A.2 Participating contract 105

A.3 Swaption pricing under DTAFNS model 107

A.3.1 Proofs 107

List of Figures

2.1 Model-implied factors and short rate time series 22

2.2 Model-implied and observed spot rate time series for 3-month and 10-year tenors 23

2.3 Model-implied and observed yield curves 26

3.1 Premium vs cost concerning the different spread factors (θ). 50

3.2 Evolution of the benefit per premium ratio for each mechanism. 61

3.3 Time evolution of the augmented account value per premium unit and policyholder. 63

3.4 Boxplots of the net amount at risk (NAAR) for each mechanism 64

3.5 Difference between the shadow reserve and the policy reserve for both mechanisms 67

3.6 Difference between the augmented reference account and the shadow reserve 68

List of Tables

2.1 Maximum likelihood estimates of the DTAFNS model parameters 21

2.2 Log-likelihood of the DTAFNS model and its benchmarks 25

2.3 Probability of observing negative short rates with the DTAFNS model or its benchmarks 27

2.4 Bivariate EGARCH equity model parameter estimates 34

2.5 Mixed fund model parameter estimates 34

3.1 Total premium for different ages, maturities and benefit increase mechanisms. 59

3.2 Total premium for different loading factors θ, minimum guarantee rates g, and mechanisms. 60

A.1 DG3 model parameter estimates 97

A.2 Dynamic Nelson-Siegel model parameters estimates. 98

A.3 Maximum likelihood estimates of the DTAFNS model with constraint $\kappa^{\mathbb{P}}_{1,1} = 0$101

A.4 Log-likelihood of the DTAFNS model with and without constraint $\kappa^{\mathbb{P}}_{1,1} = 0$ 101

A.5 Probability of observing negative short rates with the DTAFNS model with and without constraint $\kappa^{\mathbb{P}}_{1,1} = 0$ 102

A.6 Performance metrics for spot rate point predictions 103

A.7 DTAFNS model parameters 106

A.8 Bivariate EGARCH model parameters 106

A.9 Mixed fund model parameters 106

A.10 Initial financial risk factors parameters for simulations 107

Chapter 1

Introduction

Interest rate modeling plays a critical role in both finance and actuarial science. In finance, it is used to determine the value of financial instruments, such as bonds and mortgages, by taking into account the future behavior of interest rates. This information is then used to make informed investment decisions and to price financial products accurately. In the field of actuarial science, interest rate modeling is essential for calculating the present value of future cash flows, which is used to assess the financial risk associated with insurance policies, pensions, and other financial contracts. Interest rate models allow actuaries to make predictions about future interest rate trends, which can be used to design and price insurance products that are both financially sound and attractive to consumers. The importance of interest rate modeling in these fields highlights its significance in ensuring the stability of the financial sector and in protecting the financial well-being of individuals and institutions.

Modeling yield curves have been a crucial topic in finance for several decades, with a primary focus on understanding the behavior of interest rates and their impact on various financial instruments. The yield curve represents the relationship between the yield and maturity of fixed-income securities, and it is widely used in financial analysis and decision-

making.

One area where yield curve modeling has significant implications is in the pricing of participating life insurance contracts. These contracts, also known as PAR policies, provide a combination of insurance coverage and participation to internal insurance company performance, and their pricing requires a deep understanding of the interest rates environment. By incorporating yield curve models, actuaries can better assess the financial risk associated with participating life insurance contracts and ensure their sustainability over the long term.

Lastly, the use of yield curve models is also crucial in the pricing of swaptions, which are options that give the holder the right but not the obligation to enter into a swap agreement at a pre-determined strike rate. Swaps are financial instruments used to hedge against changes in interest rates, and the pricing of swaptions requires a comprehensive understanding of the yield curve and its evolution over time. Yield curve models help to identify and quantify the various risk factors associated with swaptions, allowing for more accurate pricing and risk management.

This dissertation first expands on the work of Hong, Niu, and Zeng 2016 and proposes a slightly different version of the discrete-time arbitrage-free Nelson-Siegel (DTAFNS) model, making three main contributions. Firstly, a closed-form expression for risk-free spot rates is derived under this model. Secondly, the DTAFNS model is shown to have superior distributional out-of-sample predictive ability over benchmarks when applied to historical Canadian term structure data. Lastly, a modified version of the mixed fund return model of Augustyniak, Godin, and Hamel 2021 is developed, substituting their three-factor Gaussian model with the DTAFNS model, justified by the higher performance of the DTAFNS model.

A second objective of this dissertation is to extend participating contract analysis schemes to include realistic dynamics for the main risk drivers, including stochastic interest rates and mortality. Fixed premium contracts where bonuses are paid out as

benefit increases are considered. Bonus calculations are generalized to handle general stochastic dynamics. Pricing, reserving, and risk measurement approaches are discussed. The introduction of a shadow reserve is proposed to replace conventional reserving formulas and provide additional accuracy when measuring liabilities as future benefit increases beyond the guaranteed rate are recognized. Several numerical tests based on Monte-Carlo simulations are presented to demonstrate the importance of considering such a shadow reserve when measuring risk.

This dissertation lastly outlines procedures for the pricing swaptions under the DTAFNS model. It also demonstrates how to obtain forward measure dynamics of interested under such model, which can be leveraged to obtain a faster calculation for the price of a swaption compared to conventional methods relying on the risk-neutral measure. Overall, this dissertation contributes to the literature on fixed-income securities valuation, life insurance contract analysis, and swaptions pricing, providing practical applications for the DTAFNS model in each field.

This dissertation comprises three papers, each of which has been included as a chapter in the dissertation. As a result, readers may notice some repetition in certain sections of the dissertation. This is because the content of each paper has been copied exactly into the corresponding chapter of the dissertation, with the only difference being that appendices of all papers have been regrouped into a single appendix in the thesis. Nevertheless, the dissertation as a whole provides a comprehensive and detailed exploration of the topics at hand, with each paper providing unique insights and findings.

The DTAFNS model is discussed in Chapter 2, which provides a closed-form expression for risk-free spot rates. Chapter 3 explores how stochastic interest rates and mortality dynamics impact the pricing, reserving, and risk measurement approaches for participating life insurance contracts. Finally, chapter 4 delves into the pricing of swaptions, which are often used to hedge interest rate risk.

Chapter 2

Discrete-time arbitrage-free Nelson–Siegel model

2.1 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. Superior models provide enhanced predictive power and risk evaluation ability, which are crucial to insurers and pensioners. Interest rate models often study the dynamics of the term structure specifying simultaneously spot rates for different tenors. The evolution of the term structure can be represented by various types of models, which can be of statistical, economic, or financial nature. Statistical approaches, such as these of Diebold and Li 2006, Guidolin and Timmermann 2006 or Engle, Roussellet, and Siriwardane 2017, mainly rely on time series models. In economic methods, see for instance Ang and Piazzesi 2003, Hördahl, Tristani, and Vestin 2006 or Ang et al. 2011, the evolution of the

term structure is linked to macro-economic factors, such as inflation rates, the growth rate of the economy and the central bank monetary policy. Financial approaches are based on the absence of arbitrage paradigm ensuring that pricing systems are consistent. Brigo and Mercurio 2007 is a comprehensive reference describing multiple models from that category.

Vasicek 1977 and Cox, Ingersoll, and Ross 1985 are early contributions in the stream of literature of financial models for the term structure. They propose to model the evolution of the short rate as specific Itô diffusions leading to a closed-form solution for zero-coupon prices and the corresponding spot rates. Multi-factor generalizations of such models leading to more flexible term structure shapes and to the ability to properly model correlations between future spot rates have then been proposed. See for instance Brennan and Schwartz 1979, Chen and Scott 1992, Rogers 1995 and Jamshidian 1996 for example of multi-factor models of the term structure. Factor models can be motivated for instance by the work of Litterman and Scheinkman 1991 showing that at least 95% of variation in the yield curve can be explained by three statistical factors. An important category of financial factor models for interest rates that is described in Duffie and Kan 1996 is affine term structure (ATS) models, in which any spot rate is a linear combination of underlying factors. Such models are typically obtained by assuming diffusive dynamics for the factors and deriving corresponding bond prices through common no-arbitrage arguments. By suitably choosing the drift of the factors, it is possible to obtain a model-implied term structure that exactly match the observed one, see for instance Hull and White 1990. The latter model is a particular case of the Heath, Jarrow, and Morton 1992 framework which, instead of directly modeling the short rate, uses the current term structure as initial input and applies perturbations on it to represent the evolution of the yield curve.

An issue with some multi-factor models of the term structure is that the factors are sometimes hard to interpret. The seminal work of Nelson and Siegel 1987 and the

Svensson 1994 extension specifically aim to depict the term structure as a linear combination of interpretable factors representing the level, slope and humps in the yield curve. Whereas the latter two models were initially developed in a static setting, Diebold and Li 2006 extend the Nelson and Siegel 1987 model into a dynamic framework. However, a drawback associated with the dynamic version of the Nelson and Siegel 1987 model is that it is shown to be incompatible with no-arbitrage theory in Filipović 1999. The work of Christensen, Diebold, and Rudebusch 2011 successfully attempts to reconcile both the no-arbitrage strand of literature and the Nelson and Siegel 1987 model by developing the so-called arbitrage-free Nelson-Siegel (AFNS) model. Such model is an affine term structure model 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. There are debates in the literature about whether enforcing absence of arbitrage in interest rate models is necessary for predictive purposes: Duffee 2002 argues that ATS models often have a poor performance when used for forecasting future Treasury yields. Nevertheless, when considering asset pricing applications, enforcing absence of arbitrage is often required to obtain a meaningful model.

Many of the aforementioned financial models of the term structure consider continuous-time settings. However, most of these models have analog discrete-time counterparts, many of which are described in Wüthrich and Merz 2013. Discrete-time models often possess the advantage of being simpler to manipulate, which is favorable in applications. A discrete-time multi-factor Gaussian term structure model is illustrated in Augustyniak, Godin, and Hamel 2021, in which it serves as a building block for an econometric model for mixed equity and fixed income fund returns. That paper uses the mixed fund model to price variable annuities. The discrete-time analog of the multi-factor Cox-Ingersoll-Ross

model is developed in the work of Le, Singleton, and Dai 2010, which uses auto-regressive Gamma processes introduced by Gouriéroux and Jasiak 2006. A discrete-time version of the AFNS model has been developed in Hong, Niu, and Zeng 2016. As mentioned above, the AFNS model possesses both the favorable attributes of being consistent with no-arbitrage considerations and of embedding interpretable risk factors; having a discrete-time version of such model available is therefore highly relevant for practitioners who desire avoiding technicalities associated with continuous-time models and who wish to retain positive features of the AFNS model.

This paper further expands on the work Hong, Niu, and Zeng 2016 and revisits the discrete-time arbitrage-free Nelson-Siegel (DTAFNS) model, providing the following three main contributions: (i) obtaining a closed-form expression for risk-free spot rates under such model, (ii) presenting evidence of superior distributional out-of-sample predictive ability of the DTAFNS model over benchmarks when applied on historical Canadian term structure data, and (iii) developing a modified version of the mixed fund return model of Augustyniak, Godin, and Hamel 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.

The remainder of this paper is organized as follows. Section 2.2 specifies the proposed interest rate model and provides closed-form expressions for risk-free zero-coupon bond prices and spot rates. In Section 2.3, the model calibration is illustrated and numerical experiments assessing predictive performance are provided. Section 2.4 adapts the mixed fund model of Augustyniak, Godin, and Hamel 2021 to include the DTAFNS model as a building block for the term structure model. Section 2.5 concludes.

2.2 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, Niu, and Zeng 2016.

2.2.1 The Nelson-Siegel term structure representation

Nelson and Siegel 1987 propose a parametric representation of the yield curve of the form

$$y(t,T) = X^{(1)} + X^{(2)}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + X^{(3)}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right), \tag{2.2.1}$$

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 and 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 and Li 2006 is its incompatibility with no-arbitrage theory as discussed Filipović 1999. Christensen, Diebold, and Rudebusch 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:

$$r_t \equiv X_t^{(1)} + X_t^{(2)}, \tag{2.2.2}$$

$$\begin{bmatrix} dX_t^{(1)} \\ dX_t^{(2)} \\ dX_t^{(3)} \end{bmatrix} \equiv \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}}_{\kappa^{\mathbb{Q}}} \underbrace{\begin{bmatrix} \theta_1^{\mathbb{Q}} - X_t^{(1)} \\ \theta_2^{\mathbb{Q}} - X_t^{(2)} \\ \theta_3^{\mathbb{Q}} - X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{Q}} - \mathbf{X}_t} dt + \Sigma \begin{pmatrix} dW_{t,1}^{\mathbb{Q}} \\ dW_{t,2}^{\mathbb{Q}} \\ dW_{t,3}^{\mathbb{Q}} \end{pmatrix} \tag{2.2.3}$$

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, Diebold, and Rudebusch 2011, this leads to the following formulas for spot rates:

$$y(t,T) = -\frac{1}{\tau}\sum_{j=1}^{3} \tilde{B}^{(j)}(t,T) X_t^{(j)} - \frac{\tilde{C}(t,T)}{\tau}, \tag{2.2.4}$$

$$\tilde{C}(t,T) = \sum_{j=2}^{3} \left(\kappa^{\mathbb{Q}}\theta^{\mathbb{Q}}\right)_j \int_t^T \tilde{B}^{(j)}(s,T)ds + \frac{1}{2}\sum_{j=1}^{3}\int_t^T \left(\Sigma^\top \tilde{B}(s,T)\tilde{B}(s,T)^\top \Sigma\right)_{j,j} ds \tag{2.2.5}$$

with

$$\begin{aligned} \tilde{B}(s,T) &\equiv \left[\tilde{B}^{(1)}(s,T),\ \tilde{B}^{(2)}(s,T),\ \tilde{B}^{(3)}(s,T)\right]^\top \\ &\equiv \left[-\tau,\ -\left(\frac{1-e^{-\lambda\tau}}{\lambda}\right),\ \left(\tau e^{-\lambda\tau} - \frac{1-e^{-\lambda\tau}}{\lambda}\right)\right]^\top \end{aligned} \tag{2.2.6}$$

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

Striking similarities can be observed when comparing (2.2.1) with (2.2.4) and (2.2.6). 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 is only caused by the corrective term $-\frac{\tilde{C}(t,T)}{\tau}$ appearing in (2.2.4). This leads Christensen, Diebold, and Rudebusch 2011 to describe the arbitrage-free Nelson-Siegel model (2.2.2)-(2.2.3) as the "*closest match to the Nelson-Siegel yield function*" within some class of exponential affine term structure models.

2.2.2 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)-(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

$$r_t = X_t^{(1)} + X_t^{(2)}, \tag{2.2.7}$$

$$X_{t+1} = X_t + \kappa^{\mathbb{P}}(\theta^{\mathbb{P}} - X_t) + \Sigma Z_{t+1}^{\mathbb{P}}, \tag{2.2.8}$$

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 $Z^{\mathbb{P}} = \{Z_t^{\mathbb{P}}\}_{t=1}^{T}$ 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 $\rho_{i,j}^{(Z)} \equiv \text{corr}(Z_{t,i}^{\mathbb{P}}, Z_{t,j}^{\mathbb{P}})$, $t = 1, \ldots, T$ 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 (2.2.7)-(2.2.8) are meant to mimic (2.2.2)-(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 and 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, Godin, and Hamel 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. 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 $Z^{\mathbb{Q}} = \{Z_t^{\mathbb{Q}}\}_{t=1}^T$ 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 (2.2.8) leads to

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

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

$$\begin{cases} \kappa^{\mathbb{Q}}\theta^{\mathbb{Q}} = \kappa^{\mathbb{P}}\theta^{\mathbb{P}}, \\ \kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} - \Sigma\gamma, \end{cases} \tag{2.2.9}$$

then

$$X_{t+1} = X_t + \kappa^{\mathbb{Q}}(\theta^{\mathbb{Q}} - X_t) + \Sigma Z^{\mathbb{Q}}_{t+1} \tag{2.2.10}$$

and thus the factor process $\{X_t\}_{t=0}^T$ preserves its auto-regressive structure under the risk-neutral measure.

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.1.1.

Lemma 2.2.1. *For some integer $\tau > 0$ and a real number $r \neq 1$,*

$$\zeta_0(r,\tau) \equiv \sum_{u=1}^{\tau-1} r^u = \frac{r - r^\tau}{1-r}, \tag{2.2.11}$$

$$\zeta_1(r,\tau) \equiv \sum_{u=1}^{\tau-1} u r^u = \frac{r - \tau r^\tau + (\tau-1)r^{\tau+1}}{(1-r)^2}, \tag{2.2.12}$$

$$\zeta_2(r,\tau) \equiv \sum_{u=1}^{\tau-1} u^2 r^u = \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1)r^{\tau+1} - \tau^2 r^\tau + r^2 + r}{(1-r)^3}. \tag{2.2.13}$$

Proposition 2.2.1. *Consider $\lambda \in (0,1)$ and assume*

$$\kappa^{\mathbb{Q}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix}. \tag{2.2.14}$$

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

dynamics:

$$\begin{pmatrix} X_{t+1}^{(1)} - X_t^{(1)} \\ X_{t+1}^{(2)} - X_t^{(2)} \\ X_{t+1}^{(3)} - X_t^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{bmatrix} \theta_1^{\mathbb{Q}} - X_t^{(1)} \\ \theta_2^{\mathbb{Q}} - X_t^{(2)} \\ \theta_3^{\mathbb{Q}} - X_t^{(3)} \end{bmatrix} + \begin{pmatrix} \Sigma_{1,1} & 0 & 0 \\ 0 & \Sigma_{2,2} & 0 \\ 0 & 0 & \Sigma_{3,3} \end{pmatrix} \begin{pmatrix} Z_{t+1,1}^{\mathbb{Q}} \\ Z_{t+1,2}^{\mathbb{Q}} \\ Z_{t+1,3}^{\mathbb{Q}} \end{pmatrix}.$$

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

$$P(t,T) = A_\tau \exp\left[-\Delta B_\tau^\top X_t\right], \tag{2.2.15}$$

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

$$\begin{aligned} B_\tau^{(1)} &= \tau, \quad B_\tau^{(2)} = \frac{1-(1-\lambda)^\tau}{\lambda}, \quad B_\tau^{(3)} = \frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}, \\ \log A_\tau &= -\Delta\theta_2^{\mathbb{Q}}\left(B_\tau^{(1)} - B_\tau^{(2)}\right) + \Delta\theta_3^{\mathbb{Q}} B_\tau^{(3)} + \frac{1}{2}\Delta^2 \upsilon_\tau \end{aligned}$$

with

$$\begin{aligned} \upsilon_\tau &= \sum_{i=1}^{3}\sum_{j=1}^{3} \upsilon_\tau^{(i,j)} \\ \upsilon_\tau^{(1,1)} &= \Sigma_{1,1}^2 \frac{\tau(\tau-1)(2\tau-1)}{6}, \\ \upsilon_\tau^{(2,2)} &= \frac{\Sigma_{2,2}^2}{\lambda^2}\left(\tau - 2\left[\frac{1-(1-\lambda)^\tau}{\lambda}\right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2}\right), \\ \upsilon_\tau^{(3,3)} &= \frac{\Sigma_{3,3}^2}{\lambda^2}\Bigg[\tau - 2 + \zeta_0\left((1-\lambda)^2, \tau-1\right) + \lambda^2\zeta_2\left((1-\lambda)^2, \tau-1\right) \\ &\qquad - 2\zeta_0\left((1-\lambda), \tau-1\right) - 2\lambda\zeta_1\left((1-\lambda), \tau-1\right) + 2\lambda\zeta_1\left((1-\lambda)^2, \tau-1\right)\Bigg], \\ \upsilon_\tau^{(1,2)} &= \upsilon_\tau^{(2,1)} = \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda), \tau)\right), \\ \upsilon_\tau^{(1,3)} &= \upsilon_\tau^{(3,1)} = \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0\left((1-\lambda), \tau-1\right) - (\lambda+1)\zeta_1\left((1-\lambda), \tau-1\right) \\ &\qquad - \lambda\zeta_2\left((1-\lambda), \tau-1\right)\Bigg], \end{aligned}$$

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

This entails that spot rates have the form

$$y(t,T) = X_t^{(1)} + \frac{(1-(1-\lambda)^\tau)}{\tau\lambda}X_t^{(2)} + \left[\frac{(1-(1-\lambda)^{\tau-1})}{\tau\lambda} - \frac{\tau-1}{\tau}(1-\lambda)^{\tau-1}\right]X_t^{(3)} - \frac{1}{\Delta\tau}\log A_\tau. \tag{2.2.16}$$

Remark 2.2.1. *In Proposition 2.2.1, parameter $\theta_1^{\mathbb{Q}}$ does not appear in the zero-coupon price formula* (2.2.15) *and thus can be set to $\theta_1^{\mathbb{Q}} = 0$.*

To obtain risk-neutral dynamics of the form proposed in Proposition 2.2.1, 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 (2.2.9) with

$$\gamma = \begin{pmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{pmatrix}, \quad \gamma_i = (\kappa_{i,i}^{\mathbb{P}} - \mathbb{1}_{\{i>1\}}\lambda)/\Sigma_{i,i},\ i = 1,2,3. \tag{2.2.17}$$

Furthermore, substituting (2.2.17) in the first equation of (2.2.9) leads to

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

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

$$\kappa^{\mathbb{P}}_{i,i} = \mathbb{1}_{\{i>1\}}\lambda + \Sigma_{i,i}\gamma_i,\ i = 1,2,3, \quad \theta^{\mathbb{P}}_3 = \lambda\theta^{\mathbb{Q}}_3/\kappa^{\mathbb{P}}_{3,3}, \quad \theta^{\mathbb{P}}_2 = \frac{\lambda}{\kappa^{\mathbb{P}}_{2,2}}\left[\theta^{\mathbb{Q}}_2 - \frac{\theta^{\mathbb{Q}}_3}{\kappa^{\mathbb{P}}_{3,3}}(\kappa^{\mathbb{P}}_{3,3} - \lambda)\right]. \tag{2.2.18}$$

Remark 2.2.2. *Since $\kappa^{\mathbb{Q}}_{1,1} = 0$, an alternative specification enforcing $\kappa^{\mathbb{P}}_{1,1} = 0$ has been tested, but it was ultimately not retained due to lower performance. See Appendix A.1.4 for more details. The reason why such alternative specification has been tested is because it provides for non-stationary dynamics of the first risk factor under $\mathbb{P}$, which mimics its $\mathbb{Q}$-dynamics.*

Remark 2.2.3. *The DTANFS model of this paper is specified slightly differently than in Hong, Niu, and Zeng 2016. In the latter work, the matrix $\kappa^{\mathbb{Q}}$ is chosen such that factor loadings B_τ exactly match that of the Nelson-Siegel model* (2.2.1)*. Conversely, this paper emphasizes an auto-regressive representation of factors, i.e. Equation* (2.2.10)*, which is the discrete-time counterpart of the diffusive dynamics of the AFNS model* (2.2.3)*. Nevertheless, both specifications of the model are quite similar conceptually. In Hong, Niu, and Zeng 2016, 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.*

2.3 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, Godin, and Hamel 2021 and a version of the Dynamic Nelson-Siegel model of Diebold and Li 2006.

2.3.1 Estimation framework

Several methods have been considered in the literature to estimate factor models of interest rates, such as conventional maximum likelihood estimation (Chen and Scott 1993), maximum likelihood with Kalman filters (Duan and Simonato 1999; Lemke 2006; Park 2014; Augustyniak, Godin, and Hamel 2021), the method of simulated moments (Dai and Singleton 2000) or Bayesian approaches (Hong, Niu, and Zeng 2016). 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

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

where (2.2.16) leads to

$$y(t, t+n) = \frac{-1}{n\Delta} \log A_n + \frac{1}{n} B_n^\top X_t. \tag{2.3.1}$$

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

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

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

$$\hat{\mathsf{y}}(t) = \mathsf{a} + \mathsf{B}X_t + \eta_t, \tag{2.3.2}$$

where

$$\mathsf{a} \equiv \left(\frac{-1}{n_1 \Delta} \log A_{n_1}, \ldots, \frac{-1}{n_M \Delta} \log A_{n_M} \right)^\top, \tag{2.3.3}$$

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

$$\mathsf{B} \equiv (\beta_{n_1}, \ldots, \beta_{n_M})^\top. \tag{2.3.4}$$

Furthermore, the transition equation describing the dynamics of the latent factors X_t is obtained from (2.2.8):

$$X_{t+1} = \mathbf{b} + \mathbf{D}X_t + \xi_{t+1}, \tag{2.3.5}$$

where

$$\mathbf{b} \equiv \kappa^{\mathbb{P}}\theta^{\mathbb{P}}, \quad \mathbf{D} \equiv I - \kappa^{\mathbb{P}}, \quad \xi_{t+1} \equiv \Sigma Z_{t+1}^{\mathbb{P}}. \tag{2.3.6}$$

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

$$\mathbf{Q} = \Sigma\rho\Sigma. \tag{2.3.7}$$

The representation (2.3.2)-(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 A.1.3. 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, Godin, and Hamel 2021, an optimization of the likelihood is conducted with the `R` package `Rsolnp` (Ghalanos and 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.

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 $L(\hat{y}) = \log\left(p\left(\hat{\mathsf{y}}(T), \hat{\mathsf{y}}(T-1), ..., \hat{\mathsf{y}}(1)\right)\right)$ and smoothed density moments $\bar{x}_{t|T}, P_{t|T}$, $t = 1, \ldots, T$.

Calculating the log-likelihood:

Calculate $\mathsf{a}, \mathbf{b}, \mathbf{D}, \mathbf{Q}$ *and* B *with* (2.2.18), (2.3.3), (2.3.4), (2.3.6), (2.3.7) *and Proposition 2.2.1.*

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* $P_{t|t}$*:*

$\bar{x}_{t|t} \leftarrow \bar{x}_{t|t-1} + P_{t|t-1}\mathsf{B}^\top\bar{\Sigma}_t^{-1}(\hat{\mathsf{y}}(t) - \bar{\mathsf{y}}(t))$

$P_{t|t} \leftarrow P_{t|t-1} - P_{t|t-1}\mathsf{B}^\top\bar{\Sigma}_t^{-1}\mathsf{B}P_{t|t-1}$

$\bar{x}_{t+1|t} \leftarrow \mathbf{b} + \mathbf{D}\bar{x}_{t|t}$

$P_{t+1|t} \leftarrow \mathbf{D}P_{t|t}\mathbf{D}^\top + \mathbf{Q}$

end for

$L(\hat{y}) \leftarrow \sum_{t=1}^{T} \log p\left(\hat{\mathsf{y}}(t)|\hat{\mathsf{y}}(1), ..., \hat{\mathsf{y}}(t-1)\right)$

Calculating the smoothed state density:

for $t \in \{T-1, \ldots, 1\}$ **do**

$\mathbf{J}_t \leftarrow P_{t|t}\mathbf{D}^\top P_{t+1|t}^{-1}$,

$\bar{x}_{t|T} \leftarrow \bar{x}_{t|t} + \mathbf{J}_t(\bar{x}_{t+1|T} - \bar{x}_{t+1|t})$,

$P_{t|T} \leftarrow P_{t|t} + \mathbf{J}_t(P_{t+1|T} - P_{t+1|t})\mathbf{J}_t^\top$.

end for

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 and Stoffer 2017) which is also shown in Algorithm 1.

2.3.2 Model estimation and analysis

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

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[1] and are constructed through an exponential smoothing methodology described in Bolder, Metzler, and Johnson 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.[2]

Estimated parameters and performance

Table 2.1 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 $X^{(3)}$

1. Source: https://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/
2. 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.

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 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 $\kappa^{\mathbb{P}}_{i,i}$ 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 $\theta^{\mathbb{P}}_2 = 0.0301$ under physical dynamics.

Table 2.1: Maximum likelihood estimates of the DTAFNS model parameters

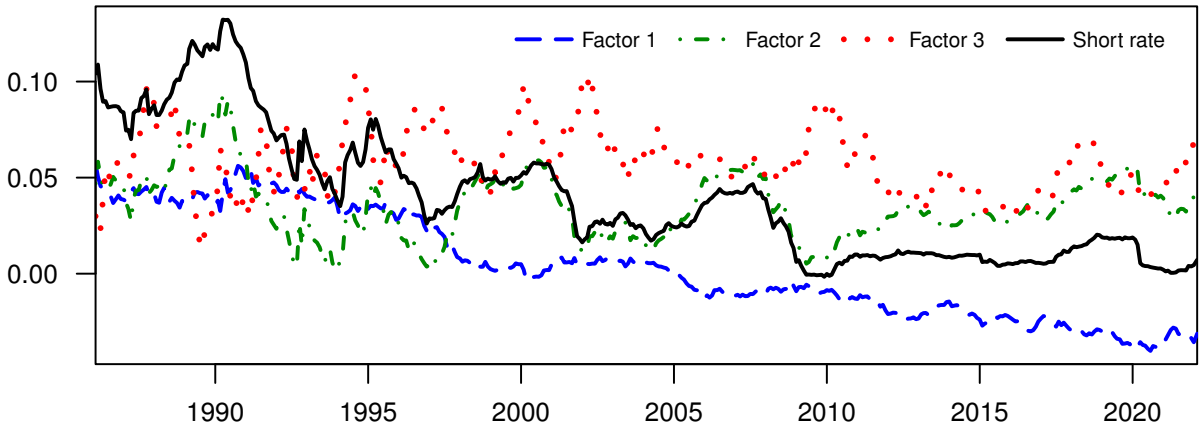
i	$\kappa^{\mathbb{P}}_{i,i}$	γ_i	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_i$	$\theta^{\mathbb{Q}}_i$	$\bar{x}_{i,1\|0}$	λ	h
1	0.0075	2.7923	0.0027	0	0	0.0491		
2	0.0288	1.2016	0.0045	0.0301	0.0633	0.0391	0.0233	3.76×10^{-6}
3	0.0354	1.7167	0.0070	0.0505	0.0766	0.0291		

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

Notes: Maximum likelihood parameter estimates for the DTAFNS model presented in Section 2.2.2 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 2.1 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 (2.2.7). 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,

Figure 2.1: Model-implied factors and short rate time series

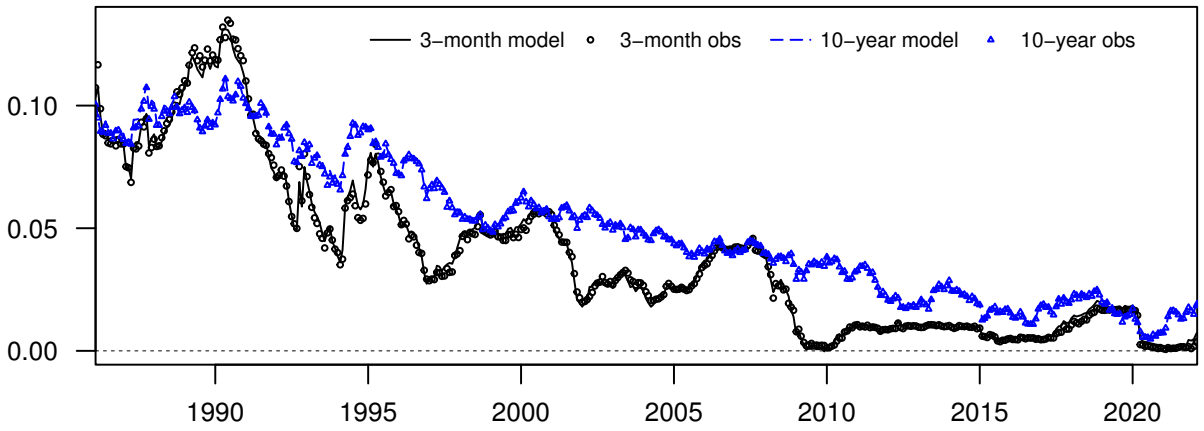


Notes: Time series of DTAFNS model-implied factors which correspond to the smoothed state inferences $E^{\mathbb{P}}\left[x_t^{(i)}|\hat{y}(1),\hat{y}(2),...,\hat{y}(T)\right]$, $i = 1,2,3$ provided by Algorithm 1, and implied short rates obtained by summing smoothed values of the first two factors.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

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 Figure 2.2. 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.

Figure 2.2: 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.

2.3.3 Performance assessment and benchmarking

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

Benchmarks

The discrete-time Gaussian three-factor model of Augustyniak, Godin, and Hamel 2021, subsequently referred to as the DG3 model, and a version of the dynamic Nelson-Siegel (DNS) model introduced by Diebold and Li 2006 are benchmarks considered in this study for comparison of performance results. See Appendix A.1.2 for a formal specification of these two models. The main differences between DG3 and the DTAFNS model is 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, Godin, and Hamel 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.

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 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 by summing the log-likelihoods associated with all the test set folds using an expanding window sequential approach. Table 2.2 provides the results. Whereas the in-sample log-likelihood of the DG3 is higher than that of the DTAFNS and the DNS models, the out-of-sample value for the DTAFNS model is higher than that of the DNS and DG3 benchmarks for the aggregate of all test years. This implies better distributional predictive ability for the DTAFNS model.

To visualize the ability of the DTAFNS model and benchmarks to replicate observed yield curves, Figure 2.3 presents the realized spot curves and model-implied counterparts for four selected dates. These four days are the same that are considered in Augustyniak, Godin, and Hamel 2021 to display fitting performance, namely December 29, 2006, December 31,

Table 2.2: Log-likelihood of the DTAFNS model and its benchmarks

Model	Out-of-sample						In-sample
	2017	2018	2019	2020	2021-22	Aggregated	
DTAFNS	2,034	2,001	2,035	2,015	2,183	10,268	65,702
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 2.2.2 and of the DG3 and DNS benchmarks described in Appendix A.1.2. 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 2.3.3. The aggregated out-of-sample log-likelihood is ultimately obtained by summing the log-likelihood for all test years.

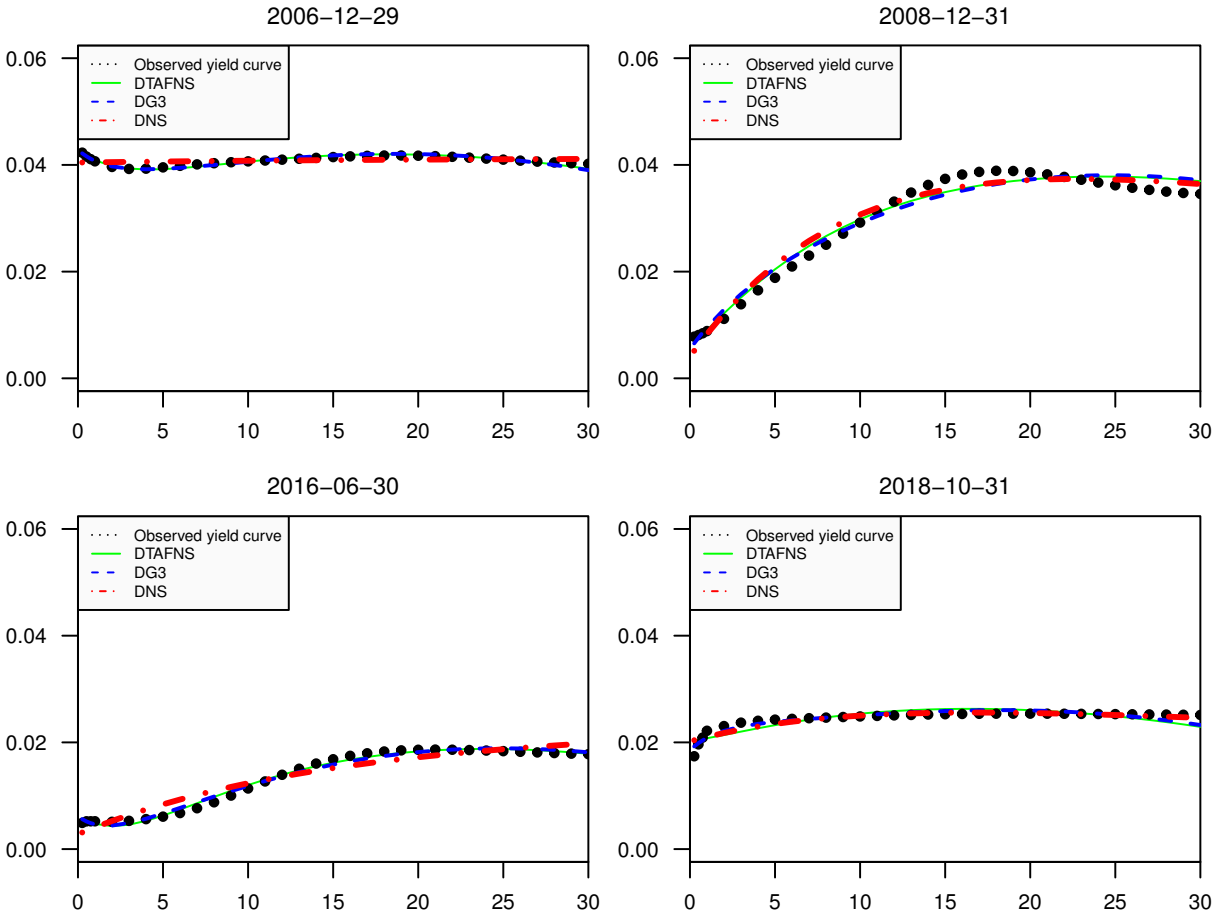
2008, June 30, 2016 and October 31, 2018. Such dates exhibit several spot curve shapes: flat, upward sloping or humped. The figure shows that the DTAFNS and both benchmarks are able to reproduce observed yield curves reasonably well.

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 Table 2.1, Table A.1 and Table A.2, respectively, and starting values for factors being their smoothed values on January 2022.

Table 2.3 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 both the DG3 and DNS models, which is desirable.

Figure 2.3: 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. The observed data are end-of-month Canadian spot rates provided by the Bank of Canada.

Table 2.3: 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
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 A.1.2, 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 Table 2.1, Table A.1, and Table A.2, respectively, with starting values of the factors being their smoothed values on January 2022.

2.4 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, Godin, and Hamel 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 improve predictive performance over the DG3 benchmark, it is relevant to adapt the Augustyniak, Godin, and Hamel 2021 mixed fund model by replacing the DG3 model with the DTAFNS model.

2.4.1 Rolling bond fund returns

The dynamics of the mixed fund model of Augustyniak, Godin, and Hamel 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, Godin, and Hamel 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 $R_{t+1}^{(\tau)}$, (2.2.15) leads to

$$\begin{aligned} R_{t+1}^{(\tau)} &= \log\left(\frac{P(t+1,t+\tau)}{P(t,t+\tau)}\right) \\ &= \log A_{\tau-1} - \log A_\tau - \Delta B_{\tau-1}^\top X_{t+1} + \Delta B_\tau^\top X_t \\ &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} B_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(\frac{B_\tau^{(i)}}{B_{\tau-1}^{(i)}}\right) X_t^{(i)} \right). \end{aligned} \tag{2.4.1}$$

Moreover, recalling that $r_t = X_t^{(1)} + X_t^{(2)}$ and using (2.2.14), (2.2.18), Lemma A.1.2 and (2.4.1), the excess return can be expressed as

$$\begin{aligned} R_{t+1}^{(\tau)} - \Delta r_t &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} B_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(\frac{B_\tau^{(i)} - 1}{B_{\tau-1}^{(i)}}\right) X_t^{(i)} \right) + \Delta \sum_{i=1}^{3} X_t^{(i)} - \Delta(X_t^{(1)} + X_t^{(2)}) \\ &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{2} B_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(\frac{B_\tau^{(i)} - 1}{B_{\tau-1}^{(i)}}\right) X_t^{(i)} \right) + \Delta X_t^{(3)} \\ &\quad - \Delta B_{\tau-1}^{(3)} \left(X_{t+1}^{(3)} - \left(\frac{B_\tau^{(3)} - 1}{B_{\tau-1}^{(3)}}\right) X_t^{(3)} \right) \\ &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{2} B_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)} \right) + \Delta X_t^{(3)} \\ &\quad - \Delta B_{\tau-1}^{(3)} \left(X_{t+1}^{(3)} - \left((1-\lambda) - \frac{(1-\lambda)^{\tau-1}}{B_{\tau-1}^{(3)}} \right) X_t^{(3)} \right) \\ &= \log A_{\tau-1} - \log A_\tau - \Delta \sum_{i=1}^{3} B_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)} \right) + \Delta \left(1 - (1-\lambda)^{\tau-1}\right) X_t^{(3)}. \end{aligned} \tag{2.4.2}$$

Mimicking Augustyniak, Godin, and Hamel 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 $R_{t+1}^{(B)}$, is approximately characterized by

$$R_{t+1}^{(B)} - \Delta r_t \approx \omega_1 R_{t+1}^{(\tau_1)} + \cdots + \omega_M R_{t+1}^{(\tau_M)} - \Delta r_t$$

$$
\begin{aligned}
&= \omega_1 \left(R_{t+1}^{(\tau_1)} - \Delta r_t \right) + \cdots + \omega_M \left(R_{t+1}^{(\tau_M)} - \Delta r_t \right) \\
&= \sum_{j=1}^{M} \omega_j \Bigg(\log A_{\tau_j - 1} - \log A_{\tau_j} - \Delta \sum_{i=1}^{3} B_{\tau_j - 1}^{(i)} \left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}} \right) X_t^{(i)} \right) \\
&\quad + \Delta \left(1 - (1-\lambda)^{\tau_j - 1} \right) X_t^{(3)} \Bigg) \\
&= \sum_{j=1}^{M} \omega_j \left(\log A_{\tau_j - 1} - \log A_{\tau_j} \right) - \Delta \sum_{i=1}^{3} \sum_{j=1}^{M} \omega_j B_{\tau_j - 1}^{(i)} \left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}} \right) X_t^{(i)} \right) \\
&\quad + \Delta \sum_{j=1}^{M} \omega_j \left(1 - (1-\lambda)^{\tau_j - 1} \right) X_t^{(3)}. \qquad (2.4.3)
\end{aligned}
$$

Relationship (2.4.3) 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, Godin, and Hamel 2021 model is the presence of the last term proportional to $X_t^{(3)}$ in (2.4.3), which is not present in their work.

2.4.2 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, Godin, and Hamel 2021. Returns of q equity assets are represented with the exponential generalized auto-regressive conditional heteroskedastic (EGARCH) model of Nelson 1991:

$$
R_{t+1,j}^{(S)} - \Delta r_t = \lambda_j^{(S)} \sqrt{h_{t,j}^{(S)}} - \frac{1}{2} h_{t,j}^{(S)} + \sqrt{h_{t,j}^{(S)}} Z_{t+1,j}^{(S)}, \qquad (2.4.4)
$$

$$
\log h_{t,j}^{(S)} = \omega_j^{(S)} + \alpha_j^{(S)} Z_{t,j}^{(S)} + \gamma_j^{(S)} \left(|Z_{t,j}^{(S)}| - \frac{2}{\sqrt{2\pi}} \right) + \beta_j^{(S)} \log h_{t-1,j}^{(S)}, \qquad (2.4.5)
$$

where $(\omega_j^{(S)}, \alpha_j^{(S)}, \gamma_j^{(S)}, \beta_j^{(S)})$ are parameters associated with the conditional volatility. The process $\{h_{t,j}^{(S)}\}_{t=1}^{T-1}$ is $\mathcal{F}$-adapted. Parameters $\lambda_j^{(S)}$ represent equity risk premia. For $j = 1, \ldots, q$, $Z_j^{(S)} \equiv \{Z_{t,j}^{(S)}\}_{t=1}^{T}$ are standard Gaussian white noises under $\mathbb{P}$, independent of Z_t, where the correlation between $Z_{t,i}^{(S)}$ and $Z_{t,j}^{(S)}$ is denoted $\Gamma_{i,j}$.

As in Augustyniak, Godin, and Hamel 2021, the following risk-neutral dynamics for the q equity assets is assumed:

$$R_{t+1,j}^{(S)} - \Delta r_t = -\frac{1}{2} h_{t,j}^{(S)} + \sqrt{h_{t,j}^{(S)}} Z_{t+1,j}^{\mathbb{Q}(S)}, \tag{2.4.6}$$

$$\log h_{t,j}^{(S)} = \omega_j^{(S)} + \alpha_j^{(S)} (Z_{t,j}^{\mathbb{Q}(S)} - \lambda_j^{(S)}) + \gamma_j^{(S)} \left(|Z_{t,j}^{\mathbb{Q}(S)} - \lambda_j^{(S)}| - \frac{2}{\sqrt{2\pi}} \right) + \beta_j^{(S)} \log h_{t-1,j}^{(S)}, \tag{2.4.7}$$

where, under the risk-neutral measure $\mathbb{Q}$, $\{Z_{t,j}^{\mathbb{Q}(S)}\}_{t=1}^T = \{Z_{t,j}^{(S)} + \lambda_j^{(S)}\}_{t=1}^T$ 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 $\{Z_{t,j}^{\mathbb{Q}(S)}\}_{t=1}^T$ and interest rate innovations $\{Z_t^{\mathbb{Q}}\}_{t=1}^T$ is assumed to also hold under $\mathbb{Q}$.

2.4.3 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

$$\begin{aligned} \tilde{R}_{t+1}^{(F)} - \Delta r_t &\approx W(R_{t+1}^{(B)} - \Delta r_t) + \sum_{j=1}^q \tilde{\psi}_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t \right) \\ &= \tilde{\psi}_0 + \sum_{i=1}^3 \tilde{\psi}_i \left(X_{t+1}^{(i)} - \left(1 - \kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)} \right) + \tilde{\psi}_3' X_t^{(3)} + \sum_{j=1}^q \tilde{\psi}_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t \right) \end{aligned} \tag{2.4.8}$$

where

$$\tilde{\psi}_0 = W \sum_{j=1}^M \omega_j \left(\log A_{\tau_j - 1} - \log A_{\tau_j} \right), \tag{2.4.9}$$

$$\tilde{\psi}_i = -W\Delta \sum_{j=1}^M \omega_j B_{\tau_j - 1}^{(i)}, \quad i = 1, 2, 3, \tag{2.4.10}$$

$$\tilde{\psi}_3' \;=\; W\Delta \sum_{j=1}^{M} \omega_j \left(1-(1-\lambda)^{\tau_j-1}\right). \tag{2.4.11}$$

Equation (2.4.8) serves as the conceptual basis for the proposed mixed fund model. However, instead of relying on (2.4.9)-(2.4.11), 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, Godin, and Hamel 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

$$R_{t+1}^{(F)} - \Delta r_t = \psi_0 + \sum_{i=1}^{3} \psi_i \left(X_{t+1}^{(i)} - \left(1-\kappa_{i,i}^{\mathbb{Q}}\right) X_t^{(i)}\right) + \psi_3' X_t^{(3)} + \sum_{j=1}^{q} \psi_j^{(S)} \left(R_{t+1,j}^{(S)} - \Delta r_t\right) + \sqrt{h_t^{(F)}} Z_{t+1}^{(F)}, \tag{2.4.12}$$

$$\log h_t^{(F)} = \omega^{(F)} + \alpha^{(F)} Z_t^{(F)} + \gamma^{(F)} \left(|Z_t^{(F)}| - \frac{2}{\sqrt{2\pi}}\right) + \beta^{(F)} \log h_{t-1}^{(F)}, \tag{2.4.13}$$

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 $(\omega^{(F)}, \alpha^{(F)}, \gamma^{(F)}, \beta^{(F)})$ and linear coefficients $(\psi_0, \psi_1, \psi_2, \psi_3, \psi_3', \psi_1^{(S)}, \ldots, \psi_q^{(S)})$ are model parameters to be estimated.

The risk-neutral dynamics can also be obtained along the lines of Augustyniak, Godin, and Hamel 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.1.1.

Proposition 2.4.1. *The risk-neutral dynamics of mixed fund excess returns are given by*

$$R_{t+1}^{(F)} - \Delta r_t = -\frac{1}{2}\left(\sigma_t^{(F)}\right)^2 + \sigma_t^{(F)}\epsilon_{t+1}^{\mathbb{Q}(F)} \tag{2.4.14}$$

where

$$\left(\sigma_t^{(F)}\right)^2 \equiv \sum_{i=1}^{3}\sum_{l=1}^{3}\psi_i\psi_l\Sigma_{i,i}\Sigma_{l,l}\rho_{i,l} + \sum_{j=1}^{q}\sum_{k=1}^{q}\psi_j^{(S)}\psi_k^{(S)}\Gamma_{j,k}\sqrt{h_{t,j}^{(S)}h_{t,k}^{(S)}} + h_t^{(F)}, \tag{2.4.15}$$

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

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

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

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

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

2.4.4 Estimation of the mixed fund model

The mixed fund model developed above is estimated on real data. The equity model (2.4.4) and (2.4.5) 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 (2.4.12)-(2.4.13), we consider the *Assumption/Louisbourg Balanced Fund A*.[3] This is a mixed bond and equity fund composed approximately of 39% of Canadian fixed income, 36% of Canadian equity, 15% of US equity,

3. https://assumption.lipperweb.com/assumplife/list#FundDetail

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 which 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.

Table 2.4 and Table 2.5 respectively show maximum likelihood estimated parameter and corresponding standard errors for the equity model (2.4.4)-(2.4.5) and the mixed fund model (2.4.12)-(2.4.13).[4] 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, Godin, and Hamel 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).

2.5 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

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

Table 2.4: 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 (2.4.4)-(2.4.5). 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 2.5: 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 (2.4.12)-(2.4.13). The estimation is performed on a monthly NAV return time serie for the Assumption/Louisbourg Balanced Fund A extending from February 1996 to January 2022 (312 returns).

Hong, Niu, and Zeng 2016. 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 reveals 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, Godin, and Hamel 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, Godin, and Hamel 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.

Chapter 3

Participating Life Insurance Contract

3.1 Introduction

Participating life insurance policies are life insurance contracts that provide some participation in the insurance company's profits. They appear under different mechanisms and were extremely popular before the appearance of equity-linked contracts. They can provide higher returns to the policyholder than standard insurance policies since they are linked, among others, to the financial markets. These higher returns are driven by better than foreseen financial returns, favorable mortality experience, and cost savings. Unlike equity-linked products which are purely tied to broad market performance through an index or a mutual fund, participating contracts are tied to the insurer's performance.

The participation received by the policyholder is often referred to as bonuses or dividends, which can take different forms (Booth et al. 2020). Policyholders can use dividends to reduce premiums. Alternatively, dividends can be invested in a risk-free savings account or be received in cash. They can also be used as premium to purchase additional one-year term insurance. Lastly, they can also serve to increase the contract benefits. Examples

of participating contracts available for sale in North America and Europe include fully variable life insurance and fixed premium variable life insurance. For both types, insurance companies share investment profits by increasing associated policy reserves at the end of each year, with a new benefit level being extracted from such augmented reserves.

Pioneering work in the participating contracts literature is Wilkie 1987 who proposes to use option pricing theory to evaluate the implicit options embedded in the policies. Key assumptions required to evaluate such options are the reference portfolio dynamics, mortality dynamics, and the financial framework. Diverse financial models are proposed in the academic literature to depict financial risks inherent to various types of life insurance policies including equity-linked insurance, variable annuities, and participating contracts, see for instance to Briys and De Varenne 2001 or Hardy 2003.

In the context of participating contracts, earlier works such as Grosen and Jørgensen 2000, Hansen and Miltersen 2002, Tanskanen and Lukkarinen 2003, Haberman, Ballotta, and Wang 2003, Bauer et al. 2006, Gatzert and Kling 2007 or Gatzert, Holzmüller, and Schmeiser 2012 often rely on simple assumptions either in their theoretical development or in their numerical implementation, such as constant interest rates, the Black-Scholes-Merton financial model for equity returns, or absence of mortality. Simple assumptions possess the advantage of enhancing the tractability of the model, while sacrificing some realism. For instance, when working with policies with long maturities, considering fluctuations in interest rates is a desirable property. The Black-Scholes-Merton model is furthermore unable to explain the heteroskedasticity, the positive excess kurtosis, and negative skewness of the asset returns. More recent works generalize such assumptions and include stochastic interest rates (Zaglauer and Bauer 2008), non-trivial mortality rates (Bacinello 2003) or non-Gaussian financial returns (Zheng et al. 2019). However, the inclusion of such realistic modeling features is much more scarce in works considering participating contracts than in the literature dedicated

to the related equity-linked or variable annuity contracts, where the inclusion of stochastic interest rate (for example Bacinello and Ortu 1994; Nielsen and Sandmann 1995; Bacinello and Persson 2002; Bernard, Le Courtois, and Quittard-Pinon 2005; Kijima and Wong 2007; Tiong 2013), or heteroskedastic/fat tailed equity returns (for instance Hardy 2003; Lin, Tan, and Yang 2009; Ng, Li, and Chan 2011; Fan 2013) is much more prevalent.

Due to the scarcity of studies on participating contracts embedding stochastic dynamics for the various risk drivers (including equity risk, interest risk, and mortality risk), the formulation of many existing quantitative frameworks applicable to the analysis of participating contracts is not sufficiently general to handle randomness in the entire set of risk factors. For instance, the variable benefits life insurance approach from Bowers et al. 1997 reflects deterministic equity returns, interest rate, and mortality assumptions. Furthermore, formulas from Bacinello 2003 based the compound reversionary bonus (CRB) method, which is typically used for participating contracts in the UK, also only reflect deterministic interest rates and mortality. An exception is the cliquet style bonus distribution scheme studied in Zaglauer and Bauer 2008, which includes a Vasicek process for the risk-free rate.

The main objective of this study is therefore to bridge that gap in the literature and to extend participating contract analysis schemes to include realistic dynamics for the main risk drivers, including stochastic interest rates, and mortality. We focus on the specific case of fixed premium contracts where bonuses are paid out as benefit increases. First, formulas for bonus calculations underlying the variable benefits life insurance approach from Bowers et al. 1997 and the CRB approach of Bacinello 2003 are generalized to handle general stochastic dynamics. Pricing, reserving and risk measurement approaches are then discussed in such context. In particular, the stochastic nature of risk drivers requires the introduction of a so-called *shadow reserve* representing the genuine actuarial

present value of future cash flows. The shadow reserve can be integrated into liabilities evaluation and risk metric calculations to replace conventional reserving formulas and provide additional accuracy. Several numerical tests based on Monte-Carlo simulations are presented to outline the importance of considering such shadow reserve instead of the conventional reserves when measuring risk, as the latter fails to consider future benefit increases beyond the guaranteed rate. In simulations, dynamics of the separate account involve returns driven by (i) a mix of equity and fixed income assets, whose modeling is borrowed from Eghbalzadeh, Godin, and Gaillardetz 2022 and reflects heteroskedasticity (through GARCH effects) and multi-factor term structure evolution, and (ii) mortality experience including idiosyncratic mortality effects. Such stylized facts are meant to improve the realism of the separate account dynamics with respect to models currently used in the participating contracts literature.

The paper is organized as follows. Section 3.2 reviews standard (i.e. non-participating) endowment life insurance contracts along with actuarial notation. The participating contract cash flow mechanisms and benefits dynamics considered herein are presented in Section 3.3. Section 3.4 focuses on the evaluation of the contract, including fair premium identification, reserving, and risk measurement. Section 3.5 introduces the stochastic financial models used to represent the reference account dynamics. Results from simulation experiments are presented in Section 3.6. Section 3.7 concludes.

3.2 Standard endowment life insurance

This section first introduces the standard actuarial notation describing mortality for a cohort of policyholders. Standard endowment contracts, which will serve as a basis for the introduction of participating contracts in subsequent sections, are then discussed.

3.2.1 Actuarial notations for mortality

Yearly time periods with yearly time points $0, \ldots, n$ are considered. For simplicity, a homogeneous cohort of N policyholders aged x at time $k = 0$ with independent and identically distributed lifetimes is considered. Let T_i be the future lifetime of insured i, $i = 1, 2, \ldots, N$. Moreover, let ${}_kp_x$ and ${}_{k|}q_x$ respectively denote the probability for any of the policyholders to survive k years, or to survive k years and die within the following year: for $i = 1, 2 \ldots, N$ and $k = 0, 1, 2, \ldots, n$,

$$ {}_kp_x = \mathbb{P}(T_i \geq k), \quad {}_{k|}q_x = \mathbb{P}(k \leq T_i < k+1). $$

Define a counting process $L = \{L_{x+k}\}_{k=0}^n$ giving the number of policyholders still alive at any time k:

$$ L_{x+k} = \sum_{i=1}^{N} \mathbb{1}_{\{T_i \geq k\}}, \quad k = 0, 1, \ldots, n $$

where $\mathbb{1}_A$ is the binary indicator function for event A. The number of deaths within the cohort in the time interval $[k, k+1)$ is represented by $D_{x+k} = L_{x+k} - L_{x+k+1}$. The filtration generated by the process L is denoted by $\{\mathcal{H}_k\}_{k=0}^n$. A financial market is also introduced in subsequent sections, whose information is carried through a filtration $\{\mathcal{F}_k\}_{k=0}^n$. The combined information from both mortality and the financial market is thus denoted by $\{\mathcal{G}_k\}_{k=0}^n$ with $\mathcal{G}_k = \mathcal{F}_k \vee \mathcal{H}_k$. We suppose independence between the financial market and mortality events in the cohort of policyholders.

3.2.2 Endowment life insurance

Standard endowment life insurance contracts are now introduced. Related concepts will allow introducing participating contracts in the subsequent section. Consider a fully discrete indexed n-year endowment insurance policy. In this case, the insurer pays a

death benefit at the end of the year of death if the policyholder i dies before age $x+n$ (i.e. before time n). Otherwise a survival benefit is paid at the time n. Let b_0 denote the initial death benefit that is paid if a death occurs in the first year. Both death and survival benefits increase by an indexing factor of $(1+g)$ per year of survival. In exchange of the future benefit, the policyholder pays annual level premiums Π at the beginning of each year before n as long as he/she is alive.

Denote the bank account numéraire by B, where $B_0 = 1$. The total time-k insurance company's prospective loss is defined as the difference between the present value of future benefits and the present value of future premiums (including that at time k):

$$\mathcal{L}_k = \sum_{l=0}^{n-k-1} b_{k+l} D_{x+l+k} \frac{B_k}{B_{k+l+1}} + b_n L_{x+n} \frac{B_k}{B_n} - \Pi \sum_{l=0}^{n-k-1} L_{x+l+k} \frac{B_k}{B_{k+l}}, \tag{3.2.1}$$

where

$$b_k = b_0(1+g)^k \tag{3.2.2}$$

is the death benefit associated with deaths during period $[k, k+1)$. For the endowment life insurance contract, in (3.2.1), the benefits amounts are deterministic as seen in (3.2.2), while the number of survivals and deaths are random.

The equivalence premium principle can be used to price these life insurance products. This approach entails setting the premium in a way to equate expected present values of the stream of premiums and of the benefit, which is $\mathbb{E}[\mathcal{L}_0] = 0$. See Bowers et al. 1997 for more information. In the present work, the valuation is performed with a risk-neutral measure $\mathbb{Q}$, as required by recent IFRS 17 norms which require market-consistent valuation (see paragraphs BC153 of the basis of conclusions IASB 2017). This leads to

$$\mathbb{E}^{\mathbb{Q}}[\mathcal{L}_0] = \mathbb{E}^{\mathbb{Q}}\left[\sum_{l=0}^{n-1} \frac{b_0(1+g)^l D_{x+l}}{B_{l+1}} + \frac{b_0(1+g)^n L_{x+n}}{B_n} - \Pi^{np} \sum_{l=0}^{n-1} \frac{L_{x+l}}{B_l}\right] = 0. \tag{3.2.3}$$

with the superscript in Π^{np} emphasizing that such premium is the fair premium for the non-participating contract. Let $P(k,k+l) = \mathbb{E}^{\mathbb{Q}}\left[B_k/B_{k+l}|\mathcal{F}_k\right]$ be the time-k price of a risk-free zero-coupon bond with unit face value maturing at time $k+l$. Furthermore, for a policyholder aged $x+k$ at time k, denote by

$$\ddot{a}_{x+k:\overline{n-k}|} = \sum_{l=0}^{n-k-1} P(k,k+l)_l p_{x+k}, \tag{3.2.4}$$

the acturial present value (APV) of a $(n-k)-$year temporary life annuity and by

$$A_{x+k:\overline{n-k}|g} = \sum_{l=0}^{n-k-1} (1+g)^l P(k,k+l+1)_{l|}q_{x+k} + (1+g)^{n-k}P(k,n)_{n-k}p_{x+k}, \tag{3.2.5}$$

the APV of an indexed $(n-k)-$year endowment insurance with indexing rate g. Using the independence assumption between the mortality and financial risk, and substituting (3.2.4) and (3.2.5) into (3.2.3), allows obtaining the fair premium:

$$\Pi^{np} = b_0 \frac{A_{x:\overline{n}|g}}{\ddot{a}_{x:\overline{n}|}}. \tag{3.2.6}$$

The time-k *policy reserve* ${}_kV(\Pi, b_k)$ for the premium and current benefit levels Π and b_k, respectively, is defined as the APV of the corresponding prospective total loss $\mathcal{L}_k$, which can be obtained through

$$\begin{aligned} L_{x+k}\ {}_kV(\Pi,b_k) &= \mathbb{E}^{\mathbb{Q}}\Bigg[\sum_{l=0}^{n-k-1} b_k(1+g)^l D_{x+k+l}\frac{B_k}{B_{k+l+1}} + b_k(1+g)^{n-k}L_{x+n}\frac{B_k}{B_n} \\ &\qquad -\Pi \sum_{l=0}^{n-k-1} L_{x+k+l}\frac{B_k}{B_{k+l}}\Bigg|\mathcal{G}_k\Bigg], \quad k=0,1,\cdots,n. \end{aligned}$$

This leads to

$${}_kV(\Pi,b_k) = b_k A_{x+k:\overline{n-k}|g} - \Pi\ddot{a}_{x+k:\overline{n-k}|}. \tag{3.2.7}$$

Due to (3.2.3), ${}_0V(\Pi^{np}, b_0) = 0$. Moreover for any value of Π, ${}_nV(\Pi, b_n) = b_n$.

3.3 Participating policies

This section introduces participating life insurance contracts. Participation to insurers performance can be embedded in various ways into the contracts, with the participation dividends having several possible forms as mentioned by Booth et al. 2020: decrease in premiums, immediate dividend payoffs or increased benefits and coverage. This study focuses on the latter case, with premiums being fixed and benefits being adjusted yearly through either one of the two following mechanisms that are considered: a standard approach, which is a generalization of the variable benefits life insurance approach from Bowers et al. 1997 that we propose, and the compound reversionary bonus approach detailed for instance in Booth et al. 2020 or Bacinello 2003. In the proposed setting, dividends are permanently integrated into the contract; benefits can only be increased and can never be decreased.

For participating contracts considered in this work, the basic insurance component underlying each participating contract is a discrete $n-$year endowment insurance policy with annual premiums. Death and survival benefits increase yearly through the mechanisms described subsequently, subject to a minimum rate of increase g that is guaranteed. The total premium charged to the policyholder is split into two parts: $\Pi^{tot} = \Pi + \Pi^*$. The portion Π is taken into consideration when determining benefits increases, while the second part Π^* is not. The rationale for considering such subdivision of the premium is explained in further detail below, where it is shown that enforcing $\Pi^* = 0$ can sometimes lead to the non-existence of a fair premium.

3.3.1 Standard variable benefit increase mechanism

A first benefit increase mechanism that we propose, referred to subsequently as the *standard mechanism*, is a generalization of the variable benefits life insurance approach from Bowers et al. 1997, which we design to handle randomness in financial returns, interest rates, and mortality. Indeed, the setting described in Bowers et al. 1997 is fully deterministic, which therefore requires adaptation to be applicable in the stochastic context of the present work.

For the pool of participating contracts, the portion Π of premiums as well as the policy reserves are invested in a separate account called the *reference account*. When using the standard mechanism, adjustments to the policyholder benefits are determined based on the reference account performance, which is impacted among others by mortality experience and reserve levels for the pool of policyholders. On each year, the reference account earns investment returns. Participation to the reference account performance by policyholders is only partial: on each period, the insurer deducts a portion θ from the account return to cover participation guarantees and different expenses.

The reference account value thus evolves according to the following dynamics. At the beginning of year k, the insurer collects the premiums for all policyholders that are alive: the total amount of $L_{x+k}\Pi$ is therefore added to the reference account. Then, the reference account earns log-return $R_{k+1}^{(F)}$, from which a portion θ is removed. Finally, the death benefits b_k are withdrawn from the reference account at the end of the year. Denoting by $\mathcal{A}_k$ the time-k reference account value, this leads to[1]

$$\mathcal{A}_{k+1} = (\mathcal{A}_k + L_{x+k}\Pi)\left\{1 + \left(\exp\left(R_{k+1}^{(F)}\right) - 1 - \theta\right)\right\} - D_{x+k}b_k, \quad k = 0, 1, ..., n-1. \quad (3.3.1)$$

Then, at the end of year k, the updated account value $\mathcal{A}_{k+1}$ is used as input to the

1. Although the account value $\mathcal{A}_{k+1}$ given by (3.3.1) is not bounded below by zero and could in theory become negative, such situation is not encountered in the simulation experiments present in subsequent sections. Thus, modifications to (3.3.1) would ensure non-negative account values are not considered.

calculation of next period's benefit b_{k+1}. The benefit adjustment is meant to redistribute to all policyholders the amounts accrued into the reference account while imposing a minimum rate of increase g. The benefit update formula is

$$b_{k+1} = \max\left\{ \frac{\mathcal{A}_{k+1}}{L_{x+k+1}} + \Pi\ddot{a}_{x+k+1:\overline{n-k-1}|} \over A_{x+k+1:\overline{n-k-1}|g}, (1+g)b_k \right\}, \quad k = 0, 1, ..., n-1. \tag{3.3.2}$$

Indeed, when $b_{k+1} > (1+g)b_k$ for $k \leq n-1$, this ensures that

$$b_{k+1}A_{x+k+1:\overline{n-k-1}|g} = \frac{\mathcal{A}_{k+1}}{L_{x+k+1}} + \Pi\ddot{a}_{x+k+1:\overline{n-k-1}|}, \tag{3.3.3}$$

which reflects that the sum of the account value and APV of premiums is equal to the APV of benefits, assuming these would only increase at the guaranteed rate in the future.

Such benefit increase does not lead to actuarial equivalence between the outflows and the inflows plus reserve due to (i) the presence of a minimum benefit increase rate g, (ii) the fact that benefits might increase at a higher rate than g in the future which is disregarded when considering $A_{x+k+1:\overline{n-k-1}|g}$ on the left-hand side of (3.3.3), (iii) the portion θ of returns that will be deducted from future account returns, and (iv) the portion Π^* of the premium not accounted for in the revised benefit calculation.

When the benefit increases beyond its minimum rate of increase g, relation (3.3.3) implies that the difference between the new benefit b_{k+1} and that which would have been provided by a non-participating contract $b'_{k+1} \equiv b_0(1+g)^{k+1}$ is, using (3.2.7), given by

$$(b_{k+1} - b'_{k+1}) = \frac{1}{A_{x+k+1:\overline{n-k-1}|g}} \left(\frac{\mathcal{A}_{k+1}}{L_{x+k+1}} - {}_{k+1}V(\Pi, b'_{k+1}) \right). \tag{3.3.4}$$

Benefits are therefore dependent on the insurer performance through the difference between the restrospective reserve and the prospespective policy reserve, namely

$$\frac{\mathcal{A}_{k+1}}{L_{x+k+1}} - \ {}_{k+1}V(\Pi, b'_{k+1}).$$

3.3.2 Compound Reversionary Bonus Methods

Booth et al. 2020 describes an alternative mechanism to determine dividends based on reserve levels. This approach is referred to as the *compound reversionary bonus* (CRB) method and is commonly used, for instance, in the United Kingdom (Bacinello 2003). An extension of such mechanism incorporating for instance stochastic interest rates is considered in the present paper. This extension is used for comparison with the standard approach (3.3.2) in numerical experiments of subsequent sections.

Regular benefit adjustment for the compound reversionary bonus method

The CRB approach relies on the idea that updated benefits should be increased in a way that the policy reserve grows at some bonus rate reflecting reference account performance. A bonus rate with a similar structure than in Bacinello 2003 is considered,[2] which consists in the adjustment rate δ_k defined as

$$\delta_k = \max\left(\frac{\exp\left(R_k^{(F)}\right) + m_k - \theta - B_k/B_{k-1}}{B_k/B_{k-1}}, 0\right),$$

where m_k is the mortality excess return. δ_k reflects the excess rate of the return over the risk-free rate obtained by the reference account due to financial returns and mortality experience, from which a portion θ is deducted.[3] The excess of expected mortality costs

2. Here we assume that the risk-free rate is used as the so-called technical rate in Bacinello 2003.

3. Generally, the adjustment rate δ_k is defined by the excess of the return on the reference account over the expected return plus the excess of the expenses over the occurred ones (Booth et al. 2020). However, consistently with (Bacinello 2003), expenses are not considered in the present work.

over actual costs is given by

$$m_k = \frac{b_{k-1}(q_{x+k-1}L_{x+k-1} - D_{x+k-1}) + \ {}_kV(\Pi, b_{k-1})(p_{x+k-1}L_{x+k-1} - L_{x+k})}{{}_kV(\Pi, b_{k-1})L_{x+k}} \mathbb{1}_{\{{}_kV(\Pi, b_{k-1})>0\}},$$

where ${}_kV(\Pi, b_{k-1})$ is the pre-adjustment reserve (the policy reserve right before the benefit annual adjustment). It corresponds to the reserve assuming that there is no benefit increase at time k.

In the CRBM, the new benefit b_k is set in a way to reflect the application of the adjustment rate δ_k to the pre-adjustment reserve ${}_kV(\Pi, b_{k-1})$. Such adjustment is achieved through the following updating rule for the benefit:

$$b_k = b_{k-1} \max\left((1 + \Psi_k \delta_k),\ 1 + g\right), \quad \text{where} \tag{3.3.5}$$

$$\Psi_k = \frac{\max\left(0, {}_kV(\Pi, b_{k-1})\right)}{b_{k-1} A_{x+k:\overline{n-k}|g}}. \tag{3.3.6}$$

Indeed, if the benefit increase is above the minimal rate g, i.e. $b_k > b_{k-1}(1+g)$, and the pre-adjustment reserve is positive, i.e. ${}_kV(\Pi, b_{k-1}) > 0$, the reserve adjustment is, using (3.2.7) and (3.3.5),

$$\begin{aligned}
{}_kV(\Pi, b_k) - \ {}_kV(\Pi, b_{k-1}) &= b_k A_{x+k:\overline{n-k}|g} - \Pi \ddot{a}_{x+k:\overline{n-k}|} - b_{k-1} A_{x+k:\overline{n-k}|g} + \Pi \ddot{a}_{x+k:\overline{n-k}|} \\
&= (b_k - b_{k-1})\, A_{x+k:\overline{n-k}|g} \\
&= b_{k-1} \Psi_k \delta_k A_{x+k:\overline{n-k}|g} \\
&= \delta_k \ {}_kV(\Pi, b_{k-1}).
\end{aligned}$$

An important difference between the CRB method and the proposed standard variable increase mechanism is that the former uses the prospective policy reserve as a basis for applying the bonus rate, whereas the latter uses the account value calculated retrospectively to calculate the redistribution of the surplus over the policy reserve.

3.4 Pricing, reserving, and risk evaluation for participating contracts

3.4.1 The shadow reserve

As mentioned above, the policy reserve (3.2.7) is not a good depiction of the insurer's liability for participating contracts since such formula does not reflect future participation bonuses embedded in benefits and the portion of the premium Π^* not reflected in such calculation. This leads to considering a quantity referred to as the *shadow reserve*, which is defined as the APV of all future cash flows to the policyholder related to the participating contract. The shadow reserve differs from the policy reserve in the way the benefits are considered; the true future benefits are considered in the former instead of the minimum guaranteed benefit reflected in the latter. The shadow reserve represents a best estimate of the amount that the insurance company needs to set aside to fulfill its future obligations related to the contracts.

Definition 3.4.1. *For premium parts* Π *and* Π^**, the time-k shadow reserve, denoted* $C_k(\Pi, \Pi^*)$*, is the risk-neutral conditional expectation of discounted future benefits minus discounted future premiums: for* $k = 0, ..., n-1$*,*

$$C_k(\Pi, \Pi^*) = \mathbb{E}^{\mathbb{Q}}\left[\sum_{l=0}^{n-k-1} b_{l+k} D_{x+l+k} \frac{B_k}{B_{k+l+1}} + b_n L_{x+n} \frac{B_k}{B_n} - \sum_{l=0}^{n-k-1} (\Pi + \Pi^*) L_{x+k+l} \frac{B_k}{B_{k+l}} \middle| \mathcal{G}_k \right], \tag{3.4.1}$$

with benefits only being updated considering the premium part Π *as exhibited in above sections.*

3.4.2 The fair premium of a participating contract

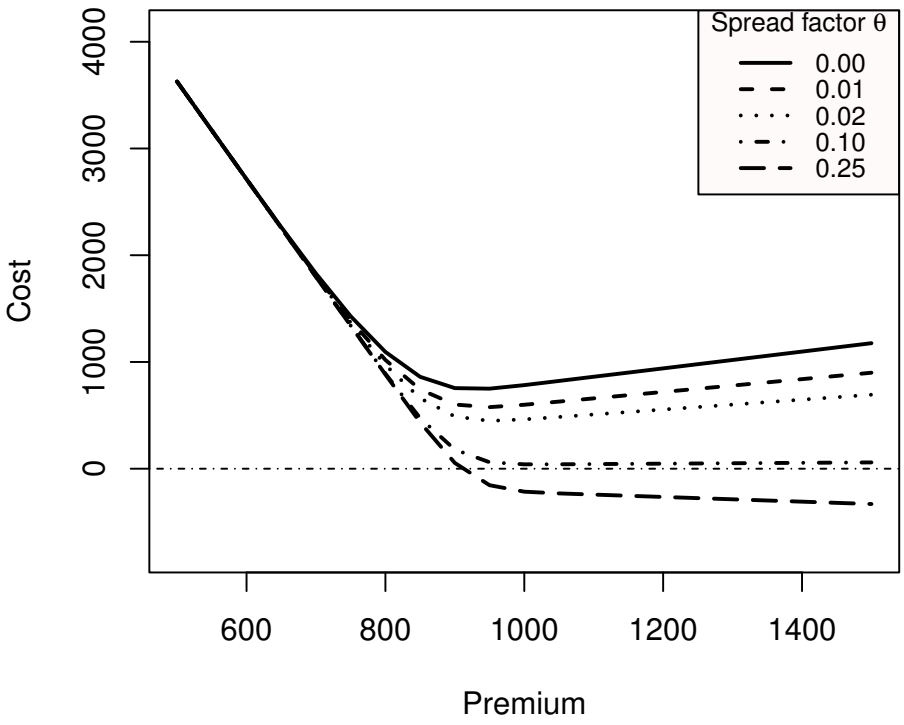
Participating contracts are also typically priced using the equivalence premium principle (Booth et al. 2020). A fair premium of the participating contract is a duplet (Π, Π^*) for which the initial shadow reserve is set to zero, i.e. $C_0(\Pi, \Pi^*) = 0$. Because participating mechanisms (3.3.2) or (3.3.5) entail a complex non-linear relationship between benefits and premiums, isolating Π in the equation $C_0(\Pi, \Pi^*) = 0$ cannot be done and fair premiums must be calculated numerically.

A natural choice would consist in setting $\Pi^* = 0$ and having the full premium $\Pi^{tot} = \Pi$ being recognized within benefits increases. Unfortunately, the fair premium defined under this assumption does not always exist, e.g. when the participation deduction parameter θ is too small. Figure 3.1 shows the initial shadow reserve $C_0(\Pi, 0)$ estimated through Monte-Carlo simulations versus the premium Π for different spreads θ under the standard benefit increase mechanism detailed in Section 3.3.1. The financial model used to simulate cash flows is explained in Section 3.5, and associated parameters are given in Appendix A.2. For small values of θ, the shadow reserve does not hit zero, implying non-existence of the fair premium.

When using the standard benefit increase mechanism from Section 3.3.1, an alternative procedure is therefore considered. We set $\Pi = \Pi^{np}$ as defined in (3.2.6), i.e. benefits are only incremented based on what the premium would be if the contract was non-participating. Then the excess part of the premium Π^* is adjusted to obtain a fair premium: $C_0(\Pi^{np}, \Pi^*) = 0$.

Conversely, when using the CRB approach from Section 3.3.2 for benefit increases, we instead set $\Pi^{tot} = \Pi$, and therefore $\Pi^* = 0$. Indeed, cases of non-existence of a fair premium were never encountered in numerical experiments that were performed with the CRB mechanism.

Figure 3.1: Premium vs cost concerning the different spread factors (θ).



Notes: For a participating contract based on the standard benefit increase mechanism detailed in Section 3.3.1, this figure reports the initial shadow reserve $C_0(\Pi, 0)$ estimated through Monte-Carlo simulations ($100,000$ simulated paths) versus the premium Π for different spread values θ. In this example, the initial benefit is $b_0 = 10{,}000$. There are $L_x = 1000$ policyholders at the onset. The term of the contract is $n = 10$ years. The guaranteed benefit increase rate is $g = 0$. The financial model is explained in Section 3.5 and parameters are reported in Appendix A.2.

3.4.3 Risk measurement

From the insurer's point of view, participating contracts are risky and need to be well managed, which means the risk level to the insurer needs to be evaluated.

A first way to do this consists in comparing the prospective reserve (i.e. the shadow reserve) with a retrospective reserve representing the cash flows accumulated by the insurer to cover future obligations. Such retrospective reserve is represented by a modified version of the account value (3.3.1) that includes additional amounts which are charged by the insurers but which are not reflected when calculated benefit increases. We therefore introduce an *augmented reference account* process $\mathcal{A}^*$ represented by the following for $k = 0, 1, \ldots, n-1$:

$$\mathcal{A}^*_{k+1} = \left(\mathcal{A}^*_k + L_{x+k}\Pi^{tot}\right)\left(\exp\left(R^{(F)}_{k+1}\right) - \theta^*\right) - D_{x+k}b_k - L_{x+n}b_n\mathbb{1}_{\{k=n-1\}}, \tag{3.4.2}$$

with $\mathcal{A}^*_0 = 0$ and $\theta^* \leq \theta$ being the portion of the participation deduction serving company expenses. The gap between the shadow reserve and the augmented reference account represents the current losses (gains when negative) associated with the issue of the policy.

A metric commonly used for measuring the contract risk is the *net amount at risk* (NAAR). In the literature, the NAAR is often defined for non-stochastic setting (e.g. Bowers et al. 1997). Hence, we extend the definition for a stochastic framework.

Definition 3.4.2. *The net amount at risk is defined as the difference between benefits and the augmented reference account: for* $k = 0, 1, \ldots, n$,

$$_kNAAR \equiv b_k - \frac{\mathcal{A}^*_k}{L_{x+k}} - \Pi^{tot}\mathbb{1}_{\{k<n\}}. \tag{3.4.3}$$

The NAAR can be understood as a one-year term risk exposure.

3.5 Financial model

The financial market model considered in subsequent simulation experiments, which is drawn directly from Eghbalzadeh, Godin, and Gaillardetz 2022, is hereby provided. Unlike in previous sections where yearly time periods are considered, the models from the latter work were developed for monthly periods. Relations tying yearly returns to monthly values generated by the models are provided.

3.5.1 Interest rate term structure model

The year-k bank account numéraire is obtained by accruing risk-free interest over $12k$ months:

$$B_k = \exp\left(\Delta \sum_{t=0}^{12k-1} r_t\right), \quad k = 1, \ldots, n$$

with $\Delta = 1/12$ and r_t being the (annualized) risk-free short rate for the $t+1^{th}$ monthly period. The interest rate term structure dynamics is modeled by the discrete-time arbitrage-free Nelson-Siegel (DTAFNS) model of Eghbalzadeh, Godin, and Gaillardetz 2022 which is numerically tractable, easily interpretable and which provides an accurate depiction of realized term structures. Model parameters considered in subsequent numerical experiments, which are presented in Appendix A.2 are also drawn from Eghbalzadeh, Godin, and Gaillardetz 2022. Such parameters were calibrated on historical Canadian term structure data. See Appendix A.2 for more details. The DTAFNS model is a three-factor stochastic short-rate model. The short rate is the sum of the first two:

$$r_t \equiv X_t^{(1)} + X_t^{(2)}, \tag{3.5.1}$$

where the state variable process $\{X_t\}_{t=0}^{12n}$ defined with time-t factors $X_t = [X_t^{(1)}, X_t^{(2)}, X_t^{(3)}]$ follows the following auto-regressive dynamics under the physical measure $\mathbb{P}$:

$$\begin{bmatrix} X_{t+1}^{(1)} - X_t^{(1)} \\ X_{t+1}^{(2)} - X_t^{(2)} \\ X_{t+1}^{(3)} - X_t^{(3)} \end{bmatrix} \equiv \underbrace{\begin{bmatrix} -\Sigma_{11}\gamma_1 & 0 & 0 \\ 0 & \lambda - \Sigma_{22}\gamma_2 & -\lambda \\ 0 & 0 & \lambda - \Sigma_{33}\gamma_3 \end{bmatrix}}_{\kappa^{\mathbb{P}}} \underbrace{\begin{bmatrix} 0 - X_t^{(1)} \\ \theta_2^{\mathbb{P}} - X_t^{(2)} \\ \theta_3^{\mathbb{P}} - X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{P}} - X_t} + \tag{3.5.2}$$
$$\underbrace{\begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}}_{\Sigma} \begin{pmatrix} Z_{t+1,1}^{\mathbb{P}} \\ Z_{t+1,2}^{\mathbb{P}} \\ Z_{t+1,3}^{\mathbb{P}} \end{pmatrix},$$

with $(\theta^{\mathbb{P}}, \kappa^{\mathbb{P}}, \Sigma)$ being model parameters, $\{Z_{t,i}^{\mathbb{P}}\}_{t=1}^{12n}, i = 1, 2, 3$ being $\mathcal{F}$-adapted standard Gaussian white noises with contemporaneous correlation parameters $\rho_{ij} = \text{corr}(Z_{t,i}^{\mathbb{P}}, Z_{t,j}^{\mathbb{P}})$, $t = 1, \ldots, T$ and $i, j = 1, 2, 3$. Risk-neutral dynamics of the factor process are

$$X_{t+1} = X_t + \kappa^{\mathbb{Q}}(\theta^{\mathbb{Q}} - X_t) + \Sigma Z_{t+1}^{\mathbb{Q}}, \tag{3.5.3}$$

where

$$\kappa^{\mathbb{Q}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix} \text{ and } \theta^{\mathbb{Q}} = \begin{pmatrix} 0 \\ (\theta_2^{\mathbb{P}}\kappa_{22}^{\mathbb{P}} + \theta_3^{\mathbb{P}}\Sigma_{33}\gamma_3)/\lambda \\ \theta_3^{\mathbb{P}}\kappa_{33}^{\mathbb{P}}/\lambda \end{pmatrix}. \tag{3.5.4}$$

and processes $Z_i^{\mathbb{Q}} = \{Z_{t,i}^{\mathbb{Q}}\}_{t=1}^{12n}$, $i = 1, 2, 3$ defined through $Z_{t+1,i}^{\mathbb{Q}} = Z_{t+1,i}^{\mathbb{P}} - \gamma_i X_t^{(i)}$ also are $\mathcal{F}$-adapted standard Gaussian white noises under $\mathbb{Q}$ with contemporaneous dependence parameters ρ_{ij}, $i, j = 1, 2, 3$.

Eghbalzadeh, Godin, and Gaillardetz 2022 show that under this model the month-t price

of a risk-free zero-coupon bond paying one dollar on month T is

$$P(t,T) = A(t,T)\exp\left[-\Delta\mathcal{B}(t,T)^\top X_t\right], \tag{3.5.5}$$

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

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

$$\log A(t,T) = -\Delta\theta_2^{\mathbb{Q}}\left(\mathcal{B}^{(1)}(t,T) - \mathcal{B}^{(2)}(t,T)\right) + \Delta\theta_3^{\mathbb{Q}}\mathcal{B}^{(3)}(t,T) + \frac{1}{2}\Delta^2\upsilon_\tau,$$

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

$$\upsilon_\tau^{(1,1)} = \Sigma_{11}^2\frac{\tau(\tau-1)(2\tau-1)}{6},$$

$$\upsilon_\tau^{(2,2)} = \frac{\Sigma_{22}^2}{\lambda^2}\left(\tau - 2\left[\frac{1-(1-\lambda)^\tau}{\lambda}\right] + \frac{1-(1-\lambda)^{2\tau}}{1-(1-\lambda)^2}\right),$$

$$\upsilon_\tau^{(3,3)} = \frac{\Sigma_{33}^2}{\lambda^2}\Bigg[\tau - 2 + \zeta_0\left((1-\lambda)^2,\tau-1\right) + \lambda^2\zeta_2\left((1-\lambda)^2,\tau-1\right)$$
$$- 2\zeta_0\left((1-\lambda),\tau-1\right) - 2\lambda\zeta_1\left((1-\lambda),\tau-1\right) + 2\lambda\zeta_1\left((1-\lambda)^2,\tau-1\right)\Bigg],$$

$$\upsilon_\tau^{(1,2)} = \upsilon_\tau^{(2,1)} = \rho_{12}\Sigma_{11}\Sigma_{22}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda),\tau)\right),$$

$$\upsilon_\tau^{(1,3)} = \upsilon_\tau^{(3,1)} = \rho_{13}\Sigma_{11}\Sigma_{33}\frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0\left((1-\lambda),\tau-1\right) - (\lambda+1)\zeta_1\left((1-\lambda),\tau-1\right)$$
$$- \lambda\zeta_2\left((1-\lambda),\tau-1\right)\Bigg],$$

$$\upsilon_\tau^{(2,3)} = \upsilon_\tau^{(3,2)} = \rho_{23}\Sigma_{22}\Sigma_{33}\Bigg(\frac{\tau - 2 - (2-\lambda)\zeta_0\left((1-\lambda),\tau-1\right) + (1-\lambda)\zeta_0\left((1-\lambda)^2,\tau-1\right)}{\lambda^2}$$
$$+ \frac{-\zeta_1\left((1-\lambda),\tau-1\right) + (1-\lambda)\zeta_1\left((1-\lambda)^2,\tau-1\right)}{\lambda}\Bigg)$$

and

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

$$\zeta_1(r,\tau) \equiv \sum_{u=1}^{\tau-1} u r^u = \frac{r - \tau r^\tau + (\tau-1) r^{\tau+1}}{(1-r)^2},$$
$$\zeta_2(r,\tau) \equiv \sum_{u=1}^{\tau-1} u^2 r^u = \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1) r^{\tau+1} - \tau^2 r^\tau + r^2 + r}{(1-r)^3}.$$

3.5.2 Reference account dynamics

The reference account of participating life insurance contracts in which reserves are invested is typically composed of a mix of equity and fixed income securities. The mixed fund model of Eghbalzadeh, Godin, and Gaillardetz 2022 which represents dynamics of funds invested in both these two classes of asset is therefore used in this work. Such model is an adaptation of the mixed fund model of Augustyniak, Godin, and Hamel 2021 which uses the DTAFNS term structure model instead of their conventional three-factor Gaussian model.

Again, the considered model parameters are borrowed from Eghbalzadeh, Godin, and Gaillardetz 2022, who use monthly time periods in their study. The year-k reference account annual log-return $R_k^{(F)}$ is thus represented as the sum of the monthly log-returns $\tilde{R}_k^{(F)}$:

$$R_k^{(F)} = \sum_{t=1+12(k-1)}^{12k} \tilde{R}_t^{(F)}.$$

The model first considers the dynamics of q equity indices, where index j's monthly excess log-return dynamics are driven by an EGARCH process:

$$\tilde{R}_{t+1,j}^{(S)} - \Delta r_t = \lambda_j^{(S)} \sqrt{h_{t,j}^{(S)}} - \frac{1}{2} h_{t,j}^{(S)} + \sqrt{h_{t,j}^{(S)}} Z_{t+1,j}^{(S)}, \tag{3.5.6}$$
$$\log h_{t,j}^{(S)} = \omega_j^{(S)} + \alpha_j^{(S)} Z_{t,j}^{(S)} + \gamma_j^{(S)} \left(|Z_{t,j}^{(S)}| - \frac{2}{\sqrt{2\pi}} \right) + \beta_j^{(S)} \log h_{t-1,j}^{(S)}, \tag{3.5.7}$$

where $(\omega_j^{(S)}, \alpha_j^{(S)}, \gamma_j^{(S)}, \beta_j^{(S)})$ are parameters associated with the conditional volatility, $\lambda_j^{(S)}$ is the equity risk premium parameter, $\{h_{t,j}^{(S)}\}_{t=0}^{12n-1}$ is the GARCH volatility process, and $Z_j^{(S)} = \{Z_{t,j}^{(S)}\}_{t=1}^{12n}$, $j = 1, \ldots, q$ are standard Gaussian white noise under $\mathbb{P}$ independent of

Z_t whose contemporaneous correlation is $\Gamma_{ij} = \text{corr}\left(Z_{t,i}^{(S)}, Z_{t,j}^{(S)}\right)$. The EGARCH process was introduced by Nelson 1991, which has the advantageous properties of embedding Black 1976's leverage effect, of not requiring constraints on parameters during the calibration to ensure positive volatilities and of better capturing the persistence of the volatility shocks than conventional GARCH processes.

The month-t reference account returns $\tilde{R}_{t+1}^{(F)}$ is then represented through an econometric model driven by (i) a linear relationship with respect to term structure shocks and excess returns on the equity indices, and (ii) a basis risk term independent of above terms which also follows an EGARCH process:

$$\tilde{R}_{t+1}^{(F)} - \Delta r_t = \psi_0 + \sum_{i=1}^{3} \psi_i \left(X_{t+1}^{(i)} - \left(1 - \kappa_{ii}^{\mathbb{Q}}\right) X_t^{(i)} \right) + \psi_3' X_t^{(3)} + \sum_{j=1}^{q} \psi_j^{(S)} \left(\tilde{R}_{t+1,j}^{(S)} - \Delta r_t \right) + \sqrt{h_t^{(F)}} Z_{t+1}^{(F)}, \tag{3.5.8}$$

$$\log h_t^{(F)} = \omega^{(F)} + \alpha^{(F)} Z_t^{(F)} + \gamma^{(F)} \left(|Z_t^{(F)}| - \frac{2}{\sqrt{2\pi}} \right) + \beta^{(F)} \log h_{t-1}^{(F)}, \tag{3.5.9}$$

where $(\psi_0, \psi_1, \psi_2, \psi_3, \psi_3', \psi_1^{(S)}, \ldots, \psi_q^{(S)})$ are parameters associated with the linear mapping between the account returns and risk factors, $(\omega^{(F)}, \alpha^{(F)}, \gamma^{(F)}, \beta^{(F)})$ are parameters associated with the conditional volatility process $\{h_t^{(F)}\}_{t=1}^{12n-1}$, and $Z_t^{(F)} := \{Z_t^{(F)}\}_{t=1}^{12}$ is a standard Gaussian white noise under $\mathbb{P}$ independent of Z_t and $Z_j^{(S)}$, $j = 1, \ldots, q$.

The considered risk-neutral dynamics of the equity indices and account returns, also taken from Eghbalzadeh, Godin, and Gaillardetz 2022, are

$$R_{t+1}^{(F)} - \Delta r_t = -\frac{1}{2}\left(\sigma_t^{(F)}\right)^2 + \sigma_t^{(F)} \epsilon_{t+1}^{\mathbb{Q}(F)}, \tag{3.5.10}$$

where

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

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

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

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

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

and processes $\{Z_{t,j}^{\mathbb{Q}(S)}\}_{t=1}^{12n} = \{Z_{t,j}^{(S)} + \lambda_j^{(S)}\}_{t=1}^{12n}$, $j = 1, \ldots, q$ and $\{Z_t^{\mathbb{Q}(F)}\}_{t=1}^{12n} = \{Z_t^{(F)} + \lambda_{t-1}^{(F)}\}_{t=1}^{12n}$ are also standard Gaussian white noises under $\mathbb{Q}$ retaining the same dependence structure than their physical measure counterpart.

For numerical experiments of the subsequent sections, parameters considered are again drawn from Eghbalzadeh, Godin, and Gaillardetz 2022 and are presented in Appendix A.2. Two equity indices are considered, namely the S&P TSX and S&P 500. The account return $R_{t+1}^{(F)}$ parameters were estimated on a time series of NAV returns of the *Assumption/Louisbourg Balanced Fund A* mixed fund. See Appendix A.2 for details about the time series considered.

3.6 Numerical analyses

This section performs numerical experiments and implements both dividend mechanisms for participating contracts within the proposed financial framework. The objective is to obtain insight about the level of premiums, the evolution of benefits, and the exposure to risk of the insurer for the two benefit increase mechanisms and various contract specifications.

Our examples involve fully discrete ten-year ($n = 10$) fixed premium endowment

participating contracts issued to $L_x = 1000$ individuals aged $x = 50$-years-old. Unless specified, a baseline case is considered in which the initial benefit is $b_0 = \$10{,}000$, the guaranteed benefit increase is nil ($g = 0$), and no participation spread is considered ($\theta = \theta^* = 0$). The mortality follows the CPM2014 male mortality table (CIA 2014). The results provided subsequently are reported in terms of "per policyholder" ratios, i.e. they are normalized by L_x. The Monte Carlo simulation uses 100,000 generated paths. We respectively denote by Π_{st}^{tot} and Π_{CRB}^{tot} total premiums (i.e. $\Pi + \Pi^*$) obtained with the standard mechanism (3.3.1) and the compound reversionary bonus method (3.3.5), respectively.

3.6.1 Fair premiums

For the aforementioned baseline scenario, the fair premium for the non-participating contract is $\Pi^{np} = 896.53$, whereas it is $\Pi_{st}^{tot} = 979.44$ or $\Pi_{CRB}^{tot} = 1162.15$ for the participating contracts. As a result, the participation feature raises the fair premium by more than 9% or 29% over that of the non-participating contract when using the standard or the CRB mechanism, respectively.

We now analyze the impact on the premium of the policyholder age at initiation and the contract maturity. Table 3.1 provides premiums for maturities $n = 10, 15, 20, 30$ and policyholder ages $x = 50, 51, \ldots, 60$, and 70. As expected, the premium increases with age and decreases with the maturity. In all cases, premiums obtained with the standard mechanism are smaller than those obtained with the CRB approach. This reflects the fact that dividends are generally smaller for the standard contracts than for contracts using the CRB mechanism.

Table 3.2 shows the fair premium for different values of the guaranteed rate of benefit increase ($g = 0$, 1%, 2%, 3% or 5%) and of the loading factor ($\theta = 0$, 0.5%, 1%, 1.5%, 2% or 5%). As expected, the premium increases with the guaranteed rate g and decreases

Table 3.1: Total premium for different ages, maturities and benefit increase mechanisms.

x \ n	Non-participating				Standard				CRB			
	10	15	20	30	10	15	20	30	10	15	20	30
50	897	567	411	281	979	650	498	379	1162	819	643	469
51	898	568	413	284	981	651	500	381	1163	820	645	471
52	900	570	415	287	982	653	502	384	1165	822	646	474
53	902	572	418	291	984	655	504	386	1166	823	648	477
54	904	575	420	295	986	657	506	389	1168	825	650	480
55	906	577	423	300	988	659	509	392	1169	826	652	484
56	908	580	426	305	990	662	511	396	1171	828	654	489
57	910	582	430	311	992	664	514	400	1173	830	656	494
58	912	585	433	317	994	666	517	405	1174	832	659	500
59	914	588	437	325	996	669	521	411	1176	834	662	506
60	916	591	442	333	998	672	524	417	1178	836	665	514
70	966	664	544	491	1045	738	616	558	1216	890	744	667

Notes: Total premiums (Π^{tot}) for different benefit increase mechanisms (Standard, CRB), maturities ($n = 10, 15, 20, 30$) and ages ($x = 50, 51, ..., 60, 70$), where b_0 = 10,000, $\theta = 0$, and $g = 0$. The premiums are obtained by equating the initial cost C_0 to 0, using Monte Carlo simulations with 100,000 generated paths. They are rounded to the nearest integer. The financial variables dynamics are described in Section 3.5.

with the loading factor θ. The premium obtained with the standard participation contract mechanism is more sensitive to changes in g than that obtained with the CRB mechanism. Indeed, in the baseline case, when g goes from 0% to 1%, the standard mechanism premium increases by 10.3%, but such increase is nearly halved (5.4%) for the CRB mechanism. The evolution of dividends explains such behavior. On average, the first dividend from the standard approach is higher than that of the CRB. Even though CRB dividends begin at a lower level, the mechanism provides higher dividend increases than the standard approach over the course of the contract. Since the minimum guarantee imposes a lower limit to these dividend increases, the standard mechanism dividends are more impacted than these from the CRB, which leads to a higher impact on the premium.

Table 3.2: Total premium for different loading factors θ, minimum guarantee rates g, and mechanisms.

θ / g	Standard participating						CRB					
	0	0.005	0.01	0.015	0.02	0.05	0	0.005	0.01	0.015	0.02	0.05
0	979	970	963	956	951	931	1162	1147	1132	1117	1103	1030
0.01	1080	1070	1061	1054	1048	1026	1225	1210	1195	1182	1168	1100
0.02	1190	1179	1169	1161	1155	1130	1299	1285	1272	1259	1246	1184
0.03	1309	1297	1287	1278	1271	1244	1386	1373	1360	1347	1336	1279
0.05	1582	1568	1555	1545	1536	1503	1595	1583	1572	1562	1552	1506

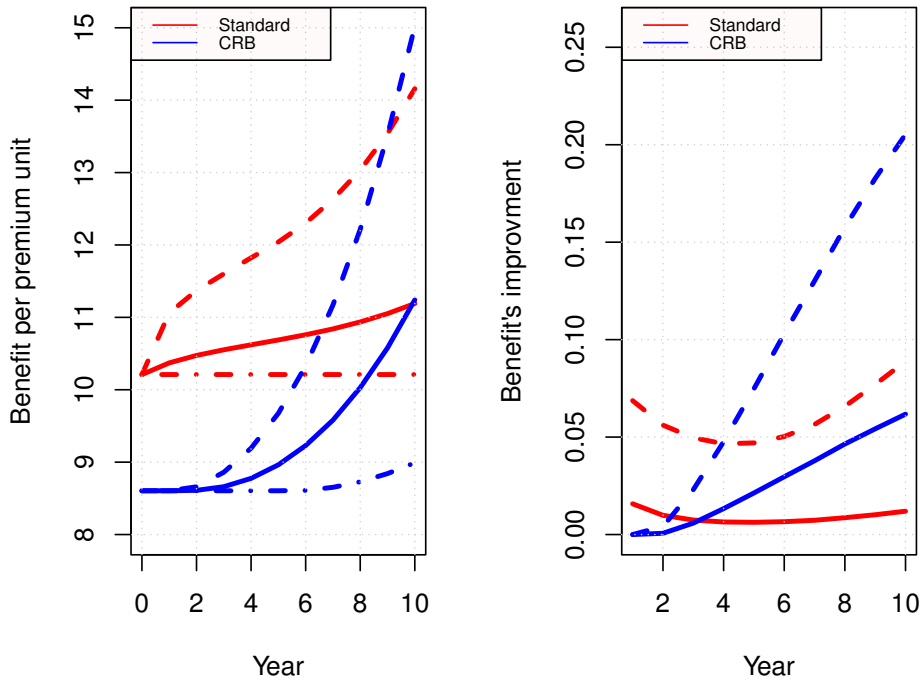
Notes: Total premiums (Π^{tot}) for different benefit increase mechanisms (Standard, CRB), guarantee rates ($g = 0, 1\%, 2\%, 3\%, 5\%$), and loading factors ($\theta = 0, 0.5\%, 1\%, 1.5\%, 2\%, 5\%$), where $b_0 = 10{,}000$, $x = 50$, and $n = 10$. They are evaluated using Monte Carlo simulations with 100,000 paths and are rounded to the nearest integer. The financial variables dynamics are described in Section 3.5.

3.6.2 Analysis of benefits

We now analyze the benefits generated by each mechanism. For comparison purposes, the ratios of benefits over premiums are considered. Figure 3.2 illustrates the average and the corresponding 95%-level confidence interval across the simulated paths of such ratios (left-hand side) and of benefit improvements $\frac{b_k - b_{k-1}}{b_{k-1}}$ (right-hand side) versus time t for each mechanism. The standard approach provides higher benefits during the first years compared to the CRB mechanism. On average, the survival benefit is approximately \$11 per unit of premium for both mechanisms. It represents an average return on investment of 1.73% per year for both approaches, assuming survival at maturity. However, the CRB benefits are more volatile than these of the standard approach when the contracts approach maturity. The investment returns contingent on survival can go up to 7.26% (or as low as 0.4%) if we consider the best (or worst) 2.5% scenarios.

The standard participation mechanism leads to higher benefit increases in the first year (about 1.6% on average), with the rate of increase declining in subsequent years. In contrast, the CRB mechanism has the least benefit growth in the initial years (close to

Figure 3.2: Evolution of the benefit per premium ratio for each mechanism.



Notes: Left-hand side figure: *Solid lines* are the mean (across generated paths) of the benefit per premium ratio, and *Dash lines* are the 95% confidence intervals. Right-hand side: *Solid lines* are the benefit improvement average, and *Dash lines* are 97.5%-level quantiles. Note that the 2.5%-level quantiles are 0 for both mechanisms. All numbers are obtained using $b_0 = 10{,}000$, $x = 50$, $g = \theta = 0$, $n = 10$, and 100,000 simulated paths. The financial variables dynamics are described in Section 3.5.

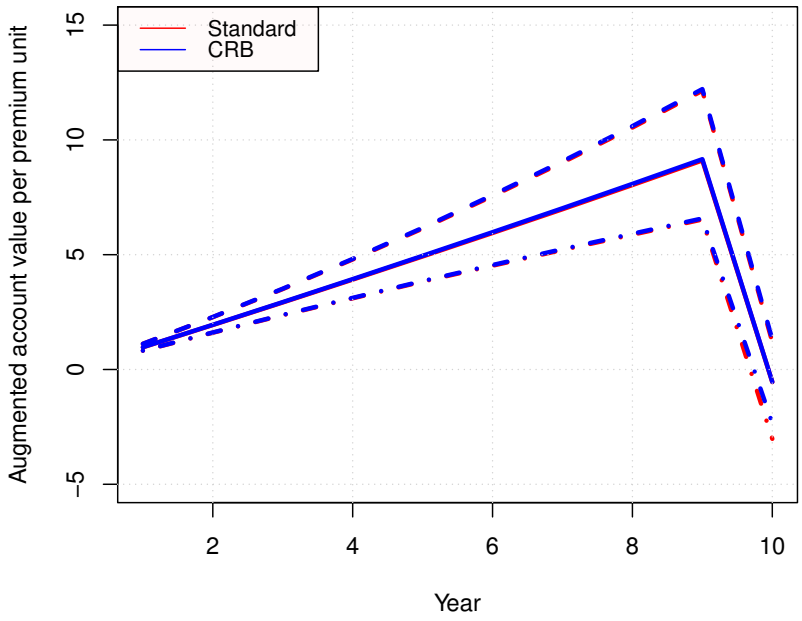
0%), with the rate of increase climbing up quickly up to 6.2% in average at contract maturity. In general, the benefit growth rate for the standard participation mechanism is higher than that of the CRB method in the first three years, whereas it becomes lower in the following years.

3.6.3 Risk exposure

We now analyze implications about risk associated with each mechanism for the insurer. The reference account evolution can lead to possible future deficits for the participating contracts line of business. Figure 3.3 depicts the evolution over the contract's life of the sample average and 5% confidence intervals, across simulated paths, of the augmented reference account value (3.4.2) per premium unit. Negative values for the augmented reference accounts imply that issuers need to inject money into the line of business, which represents potential liquidity risk. These ratios behave similarly for both benefit increase mechanisms. They have a positive trend at the beginning, and they decrease sharply at maturity when survival benefits are paid. In fact, due to low mortality rates, these ratios are almost identical for both mechanisms since little benefits are paid in the early years. Negative values for the ratio are more probable towards the end of the contract. Such negative values are expected since a fair premium principle and the absence of participation charges ($\theta = \theta^* = 0$) are considered in our simulation. For the standard approach, the probability of observing a shortfall at maturity is $\mathbb{P}(\mathcal{A}_n^* < 0) = 0.599$, whereas the expected shortfall per premium unit (and per policyholder) is $\mathbb{E}^{\mathbb{P}}[\mathcal{A}_n^* | \mathcal{A}_n^* < 0]/(\Pi^{tot} L_x) = -1.449$. These statistics for the CRB mechanism are 0.692 and -1.106, respectively.

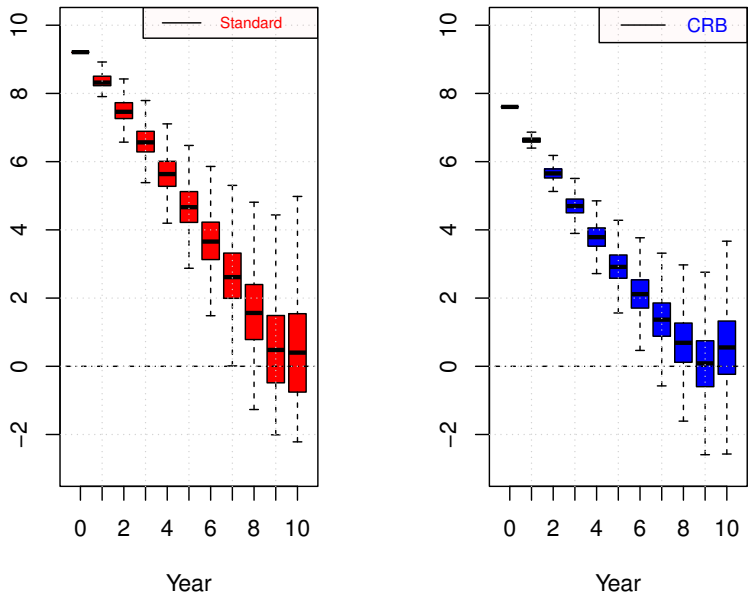
We also analyze the NAAR evolution (see Definition 3.4.2) to examine the insurer's exposure to losses caused by sudden deaths. Figure 3.4 shows the boxplots representing the NAAR unconditional distributions obtained over the various simulated paths, versus time. Both benefit increase mechanisms exhibit similar patterns with a decreasing trend for the NAAR evolution. Indeed, at inception, the insurer is exposed to the initial death benefit $b_0 = 10{,}000$ while it has only received the initial premium. The collection of annual premiums and investment returns over time progressively dampens the NAAR.

Figure 3.3: Time evolution of the augmented account value per premium unit and policyholder.



Notes: Sample distribution versus time t of $\mathcal{A}_t^*/(\Pi^{tot} L_x)$, the augmented account value normalized per premium unit and per policyholder. Results are presented for the standard and CRB benefit increase mechanisms. *Solid lines*: average. *Dash lines*: level-97.5% quantile, *Point and dash lines*: level-2.5% quantile. All numbers are obtained using $b_0 =$ 10,000, $x = 50$, $g = \theta = 0$, $n = 10$, and 100,000 simulated paths. The financial variables dynamics are described in Section 3.5.

Figure 3.4: Boxplots of the net amount at risk (NAAR) for each mechanism



Notes: NAAR sample distribution versus time t. The sample scenarios of NAAR values for both standard and CRB mechanisms are depicted. The NAAR is computed using Monte Carlo simulation, as described in Definition 3.4.2. All numbers are obtained using $b_0 = 10{,}000$, $x = 50$, $g = \theta = 0$, $n = 10$, and 100,000 simulated paths. The financial variables dynamics are described in Section 3.5.

3.6.4 Analysis of reserves

As mentioned above, the policy reserve (3.2.7) is insufficient to cover the participating life insurance liabilities as a result of potential benefit enhancements that are not recognized. This section analyses gaps between the shadow reserve (3.4.1) and the policy reserve to determine the extent of underestimation of liabilities when using the policy reserve. The shadow reserve is also compared to the augmented account value defined in (3.4.2) to assess the magnitude of potential actuarial shortfalls.

Stochastic on stochastic framework

Evaluating the shadow reserve at any given point in time requires performing a Monte Carlo simulation. Therefore, depicting the evolution of the shadow reserves across multiple simulated paths requires applying a so-called *stochastic on stochastic* framework involving nested simulations. Under this approach, multiple paths of state variables are first generated under the physical probability measure, each of these being referred to as *outer loops*. Then, on each time point of an outer loop, a nested Monte Carlo simulation referred to an *inner loop* is conducted to evaluate the shadow reserve. The inner loop is performed under the risk-neutral measure, conditionally on the current values of the state variables provided by the outer loop.[4] State variables summarizing the relevant information up to time k include the number of policyholders alive L_{x+k}, the current benefit level b_k, the account value $\mathcal{A}_k$, equity indices conditional variances $h^{(S)}_{12k,j}$, $j = 1, 2$, the fund-related variance $h^{(F)}_{12k}$ and the term structure factors $X^{(i)}_{12k}, i = 1, 2, 3$. Our simulations involve 10,000 outer loops, each with 10,000 inner loops on any time step of an outer loop.

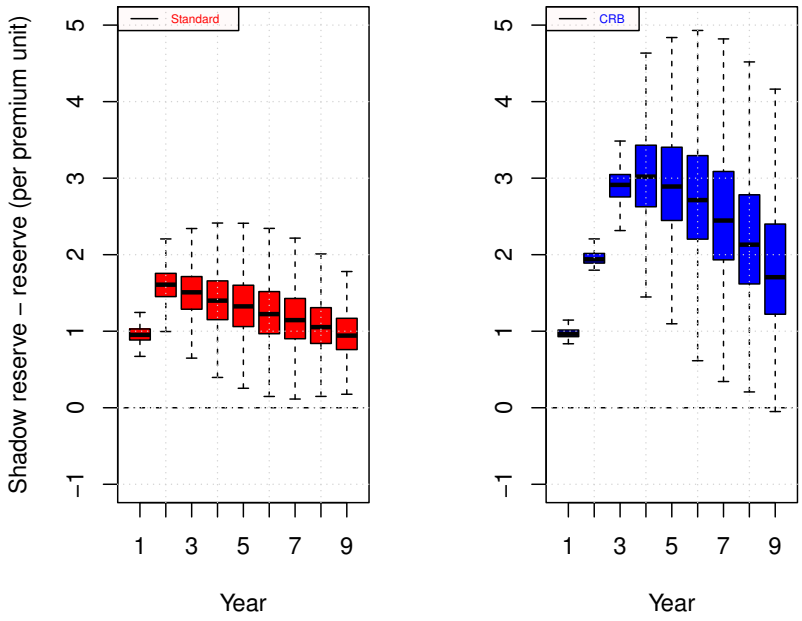
4. The shadow reserve, therefore, depicts the market value of future cash flows to the insurer, not their actual expected value.

Stochastic on stochastic simulation results

Figure 3.5 exhibits distributions across the paths of differences between shadow and policy reserves per premium and policyholder for each contract year. As expected, the policy reserve underestimates future liabilities in most scenarios. Therefore, insurance companies relying on policy reserves instead of shadow reserves to assess future liabilities expose themselves to possible shortfalls. For both mechanisms, gaps between reserves are shrinking as time passes, which can be explained by smaller potential for benefit increases when approaching maturity. The main distinction between the mechanisms is the size of gaps between the augmented account value and the shadow reserve. In the standard participating mechanism, the average gap increases from 1 after year 1 to 1.59 after year 2, before decreasing to 0.91 after year 9. Conversely, for the CRB mechanism, the average gap first increases from 0.98 to 3 between the first and the fourth year, and then gradually decreases to 1.43 at time $t = 9$. Higher gaps associated with the CRB mechanism are caused by its higher rate of increase for the benefit, as outlined above by Figure 3.2. Such higher benefit increases are not fully reflected by the policy reserve, hence causing higher gaps. For both mechanisms, the gap rarely reaches zero, indicating that the misrepresentation of liabilities stemming from the use of the policy reserve is systematic.

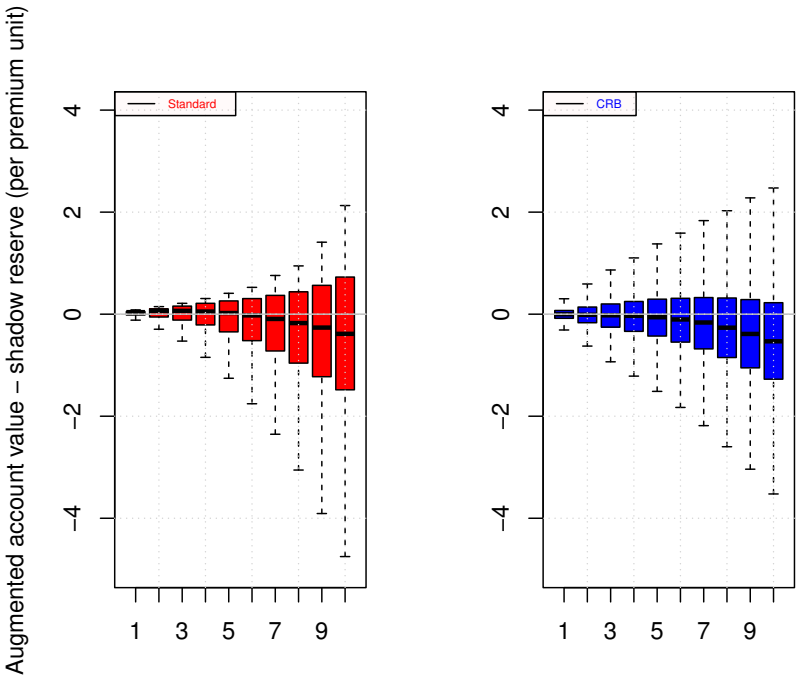
We also examine to what extent the augmented reference account (3.4.2) is adequate to cover liabilities by comparing it to the shadow reserve (3.4.1). Figure 3.6 gives the distribution of the difference between the augmented reference account and the shadow reserve, versus time, again reported as a ratio per premium unit. While the augmented reference account seems sufficient to cover the future liabilities at the beginning for both mechanisms, the probability of a shortfall increases as the contract approaches maturity. Indeed, the shortfall could be as large as -0.5 per premium unit on average at maturity. Of course, in practice insurers would charge a premium higher than the fair premium to

Figure 3.5: Difference between the shadow reserve and the policy reserve for both mechanisms



Notes: Distributions of the difference between the shadow reserve and the policy reserve are evaluated using the stochastic on stochastic Monte Carlo simulation approach described in Section 3.6.4 with 10,000 inner loops and 10,000 outer loops. The shadow reserve is defined in (3.4.1), whereas the policy reserve is given by (3.2.7). All numbers are obtained using $b_0 = 10{,}000$, $x = 50$, $g = \theta = 0$ and $n = 10$. The financial variables dynamics are described in Section 3.5.

Figure 3.6: Difference between the augmented reference account and the shadow reserve



Notes: The distributions of the difference between the augmented reference account (3.4.2) and the shadow reserve (3.4.1) are assessed using the stochastic on stochastic Monte Carlo simulation technique described in Section 3.6.4 with 10,000 inner loops and 10,000 outer loops. All numbers are obtained using $b_0 = 10{,}000$, $x = 50$, $g = \theta = \theta^* = 0$ and $n = 10$. The financial variables dynamics are described in Section 3.5.

offset such shortfalls and make a profit.

3.7 Conclusion

This paper generalizes existing approaches from the literature to do the pricing, reserving, and risk assessment for participating contracts issued by life insurers. In particular, the variable benefits life insurance approach from Bowers et al. 1997 (which we refer to as the standard approach) and the compound reversionary bonus method presented in Bacinello 2003 are adapted to integrate randomness in interest rates, equity returns, and idiosyncratic mortality. The latter inclusion is necessary for participation contract bonuses to reflect insurer performance rather than simply market performance as in equity-linked contracts.

Fixed premium endowment participating contracts are considered. First, we propose benefit update mechanisms in the stochastic setting. The shadow reserve is then introduced to obtain a genuine representation of the value of future cash flows, which is not provided by the conventional policy reserve formula not integrating potential benefit increases beyond the minimum rate. This allows for the calculation of a fair premium for the contract. We highlight that unless the portion of the bonus rate retained by the insurer is sufficiently large, the existence of a fair premium is not guaranteed, at least in the standard approach. This led to proposing only recognizing a portion of the full premium when determining participation bonuses.

Numerical experiments based on Monte Carlo simulations are conducted to compare the behavior of fair premiums, the dynamic benefit per premium ratios, and the risk associated with participating contracts for the two benefit increase mechanisms. The interest rate model (the discrete-time arbitrage-free Nelson-Siegel model) and the mixed fund reference account return dynamics considered in simulations are borrowed from Eghbalzadeh, Godin, and Gaillardetz 2022. The following observations are made. First, the standard approach

leads to a slower rate of increase in the benefit than the compound reversionary bonus method, the former leading to larger benefit to premium ratios when closer to the issue date. Moreover, the comparison between the shadow and policy reserves shows that the policy reserves cannot efficiently reflect future liabilities. Indeed, the benefits are stochastic in the participating contracts, while the policy reserve assumes deterministic benefits. Therefore, replacing the policy reserves with the shadow reserves leads to a more reliable estimation of future liabilities.

Chapter 4

Swaption pricing under the discrete-time arbitrage-free Nelson-Siegel model

4.1 Introduction

Interest rate risk management is of paramount importance for financial institutions such as banks and insurance companies. An interest rate swap is a financial contract between two parties stipulating that a stream of variable payments proportional to a floating interest rate is exchanged for a series of fixed payments. Using such an instrument allows modifying the risk exposure of interest-rate-sensitive portfolios, for instance by altering its dependence on interest rate fluctuations. A swaption is an option conferring the right, but not the obligation, to enter a swap of pre-determined maturity and payment rate at the maturity of the option. Such derivatives are useful devices to protect against large increases or decreases of interest rates.

Pricing swaptions is a sophisticated and difficult undertaking that necessitates a thorough grasp of the interest rates environment. In particular, such an endeavor requires considering a stochastic model to characterize future interest rate movements. Such model typically

depends on parameters that represent, for instance, the drift and volatility of interest rates. A characteristic specific to swaptions is that, unlike other interest rate derivatives such as caps and floors, they strongly depend on correlations between future spot rates (Brigo and Mercurio 2007). For that reason, the incorporation of swaption prices into calibration exercises for interest rate models is crucial as they contain information that is not provided by other fixed-income instruments and derivatives.

Single-factor models of the yield curve typically generate comonotonic movements in the term structure and are therefore not adequate to determine swaption prices. Multi-factor models are thus more appropriate as they are sufficiently flexible to generate de-correlated movements of spot rates of various tenors. A common multi-factor framework used to study term structure fluctuations and price interest rate derivatives is affine term structure (ATS) models, see Duffie and Kan 1996. Such setting offers the convenience of very high tractability since spot rates depend linearly on a set of risk factors. Moreover, it is consistent with no-arbitrage principles.

Several works study the pricing of swaptions in the context of multi-factor or ATS models. Munk 1999 demonstrates that the price of a European option on a coupon bond (e.g. a swaption) is roughly equal to some multiple of the price of a European option on a zero-coupon bond with maturity equal to the coupon bond's stochastic duration. Collin-Dufresne and Goldstein 2002 propose to apply an Edgeworth expansion to approximate the density of the coupon bond price and obtain the price of a swaption. Singleton and Umantsev 2002 rely on Fourier inversion methods to calculate swaption prices in the ATS framework. Schrager and Pelsser 2006 propose to approximate the swap rate volatility under the swap measure, which is a low-variance martingale, by its time-zero value. Such strategy leads to a closed-form formula for the swaption price.

A common difficulty with several multi-factor term structure models is that the factors might be difficult to interpret. For this reason, the present work considers the specific

case of the arbitrage-free Nelson-Siegel (AFNS) model developed in Christensen, Diebold, and Rudebusch 2011, or more precisely its discrete-time version (the DTANFS model) outlined in Eghbalzadeh, Godin, and Gaillardetz 2022. Such three-factor model has the advantage of having a clear interpretation associated with each of the factors, namely the term structure level, slope and curvature, respectively. The AFNS model builds on the seminal work Nelson and Siegel 1987 characterizing yield curve shapes and on its dynamic counterpart, Diebold and Li 2006, by applying corrections to short rate dynamics to ensure compliance with no-arbitrage pricing principles.

The paper's main contribution is to provide a fast pricing approach for swaption prices under the DTAFNS model. The approach relies on Monte-Carlo simulation, which is straightforward to implement. This paper derives the dynamics of the DTAFNS model's risk factors under the forward measure, which leads to a semi-analytic expression for the price of a swaption. Such formula consists in the expectation of a non-linear function of a single tri-dimensional Gaussian variable, which can be computed quickly and circumvents the need to simulate entire paths of the risk-free rate.

The paper is structured as follows. Section 4.2 provides a description of the DTAFNS term structure model, as well as a review of notions about European swaptions pricing, including a simulation algorithm relying on the risk-neutral measure. Section 4.3 focuses on extracting the state variable dynamics under the forward measure, which leads to an efficient algorithm to obtain swaption prices. Section 4.4 concludes.

4.2 Preliminaries: the DTAFNS model and swaptions

This section begins by discussing the DTAFNS model established by Eghbalzadeh, Godin, and Gaillardetz 2022, which characterizes dynamics of zero coupon bond prices and associated discount factors. Then, European swaptions and their pricing under a

risk-neutral measure are explained.

4.2.1 DTAFNS model

This section provides a description the DTAFNS interest rate term structure model of Eghbalzadeh, Godin, and Gaillardetz 2022. Consider a discrete-time setting with monthly time points $t = 0, \ldots, T$ and time elapse Δ year between each point. The filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^{T}$ characterizes the information flow in the market. The DTAFNS model assumes that the term structure of interest rates is determined by three factors: the long-term level of interest rates, the slope of the yield curve and its curvature. The time-t short rate applying over period $[t, t+1)$ is

$$r_t = X_t^{(1)} + X_t^{(2)}, \tag{4.2.1}$$

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

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

$$\begin{pmatrix} X_{t+1}^{(1)} - X_t^{(1)} \\ X_{t+1}^{(2)} - X_t^{(2)} \\ X_{t+1}^{(3)} - X_t^{(3)} \end{pmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}}_{\kappa^{\mathbb{Q}}} \underbrace{\begin{bmatrix} \theta_1^{\mathbb{Q}} - X_t^{(1)} \\ \theta_2^{\mathbb{Q}} - X_t^{(2)} \\ \theta_3^{\mathbb{Q}} - X_t^{(3)} \end{bmatrix}}_{\theta^{\mathbb{Q}} - X_t} + \underbrace{\begin{pmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{pmatrix}}_{\Sigma} \begin{pmatrix} Z_{t+1,1}^{\mathbb{Q}} \\ Z_{t+1,2}^{\mathbb{Q}} \\ Z_{t+1,3}^{\mathbb{Q}} \end{pmatrix}, \tag{4.2.2}$$

with scalar $\lambda \in (0, 1)$ and matrices $\theta^{\mathbb{Q}}$, $\kappa^{\mathbb{Q}}$ and Σ, with $\Sigma_{ii} > 0$, representing model parameters, and $\{Z_{t,i}^{\mathbb{Q}}\}_{t=1}^{n}$, $i = 1, 2, 3$ being $\mathcal{F}$-adapted standard Gaussian white noises with contemporaneous with correlation $\text{Corr}[Z_{t_1,i}^{\mathbb{Q}}, Z_{t_2,j}^{\mathbb{Q}}] = \mathbb{1}_{\{t_1=t_2\}}\rho_{ij}$ represented by correlation matrix $\rho = [\rho_{ij}]_{i,j=1}^{3}$.

According to Eghbalzadeh, Godin, and Gaillardetz 2022, under the such model, the time-t

price of a risk-free zero-coupon bond paying one dollar on maturity $\mathcal{T}$ is

$$P(t,\mathcal{T}) = A_\tau \exp\left[-\Delta \mathcal{B}_\tau^\top X_t\right], \tag{4.2.3}$$

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

$$\mathcal{B}_\tau^{(1)} = \tau, \quad \mathcal{B}_\tau^{(2)} = \frac{1-(1-\lambda)^\tau}{\lambda}, \quad \mathcal{B}_\tau^{(3)} = \frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}, \tag{4.2.4}$$

$$\log A_\tau = -\Delta\theta_2^{\mathbb{Q}}\left(\mathcal{B}_\tau^{(1)} - \mathcal{B}_\tau^{(2)}\right) + \Delta\theta_3^{\mathbb{Q}}\mathcal{B}_\tau^{(3)} + \frac{1}{2}\Delta^2 \upsilon_\tau, \tag{4.2.5}$$

with

$$\upsilon_\tau = \left(\sum_{i=1}^{3}\sum_{j=1}^{3}\upsilon_\tau^{(i,j)}\right), \tag{4.2.6}$$

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

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

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

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

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

$$\upsilon_\tau^{(2,3)} = \upsilon_t^{(3,2)} = \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\Bigg(\frac{\tau - 2 - (2-\lambda)\zeta_0\left((1-\lambda),\tau-1\right) + (1-\lambda)\zeta_0\left((1-\lambda)^2,\tau-1\right)}{\lambda^2}$$
$$+ \frac{-\zeta_1\left((1-\lambda),\tau-1\right) + (1-\lambda)\zeta_1\left((1-\lambda)^2,\tau-1\right)}{\lambda}\Bigg),$$

and

$$\zeta_0(r,\tau) \equiv \sum_{u=1}^{\tau-1} r^u = \frac{r - r^\tau}{1-r}, \tag{4.2.7}$$

$$\zeta_1(r,\tau) \equiv \sum_{u=1}^{\tau-1} u r^u = \frac{r - \tau r^\tau + (\tau-1) r^{\tau+1}}{(1-r)^2}, \tag{4.2.8}$$

$$\zeta_2(r,\tau) \equiv \sum_{u=1}^{\tau-1} u^2 r^u = \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1) r^{\tau+1} - \tau^2 r^\tau + r^2 + r}{(1-r)^3}. \tag{4.2.9}$$

4.2.2 European swaptions

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

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

$$PS\left[t; T_\alpha; T_\beta; K; N\right] = \mathbb{E}^{\mathbb{Q}}\left[D(t,T_\alpha) \left(N \left(S_{\alpha,\beta}(T_\alpha) - K \right)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i) \right) \middle| \mathcal{F}_t \right] \tag{4.2.10}$$

where $\tau_i = T_i - T_{i-1}$. The discount factor $D(t,T_\alpha)$ between times t and T_α and the time-t

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

forward swap rate $S_{\alpha,\beta}(t)$ are

$$D(t, T_\alpha) = \exp\left(-\Delta \sum_{j=t}^{T_\alpha - 1} r_j\right), \tag{4.2.11}$$

$$S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t, T_i)}. \tag{4.2.12}$$

The swap rate $S_{\alpha,\beta}(t)$ corresponds to a value of the fixed rate which would make the time-t value of the swap nil. The rationale underlying (4.2.10) is that a market participant could, while exercising the option, enter without fee into a receiver swap with the swap rate as the fixed rate. Combining both positions would lead to a net payment being the difference between the swap rate and the strike rate at each payment date.

A straightforward approach to obtain the swaption price via (4.2.10) is to conduct a Monte-Carlo simulation of the term structure factors under the risk-neutral measure and to average discounted cash flows to approximate the expectation in (4.2.10). Algorithm 2 summarizes such process.

4.3 Pricing swaptions under the forward measure

Calculating European swaption prices using Algorithm 2 requires simulating the entire path of risk-free rate factors, which might be numerically cumbersome in some situations. For instance, in the context of calibration, swaption prices need to be recomputed for a wide variety of parameters, and the approach underlying Algorithm 2 might not be feasible in a reasonable time frame. By applying a change of numéraire, we can obtain a pricing approach which is more time-efficient. Detailing such an approach in the context of the DTAFNS model is the objective of this section.

The probability measure using the risk-free zero-coupon bond maturing at time $\mathcal{T}$ as a numéraire is known as the $\mathcal{T}$-forward measure and is denoted by $\mathbb{Q}^{\mathcal{T}}$. The

Algorithm 2 Calculating the European swaptions price under the risk-neutral measure

Input: t, $X_t^{(1)}$, $X_t^{(2)}$, $X_t^{(3)}$, K, N, α, β, Δ, $\kappa^{\mathbb{Q}}$, $\theta^{\mathbb{Q}}$, ρ **and** Σ
for $m \in \{1, \ldots, M\}$ **do** ▷ M is the number of simulated paths
Simulate the m^{th} *path of the state variables,* X*, denoted by* $\{ {}_mX_{t'}^{(i)}\}_{t'=t+1}^{T_\alpha}, i = 1, 2, 3$ *under the risk neural measure* $\mathbb{Q}$
for $t' \in \{t+1, \ldots, T_\alpha\}$ **do**
Calculate the interest rate, $r({}_mX_{t'})$*, through* (4.2.1)
end for
Calculate $D(t, T_\alpha)$ *through* (4.2.11)
for $t'' \in \{T_{\alpha+1}, \ldots, T_\beta\}$ **do**
Calculate ${}_mP(T_\alpha, t'')$ *through* (4.2.3)
end for
Calculate ${}_mS_{\alpha,\beta}(T_\alpha)$ *through* (4.2.12)
Calculate the payoff for m^{th} *path*

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

end for

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

Radon-Nikodym derivative allowing to pass from the risk-neutral to the $\mathcal{T}$-forward measure, which is provided by Jamshidian 1996 or Brigo and Mercurio 2007, is

$$ \frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}} = \frac{B(0)P(\mathcal{T}, \mathcal{T})}{P(0, \mathcal{T})B(\mathcal{T})} = \frac{D(0, \mathcal{T})}{P(0, \mathcal{T})}, \tag{4.3.1} $$

where, $B(t) = \exp(\Delta \sum_{s=0}^{t-1} r_s)$ is the year-t bank account numéraire under the risk-neutral measure. Note that the Radon-Nykodyn derivative allowing to go from the forward measure to the risk-neutral measure is $\frac{d\mathbb{Q}}{d\mathbb{Q}^{\mathcal{T}}} = \left(\frac{d\mathbb{Q}^{\mathcal{T}}}{d\mathbb{Q}}\right)^{-1}$.

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

an asset providing a payoff $H_{\mathcal{T}}$ at time $\mathcal{T}$ is

$$H_t = P(t,\mathcal{T})\mathbb{E}^{\mathcal{T}}\left[H_{\mathcal{T}}|\mathcal{F}_t\right]. \tag{4.3.2}$$

Considering the zero-coupon bond maturing at $\mathcal{T} = T_\alpha$ as the new numéraire makes the computation of the swaption price much more convenient. In such case, the payer swaption price may therefore be rewritten using (4.2.10), (4.2.12) and (4.3.2) as

$$\begin{aligned} PS\left[t;T_\alpha;T_\beta;K;N\right] &= P(t,T_\alpha)\mathbb{E}^{T_\alpha}\left[\left(N\left(S_{\alpha,\beta}(T_\alpha)-K\right)^+\sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right)\Big|\mathcal{F}_t\right] \\ &= P(t,T_\alpha)\mathbb{E}^{T_\alpha}\left[\left(N\left(1-P(T_\alpha,T_\beta)-K\sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right)^+\right)\Big|\mathcal{F}_t\right]. \end{aligned} \tag{4.3.3}$$

Equation (4.3.3) involves the t-conditional expectation of a function of time-T_α zero-coupon bonds, which are fully characterized by term structure factors $X_{T_\alpha} = \left[X^{(1)}_{T_\alpha}, X^{(2)}_{T_\alpha}, X^{(3)}_{T_\alpha}\right]^\top$. As a result, in order to calculate (4.3.3), the dynamics of the state variables under the forward measure must be understood.

4.3.1 Dynamics of the state variables under the forward measure

To simplify the statements of this section, we assume $T_i = i$ for $i = \alpha, \ldots, \beta$. The following proposition defines so-called *forward measure innovations* and outlines their dynamics.

Proposition 4.3.1. *For any $\mathcal{T} \in \{1,\ldots,T\}$ and $t < \mathcal{T}$, conditional on $\mathcal{F}_t$, the forward measure innovation defined as $Z^{\mathcal{T}}_{t+1} = Z^{\mathbb{Q}}_{t+1} + \Delta\rho\Sigma\mathcal{B}_{\tau-1}$ follows the multivariate Gaussian distribution with mean vector zero and covariance matrix ρ under the $\mathcal{T}$-forward measure.*
Proof. See Appendix A.3.1.

Corollary 4.3.1. *Since the conditional distribution of $Z^{\mathcal{T}}_{t+1}$ with respect $\mathcal{F}_t$ does not depend*

on $Z_1^{\mathcal{T}}, \ldots, Z_t^{\mathcal{T}}$, and since the latter variables characterize the information contained in $\mathcal{F}_t$, elements of the sequence $\{Z_j^{\mathcal{T}}\}_{j=1}^{\mathcal{T}}$ are independent.

Based on the above results, we now provide an expression of the dynamics of term structure factors $\{X_t\}_{t=0}^{T}$ analogous to (4.2.2), but using instead the $\mathcal{T}$-forward measure innovations. Define

$$\theta^{\mathcal{T}} = \theta^{\mathbb{Q}}, \quad \kappa^{\mathcal{T}} = \kappa^{\mathbb{Q}}, \quad \eta_t = \Delta\Sigma\rho\Sigma\mathcal{B}_{\tau-1} = \Delta\Sigma\rho\Sigma\mathcal{B}_{\mathcal{T}-t-1}. \tag{4.3.4}$$

A direct consequence of the application of Proposition 4.3.1 into (4.2.2) is that

$$X_{t+1} = X_t - \eta_t + \kappa^{\mathcal{T}}(\theta^{\mathcal{T}} - X_t) + \Sigma Z_{t+1}^{\mathcal{T}}. \tag{4.3.5}$$

Representation (4.3.5), along with some additional lemmas provided in Appendix A.3.1, allow obtaining the t-conditional distribution of $X_{\mathcal{T}}$ under the $\mathcal{T}$-forward measure.

Proposition 4.3.2. *Under the $\mathcal{T}$-forward measure, conditionally on $\mathcal{F}_t$, factors X_{t+n} follow the multivariate Gaussian distribution with mean vector $\mathcal{M}_{t,n} = \left[\mathcal{M}_{t,n}^{(i)}\right]_{i=1}^{3}$ and covariance matrix $\mathcal{V}_n = \left[\mathcal{V}_n^{(i,j)}\right]_{i,j=1}^{3}$ where*

$$\begin{aligned}
\mathcal{M}_{t,n}^{(1)} =& X_t^{(1)} - \sum_{l=0}^{n-1} \eta_{t+l}^{(1)}, \\
\mathcal{M}_{t,n}^{(2)} =& X_t^{(2)}(1-\lambda)^n + (\theta_2^{\mathcal{T}} - \theta_3^{\mathcal{T}})\left(1 - (1-\lambda)^n\right) - \sum_{l=0}^{n-1} \eta_{t+l}^{(2)}(1-\lambda)^{n-1-l} \\
&+ \lambda\left(nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\left(\frac{\zeta_0(1-\lambda, n+1)}{1-\lambda} - n(1-\lambda)^{n-1}\right) - \sum_{l=0}^{n-1}(n-l-1)\eta_{t+l}^{(3)}(1-\lambda)^{n-l-2}\right), \\
\mathcal{M}_{t,n}^{(3)} =& X_t^{(3)}(1-\lambda)^n + \theta_3^{\mathcal{T}}\left(1 - (1-\lambda)^n\right) - \sum_{l=0}^{n-1} \eta_{t+l}^{(3)}(1-\lambda)^{n-1-l}, \\
\mathcal{V}_n^{(1,1)} =& n\Sigma_{1,1}^2, \\
\mathcal{V}_n^{(2,2)} =& \Sigma_{2,2}^2\left(1 + \zeta_0((1-\lambda)^2, n)\right) + \lambda^2\Sigma_{3,3}^2(1-\lambda)^{-2}\zeta_2((1-\lambda)^2, n) + 2\Sigma_{2,2}\lambda\Sigma_{3,3}\rho_{2,3}(1-\lambda)^{-1}\zeta_1\left((1-\lambda)^2, n\right)
\end{aligned}$$

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

Proof. See Appendix A.3.1.

The following quantities appearing in Proposition 4.3.2 can be further simplified.

Lemma 4.3.1.

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

Moreover, for $i = 2, 3$,

$$
\begin{aligned}
\sum_{l=0}^{n-1}\eta_{t+l}^{(i)}(1-\lambda)^{n-1-l} =& \Delta\Sigma_{i,i}(1-\lambda)^n\Bigg[\Sigma_{1,1}\rho_{i,1}\left((\mathcal{T}-t)\zeta_0\left((1-\lambda)^{-1}, n+1\right)-\zeta_1\left((1-\lambda)^{-1}, n+1\right)\right) \\
&+\frac{\Sigma_{2,2}\rho_{i,2}}{\lambda}\left(\zeta_0\left((1-\lambda)^{-1}, n+1\right)-(1-\lambda)^{\mathcal{T}-t}\zeta_0\left((1-\lambda)^{-2}, n+1\right)\right) \\
&+\Sigma_{3,3}\rho_{i,3}\Bigg(\frac{\zeta_0\left((1-\lambda)^{-1}, n+1\right)-(1-\lambda)^{\mathcal{T}-t-1}\zeta_0\left((1-\lambda)^{-2}, n+1\right)}{\lambda} \\
&-(1-\lambda)^{\mathcal{T}-t-1}\left[(\mathcal{T}-t-1)\zeta_0\left((1-\lambda)^{-2}, n+1\right)-\zeta_1\left((1-\lambda)^{-2}, n+1\right)\right]\Bigg)\Bigg].
\end{aligned} \tag{4.3.6}
$$

Lastly,

$$
\sum_{l=0}^{n-1}(n-l-1)\eta_{t+l}^{(3)}(1-\lambda)^{n-l-2} = \frac{n-1}{1-\lambda}\sum_{l=0}^{n-1}\eta_{t+l}^{(3)}(1-\lambda)^{n-l-1}
$$

$$
\begin{aligned}
-\Delta\Sigma_{3,3}(1-\lambda)^{n-2}\Bigg[&\Sigma_{1,1}\rho_{3,1}\left((\mathcal{T}-t-1)\zeta_1\left((1-\lambda)^{-1},n\right)-\zeta_2\left((1-\lambda)^{-1},n\right)\right)\\
&+\frac{\Sigma_{2,2}\rho_{3,2}}{\lambda}\left(\zeta_1\left((1-\lambda)^{-1},n\right)-(1-\lambda)^{\mathcal{T}-t-1}\zeta_1\left((1-\lambda)^{-2},n\right)\right)\\
&+\Sigma_{3,3}\Bigg(\frac{\zeta_1\left((1-\lambda)^{-1},n\right)-(1-\lambda)^{\mathcal{T}-t-2}\zeta_1\left((1-\lambda)^{-2},n\right)}{\lambda}\\
&-(1-\lambda)^{\mathcal{T}-t-2}\left[(\mathcal{T}-t-2)\zeta_1\left((1-\lambda)^{-2},n\right)-\zeta_2\left((1-\lambda)^{-2},n\right)\right]\Bigg)\Bigg],
\end{aligned}
$$

with the summation appearing in the first term of the right-hand side of the above equation being given by (4.3.6).

Proof. See Appendix A.3.1.

Proposition 4.3.2 is the key to compute (4.3.3) in an efficient way using Monte Carlo, as it expresses swaption prices as expectations of a non-linear function of a single tri-dimensional Gaussian random variable. Algorithm 3 highlights the procedure to price swaptions using such an approach. Such $\mathcal{T}$-forward measure simulation is much quicker than the risk-neutral counterpart from Algorithm 2 which requires computing expectations over entire paths of the term structure factors.

4.4 Conclusion

This paper describes how to calculate swaption prices with Monte-Carlo simulations under the DTAFNS model. Information contained within swaption prices is important when calibrating term structure models so as to obtain market-consistent pricing for instruments depending on interest rates. Whereas the conventional risk-neutral simulation relies on the simulation of entire paths of the risk-free rate, reliance on the forward measure allows for much quicker computations. Indeed, this paper derive forward-measure dynamics of term structure factors and then expresses swaption prices as the expectation of a non-linear function of a tri-dimensional Gaussian variable, which is very quick to estimate in practice.

Algorithm 3 Calculating the European swaptions price under the forward measure

Input: t, $X_t^{(1)}$, $X_t^{(2)}$, $X_t^{(3)}$, K, N, α, β, Δ, $\kappa^{\mathbb{Q}}$, $\theta^{\mathbb{Q}}$, ρ, **and** Σ

Calculate $P(t,\alpha)$ *by* (4.2.3) *and* $\kappa^{\mathcal{T}}$, $\theta^{\mathcal{T}}$ *using* (4.3.4)

for $m \in \{1,\dots,M\}$ **do** ▷ M is the number of simulated paths

Simulate ${}_mX_\alpha^{(1)}$, ${}_mX_\alpha^{(2)}$, ${}_mX_\alpha^{(3)}$ *from multivariate Gaussian distribution with mean vector*

$\mathcal{M}_{t,\alpha-t}$ *and matrix variance-covariance* $\mathcal{V}_{\alpha-t}$ *defined in Proposition 4.3.2*

for $i \in \{\alpha+1,\dots,\beta\}$ **do**

Calculate the zero coupon bond price, ${}_mP(\alpha,i)$ *through* (4.2.3)

end for

Calculate the payoff for m^{th} *path*

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

end for

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

Chapter 5

Conclusion

The modeling of interest rates is of critical importance in the fields of finance and actuarial science, as it enables accurate pricing and risk management of various financial instruments. In particular, the modeling of yield curves is essential for understanding the behavior of interest rates and their impact on fixed-income securities, participating life insurance contracts, and swaptions. This dissertation builds upon the work of Hong, Niu, and Zeng 2016 by proposing a slightly modified version of the discrete-time arbitrage-free Nelson-Siegel model and making several contributions, including providing a closed-form expression for risk-free spot rates, demonstrating its superior predictive ability over benchmarks, and incorporating it into a modified mixed fund return model.

The dissertation also extends participating contract analysis schemes to include realistic dynamics for the main risk drivers, such as stochastic interest rates and mortality. It focuses on fixed premium contracts with bonus calculations, whose setting is generalized to handle general stochastic dynamics. The dissertation proposes a shadow reserve as a replacement for conventional reserving formulas, providing additional accuracy in measuring liabilities by recognizing future benefit increases beyond the guaranteed rate. Several numerical tests based on Monte-Carlo simulations are applied to demonstrate the importance of the

shadow reserve when measuring risk.

Additionally, the dissertation outlines procedures for pricing swaptions under the DTAFNS model. In particular, it illustrates derivations of interest rates dynamics under the forward measure, resulting in faster calculations for the determination of swaption prices compared to when risk-neutral measure dynamics are relied upon. Overall, this dissertation provides practical applications for the DTAFNS model in fixed-income securities valuation, life insurance contract analysis and swaptions pricing, contributing to the literature on these topics.

Appendix A

Appendice

A.1 Discrete-time arbitrage-free Nelson-Siegel model

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

A.1.1 Proofs

To prove Proposition 2.2.1, the following lemma is needed.

Lemma A.1.1. *Assume* (2.2.10) *holds. For* $t = 0, \ldots, T$ *and* $n = 0, \ldots, T - t$,

$$\begin{aligned} X_{t+n}^{(i)} =& X_t^{(i)}(1-\kappa_{i,i}^{\mathbb{Q}})^n + \kappa_{i,i}^{\mathbb{Q}}\theta_i^{\mathbb{Q}} \sum_{l=1}^{n}(1-\kappa_{i,i}^{\mathbb{Q}})^{(n-l)} + \Sigma_{i,i} \sum_{l=1}^{n} Z_{t+l,i}^{\mathbb{Q}}(1-\kappa_{i,i}^{\mathbb{Q}})^{(n-l)} \\ &+ \sum_{l=1}^{n}\sum_{j\neq i}^{3} \kappa_{i,j}^{\mathbb{Q}}(\theta_j^{\mathbb{Q}} - X_{t+l-1}^{(j)})(1-\kappa_{i,i}^{\mathbb{Q}})^{(n-l)}. \end{aligned} \tag{A.1.1}$$

Proof of Lemma A.1.1: 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.1.1) holds for some $n < T - t - 1$. Using (2.2.10), for any $i = 1, 2, 3$,

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

Applying (A.1.1) in the latter equality yields

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

hence completing the induction. □

Remark A.1.1. *The result of Lemma A.1.1 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 2.2.1: Equation (2.2.11) is trivial. To show (2.2.12),

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

Finally, to show (2.2.13),

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

$$= \frac{-(\tau-1)^2 r^{\tau+2} + (2\tau^2 - 2\tau - 1) r^{\tau+1} - \tau^2 r^\tau + r^2 + r}{(1-r)^3}. \quad \square$$

Proof of Proposition 2.2.1: From (2.2.7), for any $\tau = 1, \ldots, T-t$,

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

Furthermore, since $\kappa_{1,j}^{\mathbb{Q}} = 0$ for $j = 1, \ldots, 3$, Equation (A.1.1) leads to

$$\begin{aligned}
\sum_{s=0}^{\tau-1} X_{t+s}^{(1)} =& \sum_{s=0}^{\tau-1} \left[X_t^{(1)} + \Sigma_{1,1} \sum_{l=1}^{s} Z_{t+l,1}^{\mathbb{Q}} \right] \\
=& \tau X_t^{(1)} + \Sigma_{1,1} \sum_{s=1}^{\tau-1} \sum_{l=1}^{\tau-1} \mathbb{1}_{\{l \le s\}} Z_{t+l,1}^{\mathbb{Q}} \\
=& \tau X_t^{(1)} + \Sigma_{1,1} \sum_{l=1}^{\tau-1} \sum_{s=l}^{\tau-1} Z_{t+l,1}^{\mathbb{Q}} \\
=& \tau X_t^{(1)} + \Sigma_{1,1} \sum_{l=1}^{\tau-1} (\tau - l) Z_{t+l,1}^{\mathbb{Q}}. \qquad \text{(A.1.2)}
\end{aligned}$$

Moreover, relying again on Equation (A.1.1) and recalling that $\kappa_{2,1}^{\mathbb{Q}} = 0$, $\kappa_{2,2}^{\mathbb{Q}} = \lambda$ and $\kappa_{2,3}^{\mathbb{Q}} = -\lambda$,

$$\begin{aligned}
\sum_{s=0}^{\tau-1} X_{t+s}^{(2)} =& \sum_{s=0}^{\tau-1} \Bigg\{ X_t^{(2)}(1-\lambda)^s + \lambda\theta_2^{\mathbb{Q}} \sum_{l=1}^{s}(1-\lambda)^{s-l} + \Sigma_{2,2} \sum_{l=1}^{s} Z_{t+l,2}^{\mathbb{Q}}(1-\lambda)^{s-l} \\
& - \lambda \sum_{l=1}^{s}(1-\lambda)^{s-l}\theta_3^{\mathbb{Q}} + \lambda \sum_{l=1}^{s}(1-\lambda)^{s-l} X_{t+l-1}^{(3)} \Bigg\} \\
=& X_t^{(2)} \sum_{s=0}^{\tau-1}(1-\lambda)^s + \lambda\theta_2^{\mathbb{Q}} \sum_{s=1}^{\tau-1}\sum_{l=1}^{s}(1-\lambda)^{s-l} + \Sigma_{2,2} \sum_{s=1}^{\tau-1}\sum_{l=1}^{s}(1-\lambda)^{s-l} Z_{t+l,2}^{\mathbb{Q}} \\
& - \lambda\theta_3^{\mathbb{Q}} \sum_{s=1}^{\tau-1}\sum_{l=1}^{s}(1-\lambda)^{s-l} + \lambda \sum_{s=1}^{\tau-1}\sum_{l=1}^{s}(1-\lambda)^{s-l} X_{t+l-1}^{(3)} \\
=& X_t^{(2)} \sum_{s=0}^{\tau-1}(1-\lambda)^s + \lambda\left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}\right) \sum_{s=1}^{\tau-1}\sum_{l=0}^{s-1}(1-\lambda)^l + \sum_{s=1}^{\tau-1}\sum_{l=1}^{\tau-1} \mathbb{1}_{\{l\le s\}}(1-\lambda)^{s-l}\left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)}\right) \\
=& X_t^{(2)} \sum_{s=0}^{\tau-1}(1-\lambda)^s + \lambda\left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}\right) \sum_{s=1}^{\tau-1} \frac{1-(1-\lambda)^s}{\lambda} \\
& + \sum_{l=1}^{\tau-1}\sum_{s=1}^{\tau-1} \mathbb{1}_{\{l\le s\}}(1-\lambda)^{s-l}\left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)}\right) \\
=& X_t^{(2)} \frac{1-(1-\lambda)^\tau}{\lambda} + \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}\right)\left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^\tau}{\lambda}\right) \\
& + \sum_{l=1}^{\tau-1}\left(\Sigma_{2,2} Z_{t+l,2}^{\mathbb{Q}} + \lambda X_{t+l-1}^{(3)}\right) \sum_{s=1}^{\tau-1} \mathbb{1}_{\{l\le s\}}(1-\lambda)^{s-l}. \qquad \text{(A.1.3)}
\end{aligned}$$

Note that

$$\sum_{s=1}^{\tau-1} \mathbb{1}_{\{l\le s\}}(1-\lambda)^{s-l} = \sum_{s=l}^{\tau-1}(1-\lambda)^{s-l} = \sum_{s=0}^{\tau-1-l}(1-\lambda)^s = \frac{1-(1-\lambda)^{\tau-l}}{\lambda}. \qquad \text{(A.1.4)}$$

Placing (A.1.4) in (A.1.3) yields

$$\begin{aligned}
\sum_{s=0}^{\tau-1} X_{t+s}^{(2)} =& X_t^{(2)} \frac{1-(1-\lambda)^\tau}{\lambda} + \left(\theta_2^{\mathbb{Q}} - \theta_3^{\mathbb{Q}}\right)\left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^\tau}{\lambda}\right) \\
& + \Sigma_{2,2} \sum_{l=1}^{\tau-1} Z_{t+l,2}^{\mathbb{Q}} \frac{1-(1-\lambda)^{\tau-l}}{\lambda} + \sum_{l=1}^{\tau-1}\left[1-(1-\lambda)^{\tau-l}\right] X_{t+l-1}^{(3)}. \qquad \text{(A.1.5)}
\end{aligned}$$

Additionally, combining (A.1.1) with $\kappa^{\mathbb{Q}}_{3,1} = \kappa^{\mathbb{Q}}_{3,2} = 0$ and $\kappa^{\mathbb{Q}}_{3,3} = \lambda$ leads to

$$
\begin{aligned}
&\sum_{l=1}^{\tau-1} \left[1-(1-\lambda)^{\tau-l}\right] X^{(3)}_{t+l-1} \\
&= \sum_{l=0}^{\tau-2} \left[1-(1-\lambda)^{\tau-l-1}\right] X^{(3)}_{t+l} \\
&= \sum_{l=0}^{\tau-2} \left[1-(1-\lambda)^{\tau-l-1}\right] \left[X^{(3)}_t (1-\lambda)^l + \lambda\theta^{\mathbb{Q}}_3 \sum_{k=1}^{l} (1-\lambda)^{(l-k)} + \Sigma_{3,3} \sum_{k=1}^{l} Z^{\mathbb{Q}}_{t+k,3} (1-\lambda)^{(l-k)} \right] \\
&= X^{(3)}_t \left(\sum_{l=0}^{\tau-2} (1-\lambda)^l - \sum_{l=0}^{\tau-2} (1-\lambda)^{\tau-1} \right) + \lambda\theta^{\mathbb{Q}}_3 \sum_{l=0}^{\tau-2} \frac{1-(1-\lambda)^l - (1-\lambda)^{\tau-l-1} + (1-\lambda)^{\tau-1}}{\lambda} \\
&+ \Sigma_{3,3} \sum_{l=0}^{\tau-2} \sum_{k=1}^{\tau-2} \mathbb{1}_{\{k\le l\}} Z^{\mathbb{Q}}_{t+k,3} \left[(1-\lambda)^{(l-k)} - (1-\lambda)^{\tau-k-1} \right] \\
&= X^{(3)}_t \left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1} \right) + \theta^{\mathbb{Q}}_3 \sum_{l=0}^{\tau-2} \left(1-(1-\lambda)^l - (1-\lambda)^{\tau-l-1} + (1-\lambda)^{\tau-1} \right) \\
&\quad + \Sigma_{3,3} \sum_{k=1}^{\tau-2} Z^{\mathbb{Q}}_{t+k,3} \sum_{l=k}^{\tau-2} \left[(1-\lambda)^{(l-k)} - (1-\lambda)^{\tau-k-1} \right] \\
&= X^{(3)}_t \left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1} \right) + \theta^{\mathbb{Q}}_3 \Bigg((\tau-1)\left(1+(1-\lambda)^{\tau-1}\right) \\
&\quad - \frac{2-(1-\lambda)^{\tau-1} - \lambda - (1-\lambda)^{\tau}}{\lambda} \Bigg) + \Sigma_{3,3} \sum_{k=1}^{\tau-2} Z^{\mathbb{Q}}_{t+k,3} \left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k) \right].
\end{aligned}
\tag{A.1.6}
$$

Placing (A.1.6) in (A.1.5) given

$$
\begin{aligned}
\sum_{s=0}^{\tau-1} X^{(2)}_{t+s} =& X^{(2)}_t \frac{1-(1-\lambda)^{\tau}}{\lambda} + X^{(3)}_t \left(\frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1} \right) \\
&+ \theta^{\mathbb{Q}}_2 \left(\tau - 1 - \frac{1-\lambda-(1-\lambda)^{\tau}}{\lambda} \right) + \theta^{\mathbb{Q}}_3 \left((\tau-1)(1-\lambda)^{\tau-1} - \frac{1-(1-\lambda)^{\tau-1}}{\lambda} \right) \\
&+ \Sigma_{2,2} \sum_{l=1}^{\tau-1} Z^{\mathbb{Q}}_{t+l,2} \frac{1-(1-\lambda)^{\tau-l}}{\lambda} + \Sigma_{3,3} \sum_{k=1}^{\tau-2} Z^{\mathbb{Q}}_{t+k,3} \left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k) \right]
\end{aligned}
\tag{A.1.7}
$$

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.1.2) and (A.1.7), given $\mathcal{F}_t$, $-R_{t,T}$ has a Gaussian distribution with mean

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

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

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

with

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

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

$$= \frac{\Sigma_{3,3}^2}{\lambda^2} \sum_{k=1}^{\tau-2} \left[1 + (1-\lambda)^{2k} + \lambda^2(1-\lambda)^{2k}k^2 - 2(1-\lambda)^k - 2k\lambda(1-\lambda)^k + 2\lambda k(1-\lambda)^{2k}\right]$$
$$= \frac{\Sigma_{3,3}^2}{\lambda^2} \Bigg[\tau - 2 + \zeta_0\left((1-\lambda)^2, \tau-1\right) + \lambda^2\zeta_2\left((1-\lambda)^2, \tau-1\right)$$
$$- 2\zeta_0\left((1-\lambda), \tau-1\right) - 2\lambda\zeta_1\left((1-\lambda), \tau-1\right) + 2\lambda\zeta_1\left((1-\lambda)^2, \tau-1\right) \Bigg]$$

and

$$\upsilon_\tau^{(1,2)} = \upsilon_\tau^{(2,1)} = \mathrm{Cov}^{\mathbb{Q}}\left(\Sigma_{1,1}\sum_{l=1}^{\tau-1}(\tau-l)Z_{t+l,1}^{\mathbb{Q}}, \Sigma_{2,2}\sum_{l=1}^{\tau-1} Z_{t+l,2}^{\mathbb{Q}}\frac{1-(1-\lambda)^{\tau-l}}{\lambda}\right)$$
$$= \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\sum_{l=1}^{\tau-1} l\frac{1-(1-\lambda)^l}{\lambda}$$
$$= \rho_{1,2}\Sigma_{1,1}\Sigma_{2,2}\frac{1}{\lambda}\left(\frac{\tau(\tau-1)}{2} - \zeta_1((1-\lambda),\tau)\right),$$
$$\upsilon_\tau^{(1,3)} = \upsilon_\tau^{(3,1)} = \mathrm{Cov}^{\mathbb{Q}}\left(\Sigma_{1,1}\sum_{k=1}^{\tau-1}(\tau-k)Z_{t+k,1}^{\mathbb{Q}}, \Sigma_{3,3}\sum_{k=1}^{\tau-2} Z_{t+k,3}^{\mathbb{Q}}\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]\right.$$
$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\sum_{k=1}^{\tau-2}(\tau-k)\left[\frac{1-(1-\lambda)^{(\tau-1-k)}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-1-k)\right]$$
$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\sum_{k=1}^{\tau-2}(k+1)\left[\frac{1-(1-\lambda)^k}{\lambda} - (1-\lambda)^k k\right]$$
$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\sum_{k=1}^{\tau-2}(k+1)\left[1-(1+\lambda k)(1-\lambda)^k\right]$$
$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\left(\frac{(\tau-1)\tau}{2} - 1 - \sum_{k=1}^{\tau-2}(1+(\lambda+1)k+\lambda k^2)(1-\lambda)^k\right)$$
$$= \rho_{1,3}\Sigma_{1,1}\Sigma_{3,3}\frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0\left((1-\lambda),\tau-1\right) - (\lambda+1)\zeta_1\left((1-\lambda),\tau-1\right)$$
$$- \lambda\zeta_2\left((1-\lambda),\tau-1\right)\Bigg],$$

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

$$
\begin{aligned}
&\quad - (1-\lambda)^{\tau-k-1}(\tau-1-k)\Big]\Big) \\
&= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\sum_{k=1}^{\tau-2}\left(\frac{1-(1-\lambda)^{\tau-k}}{\lambda}\right)\left[\frac{1-(1-\lambda)^{\tau-k-1}}{\lambda} - (1-\lambda)^{\tau-k-1}(\tau-k-1)\right] \\
&= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\sum_{k=1}^{\tau-2}\left(\frac{1-(1-\lambda)^{k+1}}{\lambda}\right)\left[\frac{1-(1-\lambda)^{k}}{\lambda} - (1-\lambda)^{k}k\right] \\
&= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3} \\
&\quad \times \sum_{k=1}^{\tau-2}\left(\frac{1-(2-\lambda)(1-\lambda)^{k}+(1-\lambda)(1-\lambda)^{2k}}{\lambda^2} + \frac{-(1-\lambda)^{k}k+(1-\lambda)(1-\lambda)^{2k}k}{\lambda}\right) \\
&= \rho_{2,3}\Sigma_{2,2}\Sigma_{3,3}\Bigg(\frac{\tau-2-(2-\lambda)\zeta_0\left((1-\lambda),\tau-1\right)+(1-\lambda)\zeta_0\left((1-\lambda)^2,\tau-1\right)}{\lambda^2} \\
&\quad + \frac{-\zeta_1\left((1-\lambda),\tau-1\right)+(1-\lambda)\zeta_1\left((1-\lambda)^2,\tau-1\right)}{\lambda}\Bigg). \square
\end{aligned}
$$

Lemma A.1.2. *Consider assumptions of Proposition 2.2.1. Then, for $i = 1, 2, 3$,*

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

Proof of Lemma A.1.2: Based on Proposition 2.2.1,

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

Furthermore,

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

Finally,

$$
B_\tau^{(3)} = \frac{1-(1-\lambda)^{\tau-1}}{\lambda} - (\tau-1)(1-\lambda)^{\tau-1}
$$

$$= (1-\lambda)\left(\frac{1-(1-\lambda)^{\tau-2}}{\lambda} + \frac{1}{1-\lambda} - (\tau-2)(1-\lambda)^{\tau-2} - (1-\lambda)^{\tau-2}\right)$$
$$= B^{(3)}_{\tau-1}(1-\lambda) + 1 - (1-\lambda)^{\tau-1}. \square$$

Proof of Proposition 2.4.1: Using (2.2.10), (2.2.14), (2.4.6) and (2.4.12),

$$\begin{aligned}
R^{(F)}_{t+1} - \Delta r_t &= \psi_0 + \sum_{i=1}^{3} \psi_i \left(X^{(i)}_{t+1} - \left(1-\kappa^{\mathbb{Q}}_{i,i}\right) X^{(i)}_t\right) + \psi_3' X^{(3)}_t + \sum_{j=1}^{q} \psi^{(S)}_j \left(R^{(S)}_{t+1,j} - \Delta r_t\right) + \sqrt{h^{(F)}_t} Z^{(F)}_{t+1} \\
&= \psi_0 + \sum_{i=1}^{3} \psi_i \left(\Sigma_{i,i} Z^{\mathbb{Q}(i)}_{t+1} + \lambda\theta^{\mathbb{Q}}_i \mathbb{1}_{\{i>1\}} - \lambda\theta^{\mathbb{Q}}_3 \mathbb{1}_{\{i=2\}} + \lambda \mathbb{1}_{\{i=2\}} X^{(3)}_t\right) + \psi_3' X^{(3)}_t \\
&\quad + \sum_{j=1}^{q} \psi^{(S)}_j \left(R^{(S)}_{t+1,j} - \Delta r_t\right) + \sqrt{h^{(F)}_t} Z^{(F)}_{t+1} \\
&= \underbrace{\psi_0 + \sum_{i=1}^{3} \psi_i \left(\kappa^{\mathbb{Q}}_{i,i}\theta^{\mathbb{Q}}_i - \lambda\theta^{\mathbb{Q}}_3 \mathbb{1}_{\{i=2\}}\right) + (\psi_2\lambda + \psi_3') X^{(3)}_t - \sum_{j=1}^{q} \psi^{(S)}_j \left(\frac{1}{2} h^{(S)}_{t,j}\right)}_{=\phi_t} - \sqrt{h^{(F)}_t}\lambda^{(F)}_t \\
&\quad + \sum_{i=1}^{3} \psi_i \Sigma_{i,i} Z^{\mathbb{Q}(i)}_{t+1} + \sum_{j=1}^{q} \psi^{(S)}_j \sqrt{h^{(S)}_{t,j}} Z^{\mathbb{Q}(S)}_{t+1,j} + \sqrt{h^{(F)}_t} Z^{\mathbb{Q}(F)}_{t+1}.
\end{aligned}$$

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

$$\mathrm{Var}^{\mathbb{Q}}\left[R^{(F)}_{t+1} - \Delta r_t \middle| \mathcal{F}_t\right] = \left(\sigma^{(F)}_t\right)^2 \tag{A.1.8}$$

with $\sigma^{(F)}_t$ given by (2.4.15). Enforcing that the discounted mixed fund value is a $\mathbb{Q}$-martingale requires

$$\begin{aligned}
1 &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(R^{(F)}_{t+1} - \Delta r_t\right) \middle| \mathcal{F}_t\right] \\
&= \exp\left(\phi_t - \sqrt{h^{(F)}_t}\lambda^{(F)}_t + \frac{1}{2}\left(\sigma^{(F)}_t\right)^2\right)
\end{aligned}$$

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.1.8), 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. □

A.1.2 Benchmarks

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

Discrete-time Gaussian three-factor model

Augustyniak, Godin, and Hamel 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:

$$\begin{aligned} r_t =& X_t^{(1)} + X_t^{(2)} + X_t^{(3)}, \\ X_{t+1}^{(i)} =& X_t^{(i)} + \kappa_i(\mu_i - X_t^{(i)}) + \sigma_i Z_{t+1}^{(i)}, \qquad i = 1,2,3, \end{aligned} \tag{A.1.9}$$

where $(\kappa_i, \mu_i, \sigma_i)$ are the model parameters and $\{Z_t^{(1)}, Z_t^{(2)}, Z_t^{(3)}\}_{t=1}^{T}$ 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 $Z_i^{\mathbb{Q}} = \{Z_{t,i}^{\mathbb{Q}}\}_{t=1}^{T}$, $i = 1,2,3$ 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,

$$X_{t+1}^{(i)} = X_t^{(i)} + \kappa_i^{\mathbb{Q}}(\mu_i^{\mathbb{Q}} - X_t) + \sigma_i Z_{t+1}^{\mathbb{Q}}, \quad i = 1, 2, 3 \tag{A.1.10}$$

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, Godin, and Hamel (2021) show that under (A.1.10) 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

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

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, $\upsilon_\tau^{i,l}$, is

$$\upsilon_\tau^{i,l} = \frac{\sigma_i \sigma_l}{\kappa_i^{\mathbb{Q}} \kappa_l^{\mathbb{Q}}} \Gamma_{i,l} \left[\tau - B_\tau^{(i)} - B_\tau^{(l)} + \frac{1 - (1 - \kappa_i^{\mathbb{Q}})^\tau (1 - \kappa_l^{\mathbb{Q}})^\tau}{1 - (1 - \kappa_i^{\mathbb{Q}})(1 - \kappa_l^{\mathbb{Q}})}\right].$$

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 A.1.

Table A.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, Godin, and Hamel 2021 given by Equations (A.1.9)-(A.1.10). 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.

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 (2.2.15).

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 A.2.

Table A.2: Dynamic Nelson-Siegel model parameters estimates.

i	$\kappa^{\mathbb{P}}_{i,i}$	γ_i	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_i$	$\theta^{\mathbb{Q}}_i$	$\bar{x}_{i,1\|0}$	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		

$$\rho = \begin{bmatrix} 1 & -0.8397 & -0.9202 \\ 0.8397 & 1 & 0.7106 \\ -0.9202 & 0.7106 & 1 \end{bmatrix}, \; P_{1|0} = \begin{bmatrix} 4\times10^{-6} & 0 & 0 \\ 0 & 4\times10^{-6} & 0 \\ 0 & 0 & 4\times10^{-6} \end{bmatrix}$$

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.

A.1.3 Kalman filter

Filtering techniques are statistical methods applied to perform inference on latent variables, referred to as the signal, based on observations linked to such signal. The most popular filtering method is the Kalman filter which can be applied under the assumption of linear and Gaussian dynamics for the evolution of the signal and the relationship between the signal and observations. More precisely, as presented in Rémillard 2013, the such dynamics for the signal $\{Z_i\}_{i=1}^n$ and observations $\{Y_i\}_{i=1}^n$ is represented by

$$Z_i = \mu_i + F_i Z_{i-1} + G_i w_i,$$
$$Y_i = d_i + U_i Z_i + \epsilon_i,$$

where $\mu_i \in R^m$, $F_i \in R^{m\times m}$, $G_i \in R^{m\times q}$, $d_i \in R^r, U_i \in R^{r\times m}$, $w_i \sim N_q(0, W_i)$ and $\epsilon_i \sim N_r(0, R_i)$. Independent processes $\{w_i\}_{i=1}^n$ and $\{\epsilon_i\}_{i=1}^n$ are sequences of independent

and identically distributed random variables. For $i = 1, \ldots, n$, set

$$\begin{aligned}
Y_{i|i-1} &= \mathbb{E}[Y_i|Y_{i-1}, ..., Y_1], \\
Z_{i|i-1} &= \mathbb{E}[Z_i|Y_{i-1}, ..., Y_1], \\
Z_{i|i} &= \mathbb{E}[Z_i|Y_i, ..., Y_1], \\
P_{i|i-1} &= \mathbb{E}[(Z_i - Z_{i|i-1})(Z_i - Z_{i|i-1})^\top|Y_{i-1}, ..., Y_1], \\
P_{i|i} &= \mathbb{E}[(Z_i - Z_{i|i})(Z_i - Z_{i|i})^\top|Y_i, ..., Y_1], \\
e_i &= Y_i - Y_{i|i-1}, \\
V_i &= \mathbb{E}[e_i e_i^T|Y_{i-1}, ..., Y_1].
\end{aligned} \tag{A.1.11}$$

Then, the following proposition holds, see Rémillard 2013.

Proposition A.1.1. *Suppose that* $Z_0 \sim N_m(Z_{0|0}, P_{0|0})$. *Given* $Y_1, \ldots, Y_i$, *the conditional distribution*

- *of* Z_i *is Gaussian, with mean* $Z_{i|i}$ *and covariance matrix* $P_{i|i}$,
- *of* Z_{i+1} *is Gaussian with mean* $Z_{i+1|i}$ *and covariance matrix* $P_{i+1|i}$,
- *of* Y_{i+1} *is Gaussian with mean* $Y_{i+1|i}$ *and covariance matrix* V_{i+1}.

Kalman 1960 provides a recursive algorithm to calculate conditional moments of the signal and observations outlined in Proposition A.1.1. For given initial moments $Z_{0|0}$ and $P_{0|0}$, the algorithm goes as follows for $i = 1, \ldots, n$:

$$\begin{aligned}
Z_{i|i-1} &= \mu_i + F_i Z_{i-1|i-1}, \\
Y_{i|i-1} &= d_i + U_i Z_{i|i-1}, \\
P_{i|i-1} &= F_i P_{i-1|i-1} F_i^\top + G_i W_i G_i^\top, \\
V_i &= U_i P_{i|i-1} U_i^\top + R_i,
\end{aligned}$$

$$Z_{i|i} = Z_{i|i-1} + P_{i|i-1}U_i^\top V_i^{-1} e_i,$$
$$P_{i|i} = P_{i|i-1} - P_{i|i-1}U_i^\top V_i^{-1} U_i P_{i|i-1}.$$

Using previous results and calculations, the joint distribution of $Y_1, ..., Y_n$ can be computed, and the corresponding log-likelihood L is given by

$$-L = -\sum_{i=1}^{n} \log p(Y_i|Y_1, \ldots, Y_{i-1}) = \frac{1}{2}\sum_{i=1}^{n} \log\left(\det V_i\right) + \frac{1}{2}\sum_{i=1}^{n} e_i^\top V_i^{-1} e_i + \frac{nr}{2}\ln(2\pi).$$

To estimate the parameters of the DTAFNS model, the following substitutions need to be applied in the above Kalman filter algorithm: $Y_i = \hat{\mathsf{y}}_i$, $Z_i = X_i$, $i = t$, $r = M$, $m = q = 3$, $\mu_i = \mathbf{b} = \kappa^{\mathbb{P}}\theta^{\mathbb{P}}$, $F = \mathbf{D} = (I - \kappa^{\mathbb{P}})$, $G_i = \Sigma$, $d_i = \mathsf{a}$, $U_i = \mathsf{B}$, $R_i = H$ and $W_i = \rho$.

A.1.4 Alternative parameters selection for DTAFNS

As explained in Remark 2.2.2, it might be possible to consider a specification of the DTAFNS model under which $\kappa^{\mathbb{P}}_{1,1} = 0$, which (i) would be consistent with $\kappa^{\mathbb{Q}}_{1,1} = 0$ and (ii) would also yield non-stationary dynamics for the first factor under physical dynamics. From (2.2.17), setting $\kappa^{\mathbb{P}}_{1,1} = 0$ leads to $\gamma_1 = 0$. Such specification is tested in this appendix. Table A.3 shows the parameter estimation results under such additional constraints.

Table A.4 shows the in-sample and out-of-sample log-likelihoods of both the original version of the DTAFNS model from Section 2.2.2 and the version with the added constraint $\kappa^{\mathbb{P}}_{1,1} = 0$. Imposing $\kappa^{\mathbb{P}}_{1,1} = 0$ deteriorates the results both in-sample and out-of-sample, which justified ultimately not retaining this specification.

Table A.3: Maximum likelihood estimates of the DTAFNS model with constraint $\kappa^{\mathbb{P}}_{1,1} = 0$

$$\rho = \begin{bmatrix} 1 & -0.6300 & -0.4101 \\ -0.6300 & 1 & 0.2991 \\ -0.4101 & 0.2991 & 1 \end{bmatrix}$$

i	$\kappa^{\mathbb{P}}_{i,i}$	γ_i	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_i$	$\theta^{\mathbb{Q}}_i$	λ	h
1	0	0	0.0027	0	0		
2	0.2966	1.3807	0.0046	0.0501	0.0637	0.0233	3.76×10^{-6}
3	0.0361	1.8247	0.0070	0.0495	0.0767		

Notes: Maximum likelihood parameters estimates for the DTAFNS model presented in Section 2.2.2 obtained with the Kalman filter-based Algorithm 1. The data sample is the Canadian end-of-month yield curve data extending from January 1986 to January 2022.

Table A.4: Log-likelihood of the DTAFNS model with and without constraint $\kappa^{\mathbb{P}}_{1,1} = 0$

Model	Out-of-sample						In-sample
	2017	2018	2019	2020	2021-22	Aggregated	
DTAFNS ($\kappa^{\mathbb{P}}_{1,1} \neq 0$)	2,034	2,001	2,035	2,015	2,183	10,268	65,702
DTAFNS ($\kappa^{\mathbb{P}}_{1,1} = 0$)	2,005	1,977	2,015	2,002	2,175	10,174	65,700

Notes: Comparison of the log-likelihood of the DTAFNS model of Section 2.2.2 and a version of the same model with the added constraint $\kappa^{\mathbb{P}}_{1,1} = 0$. The data sample consists of Canadian end-of-month yield curves. 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 2.3.3. The aggregated out-of-sample log-likelihood is obtained by summing the log-likelihoods for all test years.

A.1.5 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 $t-1$ with the expanding window approach detailed in Section 2.3.3, i.e. models are retrained yearly, with yearly out-of-sample test sets covering the period extending from January 2017 to

Table A.5: Probability of observing negative short rates with the DTAFNS model with and without constraint $\kappa^{\mathbb{P}}_{1,1} = 0$

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($\kappa^{\mathbb{P}}_{1,1} \neq 0$)	0.177	0.057	0.015	0.003	0.644	0.299	0.106	0.029
DTAFNS($\kappa^{\mathbb{P}}_{1,1} = 0$)	0.322	0.141	0.052	0.016	0.826	0.540	0.277	0.113

Notes: Within 200,000 five-year monthly simulated paths of the DTAFNS model (either with or without constraint $\kappa^{\mathbb{P}}_{1,1} = 0$), the proportion of simulated months with short rates being smaller than respectively 0, -0.01, -0.02 and -0.03, and the proportion of simulated paths with at least one month below such thresholds. Model parameters estimates are from Table A.3 and Table 2.1, with starting values of the factors being smoothed factor values on January 2022 associated with such sets of parameters.

January 2022. Define the following performance metrics:

$$\begin{aligned}
\text{Mean error} =& \frac{1}{T^*}\sum_{t=1}^{T^*}(\hat{y}(t,t+\tau) - \tilde{y}(t,t+\tau)),\\
RMSE^{(\tau)} =& \sqrt{\frac{1}{T^*}\sum_{t=1}^{T^*}(\hat{y}(t,t+\tau) - \tilde{y}(t,t+\tau))^2},\\
MAE^{(\tau)} =& \frac{1}{T^*}\sum_{t=1}^{T^*}|\hat{y}(t,t+\tau) - \tilde{y}(t,t+\tau)|,\\
rMAE_1^{(\tau)} =& \frac{MAE^{(\tau)}_{DTAFNS}}{MAE^{(\tau)}_{DG3}}, \quad rMAE_2^{(\tau)} = \frac{MAE^{(\tau)}_{DTAFNS}}{MAE^{(\tau)}_{DNS}}
\end{aligned} \tag{A.1.12}$$

where T^* is the total number of observations in the union of all test sets and $MAE^{(\tau)}_{\mathcal{M}}$ is the $MAE^{(\tau)}$ for model $\mathcal{M}$. Table A.6 reports the predictive performance results. RMSE and MAE metrics indicate that the point forecast performance of the DTAFNS model is quite similar to that of the DG3 and DNS models.

A.1.6 Fit diagnostics output for the mixed fund model

In this section, diagnostic information about mixed fund parameter estimates is presented. Parameters mu, mxreg1, mxreg2, mxreg3, mxreg4, mxreg5, mxreg6 in the output below correspond to ψ_0, ψ_1, ψ_2, ψ_3, ψ_3', $\psi_1^{(S)}$, and $\psi_2^{(S)}$, respectively. Results of several white noise

Table A.6: Performance metrics for spot rate point predictions

Tenor	Mean error (in %)			RMSE (in %)			MAE (in %)			rMAE	
(months)	DTAFNS	DG3	DNS	DTAFNS	DG3	DNS	DTAFNS	DG3	DNS	$rMAE_1$	$rMAE_2$
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 point forecasting of spot rates by the DTAFNS model and its two benchmarks explained in Appendix A.1.2. The expanding window approach detailed in Section 2.3.3 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. The rMAE metric is explained in (A.1.12). 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.

tests are also presented to verify the adequacy of the model; weighted Ljung-Box tests on standardized residuals and squared residuals, and weighted ARCH Lagrange Multiplier (LM) tests are conducted to assess the presence of auto-correlation and volatility clusters in residuals. See the `rugarch` package documentation for details of specifications of such tests.

Optimal Parameters

	Estimate	Std. Error	t value	Pr(>|t|)
mu	0.004039	0.001227	3.2907	0.000999
mxreg1	−1.724464	0.404332	−4.2650	0.000020
mxreg2	−0.769503	0.242516	−3.1730	0.001509
mxreg3	−0.445368	0.111513	−3.9939	0.000065
mxreg4	−0.055235	0.019353	−2.8540	0.004317
mxreg5	0.400758	0.014636	27.3815	0.000000
mxreg6	0.134577	0.013903	9.6800	0.000000
omega	−1.031221	0.012377	−83.3193	0.000000
alpha1	−0.093903	0.035272	−2.6623	0.007761
beta1	0.894535	0.000279	3205.6931	0.000000
gamma1	0.066528	0.020912	3.1814	0.001466

LogLikelihood : 1038.719

Weighted Ljung–Box Test on Standardized Residuals

	statistic	p–value
Lag[1]	0.6482	0.4208
Lag[2*(p+q)+(p+q)−1][2]	1.3704	0.3922
Lag[4*(p+q)+(p+q)−1][5]	2.3021	0.5490

d.o.f=0

H0 : No serial correlation

```
Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3453  0.5568
Lag[2*(p+q)+(p+q)-1][5]    1.7040  0.6898
Lag[4*(p+q)+(p+q)-1][9]    5.8407  0.3169
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
             Statistic Shape Scale P-Value
ARCH Lag[3]     0.8588 0.500 2.000  0.3541
ARCH Lag[5]     2.7687 1.440 1.667  0.3251
ARCH Lag[7]     4.7279 2.315 1.543  0.2535
```

A.2 Participating contract

This appendix reports parameters considered for the financial market model of Section 3.5. Such parameters are drawn directly from Eghbalzadeh, Godin, and Gaillardetz 2022, which were obtained by maximum likelihood on monthly historical time series.

Table A.7 provides parameters for the DTAFNS model (3.5.2)-(3.5.3), which were obtained through calibration to January 1986 to January 2022 Canadian end-of-month yield curve data.

Two equity indices are considered for the equity model (3.5.6)-(3.5.7): the S&P/TSX and the S&P 500. Parameters considered are shown in Table A.8 and were obtained from a

Table A.7: DTAFNS model parameters

i	$\kappa^{\mathbb{P}}_{i,i}$	$\gamma_{i,i}$	$\Sigma_{i,i}$	$\theta^{\mathbb{P}}_{i,i}$	$\theta^{\mathbb{Q}}_{i,i}$	λ
1	0.008	2.792	0.003	0	0	
2	0.029	1.202	0.005	0.030	0.063	0.023
3	0.035	1.717	0.007	0.051	0.077	

$$\rho = \begin{bmatrix} 1 & -0.630 & -0.410 \\ -0.630 & 1 & 0.299 \\ -0.410 & 0.299 & 1 \end{bmatrix}$$

Notes: The data sample used to obtain maximum likelihood estimates is the Canadian end-of-month yield curve data extending from January 1986 to January 2022.

return time series extending from February 1986 to January 2022.

Table A.8: Bivariate EGARCH model parameters

Stock index	j	λ^S_j	$\omega^{(S)}_j$	$\alpha^{(S)}_j$	$\gamma^{(S)}_j$	$\beta^{(S)}_j$	$\Gamma_{1,2}$
S&P/TSX	1	0.08443	-2.38375	-0.16171	0.38711	0.62836	0.75281
S&P 500	2	0.12605	-1.92871	-0.14922	0.32486	0.69715	

Table A.9 indicates parameters used for the mixed fund model (3.5.8)-(3.5.9), which were obtained based on a time series of net asset value (NAV) returns for the *Assumption/Louisbourg Balanced Fund A* extending from February 1996 to January 2022.

Table A.9: Mixed fund model parameters

Parameter	θ_0	θ_1	θ_2	θ_3	θ'_3	$\theta^{(S)}_1$
Estimate	0.0040387	-1.7244637	-0.7695032	-0.4453684	-0.0552352	0.4007580
Parameter	$\theta^{(S)}_2$	$\omega^{(F)}$	$\alpha^{(F)}$	$\gamma^{(F)}$	$\beta^{(F)}$	
Estimate	0.1345770	-1.0312207	-0.0939034	0.0665277	0.8945346	

Starting values for interest rate factors and volatilities (both for the equity indices and the mixed fund) considered in simulations correspond to inferred values on January 31, 2022 and are given in Table A.10.

These values are used as a starting point for the simulation of the various trajectories of the interest rates, equity indices and mixed fund value.

Table A.10: Initial financial risk factors parameters for simulations

Interest rate factors			S&P/TSX	S&P 500	Mixed Fund model
$X_0^{(1)}$	$X_0^{(2)}$	$X_0^{(3)}$	$\sqrt{12h_{0,1}^{(S)}}$	$\sqrt{12h_{0,2}^{(S)}}$	$\sqrt{12h_0^{(F)}}$
-3.12%	3.84%	6.88%	11.88%	16.80%	2.62%

A.3 Swaption pricing under DTAFNS model

A.3.1 Proofs

Before proving Proposition 4.3.1, several lemmas are presented.

Lemma A.3.1. *For* $i = 1, 2, 3$,

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

Proof of Lemma A.3.1: See Lemma A.2 of Eghbalzadeh, Godin, and Gaillardetz 2022.

Lemma A.3.2. *There exists the following recursive relation between time* t *and* $t+1$ *for the zero coupon bond price presented in* (4.2.3)*:*

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

Proof of Lemma A.3.2: Using (4.2.1) and (4.2.3) for the first equality, and then Lemma A.3.1 for the third one,

$$\begin{aligned}
\log\left(\frac{P(t+1, t+\tau)}{P(t, t+\tau)}\right) - \Delta r_t &= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta \sum_{i=1}^{3} \mathcal{B}_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(\frac{\mathcal{B}_\tau^{(i)} - 1}{\mathcal{B}_{\tau-1}^{(i)}}\right) X_t^{(i)}\right) + \Delta \sum_{i=1}^{3} X_t^{(i)} \\
&\quad - \Delta(X_t^{(1)} + X_t^{(2)}) \\
&= \log\left(\frac{A_{\tau-1}}{A_\tau}\right) - \Delta \sum_{i=1}^{2} \mathcal{B}_{\tau-1}^{(i)} \left(X_{t+1}^{(i)} - \left(\frac{\mathcal{B}_\tau^{(i)} - 1}{\mathcal{B}_{\tau-1}^{(i)}}\right) X_t^{(i)}\right) + \Delta X_t^{(3)}
\end{aligned}$$

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

Therefore,

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

Lemma A.3.3. *Consider any integer $\tau > 0$ and any real number $r \neq 1$. For functions $\zeta_0(r,\tau), \zeta_1(r,\tau)$ and $\zeta_2(r,\tau)$ defined in* (4.2.7)-(4.2.9)*,*

$$
\zeta_0(r, \tau-1) - \zeta_0(r,\tau) = -r^{\tau-1} \tag{A.3.1}
$$

$$
\zeta_1(r, \tau-1) - \zeta_1(r,\tau) = -r^{\tau-1}(\tau-1) \tag{A.3.2}
$$

$$
\zeta_2(r, \tau-1) - \zeta_2(r,\tau) = -r^{\tau-1}(\tau-1)^2 \tag{A.3.3}
$$

Proof of Lemma A.3.3: This result is a direct consequence of the sum representations of $\zeta_0(r,\tau), \zeta_1(r,\tau)$ and $\zeta_2(r,\tau)$ provided in (4.2.7)-(4.2.9). $\square$

Lemma A.3.4. *The following recursive connection holds for the quantity υ_τ defined in* (4.2.6)*:*

$$
\upsilon_\tau = \mathcal{B}_{\tau-1}^{\top} \Sigma \rho (\mathcal{B}_{\tau-1}^{\top} \Sigma)^{\top} + \upsilon_{\tau-1}. \tag{A.3.4}
$$

Proof of Lemma A.3.4: First,

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

Moreover, based on (4.2.6),

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

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

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

First, for $i = j = 1$,

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

Secondly, for $i = 1$ and $j = 2$, using (A.3.2),

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

Thirdly, for $i = 1$ and $j = 3$, using (A.3.1), (A.3.2), and (A.3.3),

$$a_{\tau}^{(1,3)} = \frac{1}{\lambda}\Bigg[\frac{\tau(\tau-1)}{2} - 1 - \zeta_0(1-\lambda,\tau-1) - (1+\lambda)\zeta_1(1-\lambda,\tau-1) - \lambda\zeta_2(1-\lambda,\tau-1)$$

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

Fourthly, for $i = j = 2$,

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

Fifthly, for $i = 2$ and $j = 3$, using (A.3.1) and (A.3.2),

$$
\begin{aligned}
a^{(2,3)}_{\tau} =&\frac{\tau - 2 - (2-\lambda)\zeta_0(1-\lambda, \tau-1) + (1-\lambda)\zeta_0((1-\lambda)^2, \tau-1)}{\lambda^2} \\
&+ \frac{-\zeta_1(1-\lambda, \tau-1) + (1-\lambda)\zeta_1((1-\lambda)^2, \tau-1)}{\lambda} \\
&- \frac{\tau - 3 - (2-\lambda)\zeta_0(1-\lambda, \tau-2) + (1-\lambda)\zeta_0((1-\lambda)^2, \tau-2)}{\lambda^2} \\
&- \frac{-\zeta_1(1-\lambda, \tau-2) + (1-\lambda)\zeta_1((1-\lambda)^2, \tau-2)}{\lambda} \\
=&\frac{1 - (2-\lambda)(1-\lambda)^{\tau-2} + (1-\lambda)(1-\lambda)^{2(\tau-2)}}{\lambda^2} + \frac{-(1-\lambda)^{\tau-2}(\tau-2) + (1-\lambda)(1-\lambda)^{2(\tau-2)}(\tau-2)}{\lambda} \\
=&\frac{\left(1-(1-\lambda)^{\tau-2}\right)^2 + \lambda(1-\lambda)^{\tau-2}\left(1-(1-\lambda)^{\tau-2}\right)}{\lambda^2} + \frac{(1-\lambda)^{\tau-2}(\tau-2)\left((1-\lambda)^{\tau-1} - 1\right)}{\lambda}
\end{aligned}
$$

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

Lastly, for $i=j=3$, using (A.3.1), (A.3.2), and (A.3.3),

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

Lemma A.3.5. *For $i=1,2,3$, the following recursive relationships between $\mathcal{B}_\tau^{(i)}$ and $\mathcal{B}_{\tau-1}^{(i)}$ hold:*

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

Proof of Lemma A.3.5: Based on (4.2.4), $\mathcal{B}_\tau^{(1)} = \tau - 1 + 1 = \mathcal{B}_{\tau-1}^{(1)} + 1$. Furthermore,

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

Lastly,

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

Lemma A.3.6. *The following recursive relationship holds for quantity A_τ defined in* (4.2.3)*:*

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

Proof of Lemma A.3.6: Using Lemma A.3.4, Lemma A.3.5 and (4.2.4) to substitute into (4.2.5),

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

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

Proof of Proposition 4.3.1: The proof relies on calculating the moments generating function of innovations under the $\mathcal{T}$-forward measure. Consider the row vector $\Gamma = [\Gamma_1, \Gamma_2, \Gamma_3]$ with $\Gamma_i \in \mathbb{R}$ for $i = 1, 2, 3$. Then

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

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

Define

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

Using Lemma A.3.2 therefore leads to

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

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

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

Thus,

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

Therefore, using Lemma A.3.6 leads to

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

Lemma A.3.7. *Assume* (4.3.5) *holds. For* $t = 0, \ldots, \mathcal{T}$ *and* $n = 0, \ldots, \mathcal{T} - t$,

$$\begin{aligned}X_{t+n}^{(i)} =& X_t^{(i)}(1-\kappa_{i,i}^{\mathcal{T}})^n + \kappa_{i,i}^{\mathcal{T}}\theta_i^{\mathcal{T}}\sum_{l=1}^{n}(1-\kappa_{i,i}^{\mathcal{T}})^{(n-l)} + \Sigma_{i,i}\sum_{l=1}^{n}Z_{t+l,i}^{\mathcal{T}}(1-\kappa_{i,i}^{\mathcal{T}})^{(n-l)}\\ &+ \sum_{l=1}^{n}\sum_{j\neq i}^{3}\kappa_{i,j}^{\mathcal{T}}(\theta_j^{\mathcal{T}} - X_{t+l-1}^{(j)})(1-\kappa_{i,i}^{\mathcal{T}})^{(n-l)} - \sum_{l=0}^{n-1}\eta_{t+l}^{(i)}(1-\kappa_{i,i}^{\mathcal{T}})^{n-1-l}. \qquad \text{(A.3.5)}\end{aligned}$$

Proof of Lemma A.3.7: This proof is analogous to that of Lemma A.1 from Eghbalzadeh, Godin, and Gaillardetz 2022, which is based on induction. We apply the convention $\sum_{l=1}^{0} x_l \equiv 0$. The case $n = 0$ is therefore trivial. Then, assume (A.3.5) holds for some

$n < T - t - 1$. Using (4.3.5), for any $i = 1, 2, 3$,

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

Applying (A.3.5) in the latter equality yields

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

thereby finishing the induction. □

Lemma A.3.8. *For $n = 0, \ldots, \mathcal{T} - t$, the factors $X_{t+n}^{(i)}$ can be expressed in terms of X_t and innovations $\{Z_{t+l}^{\mathcal{T}}\}_{l=1}^{\mathcal{T}-t}$ as follows:*

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

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

Proof of Lemma A.3.8:

From (4.3.4), $\kappa_{1,1}^{\mathcal{T}} = \kappa_{1,2}^{\mathcal{T}} = \kappa_{1,3}^{\mathcal{T}} = \kappa_{2,1}^{\mathcal{T}} = \kappa_{3,1}^{\mathcal{T}} = \kappa_{3,2}^{\mathcal{T}} = 0$, $\kappa_{2,2}^{\mathcal{T}} = \kappa_{3,3}^{\mathcal{T}} = \lambda$ and $\kappa_{2,3}^{\mathcal{T}} = -\lambda$. When placed into (A.3.5), this leads to

$$\begin{aligned}
X_{t+n}^{(1)} &= X_t^{(1)} + \Sigma_{1,1}\sum_{l=1}^{n} Z_{t+l,1}^{\mathcal{T}} - \sum_{l=0}^{n-1}\eta_{t+l}^{(1)}, \\
X_{t+n}^{(2)} &= X_t^{(2)}(1-\lambda)^n + \lambda\theta_2^{\mathcal{T}}\sum_{l=1}^{n}(1-\lambda)^{(n-l)} + \Sigma_{2,2}\sum_{l=1}^{n} Z_{t+l,2}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad -\lambda\sum_{l=1}^{n}(\theta_3^{\mathcal{T}} - X_{t+l-1}^{(3)})(1-\lambda)^{(n-l)} - \sum_{l=0}^{n-1}\eta_{t+l}^{(2)}(1-\lambda)^{n-1-l}, \qquad \text{(A.3.6)} \\
X_{t+n}^{(3)} &= X_t^{(3)}(1-\lambda)^n + \lambda\theta_3^{\mathcal{T}}\sum_{l=1}^{n}(1-\lambda)^{(n-l)} + \Sigma_{3,3}\sum_{l=1}^{n} Z_{t+l,3}^{\mathcal{T}}(1-\lambda)^{(n-l)} \\
&\quad -\sum_{l=0}^{n-1}\eta_{t+l}^{(3)}(1-\lambda)^{n-1-l}. \qquad \text{(A.3.7)}
\end{aligned}$$

Furthermore,

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

$$
\begin{aligned}
=&nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\left(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1}\right) \\
&+ \Sigma_{3,3}\sum_{k=1}^{n-1} Z_{t+k,3}^{\mathcal{T}} \sum_{l=k}^{n-1}(1-\lambda)^{n-k-1} - \sum_{k=0}^{n-1}\eta_{t+k}^{(3)}\sum_{l=k+1}^{n-1}(1-\lambda)^{n-k-2} \\
=&nX_t^{(3)}(1-\lambda)^{n-1} + \theta_3^{\mathcal{T}}\left(\frac{1-(1-\lambda)^n}{\lambda} - n(1-\lambda)^{n-1}\right) \\
&+ \Sigma_{3,3}\sum_{k=1}^{n-1}(n-k)(1-\lambda)^{n-k-1} Z_{t+k,3}^{\mathcal{T}} - \sum_{k=0}^{n-1}(n-k-1)\eta_{t+k}^{(3)}(1-\lambda)^{n-k-2}.
\end{aligned}
\tag{A.3.8}
$$

Using (A.3.8) in (A.3.6),

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

Moreover, using (4.2.7) in (A.3.7) leads to

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

Proof of Proposition 4.3.2:

The joint normality of X_{t+n} given $\mathcal{F}_t$ is a direct consequence of Lemma A.3.8, which expresses X_{t+n} as a linear combination of the $\mathcal{F}_t$-measurable elements of X_t and of jointly Gaussian innovations $Z_{t+1}^{\mathcal{T}}, \ldots, Z_{t+n}^{\mathcal{T}}$ that are independent of $\mathcal{F}_t$. The composition of $\mathcal{M}_{t,n}$ is also a direct consequence of Lemma A.3.8, and of the null expectation of innovations $Z_{t+1}^{\mathcal{T}}, \ldots, Z_{t+n}^{\mathcal{T}}$ given $\mathcal{F}_t$.

Components of $\mathcal{V}_n$ are also obtained through Lemma A.3.8 and (4.2.7)-(4.2.9):

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

and

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

Proof of Lemma 4.3.1: From, (4.3.4), for $i = 1, 2, 3$, $\eta_t^{(i)} = \Delta\Sigma_{i,i} \sum_{j=1}^{3} \Sigma_{j,j}\rho_{i,j}\mathcal{B}_{\mathcal{T}-t-1}^{(j)}$.

Therefore,

$$\sum_{l=0}^{n-1} \eta_{t+l}^{(1)} = \Delta\Sigma_{1,1} \sum_{l=0}^{n-1} \left(\Sigma_{1,1}\rho_{1,1}\mathcal{B}_{\mathcal{T}-t-l-1}^{(1)} + \Sigma_{2,2}\rho_{1,2}\mathcal{B}_{\mathcal{T}-t-l-1}^{(2)} + \Sigma_{3,3}\rho_{1,3}\mathcal{B}_{\mathcal{T}-t-l-1}^{(3)}\right)$$

$$
\begin{aligned}
=&\Delta\Sigma_{1,1}\Bigg[\Sigma_{1,1}\left(n(\mathcal{T}-t-1)-\frac{n(n-1)}{2}\right)+\frac{\Sigma_{2,2}\rho_{1,2}}{\lambda}\left(n-(1-\lambda)^{\mathcal{T}-t}\zeta_0\left((1-\lambda)^{-1},n+1\right)\right)\\
&+\Sigma_{3,3}\rho_{1,3}\Bigg(\frac{n-(1-\lambda)^{\mathcal{T}-t-1}\zeta_0\left((1-\lambda)^{-1},n+1\right)}{\lambda}\\
&-(\mathcal{T}-t-2)(1-\lambda)^{\mathcal{T}-t-1}\zeta_0\left((1-\lambda)^{-1},n+1\right)+(1-\lambda)^{\mathcal{T}-t-2}\zeta_1\left((1-\lambda)^{-1},n\right)\Bigg)\Bigg].
\end{aligned}
$$

Moreover, for $i=2,3$,

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

Lastly,

$$
\sum_{k=0}^{n-1}(n-k-1)\eta_{t+k}^{(3)}(1-\lambda)^{n-k-2}=\frac{n-1}{1-\lambda}\sum_{k=0}^{n-1}\eta_{t+k}^{(3)}(1-\lambda)^{n-k-1}-\sum_{k=0}^{n-1}k\eta_{t+k}^{(3)}(1-\lambda)^{n-k-2}.
$$

Furthermore,

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

$$- (1-\lambda)^{\mathcal{T}-t-2}\left[(\mathcal{T}-t-2)\zeta_1\left((1-\lambda)^{-2},n\right)-\zeta_2\left((1-\lambda)^{-2},n\right)\right]\Bigg)\Bigg]. \quad \square$$

Bibliography

Ang, A., J. Boivin, S. Dong, and R. Loo-Kung. 2011. Monetary policy shifts and the term structure. *The Review of Economic Studies* 78 (2): 429–457.

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

Augustyniak, M., F. Godin, and E. Hamel. 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.

Bacinello, A. R. 2003. Pricing guaranteed life insurance participating policies with annual premiums and surrender option. *North American Actuarial Journal* 7 (3): 1–17.

Bacinello, A. R., and F. Ortu. 1994. "Single and periodic premiums for guaranteed equity-linked life insurance under interest-rate risk: the "lognormal+ Vasicek" Case." In *Financial Modelling,* 1–25. Springer.

Bacinello, A. R., and S.-A. Persson. 2002. Design and pricing of equity-linked life insurance under stochastic interest rates. *The Journal of Risk Finance.*

Bauer, D., R. Kiesel, A. Kling, and J. Ruß. 2006. Risk-neutral valuation of participating life insurance contracts. *Insurance: Mathematics and Economics* 39 (2): 171–183.

Bernard, C., O. Le Courtois, and F. Quittard-Pinon. 2005. Market value of life insurance contracts under stochastic interest rates and default risk. *Insurance: Mathematics and Economics* 36 (3): 499–516.

Black, F. 1976. Studies of stock market volatility changes. *1976 Proceedings of the American statistical association business and economic statistics section.*

Bolder, D., A. Metzler, and G. Johnson. 2004. *An Empirical Analysis of the Canadian Term Structure of Zero-Coupon Interest Rates.* Staff Working Paper 2004-48. Bank of Canada.

Booth, P., R. Chadburn, S. Haberman, D. James, Z. Khorasanee, R. H. Plumb, and B. Rickayzen. 2020. *Modern actuarial theory and practice.* CRC Press.

Bowers, N., H. Gerber, J. Hickman, D. Jones, and C. Nesbitt. 1997. *Actuarial Mathematics.* Society of Actuaries.

Brennan, M. J., and E. S. Schwartz. 1979. A continuous time approach to the pricing of bonds. *Journal of Banking & Finance* 3 (2): 133–155.

Brigo, D., and F. Mercurio. 2007. *Interest rate models-theory and practice: with smile, inflation and credit.* Springer Science & Business Media.

Briys, E., and F. De Varenne. 2001. "Insurance: from underwriting to derivatives." In *Financial Innovations and the Welfare of Nations,* 301–314. Springer.

Chen, R.-R., and L. Scott. 1992. Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. *The Review of Financial Studies* 5 (4): 613–636.

Chen, R.-R., and L. Scott. 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., F. X. Diebold, and G. D. Rudebusch. 2011. The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics* 164 (1): 4–20.

CIA. 2014. "Canadian pensioners' mortality." Canadian Institute of Actuaries. Document 214013.

Collin-Dufresne, P., and R. S. Goldstein. 2002. Pricing swaptions within an affine framework. *The Journal of Derivatives* 10 (1): 9–26.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica* 53 (2): 385–407.

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

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

Duan, J.-C., and J.-G. Simonato. 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., and R. Kan. 1996. A yield-factor model of interest rates. *Mathematical Finance* 6 (4): 379–406.

Eghbalzadeh, R., F. Godin, and P. Gaillardetz. 2022. The discrete-time arbitrage-free Nelson–Siegel model: a closed-form solution and applications to mixed funds representation. *Available at SSRN 4263048.*

Engle, R., G. Roussellet, and E. Siriwardane. 2017. Scenario generation for long run interest rate risk assessment. *Journal of Econometrics* 201 (2): 333–347.

Fan, K. 2013. An FFT approach to price guaranteed minimum death benefit in variable annuities under a regime-switching model. *Chinese Journal of Applied Probability and Statistics* 29 (5): 531–546.

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

Gatzert, N., I. Holzmüller, and H. Schmeiser. 2012. Creating customer value in participating life insurance. *Journal of Risk and Insurance* 79 (3): 645–670.

Gatzert, N., and A. Kling. 2007. Analysis of participating life insurance contracts: a unification approach. *Journal of Risk and Insurance* 74 (3): 547–570.

Geman, H. 1989. *The importance of the forward neutral probability in a stochastic approach of interest rates.* Technical report. Working paper, ESSEC.

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

Gouriéroux, C., and J. Jasiak. 2006. Autoregressive gamma processes. *Journal of Forecasting* 25 (2): 129–152.

Grosen, A., and P. L. Jørgensen. 2000. Fair valuation of life insurance liabilities: the impact of interest rate guarantees, surrender options, and bonus policies. *Insurance: Mathematics and Economics* 26 (1): 37–57.

Guidolin, M., and A. Timmermann. 2006. Term structure of risk under alternative econometric specifications. *Journal of Econometrics* 131 (1-2): 285–308.

Haberman, S., L. Ballotta, and N. Wang. 2003. Modelling and valuation of guarantees in with-profit and unitised with profit life insurance contracts. *Cass Business School Research Paper.*

Hansen, M., and K. R. Miltersen. 2002. Minimum rate of return guarantees: the Danish case. *Scandinavian Actuarial Journal* 2002 (4): 280–318.

Hardy, M. 2003. *Investment guarantees: modeling and risk management for equity-linked life insurance.* Vol. 215. John Wiley & Sons.

Heath, D., R. Jarrow, and A. Morton. 1992. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society,* 77–105.

Hong, Z., L. Niu, and G. Zeng. 2016. Discrete-time arbitrage-free Nelson-Siegel term structure model and application. *Available at SSRN 2731041.*

Hördahl, P., O. Tristani, and D. Vestin. 2006. A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics* 131 (1-2): 405–444.

Hull, J., and A. White. 1990. Pricing interest-rate-derivative securities. *The Review of Financial Studies* 3 (4): 573–592.

IASB. 2017. *Basis for conclusion on IFRS 17 insurance contracts.* Technical report. International Accounting Standards Board.

Jamshidian, F. 1996. Bond, futures and option evaluation in the quadratic interest rate model. *Applied Mathematical Finance* 3 (2): 93–115.

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.

Kijima, M., and T. Wong. 2007. Pricing of ratchet equity-indexed annuities under stochastic interest rates. *Insurance: Mathematics and Economics* 41 (3): 317–338.

Le, A., K. J. Singleton, and Q. Dai. 2010. Discrete-Time Affine$^{\mathbb{Q}}$ Term Structure Models with Generalized Market Prices of Risk. *The Review of Financial Studies* 23 (5): 2184–2227.

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

Lin, X. S., K. S. Tan, and H. Yang. 2009. Pricing annuity guarantees under a regime-switching model. *North American Actuarial Journal* 13 (3): 316–332.

Litterman, R., and J. Scheinkman. 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1 (1): 54–61.

Munk, C. 1999. Stochastic duration and fast coupon bond option pricing in multi-factor models. *Review of Derivatives Research* 3 (2): 157–181.

Nelson, C. R., and A. F. Siegel. 1987. Parsimonious modeling of yield curves. *Journal of Business,* 473–489.

Nelson, D. B. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society,* 347–370.

Ng, A. C.-Y., J. S.-H. Li, and W.-S. Chan. 2011. Modeling investment guarantees in Japan: A risk-neutral GARCH approach. *International Review of Financial Analysis* 20 (1): 20–26.

Nielsen, J. A., and K. Sandmann. 1995. Equity-linked life insurance: A model with stochastic interest rates. *Insurance: Mathematics and Economics* 16 (3): 225–253.

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.

Rogers, L. 1995. Which model for term-structure of interest rates should one use? *Institute for Mathematics and Its Applications* 65:93.

Schrager, D. F., and A. A. Pelsser. 2006. Pricing swaptions and coupon bond options in affine term structure models. *Mathematical Finance* 16 (4): 673–694.

Shumway, R. H., and D. S. Stoffer. 2017. *Time series analysis and its applications. With R examples.* Vol. 4. Springer Texts in Statistics.

Singleton, K. J., and L. Umantsev. 2002. Pricing coupon-bond options and swaptions in affine term structure models. *Mathematical Finance* 12 (4): 427–446.

Svensson, L. E. 1994. *Estimating and interpreting forward interest rates: Sweden 1992-1994.*

Tanskanen, A. J., and J. Lukkarinen. 2003. Fair valuation of path-dependent participating life insurance contracts. *Insurance: Mathematics and Economics* 33 (3): 595–609.

Tiong, S. 2013. Pricing inflation-linked variable annuities under stochastic interest rates. *Insurance: Mathematics and Economics* 52 (1): 77–86.

Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5 (2): 177–188.

Wilkie, A. 1987. An option pricing approach to bonus policy. In memoriam Anthony P. Limb. *Journal of the Institute of Actuaries* 114 (1): 21–90.

Wüthrich, M. V., and M. Merz. 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 diss., Department of ESS, Stanford University.

Zaglauer, K., and D. Bauer. 2008. Risk-neutral valuation of participating life insurance contracts in a stochastic interest rate environment. *Insurance: Mathematics and Economics* 43 (1): 29–40.

Zheng, H., J. Hao, M. Bai, and Z. Zhang. 2019. Valuation of guaranteed unitized participating life insurance under MEGB2 distribution. *Discrete Dynamics in Nature and Society* 2019.