Rheinland-Pfälzische
Technische Universität
Kaiserslautern
Landau

Stochastic Control in the Dynamic Nelson-Siegel Framework

Indira Dhar

Fachbereich Mathematik
AG Finanzmathematik

Gutachter:

Prof. Dr. Ralf Korn
Prof. Dr. Holger Kraft

Datum der Disputation: 27.03.2026

Vom Fachbereich Mathematik der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation

DE-386

To my grandfather, Asim Kanti Biswas, who was a brilliant mathematician, a humorous friend and an exceptionally kind human.

Acknowledgements

First of all, I would like to express my sincerest and most heartfelt gratitude to Prof. Dr. Ralf Korn for agreeing to supervise me throughout this incredible journey. I thank him immensely for imparting his kindness, intelligence, and valuable wisdom from the very first day of our contact. It has been an incredible privilege to be his student.

I would also like to sincerely thank Prof. Dr. Holger Kraft for his time, patience, and willingness to review my work.

I would like to thank Dr. Falk Triebsch, and Ms. Jessica Borsche very much for their consistent and pleasant support with every minuscule matter that I brought to them. I thank the German Academic Exchange Service for their financial support.

I would like to thank my sixth-floor colleagues, both past and present, for the enjoyable working atmosphere. I thank Rana and Kezang for our shared camaraderie during this exciting phase of our lives. I am also grateful to the insightful comments and discussions during my talks at conferences that have helped me structure my research questions.

This journey would not have been possible without my loved ones residing several timezones away. My friends, Aastha and Sanskrit, thank you for your companionship and your patience especially during the last leg. My grandmother, Dida, thank you very much for the love, care, and support you have given me everyday of my life. My parents, Aruna and Sunit, thank you immensely for all of the wisdom, love, and guidance that I have received from you.

I take this opportunity to specially thank my mother, Aruna, for teaching me everything I know: Your support during the most troublesome days, and your celebration during the most rewarding days have been the greatest pillars of strength and comfort.

Last, but in no conceivable way the least, I thank my sister, Nivedita, for extending her brilliance, kindness, and exemplary friendship way beyond than what was necessary in the past years: Thank you for being a proofreader, a soundboard of ideas, and a constant positive force.

Abstract

The dynamic Nelson-Siegel (DNS) model is one of the most popular term structure models discussed by academicians and practitioners alike. Despite its immense empirical success in the fields of finance, insurance, and economics, it often suffers from theoretical criticisms. Moreover, in the current dynamic world where minute decisions impact global markets, it is possibly restrained by the deterministic nature of its underlying factors. Hence, the primary aim of this thesis is to mathematically dissect the underlying factors of the model, and thereby propose and justify the utility of a stochastic DNS framework.

One of the major contributions of this work is the detailed and elementary proof of existence of all of the possible yield curve shapes attainable by the DNS model, along with the necessary conditions required for the occurrence of each shape.

All of the other main contributions of this work are made by considering a stochastic framework of the DNS model. We, thus, consider that the three $\beta(t)$ factors of the model are driven by stochastic Ornstein-Uhlenbeck processes. This stochastic DNS framework allows for analytically tractable yield curve dynamics, with which we are able to efficiently demonstrate that our framework can reproduce economic stylized facts of the yield curve well-established in both literature and practice. Moreover, we show how the resultant bond price processes possess analytical representations that satisfy financial market interpretations.

Furthermore, we demonstrate a real-world application of our stochastic framework: the continuous-time portfolio optimization problem using the stochastic control methodology of Korn and Kraft (2002) under the stochastic DNS framework. We establish sufficient conditions for assets defined in our framework in order to obtain optimal solutions. Moreover, with an appropriate verification theorem that proves the existence and uniqueness of our optimal results, we validate the significance of employing the stochastic DNS framework to solve such portfolio optimization problems.

As an excursion, we discuss the long-standing criticism of the lack of consistent bond prices in the Nelson-Siegel family and its subsequent impact on the portfolio problem. We demonstrate that despite this absence of consistency, an investor operating in our stochastic DNS framework does not suffer from any unfair unbounded opportunities.

Keywords: dynamic Nelson-Siegel, stochastic Nelson-Siegel, stochastic control, optimal portfolios

Zusammenfassung

Das dynamische Nelson-Siegel (DNS) Modell gehört zu den am weitesten verbreiteten Modellen der Zinsstruktur in Wissenschaft und Praxis. Trotz seines ausgeprägten empirischen Erfolgs in den Bereichen Finanzwirtschaft, Versicherungswesen und Ökonomie ist es wiederholt Gegenstand theoretischer Kritik. Zudem kann die deterministische Struktur seiner zugrunde liegenden Faktoren in einer zunehmend dynamischen Marktumgebung als einschränkend angesehen werden.

Daher besteht das Hauptziel dieser Arbeit darin, die zugrunde liegenden Faktoren des Models mathematisch zu analysieren und dadurch den Nutzen eines stochastischen DNS Rahmenwerks vorzuschlagen und zu begründen.

Ein wesentlicher Beitrag dieser Arbeit ist der elementare und vollständige Existenzbeweis sämtlicher durch das DNS Modell erzeugbarer Zinskurvenformen einschließlich der hierfür notwendigen Bedingungen.

Darüber hinaus werden alle weiteren Ergebnisse im Rahmen einer stochastischen Erweiterung des DNS Modells erzielt, in der die drei $\beta(t)$ Faktoren durch stochastische Ornstein-Uhlenbeck Prozesse modelliert werden. Dieses stochastische DNS Rahmenwerk führt zu analytisch handhabbaren Zinskurvendynamiken und erlaubt den Nachweis, dass zentrale in Theorie und Praxis beobachtete stilisierte Fakten der Zinsstruktur reproduziert werden. Ferner wird gezeigt, dass die resultierenden Anleihepreisprozesse über analytische Darstellungen verfügen, die mit finanzökonomischen Interpretationen vereinbar sind.

Als Anwendung wird ein stetiges Portfoliooptimierungsproblem mithilfe stochastischer Kontrolle nach Korn and Kraft (2002) im Rahmen des stochastischen DNS Modells untersucht. Wir formulieren Bedingungen für die in unserem Rahmenwerk definierten Vermögenswerte, um optimale Lösungen zu erhalten. Mithilfe eines geeigneten Verifikationstheorems werden Existenz und Eindeutigkeit der optimalen Lösung sichergestellt und die Eignung des stochastischen DNS Rahmenwerks für Portfoliooptimierungsprobleme aufgezeigt.

Abschließend wird die bekannte Problematik inkonsistenter Anleihepreise innerhalb der Nelson-Siegel Familie diskutiert. Es wird gezeigt, dass hieraus im vorgestellten stochastischen DNS Rahmenwerk keine unfairen unbegrenzten Möglichkeiten für Investoren entstehen.

Contents

List of Figures

List of Tables

List of Symbols

t	Time
$\{\mathcal{F}_t\}_{t\geq 0}$	Filtration
$W(t)$	Brownian motion adapted to $\mathcal{F}_t$
$(\Omega, \mathcal{F}, \mathbb{P})$	Probability space
$P(X)$	Probability of the occurrence of event X
$f(a, b; c)$	Function f with input variables a and b, and parameter c
NS	Nelson-Siegel model
DNS	Dynamic Nelson-Siegel model
OU process	Ornstein-Uhlenbeck process
SDE	Stochastic differential equation

Chapter 1

Introduction

The *yield curve* is one of the most significant concepts in the world of finance, economics and insurance. In its simplest definition, the yield curve describes the relationship between the yield (or, interest rate) of a fixed-income asset and the remaining time to maturity of the asset. Interestingly, interest rates cannot be directly observed in the market as they are not traded products. As a result, the yield curve is derived from assets whose prices processes are some functions of underlying interest rates over time, and is thereby represented as a term structure of these interest rates. In finance, the yield curve often serves as an indicator of the future performance of markets. The shape of the yield curve over time (whether it is monotonically increasing, monotonically decreasing, or a function with multiple points of extrema) provides an overview of investors' expectations of future economic growth. A normal yield curve has an upward slope which is an indicator of economic expansion, whereas an inverse curve which has a downward slope is an indicator of economic recession. In insurance companies, the yield curve plays a pivotal role as it directly influences the present values of future liabilities and obligations, as well as the current value of income from a portfolio of long-term assets. All in all, the yield curve is a popular and crucial tool that contains important financial and economic information.

Due to its immense importance, a large amount of literature devoted to the theory and practice of yield curves have been amassed over time. Notable examples of literature that discuss different yield curve models and their myriad features are the works of Brigo and Mercurio (2006), Desmettre and Korn (2018), and Filipovic (2009) among others. In this thesis, we shall focus solely on one of the most popular term structure models that has been a subject of extensive interest for academicians and industry practitioners alike: the dynamic Nelson-Siegel model.

The dynamic Nelson-Siegel (DNS) model was introduced by Diebold and Li (2006) as a dynamic extension of the original Nelson-Siegel model of Nelson and Siegel (1987), thereby explaining the colloquial name of the model. The yield curve formula proposed and constructed by Nelson and Siegel (1987) was intended to be a parsimonious representation that possessed the ability to reproduce common yield curve shapes observed in the U.S Treasury market. By fitting observable yield curve data to a parametric function, Nelson and Siegel (1987) were successful in establishing a class of models that was flexible enough to reflect basic shapes of the term structure as well as strong enough to predict out-of-sample yield values.

Diebold and Li (2006) extended this Nelson-Siegel framework by dynamically varying three underlying parameters in order to model the entire yield curve over time. Furthermore, in their analysis, Diebold and Li (2006) were able to impart economic interpretations to these three parameters. The authors demonstrated how these three parameters, or factors, could be interpreted in terms of the level, slope, and curvature of the yield curve. Thus, with this re-interpretation of the Nelson-Siegel framework, Diebold and Li (2006) were able to construct a smooth yield curve model that not only succeeded in estimating long-term yield forecasts with high efficiency, but also possessed a unique feature of interpret-ability that was not often observed in multi-factor term structure models.

This desirable combination of good empirical success and reasonable economic interpretations led the dynamic Nelson-Siegel model to be widely used across central banks worldwide. Often coupled with an second curvature parameter, as introduced by the yield curve model of Svensson (1995), most central banks employ the DNS representation to measure the yield curve of their markets (see Bank for International Settlements (2014) and Gürkaynak et al. (2007)).

The widespread practical popularity of the DNS model has been met with equal interest in academic research. However, much of this research focuses primarily on a deterministic framework of the Nelson-Siegel model. For example, the original paper of Diebold and Li (2006) estimate the three factors using the method of ordinary least squares. Moreover, Diebold et al. (2008) discuss the relationship between the deterministic yield curve factors and global macroeconomic quantities of inflation and real activity. In current markets, where global dynamics are impacted by minuscule decisions, the presence of deterministic factors can restrict the framework to capture all possible realistic scenarios. Furthermore, the model often faces a strong criticism of not satisfying the no-arbitrage condition. Filipović (1999), Filipovic (2009) and Björk and Christensen (1999) discuss how the three factors of the Nelson-Siegel family cannot produce bond prices that are consistent with an arbitrage-free model. Although some popular works have sought to address parts of these theoretical criticisms, a large section of their analyses and revisions is often restricted to empirical aspects and predictive performances. For example, Krippner (2015) discuss how the Nelson-Siegel family can be represented as a three-factor Gaussian affine term structure model. However, their analysis is centered around a statistical evaluation of the yield curve. Then, Christensen et al. (2009) and Christensen et al. (2011) attempt to derive a stochastic framework that imposes the Nelson-Siegel factors to a canonical representation of an arbitrage-free affine model. However, their analysis is structured on the yield estimation and predictive performance of the framework. It is also worth mentioning that while the original papers of both Nelson and Siegel (1987) and Diebold and Li (2006) demonstrate that the model is able to reproduce shapes of the yield curve commonly observed in the market, their assertions are not supplemented with a mathematical proof.

Therefore, with this academic contribution, we aim to address the afore-mentioned research questions pertaining to the DNS family, and thereby develop a solid mathematical foundation that is consistent with established properties of the model. The scope of this research is structured to broadly answer the following questions:

Q1. Is there an elementary mathematical proof that describes all of the possible shapes attainable by the yield curve model of Diebold and Li (2006)?

Q2. Is it possible to construct a valid stochastic framework that allows the dynamic Nelson-Siegel model to retain its economic interpret-ability and empirical aptitude?

Q3. Is it justifiable to use the afore-constructed stochastic framework to solve real-world problems often researched in financial literature, such as the portfolio optimization problem?

Q4. Is there an impact of the lack of so-called consistent bond prices associated with the Nelson-Siegel family on such real-world problems?

Therefore, our primary goal is to use the answers to these questions in order to develop a framework of the DNS model that is both theoretically sound as well as practically applicable. By answering Q1, we aim to provide a mathematical justification to a statement that is not only popularly cited in literature but also observed empirically across global markets. By answering Q2, we aim to create a stochastic extension to the Nelson-Siegel model such that the resultant dynamics and simulations satisfy both analytical tractability as well as well-established economic properties. By answering Q3, we aim to demonstrate a significant use-case of such a stochastic framework. We shall choose a suitable application that can enable academicians to explore the benefits of a stochastic DNS model in further research as well as enable practitioners to optimize their investment and financial decisions. By answering Q4, we aim to comprehensively analyze whether this theoretical critique significantly impacts real-world problems.

Hence, as we can see, all of these four questions have significant relevance to the construction of a nicely defined theoretical structure to the DNS model.

With this backdrop, we shall now proceed with our research analysis. The thesis is structured as follows:

Chapter 2 establishes the groundwork of concepts and terms that shall frequently appear in this thesis. We begin with elementary definitions of the yield curve as well as other terms commonly associated with it. Then, we provide a succinct overview of popular term structure models often discussed in literature. With this, we also signify the importance of the DNS model.

Chapter 3 begins with a solid exploration of the DNS model and the advantages of constructing a stochastic framework for it. We first establish an extensive proof using elementary means to demonstrate the existence of all possible shapes attainable by the original DNS model. With this, we also provide the necessary conditions to obtain each of these shapes. Then, we begin to dissect the DNS model, extract the information provided by each of the underlying factors, and consider a suitable stochastic framework of it. We propose that each of the three underlying factors can be driven by stochastic Ornstein-Uhlenbeck processes such that the resultant yield curve dynamics have analytically tractable closed-form representations. Furthermore, we ensure that the resultant dynamics can satisfy common economic properties of the yield curve observed in the market. Given this base stochastic model, we study three different cases of it. The first case considered is the simplest wherein all three underlying factors are driven by the same Brownian motion. The second case is the analysis under the presence of three uncorrelated Brownian motions. The third and final case is the generalized scenario where the three factors are driven by three correlated Brownian motions. This thorough analysis allows us to observe the behaviour of our theoretical framework under multiple settings, thereby enabling us to account for all realistic scenarios (such as, negative interest rates, mean-reversion, more volatility on the short-end). Furthermore, we tie our earlier discussion of attainable yield curve shapes to the stochastic framework, by suggesting a methodology for calculating the probability of attaining each possible shape. We are able to develop explicit formulae for calculating these probabilities for the case of a single Brownian motion. Due to the complicated nature of extracting probabilities under a two-dimensional stochastic framework, we only demonstrate a methodological guide for the case of multiple Brownian motions. We end the chapter with some important remarks on the key results obtained in this chapter.

Chapter 4 delves intensively into a real-world application of our stochastic DNS framework. For our application, we consider the popular continuous-time portfolio optimization problem pioneered by Merton (1969). We first discuss the Merton problem and discover that due to the stochastic nature of our underlying yield curve framework, we must employ the stochastic control methodology developed by Korn and Kraft (2002) in order to solve the resultant portfolio problem. Moreover, our results must satisfy the conditions of existence and uniqueness stipulated in the theory of stochastic control. Thus, we construct an integral foundation to our problem by blending important results of stochastic control theory with their application in portfolio optimization research.

We ensure that the above foundation also satisfies the theoretical properties of our stochastic DNS framework. To this effect, we observe that solutions of the portfolio problem under our stochastic framework are obtainable if sufficient conditions are applied to assets defined within this framework. The existence and uniqueness of the optimal portfolio solutions is further validated by constructing an appropriate verification theorem for the stochastic DNS framework. This entire methodology helps us verify the mathematical existence as well as the economic interpretations of the obtained optimal portfolios. Lastly, as previously mentioned, we address the theoretical critique of the lack of consistent prices in the Nelson-Siegel family. We discuss the origin of this critique, as highlighted by Filipović (1999). Then, we apply a remedial arbitrage-free version proposed by Christensen et al. (2011) to our constructed problem, and constructively compare its performance with our original results. We discover that in the presence of similar underlying parameters, the general behaviour of optimal portfolios remain the same under both model frameworks. We conclude the chapter with an important theorem on the optimal portfolio problem under the stochastic DNS framework, and some remarks on the main results of this chapter.

Finally, Chapter 5 comprehensively summarizes the analyses performed in this thesis, and provides important conclusions and interpretations of our key results. We discuss the relevance of our results to the initial research questions (Q1 to Q4), and thereby close the chapter with a proposal on the future extensions of our work.

The appendices, Chapter A and Chapter B, contain additional calculations and proofs. The simulations depicted in this work have been performed in R (downloaded from `https://www.r-project.org/`) and can be reproduced.

Chapter 2

Preliminaries

In this chapter, we will establish the preliminary base for the discussions that follow in the subsequent chapters. First, we introduce some terms and concepts that are often used in literature pertinent to term structure models. Then, we present a short overview of selected popular term structure models. Although the field is characterized by a rich array of technical concepts and models, we only present a brief summary in order to understand the concepts that shall later appear in this thesis. A more detailed treatment can be found in Korn (2022), Desmettre and Korn (2018), and Brigo and Mercurio (2006).

In order to define a term structure model, we must first define two important ingredients, namely, the zero-coupon bond and the interest rate (or, yield).

Definition 2.1 (Zero-Coupon Bond)**.** A zero-coupon bond is a financial instrument that guarantees to pay the holder of the bond one unit of money at the time of maturity T. It is also known as a T-bond or a zero bond. At any given time $t \in [0, T]$, its price is denoted by $P(t, T)$.

At the time of maturity T, the price of the zero-coupon bond must be 1 as it ensures a guaranteed payment to its holder. In Figure 2.1, we can observe a hypothetical evolution of the price of a zero-coupon bond maturing in 30 years over time with $P(T, T) = 1$.

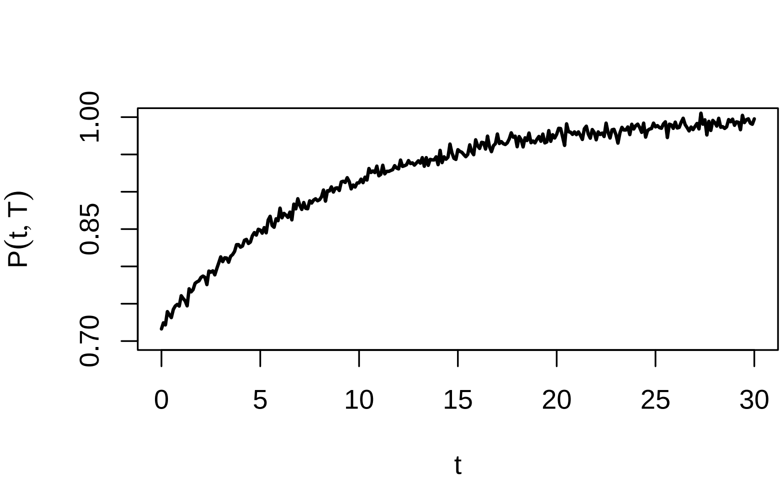


Figure 2.1: Evolution of a Zero-Coupon Bond Price in Time

The zero-coupon bond is one of the simplest financial assets in the market. Since interest rates cannot be traded directly in the financial market, the zero-coupon bond prices serve as a convenient interest rate product that provides a clear understanding of the underlying rates.

Definition 2.2 (Yield to Maturity)**.** The yield $y(t, T)$ of a zero-coupon bond is the equivalent spot interest rate that can be obtained by holding the zero-coupon bond from time t until its maturity time T. The relation between the yield and the price of a zero-coupon bond is given by the following:

$$P(t, T) = \exp\{-y(t, T) \cdot (T - t)\} \tag{2.1}$$

Definition 2.3 (Time to Maturity)**.** The time to maturity τ of a zero bond is the amount of remaining time until the maturity time T calculated from the current calendar time t. It is represented as follows:

$$\tau := T - t \tag{2.2}$$

Hence, (2.1) can be rewritten as follows:

$$P(t, t + \tau) = \exp\{-y(t, t + \tau) \cdot \tau\} \tag{2.3}$$

Now, we may define a term structure model.

Definition 2.4 (Term Structure of Bond Prices)**.** The set of all zero-coupon bond prices at time t for all possible maturities $T > t$, thereby given by $P(t, T)$, is defined as the term structure of bond prices. It can be represented as a function that maps $T \mapsto P(t, T)$.

Equivalently, we arrive at the consequent definition.

Definition 2.5 (Term Structure of Yield Curves)**.** At any fixed time t, a function that maps T to the yields $y(t, T)$ for all possible maturities T, is defined as the term structure of yield curves.

Furthermore, from the definition of yield curves, we can also obtain the following interest rates:

Definition 2.6 (Instantaneous Forward Rate)**.** The instantaneous forward rate $f(t, T)$ at time t for a maturity $T > t$ is obtained from the following relationship:

$$f(t, T) := -\frac{\partial \ln P(t, T)}{\partial T} \tag{2.4}$$

Thus, we get the following equivalence relationship:

$$P(t, T) = \exp\left\{-\int_t^T f(t, u)du\right\} = \exp\{-y(t, T) \cdot (T - t)\}$$

Definition 2.7 (Simple Forward Rate)**.** The simple forward rate is essentially the interest rate locked in at time t for an investment period between $S < T$ and T, and thereby denoted as $f(t, S, T)$. It is related to the instantaneous forward rate $f(t, T)$ by the following:

$$f(t, T) = \lim_{(T-S)\to 0} f(t, S, T)$$

Furthermore, if the investor wishes to hold an investment at time t that ends immediately, then the instantaneous forward rate can be used to obtain a short rate $r(t)$, as follows:

$$r(t) := f(t,t) \tag{2.5}$$

Remark 2.8. Both the instantaneous forward rate $f(t,T)$ and the short rate $r(t)$ are theoretical rates that cannot be observed directly in the market, due to their infinitesimal nature. This is in contrast with the yield $y(t,T)$ that can be directly derived from the market-observable prices of zero-coupon bonds $P(t,T)$.

A term structure model is often presented as a three-dimensional continuum of interest rates, with calendar time t, list of possible term to maturities τ, and the respective values of yield $y(t, t+\tau)$. It can be modeled in different ways using the different rates described so far.

The short-rate approach uses the short rate $r(t)$ to model the term structure of interest rates. Several popular models, such as the Vasicek model of Vasicek (1977), the Hull-White model of Hull and White (1990), and the Cox-Ingersoll-Ross model of Cox et al. (1985), models the short rate as stochastic processes to obtain the resultant term structure. The forward-rate approach uses the instantaneous forward rate $f(t,T)$ to model the term structure of rates. Popular forward-rate models are the Ho-Lee model of Ho and Lee (1986) and the Heath-Jarrow-Morton model of Heath et al. (1992). All of these afore-mentioned models can also be classified into no-arbitrage and equilibrium models, depending on whether they were obtained by fitting the term structure at a specific point of time to prevent arbitrage, such as the models of Hull and White (1990) and Heath et al. (1992), or by defining the underlying rates as stochastic processes whose dynamics capture economic assumptions on the term structure, such as the Vasicek (1977) and Cox et al. (1985) models.

In practice, most term structure models estimate the yield curve surface by fitting the curve using parameters and statistical techniques. One approach is proposed in Fama and Bliss (1987) wherein cubic splines are used to estimate and produce long-term forecasts. However, the most popular approach, and the primary focus of this thesis, is the approach highlighted by Diebold and Li (2006). In this model, they build on the parametric representation of yield curves established by Nelson and Siegel (1987), and extend the yield parameters dynamically. The empirical strengths of the model of Diebold and Li (2006), hereby known as the dynamic Nelson-Siegel term structure model, have been highlighted in numerous research (see Christensen et al. (2011), Diebold and Rudebusch (2013), Diez (2022)). The model, thus, serves as a successful approach to estimate and forecast the term structure of interest rates. As a result, often coupled with an extended version by Svensson (1995), the dynamic Nelson-Siegel model is used by several central banks worldwide, such as the *Deutsche Bundesbank* (Schich (1997)), the European Central Bank, and the Reserve Bank of India (Patra et al. (2021)).

With this preliminary introduction to the concept of term structure models, we may now move forward with our study on the dynamic Nelson-Siegel model.

Chapter 3

Stochastic Dynamic Nelson-Siegel

In this chapter, we extend the dynamic Nelson-Siegel (DNS) yield curve model introduced by Diebold and Li (2006), by considering a stochastic yield curve framework. Essentially, we consider that each of the three $\beta(t)$ factors of the DNS model are driven by stochastic Ornstein-Uhlenbeck processes. With this, we obtain analytically tractable representations for the yield curve and subsequent bond price processes that can be useful to solve further problems in financial mathematics.

In our analysis of the stochastic framework, we consider three main cases: the influence of a single Brownian motion on the three factors, the influence of three uncorrelated Brownian motions, and the influence of three correlated Brownian motions. We show that our framework is able to reproduce behaviours of the yield curves that are consistent with empirical observations and literature. Furthermore, not only are we able to prove the existence and attainability of all possible shapes of the yield curve (normal, inverse, humped, and dipped) in the original DNS model, we are also able to obtain all of these shapes in our stochastic framework. Additionally, we present a methodology to calculate the probabilities of occurrences of these shapes. Thus, with this stochastic framework, we ensure that our stochastic evolution of the DNS model not only satisfies mathematically tractability but also satisfies economic interpretations of the yield curves that have been established both in literature and in practice.

The chapter is structured as follows: First, in Section 3.1, we present the original DNS model of Diebold and Li (2006) and discuss the prime reasons behind developing a stochastic framework for it. Then, in Section 3.2, we present an elementary proof of the four possible shapes attainable by the DNS model. Both Nelson and Siegel (1987) and Diebold and Li (2006) asserted that the Nelson-Siegel family of yield curves were able to produce these shapes commonly seen in the economic markets. With our proof, we aim to provide a theoretical justification to their assertions. Lastly, in Section 3.3 and Section 3.4, we develop a stochastic framework of the DNS yield curve, and thereby obtain analytical representations of their dynamics, and derive explicit formulae of the probabilities of obtaining the four shapes. We conclude the chapter with some remarks on our main results.

Our main contributions in this chapter are:

1. The theoretical proof of existence of normal, inverse, humped, and dipped yield curves in the original dynamic Nelson-Siegel model.

2. The existence of analytically tractable yield curve and bond price dynamics in our stochastic dynamic Nelson-Siegel framework that also satisfy economic interpretations.

3. The formulae for the probabilities of attainable yield curve shapes in our stochastic framework.

3.1 The motivation behind a stochastic model

Nelson and Siegel (1987) proposed a class of functions that had the ability to reproduce shapes commonly associated with yield curves. They referred to these shapes as monotonic, humped, and S-shaped. The primary objective of their research was to develop a parsimonious representation for a yield curve model that could readily generate the afore-mentioned shapes. As a result, using the solution to a second-order differential equation, they arrived at the following instantaneous forward rate model calculated at a time t:

$$f(t, t+\tau) = \beta_0 + \beta_1 \cdot \exp(-\lambda\tau) + \beta_2 \cdot (\lambda\tau \exp(-\lambda\tau)) \tag{3.1}$$

where, τ is the time to maturity, and λ is some constant associated with the equation. Then, the resultant formula for the yield to maturity is given by the following:

$$y(t, t+\tau) = \beta_0 + (\beta_1 + \beta_2) \cdot \frac{\left(1 - \exp(-\lambda\tau)\right)}{\lambda\tau} - \beta_2 \cdot \exp(-\lambda\tau) \tag{3.2}$$

In Figure 3.1, we present a figure from Nelson and Siegel (1987) which depicts the possible yield curve shapes that can be generated from (3.2). While their observations were not supplemented with a formal mathematical proof, Nelson and Siegel (1987) were able to demonstrate that their formula can produce humps, S-shaped curves, and monotonic yield curves, as observable in Figure 3.1.

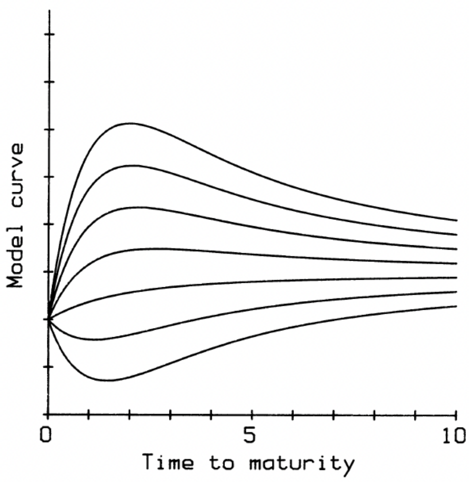


Figure 3.1: Yield Curve Shapes (from Nelson and Siegel (1987))

The dynamic Nelson-Siegel model for yield curves established by Diebold and Li (2006) restructured the original Nelson-Siegel model in the following ways. First, as the name suggests, they introduced a time-evolving dynamic nature to all of the original parameters of the Nelson-Siegel yield curve. Additionally through their model, they were able to impart economic interpretation to these new dynamic parameters.

Hence, (3.2) was redeveloped by Diebold and Li (2006) to denote the yield to maturity at time t for a bond maturing at time T as follows:

$$y(t,T) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_t(T-t)}}{\lambda_t(T-t)}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_t(T-t)}}{\lambda_t(T-t)} - e^{-\lambda_t(T-t)}\right) \tag{3.3}$$

Evidently, from the two equations above, the enhanced features of the yield curve model to display both economic interpretation as well as mathematical accuracy with dynamically evolving parameters can be observed.

1. The β parameters of the second and third terms in equation (3.2) are redistributed in (3.3). The implications are:
   (a) The new β_{1t} and β_{2t} offer a better interpretation as the slope and curvature of the yield curve, respectively, as opposed to an unclear interpretation of β_1 and β_2 from the previous NS model.
   (b) As each β parameter in the DNS model has its separate coefficient, it can be inferred that each of them have a separate identity which can be used to construct independent time-evolving dynamics for estimation - a feature that is not possible with the original NS model due to high multicollinearity between the coefficients of these parameters.
2. The introduction of the time-dependent λ_t in (3.3) allows its interpretation as an exponential decay rate that controls the level of decay. Small values of λ_t ensure slow decay and better fit of the yield curve at long maturities, while large values ensure fast decay and better fit at short maturities. It also represents the value at which the factor loading of β_{2t} reaches the maximum.

With the introduction of these three latent dynamic factors of the yield curve, Diebold and Li were able to create a parsimonious model that could preserve important stylized facts of the yield curve. In their paper Diebold and Li (2006), they presented the following stylized facts pertinent to their yield curve model. These facts concern the average shape of the yield curve, the possible shapes generated by the curve at different calendar times, the persistence of the curve at the short and long ends, and can be summarized as follows:

1. The following shapes can be attained in this model: normal (upward sloping), inverse (downward sloping), humped, inverted-humped (dipped), depending on the variations of the β_t parameters.
2. The average yield curve is concave and increasing. This means the average values of the β_t parameters produce a normal yield curve shape, with average long-term yields higher than average short-term yields.
3. The short end of the yield curve is more volatile than its long end, which can be observed in this model since the short end depends on both β_{0t} and β_{1t} whereas the long end depends only on β_{1t}.

4. Conversely, long rates are more persistent than short rates, since they depend only on one factor.

The DNS model has been able to serve central banks worldwide due to its simplicity and parsimonious nature. However, similar to other interest rate models, it has its drawbacks. The crucial drawbacks of the model are as follows:

D1. **Presence of Arbitrage:** The lack of an arbitrage-free condition in this model exemplifies the typically observed fact that yield curve models that fit practically do not satisfy this theoretical constraint, whereas yield curve models that do satisfy this theoretical no-arbitrage condition do not exhibit practical realities of the yield curve (see Chapters 5 and 26 of Lyuu (2001), and Laurini and Neto (2014)). However, several papers have since then sought to address and to some extent rectify this. A notable example is Christensen et al. (2011) which also has the contribution of one of the authors from the original DNS model, wherein they introduced an affine term-structure Nelson-Siegel class of yield curves that exhibit an arbitrage-free condition. We will discuss this topic extensively later in the thesis (see Chapter 4).

D2. **Existence of Other Shapes:** Often times, the yield curve can display shapes with more than one maxima and/or minima at different times. Since such shapes cannot be exhibited through the DNS model, this is seen as a disadvantage of the model. However, the extension of the original NS model by Svensson introduced a fourth factor β_{3t} that can also be dynamically evolved in a similar fashion, thereby creating the dynamic Nelson-Siegel-Svensson curve developed by Svensson (1995). The resultant formula becomes:

$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) \tag{3.4}$$

Here, the addition of another decay rate term λ_{2t} creates a second curvature factor in the yield curve model that is responsible for the additional maxima/minima in the generated shapes.

D3. **Deterministic Evolution of Factors:** Lastly, what remains as a major unsolved disadvantage of the model is the deterministic estimation of these important yield curve factors. In Diebold and Li (2006), the β_t factors were estimated using the univariate autoregressive (AR(1)) method. An appropriate value for the λ_t parameter was determined as the maturity at which the loading on the curvature factor (β_{2t}) achieves its maximum. They fixed λ_t to be 0.0609. The deterministic evaluation of DNS factors is not without limitations. This thesis aims to overcome such limitations, since deterministic evolution of market factors do not always produce realistic results.

Hence, the aim of this chapter is to tackle drawback D3. We shall introduce a new model of the dynamic Nelson-Siegel, which allows for a stochastic evolution of the β_t parameters. The achievements of such a stochastic model would lie in ensuring the following:

1. More realistic yield curve framework
2. Yield curve shapes consistent with empirical observations

3. Yield curve model that follows well-established economic theory
4. Sophisticated and simple model that is analytically tractable, following the "Keep it sophisticatedly simple" principle of public policy coined by Zellner (1992).

3.2 Yield Curve Shapes

The ability of a yield curve model to generate realistic and possible yield curve shapes is essential to developing any good term structure model. It is necessary to ensure that the model can produce shapes that are commonly seen in the practical yield market. Let us now define some common shapes as seen both in existing literature and in the term structure market.

Definition 3.2.1 (Yield Curve Shapes)**.** A yield to maturity curve, $y(t, t+\tau)$ observed at time t for a bond with time to maturity τ, is said to exhibit one of the following shapes:

1. flat, if $y(t, t+\tau)$ is constant $\forall \tau \in (0, \infty)$
2. normal, if $y(t, t+\tau)$ is strictly monotonically increasing $\forall \tau \in (0, \infty)$
3. inverse, if $y(t, t+\tau)$ is strictly monotonically decreasing $\forall \tau \in (0, \infty)$
4. humped, if $y(t, t+\tau)$ has exactly one extremum which is a local maximum $\forall \tau \in (0, \infty)$
5. dipped, if $y(t, t+\tau)$ has exactly one extremum which is a local minimum $\forall \tau \in (0, \infty)$

Diez and Korn (2020) proved how the shapes mentioned above can be generated by the one-factor and the two-factor Vasicek model for interest rates. At the time of writing this thesis, Keller-Ressel and Sachse (2024) proved that the Svensson family is able to produce additional shapes, some of which are listed below:

1. humped-dipped, if $y(t, t+\tau)$ has a local maxima followed by a local minima for $\tau \in (0, \infty)$
2. dipped-humped, if $y(t, t+\tau)$ has a local minima followed by a local maxima for $\tau \in (0, \infty)$
3. humped-dipped-humped, if $y(t, t+\tau)$ has a local maxima followed by a local minima followed by a local maxima for $\tau \in (0, \infty)$
4. dipped-humped-dipped, if $y(t, t+\tau)$ has a local minima followed by a local maxima followed by a local minima for $\tau \in (0, \infty)$

In the dynamic Nelson-Siegel model, the yield curves have shown to take only the shapes as defined in Definition 3.2.1. Empirical evidence across different markets across the world, for example, Germany, India, Iceland, Sweden, Norway, South Africa are testament to the ability of the model to reproduce these common shapes (see Patra et al. (2021) and Nyholm and Vidova-Koleva (2010)). These shapes are easily observable in empirical data. Figure 3.2 show the possible shapes attained by the DNS model on changing the $\beta_0(t)$, $\beta_1(t)$, and $\beta_2(t)$ factors arbitrarily with a constant λ throughout.

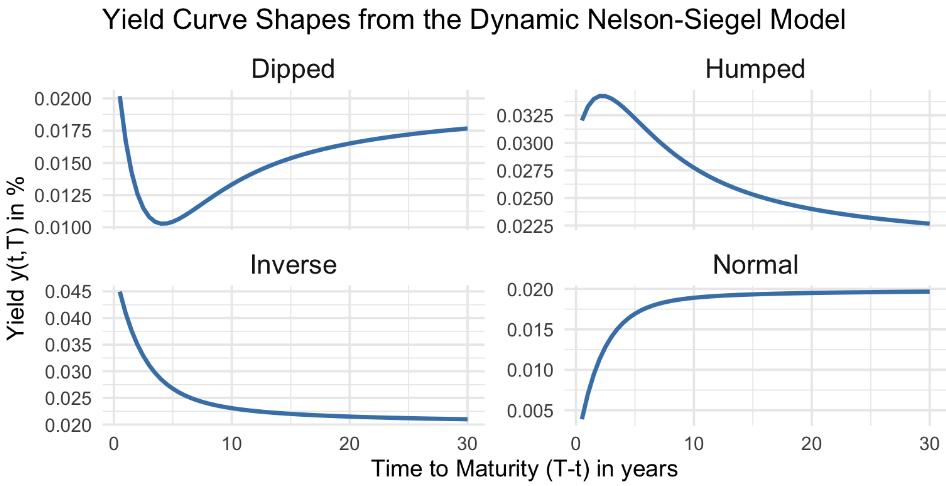


Figure 3.2: Different shapes attainable by the dynamic Nelson-Siegel model. (Values represented above are arbitrary.)

On the other hand, there are not many theoretical proofs in literature that explain the shapes generated by this model. Keller-Ressel and Sachse (2024) use the concept of total positivity with the theory of Chebyshev systems of Keller-Ressel (2021) to prove their results on the Svensson family, and use the time-derivative of the Nelson-Siegel forward curve along with a previous relation between the yield curve and forward curve stated in Keller-Ressel (2021) to demonstrate the attainability of possible shapes.

As an alternative, in the following Theorem 3.2.2, we offer a more detailed approach using elementary calculations to prove the existence of possible shapes attainable by the DNS model. We show how these shapes coincide exactly with the ones mentioned in Definition 3.2.1 and in the subsequent Corollary 3.2.4, we further specify exact conditions to generate these shapes. Our proof, therefore, offers a more intuitive approach to the problem.

3.2.1 Theorem on the Attainable Yield Curve Shapes in the Dynamic Nelson-Siegel Family

Theorem 3.2.2 (Attainable Yield Curve Shapes in the Dynamic Nelson-Siegel Model)**.** Let $\lambda \in [0, 1]$. The family of dynamic Nelson-Siegel yield curves $y(t, t + \tau)$ observed at any time $t \in (0, \infty)$ satisfying the following equation

$$y(t, t+\tau) = \beta_0(t) + \beta_1(t) \cdot \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_2(t) \cdot \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right) \tag{3.5}$$

can exhibit only one of the following possible yield curve shapes (as defined in Definition 3.2.1): flat, normal, inverse, humped, and dipped.

Proof. As stated, $\lambda \in [0, 1]$. Since we also consider the yield curve to evolve across all possible time to maturity values, τ can only take positive values.

The shapes of the yield curve can be mathematically expressed as a function that displays n local extrema. This can be observed from the signs of the derivatives of $y(t, t + \tau)$ with respect to time to maturity τ. In this case, all of the shapes mentioned in Definition 3.2.1 have $n = \{0, 1\}$ local extrema.

In this proof, we ignore the subscript of current time t under all of the β factors and λ, since we only are concerned with the partial derivative with respect to τ.

If β_1 and β_2 are zero at the same time, and if β_0 is a constant, then this would lead to a flat curve, as evident from (3.5). We now proceed to prove the existence of the other shapes.

The time-derivative of (3.5) with respect to τ is

$$\frac{\partial y}{\partial \tau} = \beta_1 \left(\frac{\lambda^2 \tau e^{-\lambda\tau} - \lambda(1 - e^{-\lambda\tau})}{\lambda^2 \tau^2} \right) + \beta_2 \left(\frac{\lambda^2 \tau e^{-\lambda\tau} - \lambda(1 - e^{-\lambda\tau})}{\lambda^2 \tau^2} + \lambda e^{-\lambda\tau} \right) \tag{3.6}$$

It is evident from (3.6) that there are four components that determine the sign and values of $\frac{\partial y}{\partial \tau}$. These are:

1. sign of β_1
2. sign of β_2
3. sign of $a(\tau) := \frac{\lambda^2 \tau e^{-\lambda\tau} - \lambda(1-e^{-\lambda\tau})}{\lambda^2\tau^2}$
4. sign of $b(\tau) := \lambda e^{-\lambda\tau}$

Essentially, (3.6) is reduced to:

$$\frac{\partial y}{\partial \tau} = \beta_1 a(\tau) + \beta_2 \Big(a(\tau) + b(\tau) \Big) =: f_{\beta_1,\beta_2}(\tau) \tag{3.7}$$

Note further that by (3.6) we have $f_{\beta_1,\beta_2}(+\infty) = 0$ as long as we have $\lambda > 0$. Therefore, (3.7) implies that the following shapes occur under the following conditions:

1. normal when $f_{\beta_1,\beta_2}(\tau) > 0$
2. inverse when $f_{\beta_1,\beta_2}(\tau) < 0$
3. humped when $f_{\beta_1,\beta_2}(0) > 0$ and one local extremum exists $\implies f_{\beta_1,\beta_2}(+\infty) \leq 0$
4. dipped when $f_{\beta_1,\beta_2}(0) < 0$ and one local extremum exists $\implies f_{\beta_1,\beta_2}(+\infty) \geq 0$

Our task is now to determine whether $f_{\beta_1,\beta_2}(\tau)$ can take these values as stated above (positive, negative, and zero). Once we are able to prove this, we can then state that the afore-mentioned shapes are attainable by the DNS model equation given by (3.5). Since we impose no restrictions on β_1 and β_2 for now, it is sufficient to analyze how $a(\tau)$ and $b(\tau)$ behave in equation (3.7), both independently and with each other. We proceed as follows.

Analyzing $b(\tau)$, we observe

$$\lim_{\tau \downarrow 0} b(\tau) = \lambda \quad \text{and} \qquad \lim_{\tau \to \infty} b(\tau) = 0$$

With our chosen assumption of $\lambda \in [0, 1]$, and $e^{-\lambda\tau} \in [0, 1]$, we observe that for any possible $\lambda > 0$, $b(\tau)$ is always positive and decreasing with maximum value of λ and minimum value of 0.

Hence, $b(\tau) \in [0, \lambda]$ and $\geq 0 \quad \forall \quad \tau > 0, \lambda > 0$

Applying l'Hospital's rule to analyze $a(\tau)$, we observe that

$$\begin{aligned}\lim_{\tau \downarrow 0} a(\tau) &= \lim_{\tau \downarrow 0} \frac{\lambda\tau e^{-\lambda\tau} + e^{-\lambda\tau} - 1}{\lambda\tau^2} \\ &= \frac{-\lambda}{2} \quad < 0\end{aligned}$$

and,

$$\begin{aligned}\lim_{\tau \to \infty} a(\tau) &= \lim_{\tau \to \infty} \frac{\lambda\tau e^{-\lambda\tau} + e^{-\lambda\tau} - 1}{\lambda\tau^2} \\ &= 0\end{aligned}$$

Now that we know the limits of $a(\tau)$, we would like to check if it is increasing or decreasing.

Conjecture 1: $a(\tau)$ is increasing.

To check this we calculate,

$$\begin{aligned}\frac{\partial a(\tau)}{\partial\tau} &= \frac{\overbrace{2 - e^{-\lambda\tau}(\lambda^2\tau^2 + 2\lambda\tau + 2)}^{=:\text{num}_1(\tau)}}{\lambda\tau^3} \\ &= \frac{\text{num}_1(\tau)}{\lambda\tau^3}\end{aligned}$$

The denominator term $\lambda\tau^3$ is always positive $\forall\tau > 0, \quad \lambda > 0$. For our conjecture to be true, we need: $\text{num}_1(\tau) > 0 \quad \forall\tau > 0$. On calculating, we observe

$$\begin{aligned}\frac{\partial \text{num}_1(\tau)}{\partial\tau} &= \frac{\partial}{\partial\tau}(2 - e^{-\lambda\tau}(\lambda^2\tau^2 + 2\lambda\tau + 2)) \\ &= e^{-\lambda\tau}\lambda^3\tau^2 \\ &> 0 \quad \text{,for } \lambda \in [0, 1]\end{aligned}$$

which, once again, fits with our assumption on λ. This implies that $\text{num}_1(\tau)$ is an increasing function. Hence, $a(\tau)$ is increasing and

$$a(\tau) \in \left[\frac{-\lambda}{2}, 0\right]$$

Analyzing $a(\tau) + b(\tau)$, we observe

$$\lim_{\tau \to \infty} (a(\tau) + b(\tau)) = 0$$

and

$$\lim_{\tau \downarrow 0} (a(\tau) + b(\tau)) = \lambda - \frac{\lambda}{2} = \frac{\lambda}{2} > 0$$

Furthermore, from (3.7) let us consider

$$\beta_1 + \beta_2 \left(\underbrace{1 + \frac{b(\tau)}{a(\tau)}}_{=:c(\tau)} \right) = h_{\beta_1,\beta_2}(\tau) \tag{3.8}$$

where, $h_{\beta_1,\beta_2}(\tau) := \frac{f_{\beta_1,\beta_2}(\tau)}{a(\tau)}$. From our analysis of $a(\tau)$ and $b(\tau)$ above, we know that

$$c(0) = -1 \quad , \quad c(\infty) = 1$$

With this, we need to further show that $c(\tau) \in [-1, 1]$.

Conjecture 2: $c(\tau)$ is increasing.

As stated,

$$c(\tau) = 1 + \frac{\lambda^2\tau^2 e^{-\lambda\tau}}{\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau}}$$

Taking its derivative, we obtain

$$\frac{\partial c(\tau)}{\partial \tau} = \frac{\overbrace{(\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau})(\lambda^2\tau e^{-\lambda\tau} - \lambda^3\tau^2 e^{-\lambda\tau}) - (\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau} + \lambda^2\tau^2 e^{-\lambda\tau})(-\lambda^2\tau e^{-\lambda\tau})}^{=:\text{num}_2(\tau)}}{\underbrace{(\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau})^2}_{=:\text{denom}(\tau)}}$$

The sign of denom(τ) will always be positive for every $\tau > 0, \quad \lambda > 0$, since the square of a number is always positive. Expanding num$_2(\tau)$, we observe

$$\begin{aligned}\text{num}_2(\tau) &= (\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau})(\lambda^2\tau e^{-\lambda\tau} - \lambda^3\tau^2 e^{-\lambda\tau}) - (\lambda\tau e^{-\lambda\tau} - 1 + e^{-\lambda\tau} + \lambda^2\tau^2 e^{-\lambda\tau})(-\lambda^2\tau e^{-\lambda\tau}) \\ &= \lambda^3\tau^2 e^{-2\lambda\tau} - \cancel{\lambda^4\tau^3 e^{-2\lambda\tau}} - 2\lambda^2\tau e^{-\lambda\tau} + \lambda^3\tau^2 e^{-\lambda\tau} + 2\lambda^2\tau e^{-2\lambda\tau} - \lambda^3\tau^2 e^{-2\lambda\tau} + \cancel{\lambda^4\tau^3 e^{-2\lambda\tau}} \\ &= \lambda^2\tau e^{-\lambda\tau} \underbrace{\Big(\lambda\tau(e^{-\lambda\tau} + 1) + 2(e^{-\lambda\tau} - 1)\Big)}_{=:\text{num}_3(\tau)}\end{aligned}$$

Now, similar to num$_1(\tau)$ in the proof of Conjecture 1, the $\lambda^2\tau e^{-\lambda\tau}$ term above is always positive $\forall\tau > 0, \quad \lambda \in [0, 1]$. Thus, the sign of num$_2(\tau)$ depends on the rest of the term denoted by num$_3(\tau)$. Taking its derivative,

$$\begin{aligned}\frac{\partial \text{num}_3(\tau)}{\partial \tau} &= \lambda e^{-\lambda\tau} - \lambda^2\tau e^{-\lambda\tau} + \lambda - 2\lambda e^{-\lambda\tau} \\ &= \underbrace{e^{-\lambda\tau}\Big(-\lambda - \lambda^2\tau\Big)}_{=:\text{num}_4(\tau)} + \lambda\end{aligned}$$

In the above, $\lambda > 0$, so the sign further depends on num$_4(\tau)$. Taking its derivative,

$$\begin{aligned}\frac{\partial \text{num}_4(\tau)}{\partial \tau} &= -\lambda e^{-\lambda\tau}(-\lambda - \lambda^2\tau) + e^{-\lambda\tau}(-\lambda^2) \\ &= e^{-\lambda\tau}\lambda^3\tau \\ &> 0 \quad ,\text{for } \lambda \in [0, 1]\end{aligned}$$

Hence, num$_4(\tau)$, num$_3(\tau)$, and num$_2(\tau)$ are all increasing functions which finally proves that $c(\tau)$ is also an increasing function of τ and

$$c(\tau) \in [-1, 1]$$

This conjecture implies that in (3.8),

$$\exists \tau^*, \text{ such that } h_{\beta_1,\beta_2}(\tau^*) = 0$$

We have even shown that there is exactly (!) one zero. Moreover, since by definition, $h_{\beta_1,\beta_2}(\tau) := \frac{f_{\beta_1,\beta_2}}{a(\tau)}$, the signs of $h_{\beta_1,\beta_2}(\tau)$ and $f_{\beta_1,\beta_2}(\tau)$ are always opposite. This is because $a(\tau)$ is always negative and an increasing function of τ, as proved.

Therefore, this finishes our proof to show that $f_{\beta_1,\beta_2}(\tau)$ in (3.7) can be either positive, negative, or zero. Hence, the listed shapes (normal, inverse, humped, and dipped) are attainable by the dynamic Nelson-Siegel family of yield curves. □

Remark 3.2.3. Theorem 3.2.2 has the following consequences:

1. The β_0 factor of the dynamic Nelson-Siegel plays no role in determining the shape. It is only responsible for the parallel shift of the entire yield curve.
2. The yield curve is flat when both $\beta_1(t)$ and $\beta_2(t)$ are zero at the same time. Then, $y(t, t + \tau) = \beta_0(t)$.
3. The probability distribution of yield curve shapes depends on the distributions of $\beta_1(t)$ and $\beta_2(t)$. We calculate this based on the chosen stochastic models for the $\beta(t)$ factors, and illustrate our results for each separate case later in this chapter.

For now, the third point of Remark 3.2.3 can be expanded to include the deterministic case to demonstrate how different values of $\beta_1(t)$ and $\beta_2(t)$ generate different shapes observed at a given point of time t. By doing so, we arrive at the following corollary to Theorem 3.2.2.

Corollary 3.2.4 (Conditions to generate the attainable yield curve shapes)**.** The following conditions are necessary to produce yield curve shapes attainable by the dynamic Nelson-Siegel model:

Shapes	$\beta_1(t)$	$\beta_2(t)$	Relation between $\beta_1(t)$ and $\beta_2(t)$
Normal	< 0	> 0	$\beta_1 < -\lvert\beta_2\rvert$
	< 0	< 0	
Inverse	> 0	> 0	$\beta_1 > \lvert\beta_2\rvert$
	> 0	< 0	
Humped	> 0	> 0	$\lvert\beta_1\rvert < \beta_2$
	< 0	> 0	
Dipped	> 0	< 0	$-\lvert\beta_1\rvert > \beta_2$
	< 0	< 0	

Proof. To prove that imposing the above-stated conditions on $\beta_1(t)$ and $\beta_2(t)$ will generate the respective shapes, we take the help of Theorem 3.2.2, specifically equations (3.7) and (3.8). In our proof of the theorem, we already observed how the behaviours of $a(\tau)$ and $b(\tau)$ independently, and together via $c(\tau)$ enable us to prove the possibility of existence of the four possible shapes. In the proof, we also stated that the signs of $\beta_1(t)$ and $\beta_2(t)$ influence the yield curve shape. We will prove now how their signs give us the exact conditions to reproduce the shapes at a given time t.

Note that, all of our assumptions for Theorem 3.2.2 hold here as well. Since we have already proved in the theorem that $c(\tau) \in [-1, 1]$ is an increasing function, we can use this further to determine the complete range of $h_{\beta_1,\beta_2}(\tau)$. We crudely obtain it as follows

$$\begin{aligned} h_{\beta_1,\beta_2}(\tau) &= \beta_1 + \beta_2 \cdot c(\tau) \quad \text{(from(3.8))} \\ \text{or, } h_{\beta_1,\beta_2}(\tau) &\in \beta_1 + [-\beta_2, \beta_2] \quad (\because c(\tau) \in [-1, 1]) \end{aligned} \tag{3.9}$$

In (3.9), we see the evolution of $h_{\beta_1,\beta_2}(\tau)$ with respect to τ, with $h_{\beta_1,\beta_2}(\tau)$ initially behaving as the sum $\beta_1 - \beta_2$, and then, as τ increases, it gradually behaves similar to the sum of $\beta_1 + \beta_2$.

Furthermore, as stated in the previous proof, by definition of $h_{\beta_1,\beta_2}(\tau)$, it is also evident that the signs of $h_{\beta_1,\beta_2}(\tau)$ and $f_{\beta_1,\beta_2}(\tau)$ are opposite. So, a positive $h_{\beta_1,\beta_2}(\tau)$ will produce a negative $f_{\beta_1,\beta_2}(\tau)$ and vice-versa.

With these information, we can now analyze how the signs of $\beta_1(t)$ and $\beta_2(t)$ influence the type of shape the yield curve takes. There are four combinations of possible signs of $\beta_1(t)$ and $\beta_2(t)$, as shown in Figure 3.3.

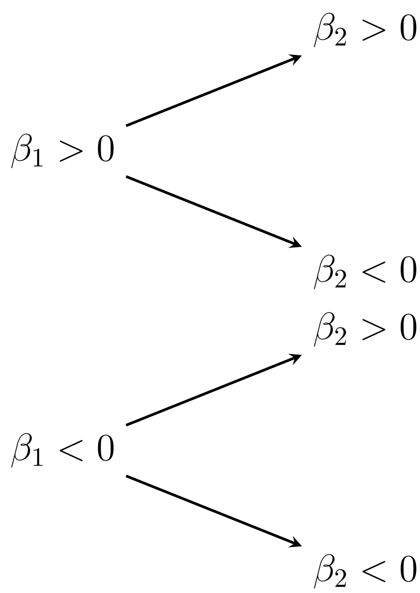


Figure 3.3: Possible combinations of $\beta_1(t)$ and $\beta_2(t)$.

Figure 3.3 is a stylistic choice and can be drawn the other way around, by switching the positions of β_1 and β_2. This does not change our interpretations or proof. In the next statements, we shall drop the subscripts of $h_{\beta_1,\beta_2}(\tau)$ and $f_{\beta_1,\beta_2}(\tau)$ for the sake of simplicity.

(i.) Let us begin with the first case: $\beta_1 > 0, \beta_2 > 0$. In this case, (3.9) shows that the values of h are always within the range between $\beta_1 - \beta_2$ and $\beta_1 + \beta_2$.

If $\beta_1 > \beta_2$, then the above relation implies that h will always be > 0, thereby also implying that $f < 0$. As $f < 0$, the yield curve has a negative slope, with the resulting shape to be inverse.

On the other hand, if $\beta_1 < \beta_2$, then the above relation shows that initially $\beta_1 - \beta_2$ will be < 0 due to a greater positive β_2. As τ increases, the h will be determined by $\beta_1 + \beta_2$

which is > 0 due to positive signs of both. Thus, h goes from $h < 0$ to $h > 0$, implying $f > 0$ to $f < 0$, giving the yield curve a humped shape.

(ii.) For the second case: $\beta_1 > 0, \beta_2 < 0$. Here, (3.9) again indicates that the values of h are within the range between $\beta_1 - \beta_2$ and $\beta_1 + \beta_2$.

Note that, here $\beta_2 < 0$. Although it is obvious that in this case, $\beta_1 > \beta_2$, their respective magnitudes might change the signs of h. Therefore, if $\beta_1 > |\beta_2|$, then β_1 always dominates β_2, resulting in $h > 0$ and $f < 0$. Thus, the yield curve always has a negative slope and is, therefore, inverse.

However if $\beta_1 < |\beta_2|$, then h is initially determined by $\beta_1 - \beta_2$ which gives a positive value due to the negative β_2. This is flipped later as τ increases, since $\beta_1 + \beta_2$ gives a negative value due to the negative β_2. To illustrate this, let us take $\beta_1 = 5$ and $\beta_2 = -8$. Then, initially $5 + (-(-8)) = 13 > 0$. With increasing time, $5 + (-8) = -3 < 0$. Thus, $f < 0$ initially and then becomes $f > 0$, yielding a dipped curve.

(iii.) For the third case: $\beta_1 < 0, \beta_2 > 0$.

It is obvious here that $\beta_2 > \beta_1$. However, similar to the previous case, their magnitudes might change the sign of h. We know from (3.9) that the values of h lie between $\beta_1 - \beta_2$ and $\beta_1 + \beta_2$.

If $|\beta_1| < \beta_2$, then initially $h = \beta_1 - \beta_2$ gives a negative value due to negative β_1 and the positive β_2 being greater in magnitude. However, with increasing time, $h = \beta_1 + \beta_2$ which is positive. Hence, $h < 0$ and $h > 0$ imply $f > 0$ initially with $f < 0$ at the end. The resultant curve is humped. Illustrating this with an example, let us take $\beta_1 = -3$ and $\beta_2 = 5$. Then at first $h = (-3) + (-5) = -8 < 0$, with $h = -3 + 5 = 2 > 0$ later on.

On the other hand, if $|\beta_1| > \beta_2$, then initially $\beta_1 - \beta_2$ will be negative due to the negative β_1 and its magnitude being larger than β_2. Hence, $h < 0$. With time, $\beta_1 + \beta_2$ will also be negative due to negative β_1 and β_2 being lesser in value. Again $h < 0$. Thus, $f > 0$ always, resulting in a normal curve.

(iv.) The final case is when $\beta_1 < 0, \beta_2 < 0$.

If $\beta_1 > \beta_2$, then h is initially positive, since it is determined by $\beta_1 - \beta_2$, with both being negative and β_1 being greater. Later, h which is determined by $\beta_1 + \beta_2$ gives a negative value. This is due to adding two negative values. Thus, h goes from $h > 0$ and $h < 0$, implying $f < 0$ initially and $f > 0$ later on. This generates a dipped curve. We illustrate this with an example, let $\beta_1 = -3$ and $\beta_2 = -5$. Initially, $h = (-3) + (-(-5)) = 2 > 0$. Then, $h = (-3) + (-5) = -8 < 0$.

If $\beta_1 < \beta_2$, then h will always be negative. This is because both $\beta_1 - \beta_2$ and $\beta_1 + \beta_2$ produces negative results due to negative β_1 and β_2, and β_2 being always greater. Hence, $h < 0$ implies $f > 0$, thereby producing a normal curve.

With these, we have shown how the four possible combinations of the signs of β_1 and β_2 can contain further subgroups depending on their absolute magnitudes with each other. These conditions generate different shapes. We also show that there is no overlap of shapes at a single point of time. Hence, the dynamic Nelson-Siegel model of yield curves can generate four possible shapes under certain conditions, as specified above. □

Moreover, the above Corollary 3.2.4 corroborates the result of Theorem 3.2 in Keller-Ressel and Sachse (2024).

We can visualize our results in Corollary 3.2.4 in the following figure.

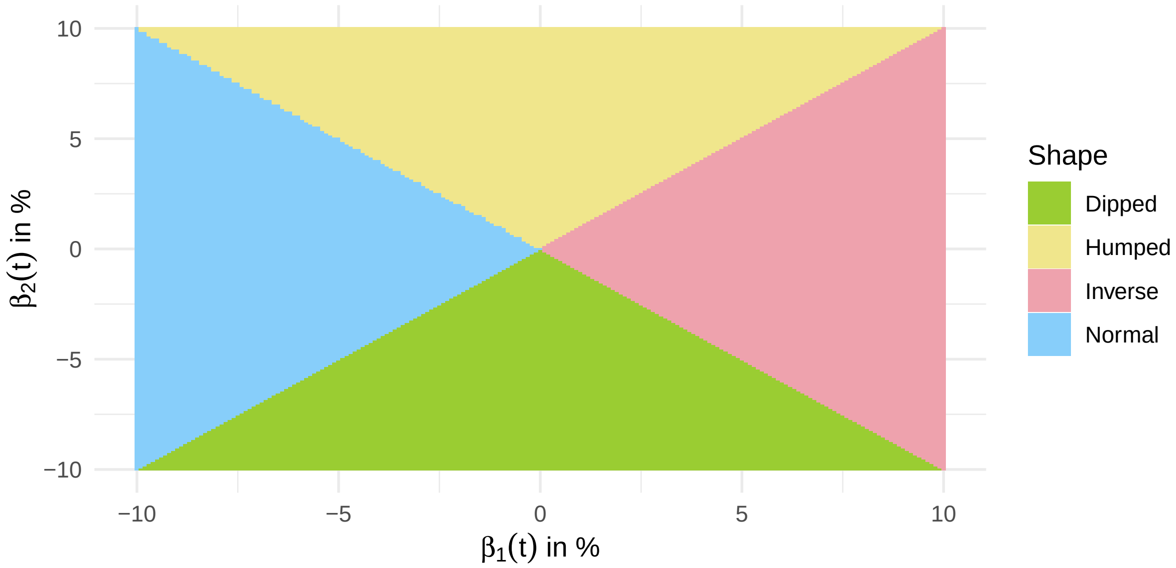


Figure 3.4: Conditions on the parametric region of $(\beta_1(t), \beta_2(t))$ of the dynamic Nelson-Siegel for generating the respective shapes.

In Figure 3.4, the occurrence of the yield curve shapes have been colored according to their respective parametric region $(\beta_1(t), \beta_2(t))$. We allow $\beta_1(t)$ and $\beta_2(t)$ to take any value in $\mathbb{R}$ between $[-10, 10]$. We choose $\lambda = 0.5$ and τ to range from 0 to 60, and run the corresponding code in R to generate dynamic Nelson-Siegel yield curves given by equation (3.5). We note the frequency of occurrence of each yield curve shape, and thereby corroborate the results with Corollary 3.2.4. The corollary holds true.

It is interesting to note that in Figure 3.4, the probability of each shape is naturally $\frac{1}{4}$. This is because each value of the respective $\{\beta_1(t), \beta_2(t)\}$ factor is equally likely to be chosen. However, when we account for stochastic dynamics of these $\beta(t)$ factors, we will obtain different probabilities of the yield curve shapes, as the frequencies of the values and signs of $\beta_1(t)$ and $\beta_2(t)$ will depend on their respective distributions. This is an interesting research objective, and also motivates our decision to introduce stochastic models of the dynamic Nelson-Siegel.

Now that we have proved the existence of the four possible shapes attained from the DNS yield model that are consistent with real world observations, we can move ahead to the next task to introduce stochastic evolution of the $\beta(t) = \{\beta_0(t), \beta_1(t), \beta_2(t)\}$ factors. Our aim is to produce a family of yield curves that can be applicable in practical scenarios and have theoretical closed form analytical solutions.

3.3 Stochastic Models of the Dynamic Nelson-Siegel

We shall now consider that the dynamic Nelson-Siegel model has stochastic factors $\beta_0(t), \beta_1(t), \beta_2(t)$. By introducing stochasticity in the model, we aim to find an analytically tractable formula to estimate realistic yield curves observed in the market. We would like our stochastic model to account for all of the shapes that can be achieved by the dynamic Nelson-Siegel and demonstrate properties that have been previously observed in the market, such as negative interest rate values. We also aim to ensure that this model exhibits important stylized facts of the yield curve that have been previously established in existing literature. Additionally, to use this stochastic model of interest rates for the purpose of portfolio optimization of interest-rate derivatives, we also need to ensure that the resultant derivative price dynamics are valid and satisfy appropriate boundary conditions.

We take the baseline DNS model for our analysis. To reiterate, in the DNS model, the yield to maturity of a zero-coupon bond maturing at time T, observed at time t, satisfies the following equation

$$y(t, t+\tau) = \beta_0(t) + \beta_1(t) \cdot \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_2(t) \left(\cdot \frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right)$$

where, $\tau := T - t$ is the remaining time to maturity.

3.3.1 Choice of Parameters

The equation above has four parameters: $\lambda_t, \beta_0(t), \beta_1(t), \beta_2(t)$. Each of them have a significance on the evolution of the yield curve. They have both economic interpretations with regards to the yield curve as well as the mathematical simplicity to be estimated in closed form solutions.

The $\beta(t)$ factors. In Diebold and Li (2006), the three $\beta(t)$ factors had separate interpretations with regard to their roles in the yield curve. $\beta_0(t)$ represented the level of the curve, $\beta_1(t)$ represented the slope of the curve, and $\beta_2(t)$ represented the curvature of the curve. These interpretations have since then been corroborated with numerical empirical evidences as well as theoretical research.

Furthermore, the factor loadings, which are nothing but the coefficients of each β factor, play a role in establishing the short-end and long-end behaviours of the yield curve. From the yield $y(t, t+\tau)$ equation above, we observe the following factor loadings:

1. $\beta_0(t)$ has the coefficient 1. Its factor loading is 1.
2. $\beta_1(t)$ has the coefficient $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$. Its factor loading is $g(\tau) := \frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$.
3. $\beta_2(t)$ has the coefficient $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}$. Its factor loading is $h(\tau) := \frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}$.

We plot the evolution of these three factor loadings over time in Figure 3.5 below. We choose $\lambda = 0.25$. The y-axis represents the values of these factor loadings over time. The time to maturity τ is denoted on the x-axis. The extreme right end is obtained by letting

τ tend to infinity. The short-term behaviour of the yield curve can be seen on the left part of the graph, while the right part shows the long-term behaviour.

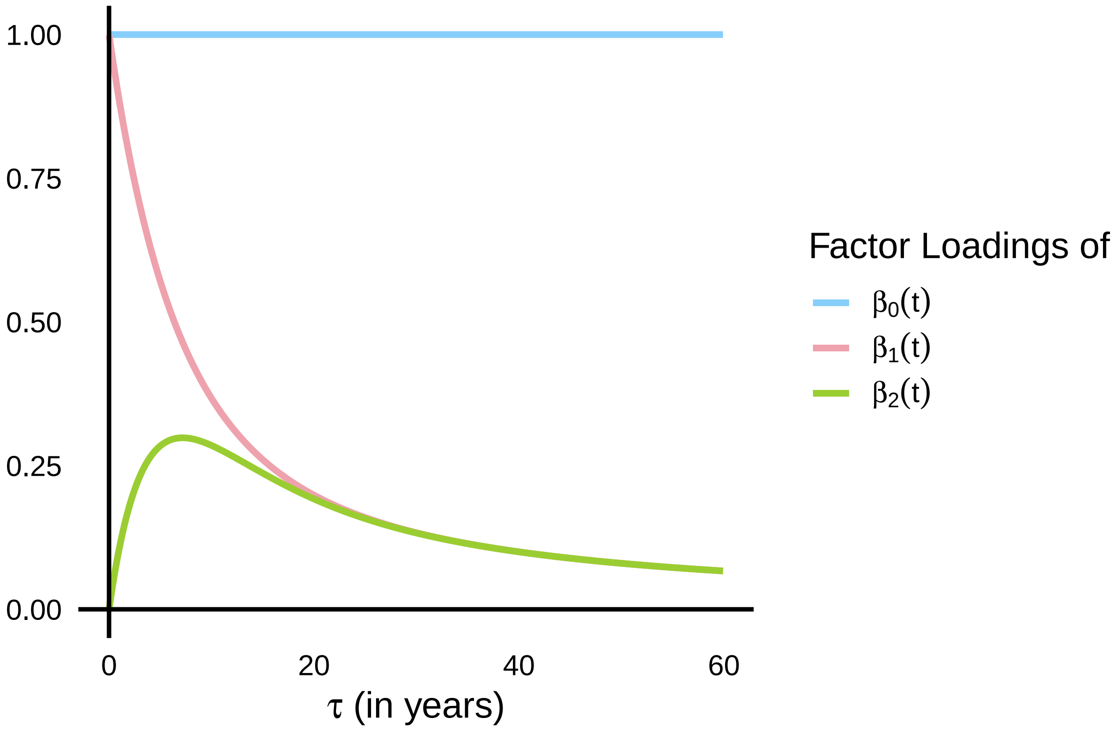


Figure 3.5: Factor loadings of the dynamic Nelson-Siegel

The implications of Figure 3.5 are:

1. The impact of $\beta_0(t)$ on the yield curve is observed throughout the yield curve. Its significance is on both the short end as well as on the long end. This can be verified with $\lim_{\tau\to\infty} y(t, t+\tau) = \beta_0(t)$. Hence, we call it the long-term factor. Diebold and Li (2006) interpret it as the level factor, governing the yield curve level. Moreover, as the factor loading is 1 for all maturities τ, it allows a parallel shift of the yield curve. This behaviour is not observed in any other popular interest rate model, for example, the Vasicek or the Hull-White models. Any change in this factor, thereby changes the yield curve $y(t, t+\tau)$ equally.

2. For $\beta_1(t)$, the factor loading decreases with increasing maturity and vanishes asymptotically. Hence, we call it the short-term factor. Moreover, an increase in $\beta_1(t)$ increases the short-term yields more than the long-term yields. This is because the short-term yields depend more on this factor which can be verified with $\lim_{\tau\to 0} y(t, t+\tau) = \beta_0(t) + \beta_1(t)$. Thus, this factor influences the slope of the yield. A higher value will yield a higher slope. Diebold and Li (2006) interpret it as the slope factor of the yield curve. Some papers, such as Frankel and Lown (1994), define the yield curve slope to be difference $\lim_{\tau\to\infty} y(t, t+\tau) - \lim_{\tau\to 0} y(t, t+\tau)$, which is calculated to be $\beta_1(t)$ in the DNS model. Additionally, due to yield curves often exhibiting mean-reversion tendencies, particularly in the long-run, we conjecture that this factor must display a mean-reverting effect in its dynamics.

3. The impact of $\beta_2(t)$ on the yield curve can only be witnessed in the middle, as seen

in the figure. Its factor loading begins at 0 and asymptotically vanishes to 0 with increasing maturity. Hence, we call it the medium-term factor. It is responsible for the "hump" of a curve, if it exists. This hump can be either upwards or downwards, depending on its value. Thereby, this denotes where the sign of the curve changes. This is also evident from an additional exponential decay term $e^{-\lambda\tau}$ in its factor loading term $h(\tau)$. Consequently, this factor influences the curvature of the yield curve, and is, thereby also, interpreted as the curvature factor by Diebold and Li (2006). Since its factor loading is equal to the factor loading of $\beta_1(t)$ minus the additional exponential decay term, we assume that it follows mean-reverting dynamics as well.

The λ parameter. The impact of the λ_t parameter is much harder to visualize. Several literature sources (including Christensen et al. (2011)) name it as the exponential decay rate and assumes a constant value throughout the yield model. The original DNS model of Diebold and Li (2006) states that λ_t determines the maturity at which the loading of the curvature factor $\beta_2(t)$ achieves its maximum value. Which means, it is the value at t where the hump on the *green* line in Figure 3.5 achieves the maximum value. They obtain the λ_t value that maximizes the factor loading of the curvature factor at exactly 30 months, and get $\lambda_t = 0.0609$. Hence, this parameter is chosen to be constant in their approach. In Diebold et al. (2006), the authors similarly fix a value of λ_t when they estimate the yield curve model using these three latent β_s factors and some additional observable macroeconomic variables. Here, they take $\lambda_t = 0.0777$. In both of these cases, the value of λ is fixed without considering its separate dynamics.

On the other hand, there have also been references in literature that emphasizes the importance of modeling the λ_t parameter dynamically. In Koopman et al. (2010), the yield curve evolves according to the DNS model, with a time-varying λ parameter. This specification allows dynamic interactions between changes in cross-maturity dependence through λ_t as well as time-series dependence of the yields through the β factors. Additionally, in Levant and Ma (2017), a Markov-switching component to the factor loading parameter λ_t was introduced in order to capture the relationship between yield curve factors and observed macroeconomic variables. Similar to the previous paper, this paper also included time-dependent volatility to model the curve. By applying a hidden regime-switching component to the factor loading parameter λ_t, they were able to show that the parameter showed significant switching instead of centering around a mean, as often noted in literature. This regime-switching extension to the DNS model was, thereby, able to explain economic activity and monetary policy changes.

While the extant literature provide justifications for both constant as well as time-dependent λ, in our work here, we shall solely consider a constant λ_t to model the yield curve. We also justify this choice. Firstly, in our work, we would like to model the yield curve accurately such that it is analytically tractable enough to subsequent chapter, the choice of either a stochastic or constant λ_t does not interfere with our interpretation of our results of the obtained optimal positions. Second, we would like to extend the baseline NS and DNS models by simply considering a stochastic evolution of the three yield curve factors, without emphasizing the relation with macroeconomic variables. For this, our choice of constant λ_t seems ideal, just as done in the works of the original Nelson-Siegel models of Nelson and Siegel (1987) and Diebold and Li (2006). By considering a constant

λ_t, we are able to remove extra complexity from our stochastic model, without losing valuable information on the interpretation of the yield curve. Finally, as always, any improvements to the extended DNS model via a stochastic evolution of λ_t can be researched as an extension to this work.

With these ideas and justifications, we have thereby established a suitable method in which the DNS model can be extended using stochastic dynamics of the three $\beta(t)$ factors, and by taking a constant λ to model the yield curve. We aim to generate all possible shapes through this model, achieve analytical closed form solutions to our yield curve dynamics, and retain economic interpretations previously established in the literature of the original DNS framework.

3.3.2 The Baseline Model

The stochastic model of the DNS yield curve must represent the assumptions and properties described above on the choice of yield curve parameters. In section 3.3.1, we justified how each of the four parameters should evolve in our yield curve model. Hence, we develop the following baseline dynamics of the overall yield curve:

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & b_2 \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix}$$
$$\text{where,} \quad \mu > 0 \tag{3.10}$$

Here, we present a base structure of the stochastic DNS model. We assume $W_0(t)$, $W_1(t)$, and $W_2(t)$ are three Brownian motions that govern each of the three yield curve dynamic factors respectively. We assume constant volatility throughout for model simplicity. We will later see in Chapter 4 how the optimal positions of our optimization problems benefit from not considering stochastic volatility. In the following subsections, we explore three possible cases of the Brownian motions and extend our analysis accordingly (see section 3.3.3 onward).

In (3.10), μ is a constant parameter that represents the mean-reversion speed of the stochastic $\beta(t)$ factors. In literature, this parameter is often taken to be equal to the factor loading decay rate λ (see Christensen et al. (2009) and Christensen et al. (2011)). Such cases can be treated as specific cases of our model, and pose no restrictions on our model. Furthermore, in the later subsections where we analyze each stochastic model, we notice that impact of increasing or decreasing the value of μ on the resultant yield curve (see Figure 3.7, Figure 3.9, Figure 3.11). We will observe how with a higher value of μ, the process will revert faster to the mean-reversion level, and thus tame the volatility caused by the Brownian motions. It is also ideal to choose both μ and λ to be in the same range to obtain realistic evolutions of the yield curve.

Additionally, in the above equation, the following observations can be made to support the initial justification of model parameters:

1. $\beta_0(t)$ is the long-term factor and enables parallel shift of the entire yield curve. Since it impacts both on the short and the long ends of the yield curve, we assume a

constant mean-reversion level of a_0. The Brownian motion $W_0(t)$ is the stochastic factor. Hence the individual dynamics of $\beta_0(t)$ is:

$$d\beta_0(t) = \mu(a_0 - \beta_0(t))dt + b_0 dW_0(t) \tag{3.11}$$

whose solution is:

$$\beta_0(t) = e^{-\mu t}\beta_0(0) + a_0(1 - e^{-\mu t}) + b_0 \int_0^t e^{-\mu(t-u)} dW_0(u) \quad , \beta_0(0) \neq 0$$

Further,

$$\begin{aligned} &\mathbb{E}(\beta_0(t)) = e^{-\mu t}\beta_0(0) + a_0(1 - e^{-\mu t}) \\ &\lim_{t\to\infty} \mathbb{E}(\beta_0(t)) = a_0 \end{aligned}$$

and

$$\begin{aligned} \text{Var}(\beta_0(t)) &= \frac{b_0^2}{2\mu}(1 - e^{-2\mu t}) \\ \lim_{t\to\infty} \text{Var}(\beta_0(t)) &= \frac{b_0^2}{2\mu} \end{aligned}$$

Hence, it can be seen that in the long run, the level of the yield curve is determined by a normal random variable with distribution $\sim N\left(a_0, \frac{b_0^2}{2\mu}\right)$.

2. $\beta_1(t)$ is the short-term factor which governs the slope of the yield curve. We conjectured in the previous subsection that its dynamics must include a mean-reversion effect. Hence, we assume μ to be the mean-reversion speed and $a_1(t)$ to be the time-dependent mean-reversion level. The Brownian motion $W_1(t)$ is responsible for the stochastic nature. The SDE dynamics for $\beta_1(t)$ is thus:

$$d\beta_1(t) = \mu\Big(a_1(t) - \beta_1(t)\Big)dt + b_1 dW_1(t) \tag{3.12}$$

whose solution can be obtained by the method of integrating factors to be:

$$\beta_1(t) = e^{-\mu t}\beta_1(0) + \mu \int_0^t e^{-\mu(t-u)} a_1(u) du + b_1 \int_0^t e^{-\mu(t-u)} dW_1(u) \quad , \beta_1(0) \neq 0$$

Note that,

$$\begin{aligned} \mathbb{E}(\beta_1(t)) &= e^{-\mu t}\beta_1(0) + \mu \int_0^t e^{-\mu(t-u)} a_1(u) du \\ \text{Var}(\beta_1(t)) &= \frac{b_1^2}{2\mu}(1 - e^{-2\mu t}) \end{aligned}$$

If $t \to \infty$, then

$$\begin{aligned} \lim_{t\to\infty} \mathbb{E}(\beta_1(t)) &= \mu \cdot \lim_{t\to\infty} \left(\int_0^t e^{-\mu(t-u)} a_1(u) du \right) \\ \lim_{t\to\infty} \text{Var}(\beta_1(t)) &= \frac{b_1^2}{2\mu} \end{aligned}$$

Since we require

$$\mu \cdot \lim_{t\to\infty}\left(\int_0^t e^{-\mu(t-u)}a_1(u)du\right) < \infty$$

this implies that $a_1(u)$ can be conveniently chosen such that

$$|a_1(t)| \leq M \quad \forall t \geq 0, \quad \text{for } M \in \mathbb{R}$$

for a constant mean-reversion speed μ. This can be easily satisfied if $a_1(t)$ is a constant, an exponentially decaying function of time (for example e^{-t}), or a periodic function of time (for example, sin and cosine functions). All of these are sensible representations of mean-reversion levels. Although a polynomial growth function of time does not satisfy the condition on $a_1(t)$, it would not have made sense as a function for mean-reversion level of interest rates, as this would imply a continuously increasing mean-reversion level. Hence, such functions shall not be considered to be suitable for our model.

3. $\beta_2(t)$ is the medium-term factor which governs the curvature or the "humpedness" of the yield curve. As mentioned previously, its dynamics must also include a mean-reversion effect to create the hump. The mean-reversion level is denoted by $a_2(t)$. We again assume μ to be the mean-reversion speed. The Brownian motion $W_2(t)$ is responsible for its stochastic nature. The SDE dynamics for $\beta_2(t)$ is thus:

$$d\beta_2(t) = \mu\Big(a_2(t) - \beta_2(t)\Big)dt + b_2 dW_2(t) \tag{3.13}$$

whose solution is:

$$\beta_2(t) = e^{-\mu t}\beta_2(0) + \mu\int_0^t e^{-\mu(t-u)}a_2(u)du + b_2\int_0^t e^{-\mu(t-u)}dW_2(u) \quad ,\beta_2(0) \neq 0$$

We have,

$$\mathbb{E}(\beta_2(t)) = e^{-\mu t}\beta_2(0) + \mu\int_0^t e^{-\mu(t-u)}a_2(u)du$$
$$\text{Var}(\beta_2(t)) = \frac{b_2^2}{2\mu}(1 - e^{-2\mu t})$$

Applying the limits at $t \to \infty$,

$$\lim_{t\to\infty}\mathbb{E}(\beta_2(t)) = \mu \cdot \lim_{t\to\infty}\left(\int_0^t e^{-\mu(t-u)}a_2(u)du\right)$$
$$\lim_{t\to\infty}\text{Var}(\beta_2(t)) = \frac{b_2^2}{2\mu}$$

We, once again, arrive at a similar conclusion that $a_2(t)$ can be chosen such that the following is satisfied:

$$|a_2(t)| \leq M \quad \forall t \geq 0, \quad \text{for } M \in \mathbb{R}$$

Hence, similarly to $a_1(t)$, $a_2(t)$ can be a constant, an exponentially decaying function of time, or a periodic function of time.

4. λ is a constant parameter in this model. Consistent with existing literature, it is the factor loading parameter. It is thus responsible for determining the value at which the curve changes it shape. As previously stated, literature concerning stochastic dynamics of the DNS model often take this to be equal to the mean-reversion speeds of one or more of the $\beta(t)$ factors, implying $\mu = \lambda$ in (3.10). We consider this specification as a special case of our base model. This allows us to not restrict our analysis to a single case, thereby enabling us to observe yield curve evolutions under different (valid) possibilities of μ. Furthermore choosing $\mu \neq \lambda$ allows us to specify our dynamics of the three $\beta(t)$ factors without any prior bias to influence the shapes of the resultant yield curve.

Thus, the base stochastic DNS model for the yield curve is constructed. Referring back to a previous comment made in the beginning of this subsection on the relation between the Brownian motions $W_0(t)$, $W_1(t)$, and $W_2(t)$ with each other, there are three possible cases evident: One, where all Brownian motions are the same. Second, where all Brownian motions are uncorrelated with each other. Third, where all Brownian motions are pairwise correlated with each other. Since each of these cases leads to a different yield curve evolution, along with different subsequent bond price dynamics and different distributions of yield curve shapes, we look at each individual case separately.

3.3.3 Case 1: Single Brownian motion

We consider the first case where: $W_0(t) = W_1(t) = W_2(t) = W(t)$. This implies that there is only a single Brownian motion $W(t)$ that affects all three $\beta(t)$ factors. In this case, (3.10) becomes

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & b_2 \end{pmatrix} dW(t) \tag{3.14}$$

As explained in Section 3.3.2, the functions $a_1(t)$ and $a_2(t)$ are deterministic and bounded in the time interval $[0, T]$, where T is a time horizon typically corresponding to the maturity term of a long-term zero coupon bond.

Yield Curve Dynamics

The Dynamic Nelson-Siegel yield curve (as seen in (3.5)) is

$$y(t, t+\tau) = \beta_0(t) + \beta_1(t) \cdot \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_2(t) \cdot \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

The above equation can be condensed to

$$y(t, t+\tau) = \beta_0(t) + \beta_1(t)g(\tau) + \beta_2(t)h(\tau)$$

where,

$$g(\tau) := \frac{1 - e^{-\lambda\tau}}{\lambda\tau}$$

$$h(\tau) := \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}$$

As we know, τ is the time to maturity and the running time (current time) is t. Since we would like to prove the existence of our derived yield curve (and subsequent bond price) dynamics at the boundary condition T, we restructure the above DNS yield curve model to include T instead of τ. This change in variable also assists in the quicker calculation of derivatives with respect to t.

$$y(t,T) = \beta_0(t) + \beta_1(t) \cdot \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}\right) + \beta_2(t) \cdot \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}\right)$$

which can be condensed to

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t)$$

where,

$$g(t) = \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}$$
$$h(t) = \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}$$

By simple substitution of values, we get the yield curve to be given as

$$\begin{aligned} y(t,T) = {} & e^{-\mu t}\beta_0(0) + a_0(1-e^{-\mu t}) + g(t)e^{-\mu t}\beta_1(0) + g(t)\mu\int_0^t e^{-\mu(t-u)}a_1(u)du \\ & + h(t)e^{-\mu t}\beta_2(0) + h(t)\mu\int_0^t e^{-\mu(t-u)}a_2(u)du \\ & + \Big(b_0 + g(t)b_1 + h(t)b_2\Big)\int_0^t e^{-\mu(t-u)}dW(u) \end{aligned}$$

Now, the yield curve dynamics is obtained as the differential form of $y(t,T)$ with respect to t and is given as follows

$$\begin{aligned} dy(t,T) &= d\beta_0(t) + g(t)d\beta_1(t) + \beta_1(t)g'(t)dt + h(t)d\beta_2(t) + \beta_2(t)h'(t)dt \\ &= (\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t) \\ &\quad + h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t))dt \\ &\quad + (b_0 + g(t)b_1 + h(t)b_2)dW(t) \end{aligned} \tag{3.15}$$

Bond Price Dynamics

From Korn (2022), the definition of the yield curve in terms of zero-coupon bond prices is given by

$$P(t,T) = e^{-y(t,T)\cdot(T-t)}$$

From the given yield curve dynamics in equation (3.15), we can further calculate the bond price dynamics using Itô's Lemma (see Itô (1951, Theorem 6)) as follows:

$$dP(t,T) = P(t,T)\Big(-dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \tag{3.16}$$

where,

$$d < y(t,T) >= \Big(b_0 + g(t)b_1 + h(t)b_2\Big)^2 dt$$

Substituting the quadratic variation computed above and (3.15) in (3.16), we get

$$\begin{aligned}
dP(t,T) = P(t,T)\Bigg(\Big[&-(T-t)\Big(\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t)\\
&+ h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t)\Big) + \beta_0(t) + g(t)\beta_1(t) + h(t)\beta_2(t)\\
&+ \frac{1}{2}(T-t)^2\Big(b_0 + g(t)b_1 + h(t)b_2\Big)^2\Big]dt\\
&- (T-t)(b_0 + g(t)b_1 + h(t)b_2)dW(t)\Bigg)\\
= P(t,T)\Bigg(\Big(&\beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big)\\
&+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big)\\
&- a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu\\
&+ \frac{1}{2}(T-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2\Big)dt\\
&- (T-t)(b_0 + g(t)b_1 + h(t)b_2)dW(t)\Bigg)
\end{aligned} \tag{3.17}$$

Therefore, the above dynamics can be simplified to

$$dP(t,T) = P(t,T)\Big(M_1(t)dt + \sigma(t)dW(t)\Big) \tag{3.18}$$

where,

$$\begin{aligned}
M_1(t) &= \beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big)\\
&+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big)\\
&- a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu\\
&+ \frac{1}{2}(T-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2\\
\text{and} \quad \sigma(t) &= -(T-t)(b_0 + g(t)b_1 + h(t)b_2)
\end{aligned}$$

Hence, this bond price process is comparable to the asset price process specified by Korn and Kraft (2002). Additionally, at every maturity time T, the bond price dynamics of (3.17) results in $P(T,T) = 1$. We summarize this in the following Theorem 3.3.1.

Theorem: Single Brownian Motion

Theorem 3.3.1 (Yield Curve Dynamics under a Single Brownian Motion)**.** We assume that all of the factors of the dynamic Nelson-Siegel model, $\beta_0(t)$ and $\beta_1(t)$ and $\beta_2(t)$, are influenced by a single Brownian motion $W(t)$ and thereby satisfy the following stochastic paths defined as:

$$\begin{aligned} d\beta_0(t) &= \mu(a_0 - \beta_0(t))dt + b_0 dW(t) \\ d\beta_1(t) &= \mu(a_1(t) - \beta_1(t))dt + b_1 dW(t) \\ d\beta_2(t) &= \mu(a_2(t) - \beta_2(t))dt + b_2 dW(t) \end{aligned}$$

Then, the resulting bond price dynamics is given by:

$$\begin{aligned} dP(t,T) = P(t,T)\Bigg(\bigg(&\beta_0(t)(1+\mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\ &+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ &- a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ &+ \frac{1}{2}(T-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2\bigg)dt \\ &- (T-t)(b_0 + g(t)b_1 + h(t)b_2)dW(t)\Bigg) \end{aligned}$$

such that $P(T,T) = 1$.

Proof. The proof follows from the Ansatz: $P(t,T) = e^{-y(t,T)\cdot(T-t)}$ and all of the above considerations made in this section before the statement of the Theorem. □

Hence, for this stochastic model, we are able to successfully obtain analytically tractable yield curve and bond price dynamics which satisfies the terminal bond price condition.

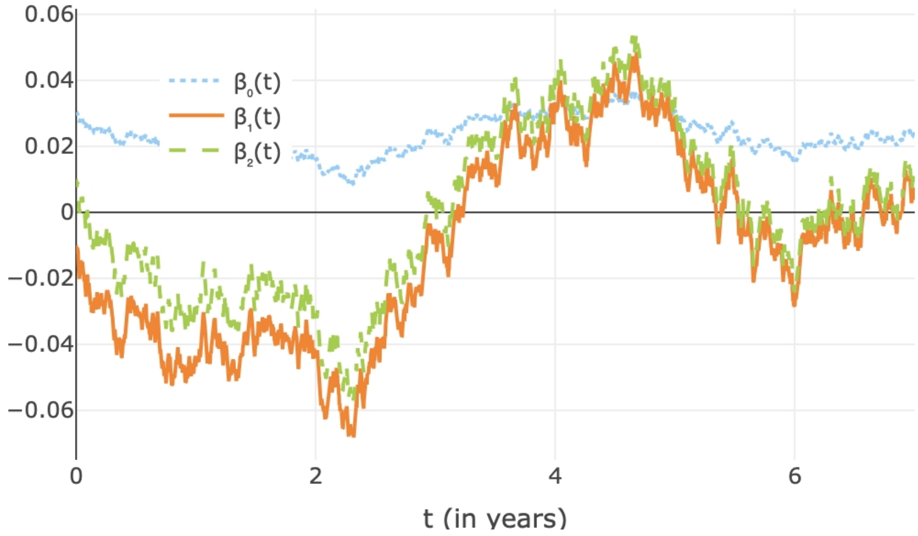
Numerical Example and Plots

Let us consider a numerical example so that we can visualize the evolution of the yield curve in this stochastic model. The simulations are carried out in R, but are easily reproducible in any other programming language. We use the Monte-Carlo approximation of stochastic differential equations highlighted in Korn et al. (2010) to simulate all SDEs.

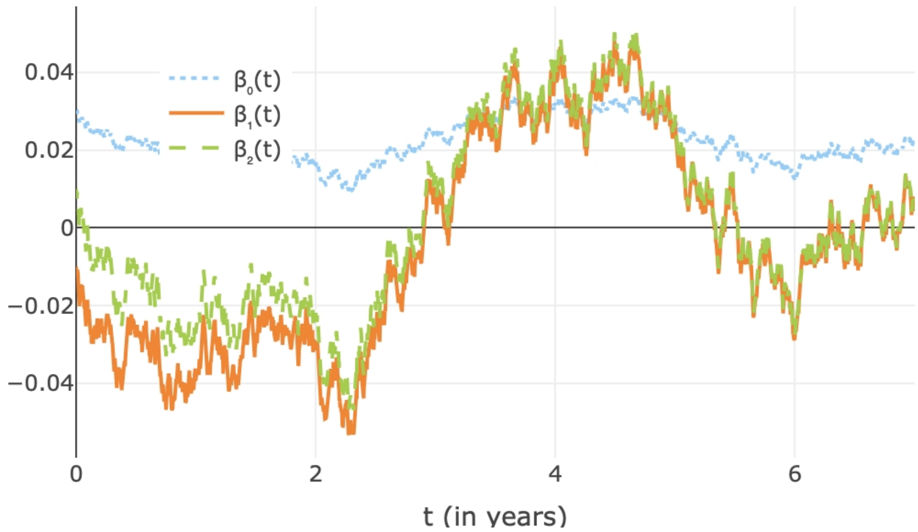
In Figure 3.6 and Figure 3.7, our choice of parameters are as follows: $\lambda = 0.5$, volatilities $b_0 = 0.008$, $b_1 = 0.03$, $b_2 = 0.03$. The drift parameters are $a_0 = 0.02$, $a_1(t) = 0.002$, $a_2(t) = 0.002$. The initial values of the $\beta(t)$ factors are $\beta_0(0) = 0.03$, $\beta_1(0) = -0.01$, and $\beta_2(0) = 0.01$. Here, the initial values $\beta_1(0)$ and $\beta_2(0)$ are chosen in a way to demonstrate how their behaviours change in comparison when the cases of uncorrelated and correlated Brownian motions are examined further in this chapter. We evolve the yield curve surface until a maximum time to maturity of $\tau = 30$. This is done as most long-term bonds have a maximum term to maturity of 30 years. The entire simulation is run daily for a time period of 7 years, thereby running from $t = 0$ to $t = 7$ with $dt = 1/365$.

To demonstrate the impact of considering the mean-reversion speed μ to be different from the factor loading decay rate λ, let us consider three different cases, namely, let

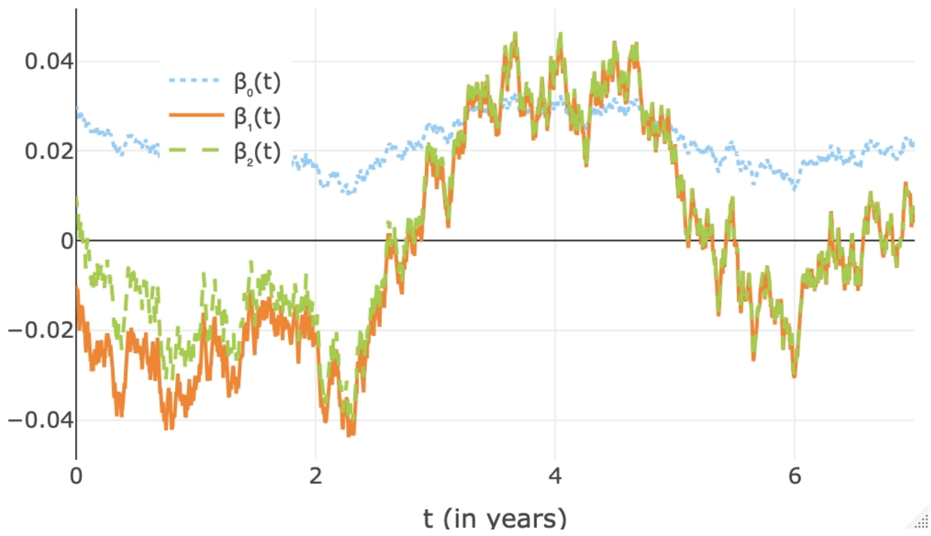
$\mu = 0.25, 0.5, 0.75$, respectively, in each sub-figure of Figure 3.6 and Figure 3.7. In Figure 3.6, on comparing each sub-figure, we see that the evolution of the $\beta_1(t)$ and $\beta_2(t)$ factors are not widely different. Their upper and lower bounds seem to decrease as the difference $\mu - \lambda$ goes from -0.25 to 0 to 0.25, respectively. This is explained with the fact that a higher mean-reversion speed causes a faster reversion to the mean-reversion level due to its impact on taming the volatility caused by the Brownian motion $W(t)$. Moreover, as μ increases, their paths seem to converge under this model specification. When comparing the different sub-figures, the difference in the paths of $\beta_0(t)$ is seemingly negligible for different μ. Hence, we further expect that the respective yield curve evolution will only exhibit a wider range in values for a larger mean-reversion speed μ. In Figure 3.7, we can observe that overall structure of the yield curve does not change drastically with respect to the relation between μ and λ. Thus, both Figure 3.6 and Figure 3.7 support this analysis, and we thereby conclude that our choice to assume μ to be different from λ is not detrimental to the model.



(a) For $\mu < \lambda$

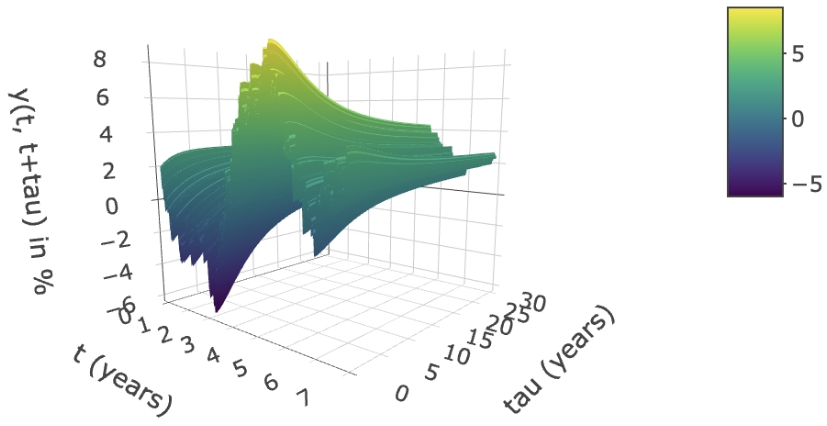


(b) For $\mu = \lambda$

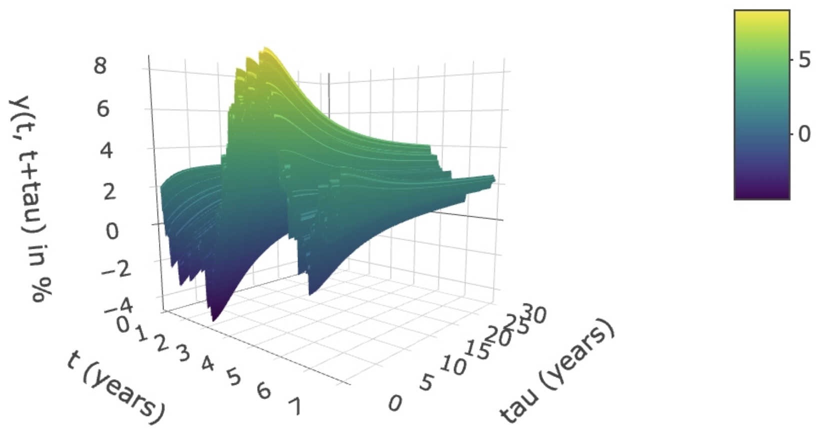


(c) For $\mu > \lambda$

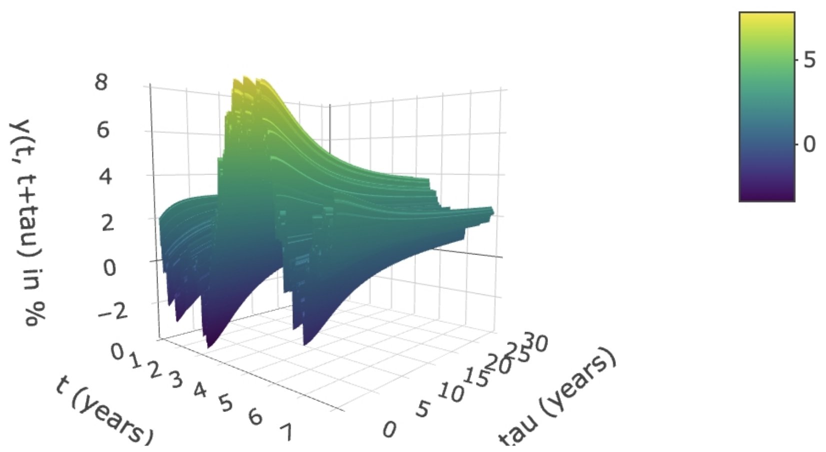
Figure 3.6: Single Brownian Motion: Stochastic Evolution of $\beta_0(t), \beta_1(t), \beta_2(t)$



(a) For $\mu < \lambda$



(b) For $\mu = \lambda$



(c) For $\mu > \lambda$

Figure 3.7: Single Brownian Motion: Stochastic Evolution of Yield Curve

To conclude the example, let us look at the descriptive statistics of the generated yield curve. Here, we consider the first sub-figure (a) of Figure 3.7, where $\mu = 0.25$. The tables for the other two sub-figures are presented in Appendix Chapter A. An overview of the descriptive statistics will provide an idea of how well our model exhibits common stylized facts of the yield curve.

	Mean	SD	CV	Min	Max
Yield (6 months)	1.3596	3.3640	2.4742	-5.7796	8.5253
Yield (10 years)	2.0667	1.6419	0.7945	-1.5848	5.6661
Yield (20 years)	2.1787	1.1307	0.5190	-0.3863	4.6839
$\beta_0(t)$	0.0229	0.0061	0.2639	0.0087	0.0365
$\beta_1(t)$	-0.0104	0.0285	-2.7358	-0.0683	0.0485
$\beta_2(t)$	-0.0010	0.0257	-26.0924	-0.0571	0.0547

Table 3.1: Descriptive statistics of generated yield curves: Single Brownian Motion with $\mu = 0.25$.

From Table 3.1, we observe the following:

1. The mean of 6-month yields are lower than that of 10-year yields, which are in turn lower than that of 20-year yields. Thus, long-term yields are on average higher than short-term yields. This is a well-known fact of the yield curve supported by economic theory, and is also satisfied by original DNS model.
2. The different standard deviation (SD) values show that short-term yields are more volatile than long-term yields due to the dependence on both $\beta_0(t)$ and $\beta_1(t)$, and on only $\beta_0(t)$, respectively. This observation is also supported by the original DNS model.
3. Diebold and Li (2006) state that $\beta_0(t)$ varies moderately relative to its mean, while $\beta_1(t)$ is quite variable relative to its mean and $\beta_2(t)$ varies the most relative to its mean. This is supported by the respective coefficient of variation (CV) values of the three factors.

Hence, we have shown that our stochastic model influenced by a single Brownian motion can fulfill requirements on analytical tractability as well as reproduce well-known economic interpretations of the yield curve.

3.3.4 Case 2: Uncorrelated Brownian Motions

In this case, we consider the case where each $\beta(t)$-factor is driven by a separate Brownian motion. There are three independent uncorrelated Brownian motions: $W_0(t)$, $W_1(t)$, and $W_2(t)$ with no pairwise instantaneous correlation.

$$dW_0(t)dW_1(t) = 0dt$$
$$dW_0(t)dW_2(t) = 0dt$$
$$dW_1(t)dW_2(t) = 0dt$$

The baseline stochastic model of (3.10) then becomes

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & b_2 \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{3.19}$$

Yield Curve Dynamics

The dynamic Nelson-Siegel yield curve

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t)$$

takes the form

$$\begin{aligned} y(t,T) = e^{-\mu t}\beta_0(0) + a_0(1 - e^{-\mu t}) + g(t)e^{-\mu t}\beta_1(0) + g(t)\mu \int_0^t e^{-\mu(t-u)} a_1(u) du \\ + h(t)e^{-\mu t}\beta_2(0) + h(t)\mu \int_0^t e^{-\mu(t-u)} a_2(u) du \\ + b_0 \int_0^t e^{-\mu(t-u)} dW_0(u) + g(t)b_1 \int_0^t e^{-\mu(t-u)} dW_1(u) + h(t)b_2 \int_0^t e^{-\mu(t-u)} dW_2(u) \end{aligned}$$

The yield curve dynamics is given as

$$\begin{aligned} dy(t,T) &= d\beta_0(t) + g(t)d\beta_1(t) + \beta_1(t)g'(t)dt + h(t)d\beta_2(t) + \beta_2(t)h'(t)dt \\ &= (\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t) \\ &\quad + h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t))dt \\ &\quad + \Big(b_0 dW_0(t) + g(t)b_1 dW_1(t) + h(t)b_2 dW_2(t)\Big) \end{aligned} \tag{3.20}$$

Equation (3.20) shows that the dynamics of the yield curve are driven by all of the three uncorrelated Brownian motions. This makes the representation of the derived bond price dynamics analytically more complex. There are two possibilities to deal with this. First, we can keep the yield curve dynamics as it is for problems that require the individual SDEs of the $\beta(t)$ factors. In such cases, the influence of all three Brownian motions are heavily significant (we see this type of problem in Chapter 4 where the individual Brownian motions of the three factors increases the dimensionality of the HJB equation). The other approach is to introduce a new Brownian motion such that the resultant yield curve dynamics and bond price dynamics are equivalent in distribution to those of the first approach. For such cases to be valid, we use the following lemma.

Lemma 3.3.2. *Let there be n stochastic processes driven by n independent Brownian motions $W_1(t), \dots, W_n(t)$ defined as follows:*

$$\begin{aligned} dX_1(t) &= a_1(t)dt + b_1(t)dW_1(t) \\ dX_2(t) &= a_2(t)dt + b_2(t)dW_2(t) \\ &\vdots \\ dX_n(t) &= a_n(t)dt + b_n(t)dW_n(t) \end{aligned}$$

where, $a_1(t), \cdots, a_n(t)$ and $b_1(t), \cdots, b_n(t)$ are deterministic functions of time, representing drift and volatility of the processes, respectively. Further, for all t values, the following is satisfied by the $b_1(t), \cdots, b_n(t)$ functions:

$$\int_0^t |b_i(s)|^2 ds < \infty \quad , i = 1, 2, \ldots, n$$

Then, the following holds at any time $t \geq 0$ for some Brownian motion $W(t)$:

$$\sum_{i=1}^{n} \int_0^t b_i(s) dW_i(s) \overset{d}{\sim} \int_0^t \sqrt{\sum_{i=1}^{n} |b_i(s)|^2} dW(s)$$

where, $\overset{d}{\sim}$ represents equality in distribution.

Proof. We know that $\int_0^t b_i(s)dW_i(s)$, for any $i = 1, 2, \ldots, n$ is a normally distributed random variable with the following parameters

$$\int_0^t b_i(s) dW_i(s) \sim N\left(0, \underbrace{\int_0^t |b_i(s)|^2 ds}_{=:\text{Var}_i}\right)$$

If we take the sum of these n random variables, we get

$$\sum_{i=1}^{n} \int_0^t b_i(s) dW_i(s) \sim N\left(0, \sum_{i=1}^{n} \int_0^t |b_i(s)|^2 ds\right) \qquad (*)$$

This is because the sum of independent normal random variables with mean $= 0$ is normally distributed with $\sim N(0, \sum_{i=1}^{n}(\text{Var}_i))$.

We also know that a normal random variable defined as $\int_0^t \sqrt{\sum_{i=1}^{n}|b_i(s)|^2} dW(s)$, driven by some Brownian motion $W(t)$ has the following parameters:

$$\int_0^t \sqrt{\sum_{i=1}^{n}|b_i(s)|^2} dW(s) \sim N\left(0, \int_0^t \sum_{i=1}^{n}|b_i(s)|^2 ds\right) \qquad (**)$$

We can now observe that $(*) = (**)$. Hence for any $t \geq 0$, the sum of stochastic processes $X_1(t) + X_2(t) + \ldots + X_n(t)$ driven by n independent Brownian motions can be represented to be equal in distribution to a stochastic process driven by a single Brownian motion $W(t)$ as follows:

$$\sum_{i=1}^{n} X_i(t) \overset{d}{\sim} \int_0^t \sum_{i=1}^{n} a_i(s) ds + \int_0^t \sqrt{\sum_{i=1}^{n} |b_i(s)|^2} dW(s)$$

□

Using this lemma, the yield curve dynamics obtained in (3.20) can be written to be equal in distribution to the following,

$$\begin{aligned} dy(t,T) \overset{d}{=} (&\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t) \\ &+ h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t))dt \\ &+ b_{\text{new}}(t)dW(t) \end{aligned} \qquad (3.21)$$

where, $W(t)$ is a single Brownian motion and,

$$b_{\text{new}}(t) = \sqrt{b_0^2 + g(t)^2 b_1^2 + h(t)^2 b_2^2}$$

Bond Price Dynamics

Using the same Ansatz for the definition of the bond price, the respective bond price dynamics is derived by substituting the condensed version of the yield curve (see (3.21)) in (3.16) to obtain the follows:

$$\begin{aligned} dP(t,T) &= P(t,T)\Big(- dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \\ &= P(t,T)\Bigg(\bigg(\beta_0(t)(1+\mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\ &\qquad + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ &\qquad - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ &\qquad + \frac{1}{2}(T-t)^2 b_{\text{new}}(t)^2\bigg)dt \\ &\quad - (T-t)b_{\text{new}}(t)dW(t)\Bigg) \end{aligned} \tag{3.22}$$

However, if we were to follow the first approach and derive the bond price dynamics directly from (3.20), keeping the three uncorrelated Brownian motions intact, we get

$$\begin{aligned} dP(t,T) &= P(t,T)\Big(- dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \\ &= P(t,T)\Bigg(\bigg(\beta_0(t)(1+\mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\ &\qquad + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ &\qquad - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ &\qquad + \frac{1}{2}(T-t)^2(b_0^2 + g(t)^2 b_1^2 + h(t)^2 b_2^2)\bigg)dt \\ &\quad - (T-t)b_0 dW_0(t) - (T-t)g(t)b_1 dW_1(t) - (T-t)h(t)b_2 dW_2(t)\Bigg) \end{aligned} \tag{3.23}$$

Similar to the derived bond price process of the previous case, these bond price processes are also comparable to the asset price process specified by Korn and Kraft (2002). Additionally, at every maturity time T, the bond price dynamics of (3.22) and (3.23) results in $P(T,T) = 1$. We leave it to the reader to decide between the appropriate equation for their suitable problem.

Thus, we arrive at the following theorem that summarizes our results for the case of uncorrelated Brownian motions.

Theorem: Uncorrelated Brownian Motions

Theorem 3.3.3 (Yield Curve Dynamics under Uncorrelated Brownian Motions)**.** If we assume all of the factors of the dynamic Nelson-Siegel model, $\beta_0(t)$ and $\beta_1(t)$ and $\beta_2(t)$, are influenced by three uncorrelated Brownian motions respectively as follows:

$$\begin{aligned} d\beta_0(t) &= \mu(a_0 - \beta_0(t))dt + b_0 dW_0(t) \\ d\beta_1(t) &= \mu(a_1(t) - \beta_1(t))dt + b_1 dW_1(t) \\ d\beta_2(t) &= \mu(a_2(t) - \beta_2(t))dt + b_2 dW_2(t) \end{aligned}$$

Then, the resulting bond price dynamics is given by:

$$dP(t,T) = P(t,T)\bigg(M_2(t)dt - (T-t)b_0 dW_0(t) - (T-t)g(t)b_1 dW_1(t) - (T-t)h(t)b_2 dW_2(t)\bigg)$$

where,

$$\begin{aligned} M_2(t)dt = \beta_0(t)(1+\mu(T-t)) &+ \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\ &+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ &- a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ &+ \frac{1}{2}(T-t)^2(b_0^2 + g(t)^2 b_1^2 + h(t)^2 b_2^2) \end{aligned}$$

such that $P(T,T) = 1$.

As an alternative, Lemma 3.3.2 may be used to derive the bond price dynamics given by (3.22) when all uncorrelated Brownian motions are condensed to a single Brownian motion $W(t)$.

Proof. The proof follows from the Ansatz: $P(t,T) = e^{-y(t,T)\cdot(T-t)}$, and all of the statements and considerations made before the theorem. If the reader prefers to condense the three independent Brownian motions into a single Brownian motion, they may use Lemma 3.3.2 and obtain the bond price dynamics as stated in (3.22). □

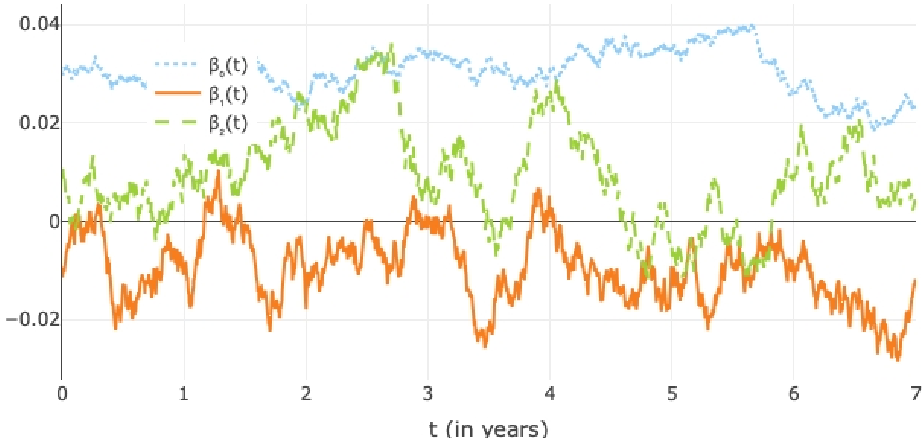
Numerical Example and Plots

Let us again consider a numerical example in order to visualize the yield curve evolution, and the evolution of the three stochastic $\beta(t)$ factors for the case of uncorrelated Brownian motions.

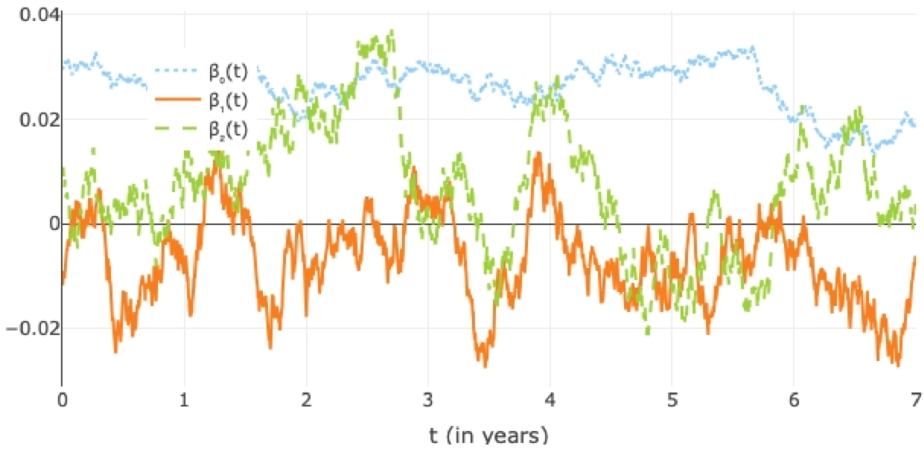
Figure 3.8 presents the evolution of all of the stochastic $\beta_0(t), \beta_1(t), \beta_2(t)$ factors where they are respectively influenced by three uncorrelated Brownian motions. Figure 3.9 depicts the corresponding resultant yield curve surfaces $y(t, t+\tau)$. We choose volatility parameters as $b_0 = 0.008, b_1 = 0.02, b_2 = 0.02$, and drift parameters as $a_0 = 0.02, a_1(t) = 0.002, a_2(t) = 0.002$. The initial values are $\beta_0(0) = 0.03, \beta_1(0) = -0.01, \beta_2(0) = 0.01$. The curve is simulated for a maximum term to maturity of $\tau = 30$ years. The calendar time is run from $t = 0$ to $t = 7$ with $dt = 1/365$. Although we concluded in Section 3.3.3 that the choice of $\mu \neq \lambda$ is not disadvantageous to model the yield curve evolution, we once again present three different yield curve surfaces based on the differences between μ and

λ in Figure 3.9 for the sake of completeness in the case of multiple Brownian motions. We choose $\lambda = 0.5$ and $\mu = 0.25, 0.5, 0.75$, respectively, to display the differences in each sub-figure of Figure 3.8 and Figure 3.9.

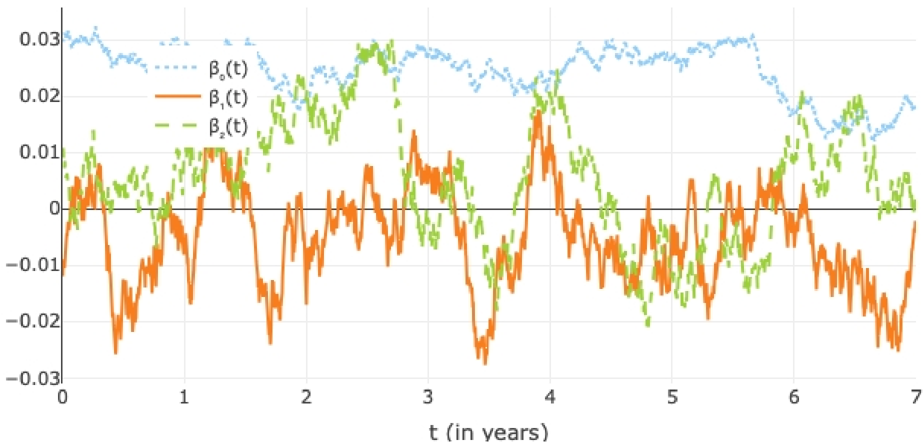
It is observable in Figure 3.9 that all of the yield curves have an overall similar structure. The range of yield values do not vary significantly, nor do the overall shape and behaviour show any drastic differences. In Figure 3.8, we notice that the $\beta(t)$ factors tend to fluctuate more closely withe each other in each sub-figure as μ increases. This can be explained with the fact that a higher mean-reversion speed μ of the three $\beta(t)$ factors result in a faster movement of their values about their respective means. Hence, we can safely conclude that it is not the difference $\mu - \lambda$, but the higher value of μ that causes the movement.



(a) For $\mu < \lambda$

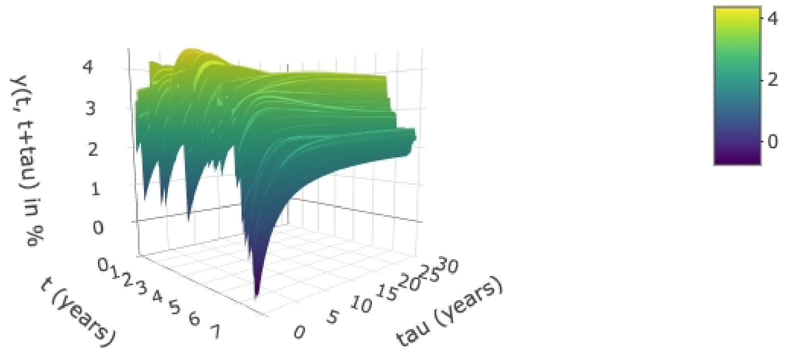


(b) For $\mu = \lambda$

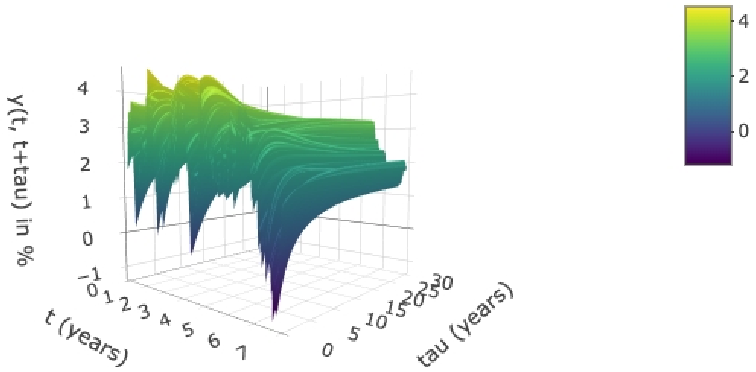


(c) For $\mu > \lambda$

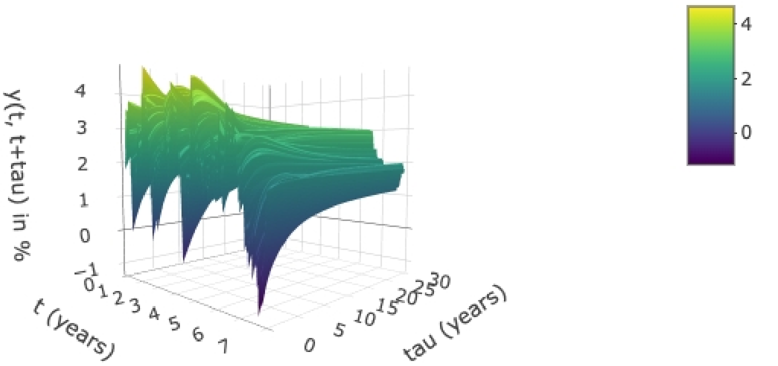
Figure 3.8: Uncorrelated Brownian Motions: Stochastic Evolution of $\beta_0(t), \beta_1(t), \beta_2(t)$



(a) For $\mu < \lambda$



(b) For $\mu = \lambda$



(c) For $\mu > \lambda$

Figure 3.9: Uncorrelated Brownian Motions: Stochastic Evolution of Yield Curve

Let us view the descriptive statistics table of the generated yield curve for this case of uncorrelated Brownian motions. Here, we only present the table for $\mu = 0.25$ (the first sub-figure of Figure 3.9). The corresponding tables for the other two sub-figures can be found in Appendix Chapter A.

	Mean	SD	CV	Min	Max
Yield (6 months)	2.2994	0.8840	0.3845	-0.4183	4.1141
Yield (10 years)	3.0414	0.5068	0.1666	1.4726	4.1140
Yield (20 years)	3.0504	0.4579	0.1501	1.6948	3.8649
$\beta_0(t)$	0.0305	0.0045	0.1459	0.0183	0.0402
$\beta_1(t)$	-0.0096	0.0071	-0.7317	-0.0285	0.0104
$\beta_2(t)$	0.0094	0.0101	1.0841	-0.0120	0.0361

Table 3.2: Descriptive statistics of generated yield curves: Uncorrelated Brownian Motions with $\mu = 0.25$.

In Table 3.2, we observe that the generated yield curve surface under this second stochastic model also satisfies the following stylized (economic) facts of yield curves:

1. The average 6-month yield is 2.29%, the average 10-year yield is 3.04%, and the average 20-year yield is 3.05%. This indicates that long-term yields are higher on average than short-term yields.
2. The volatility of the 6-month yield is 0.8840, while the volatilities of the 10-year yield and the 20-year yield are 0.5068 and 0.4579, respectively. This indicates that short-term yields are more volatile than long-term yields.
3. Since the coefficient of variation with the mean is 0.15 for $\beta_0(t)$, -0.73 for $\beta_1(t)$, and 1.08 for $\beta_2(t)$, we can infer that the level factor $\beta_0(t)$ is more persistent than the slope factor $\beta_1(t)$ which in turn is more persistent than the curvature factor $\beta_2(t)$.

3.3.5 Case 3: Correlated Brownian Motions

We close our analysis by considering the final case of the stochastic framework, where each $\beta(t)$-factor is driven by correlated Brownian motions. Let us consider three Brownian motions: $W_0(t)$, $W_1(t)$, and $W_2(t)$ with the following pairwise correlation coefficients:

$$\mathbb{E}\Big(W_0(t), W_1(t)\Big) = \rho_{01}$$
$$\mathbb{E}\Big(W_0(t), W_2(t)\Big) = \rho_{02}$$
$$\mathbb{E}\Big(W_1(t), W_2(t)\Big) = \rho_{12}$$

Hence, for a 3-dimensional Brownian motion $W(t) = (W_0(t), W_1(t), W_2(t))'$,

$$\text{Cov}(W(t)) = \Sigma t$$
$$\text{where, } \Sigma = \begin{pmatrix} 1 & \rho_{01} & \rho_{02} \\ \rho_{01} & 1 & \rho_{12} \\ \rho_{02} & \rho_{12} & 1 \end{pmatrix}$$

To extract independent Brownian motions from these, $\tilde{W}(t) = (\tilde{W}_0(t), \tilde{W}_1(t), \tilde{W}_2(t))'$, such that

$$\text{Cov}(\tilde{W}(t)) = It$$

we must have,

$$W(t) = L\tilde{W}(t)$$

where, L is the lower-triangular Cholesky factor, given by

$$L = \begin{pmatrix} L_{00} & 0 & 0 \\ L_{10} & L_{11} & 0 \\ L_{20} & L_{21} & L_{22} \end{pmatrix}$$

Using $\Sigma = LL'$ we thus obtain,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \rho_{01} & \underbrace{\sqrt{1-\rho_{01}^2}}_{=:\tilde{\rho}_{11}} & 0 \\ \rho_{02} & \underbrace{\frac{\rho_{12}-\rho_{01}\rho_{02}}{\sqrt{1-\rho_{01}^2}}}_{=:\tilde{\rho}_{21}} & \underbrace{\sqrt{\frac{1-\rho_{01}^2-\rho_{02}^2-\rho_{12}^2+2\rho_{01}\rho_{02}\rho_{12}}{1-\rho_{01}^2}}}_{=:\tilde{\rho}_{22}} \end{pmatrix}$$

The baseline stochastic model of (3.10) then becomes

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ b_1\rho_{01} & b_1\tilde{\rho}_{11} & 0 \\ b_2\rho_{02} & b_2\tilde{\rho}_{21} & b_2\tilde{\rho}_{22} \end{pmatrix} \begin{pmatrix} d\tilde{W}_0(t) \\ d\tilde{W}_1(t) \\ d\tilde{W}_2(t) \end{pmatrix} \tag{3.24}$$

with $\mu > 0$. Moreover, this representation implies that this stochastic model can be decomposed to independent Brownian motions $\tilde{W}_0(t), \tilde{W}_1(t), \tilde{W}_2(t)$, thereby making it similar to the methodology of case 2.

Yield Curve Dynamics

The dynamic Nelson-Siegel yield curve

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t)$$

takes the form

$$\begin{aligned} y(t,T) = {} & e^{-\mu t}\beta_0(0) + a_0(1-e^{-\mu t}) + g(t)e^{-\mu t}\beta_1(0) + g(t)\mu\int_0^t e^{-\mu(t-u)}a_1(u)du \\ & + h(t)e^{-\mu t}\beta_2(0) + h(t)\mu\int_0^t e^{-\mu(t-u)}a_2(u)du \\ & + (b_0 + b_1\rho_{01}g(t) + b_2\rho_{02}h(t))\int_0^t e^{-\mu(t-u)}d\tilde{W}_0(u) \\ & + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})\int_0^t e^{-\mu(t-u)}d\tilde{W}_1(u) \\ & + h(t)b_2\tilde{\rho}_{22}\int_0^t e^{-\mu(t-u)}d\tilde{W}_2(u) \end{aligned}$$

The yield curve dynamics therefore is given as

$$\begin{aligned} dy(t,T) &= d\beta_0(t) + g(t)d\beta_1(t) + \beta_1(t)g'(t)dt + h(t)d\beta_2(t) + \beta_2(t)h'(t)dt \\ &= (\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t) \\ &\quad + h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t))dt \\ &\quad + \Big((b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02})d\tilde{W}_0(t) \\ &\quad + (b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{21})d\tilde{W}_1(t) \\ &\quad + b_2 h(t)\tilde{\rho}_{22} d\tilde{W}_2(t)\Big) \end{aligned} \tag{3.25}$$

Now, for the dynamics of the yield curve and the bond prices, we can keep them as they are for our portfolio optimization problem, or once again collapse all three Brownian motions into a single Brownian motion $\tilde{W}(t)$, as done in Case 2: Uncorrelated Brownian motions. We present both resultant equations here.

Let there be a single Brownian motion $\tilde{W}(t)$. Then, using Lemma 3.3.2, we get

$$\begin{aligned} dy(t,T) &\stackrel{d}{=} (\mu(a_0 - \beta_0(t)) + g(t)\mu(a_1(t) - \beta_1(t)) + \beta_1(t)g'(t) \\ &\quad + h(t)\mu(a_2(t) - \beta_2(t)) + \beta_2(t)h'(t))dt \\ &\quad + \tilde{b}_{\text{new}}(t)d\tilde{W}(t) \end{aligned} \tag{3.26}$$

where,

$$\tilde{b}_{\text{new}}(t) = \sqrt{\Big(b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02}\Big)^2 + \Big(b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{21}\Big)^2 + \Big(b_2 h(t)\tilde{\rho}_{22}\Big)^2}$$

Bond Price Dynamics

Using the same Ansatz for the definition of the bond price, the respective dynamics is derived after substituting (3.26) in (3.16) to obtain the follows:

$$\begin{aligned} dP(t,T) &= P(t,T)\Big(-dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \\ &= P(t,T)\Bigg(\Big(\beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big) \\ &\quad + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ &\quad - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ &\quad + \frac{1}{2}(T-t)^2\tilde{b}_{\text{new}}(t)^2\Big)dt \\ &\quad - (T-t)\tilde{b}_{\text{new}}(t)d\tilde{W}(t)\Bigg) \end{aligned} \tag{3.27}$$

However, if we were to follow the first approach and derive the bond price dynamics directly from (3.25), we get,

$$dP(t,T) = P(t,T)\Big(-dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big)$$

$$
\begin{aligned}
&= P(t,T)\Bigg(\Big[\beta_0(t)(1+\mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\
&\quad + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\
&\quad - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\
&\quad + \frac{1}{2}(T-t)^2\Big((b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02})^2 \\
&\quad + (b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{12})^2 + (b_2 h(t)\tilde{\rho}_{22})^2\Big)\Big]dt \\
&\quad - (T-t)(b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02})d\tilde{W}_0(t) \\
&\quad - (T-t)(b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{21})d\tilde{W}_1(t) \\
&\quad - (T-t)h(t)b_2\tilde{\rho}_{22}d\tilde{W}_2(t)\Bigg) \qquad (3.28)
\end{aligned}
$$

Similar to the derived bond price process of case 1, these bond price processes are also comparable to the asset price process specified by Korn and Kraft (2002). Additionally, at every maturity time T, the bond price dynamics of (3.22) and (3.23) results in $P(T,T) = 1$. We leave it to the reader to decide between the appropriate equation for their suitable problem.

Theorem: Correlated Brownian Motions

Theorem 3.3.4 (Yield Curve Model Dynamics under Correlated Brownian Motions)**.** If we assume all of the factors of the dynamic Nelson-Siegel model, $\beta_0(t)$ and $\beta_1(t)$ and $\beta_2(t)$, have the following underlying stochastic processes:

$$
\begin{aligned}
d\beta_0(t) &= \mu(a_0 - \beta_0(t))dt + b_0 dW_0(t) \\
d\beta_1(t) &= \mu(a_1(t) - \beta_1(t))dt + b_1 dW_1(t) \\
d\beta_2(t) &= \mu(a_2(t) - \beta_2(t))dt + b_2 dW_2(t)
\end{aligned}
$$

where $W_0(t), W_1(t), W_2(t)$ are three pairwise correlated Brownian motions with coefficients $\rho_{01}, \rho_{02}, \rho_{12}$, and can be decomposed into independent Brownian motions $\tilde{W}_0(t), \tilde{W}_1(t), \tilde{W}_2(t)$. Then, the resulting bond price dynamics is given by:

$$
\begin{aligned}
dP(t,T) = P(t,T)\Bigg(&M_3(t)dt - (T-t)(b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02})d\tilde{W}_0(t) \\
&- (T-t)(b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{12})d\tilde{W}_1(t) \\
&- (T-t)h(t)b_2\tilde{\rho}_{22}d\tilde{W}_2(t)\Bigg)
\end{aligned}
$$

where,

$$
\begin{aligned}
M_3(t)dt = &\beta_0(t)(1+\mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu}{\lambda})\Big) \\
&+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big)
\end{aligned}
$$

$$
\begin{aligned}
&- a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\
&+ \frac{1}{2}(T-t)^2\Big((b_0 + b_1 g(t)\rho_{01} + h(t)b_2\rho_{02})^2 \\
&\qquad + (b_1 g(t)\tilde{\rho}_{11} + b_2 h(t)\tilde{\rho}_{12})^2 + (b_2 h(t)\tilde{\rho}_{22})^2\Big)dt
\end{aligned}
$$

and $\tilde{W}(t) = (\tilde{W}_0(t), \tilde{W}_1(t), \tilde{W}_2(t))'$ are the extracted independent Brownian motions, such that $P(T,T) = 1$.

Proof. The proof is the same as that of Theorem 3.3.3. It follows from the Ansatz: $P(t,T) = e^{-y(t,T)\cdot(T-t)}$, and all of the statements made in this subsection before this theorem. If the reader chooses to condense these correlated Brownian motions into a single Brownian motion, then Lemma 3.3.2 can be used to obtain the bond price dynamics as stated above in (3.27). □

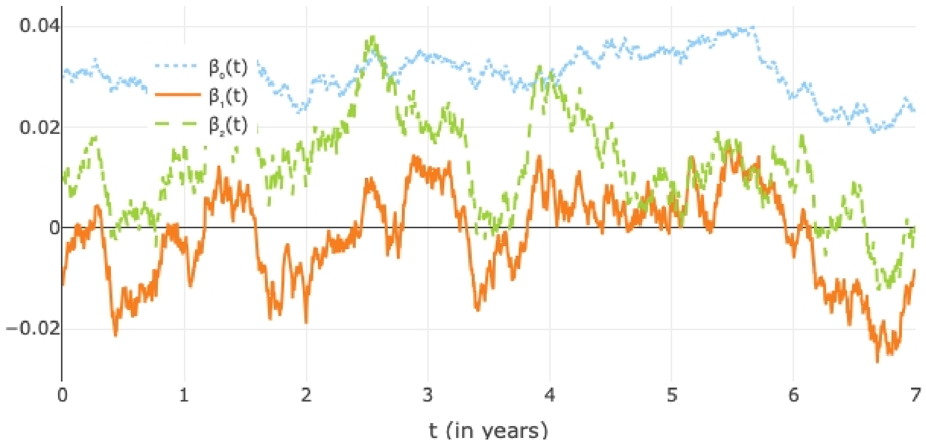
Numerical Example and Plots

In our example for correlated Brownian motions, we choose the same parameters as in the previous uncorrelated case (see Section 3.3.4). This means that the same seed is used to calculate the Brownian motions. For our analysis, we present results for (a) pairwise positive correlation with coefficients: $\rho_{01} = 0.5, \rho_{02} = 0.5, \rho_{12} = 0.5$, and (b) pairwise negative correlation with coefficients: $\rho_{01} = -0.5, \rho_{02} = -0.5, \rho_{12} = -0.3$. Since we have already showed in Figure 3.7 and Figure 3.9 that the choice of $\mu \neq \lambda$ in our model has no impact on the final yield curve surface, for this case we present only one scenario each for yield curve evolution and evolution of the stochastic $\beta(t)$ factors. Hence, we choose $\mu = 0.25$ and $\lambda = 0.5$ for both analyses.

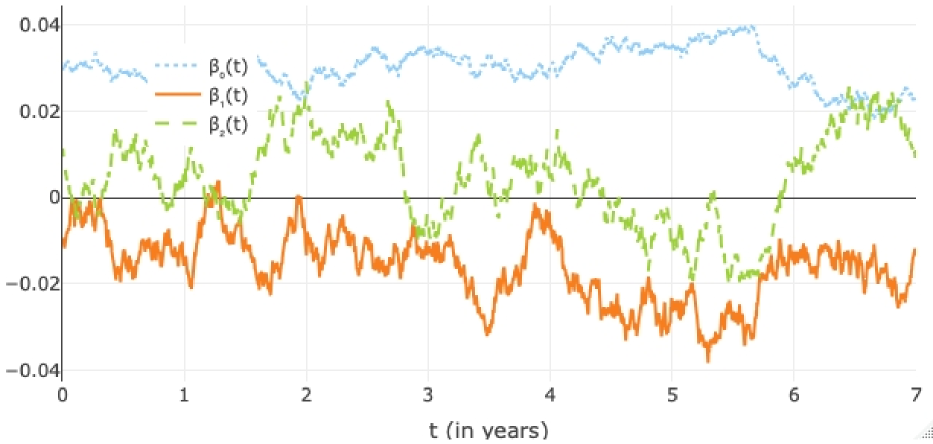
Figure 3.10 depicts the stochastic evolution of $\beta_0(t)$, $\beta_1(t)$, and $\beta_2(t)$ and Figure 3.11 depicts the corresponding yield curve surfaces $y(t, t+\tau)$.

We observe that under the same seed of Brownian motions, the stochastic evolution of the three $\beta(t)$ factors for pairwise positively correlated motions in sub-figure (a) of Figure 3.10 seemingly follow similar behaviour as their counterpart in the case of uncorrelated motions. The yield curve surface generated in sub-figure (a) of Figure 3.11 also does not differ significantly in terms of its overall range and structure when compared to the corresponding equivalent surface of uncorrelated motions (see sub-figure (a) of Figure 3.9).

On the other hand, sub-figure (b) of Figure 3.10 show that when the $\beta(t)$ factors are driven by pairwise negatively correlated Brownian motions, with the same seed, their paths behave differently than when compared to both positively correlated motions as well as uncorrelated motions. Moreover, the corresponding yield curve displayed in sub-figure (b) of Figure 3.11 differs evidently from that of the positively correlated motions. On first glance, it displays a lower range of yield values and has more normal curves. The table of descriptive statistics on the next page will enable us to infer some concrete remarks on these observations.

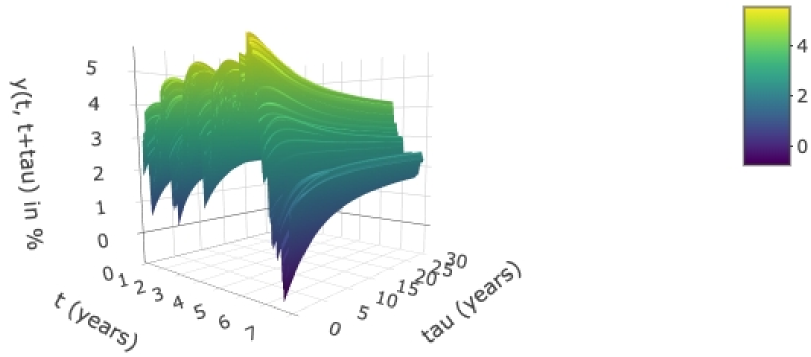


(a) With pairwise positive coefficients

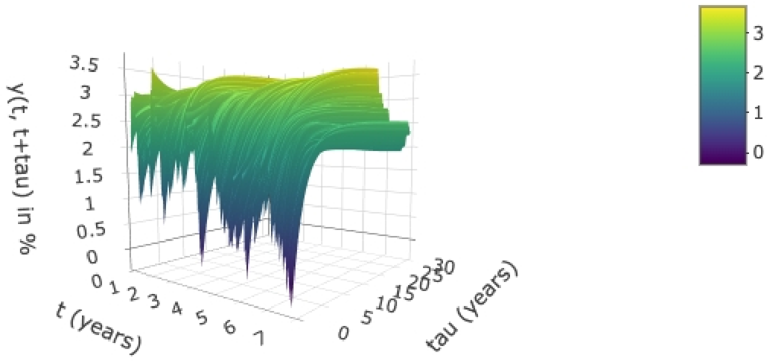


(b) With pairwise negative coefficients

Figure 3.10: Correlated Brownian Motions: Stochastic Evolution of $\beta_0(t), \beta_1(t), \beta_2(t)$



(a) With pairwise positive coefficients



(b) With pairwise negative coefficients

Figure 3.11: Correlated Brownian Motions: Stochastic Evolution of Yield Curve

	Mean	SD	CV	Min	Max
Yield (6 months)	3.0441	1.2745	0.4187	-0.5899	5.5571
Yield (10 years)	3.2597	0.7233	0.2219	1.1371	4.5988
Yield (20 years)	3.1613	0.5790	0.1831	1.5199	4.2894
$\beta_0(t)$	0.0305	0.0045	0.1459	0.0183	0.0402
$\beta_1(t)$	-0.0016	0.0093	-5.8266	-0.0268	0.0165
$\beta_2(t)$	0.0124	0.0096	0.7770	-0.0129	0.0381

Table 3.3: Descriptive statistics of generated yield curves: Positively Correlated Brownian Motions with $\mu = 0.25$.

	Mean	SD	CV	Min	Max
Yield (6 months)	1.6765	0.6259	0.3733	0.2304	3.3604
Yield (10 years)	2.8186	0.2674	0.0949	1.9921	3.3921
Yield (20 years)	2.9366	0.3337	0.1136	1.9410	3.5175
$\beta_0(t)$	0.0305	0.0045	0.1459	0.0183	0.0402
$\beta_1(t)$	-0.0161	0.0079	-0.4935	-0.0383	0.0040
$\beta_2(t)$	0.0044	0.0105	2.3581	-0.0200	0.0269

Table 3.4: Descriptive statistics of generated yield curves: Negatively Correlated Brownian Motions with $\mu = 0.25$.

Table 3.3 and Table 3.4 exhibit largely inconsistent behaviour. We make the following inferences:

1. Under the influence of negatively correlated Brownian motions, the average long-term yields are higher than the average short-term yields. This is similar to the observations made for the case of a single Brownian motion and for the case of uncorrelated Brownian motions.

   However, under the influence of highly positively correlated motions, the pattern is different. The 10-year yield is higher than both the 6-month yield and the 20-year yield. This indicates that the average yield curve is humped rather than strictly concave and increasing.

2. The coefficient of variation of $\beta_0(t), \beta_1(t), \beta_2(t)$ is in increasing order for the case of negatively correlated motions. This is in accordance with the assertion in Diebold and Li (2006) that the $\beta_0(t)$ factor has the least variation relative to its mean while the $\beta_2(t)$ factor exhibits the most variation.

   This is however not satisfied in the case of positively correlation motions. As observable, $\beta_1(t)$ displays the most variation relative to its mean with $CV = -5.8266$.

3. Lastly, under the influence of negatively correlated motions, the volatilities of the long-term yields are lower than that of the short-term yields. Surprisingly, the volatility of the 10-year yield is lower than that of the 20-year yield.

   Under the influence of positively correlated motions, this abnormality is not observed. Long-term yields are in general less volatile than short-term yields.

With these observations of the simulated yield curve surfaces, we state that it is possible that considering correlation (either positive or negative) between the three $\beta(t)$ factors is not an ideal assumption.

Interestingly, when Diebold and Li (2006) evaluated the historical yield curve data, they also made the observation that the level $\beta_0(t)$, slope $\beta_1(t)$, and the curvature $\beta_2(t)$ factors displayed negligible correlation with each other. This observation further played a crucial role in their modeling choices for estimating the dynamic Nelson-Siegel yield curves. Since our results on the stochastic model for correlated Brownian motions support their assertions without any prior bias, we thereby conclude that both stochastic as well as deterministic evolution of the dynamic Nelson-Siegel model fare better without the presence of correlated Brownian motions.

3.4 Probability Distribution of Yield Curve Shapes

From Corollary 3.2.4, we see that the occurrence of all four possible shapes are determined by certain conditions imposed on both $\beta_1(t)$ and $\beta_2(t)$. Now that we have considered models where the $\beta_1(t)$ and $\beta_2(t)$ factors are driven by stochastic processes, we can obtain explicit calculations of the probabilities of attaining each of the four yield curve shapes using these conditions.

3.4.1 Case: Single Brownian Motion

Let us construct the following remark as a result of Corollary 3.2.4 applied in the case of a single Brownian motion influencing all three $\beta(t)$ factors of the stochastic DNS framework.

Remark 3.4.1 (Probabilities of Yield Curve Shapes under a Single Brownian Motion Framework)**.** Let the stochastic processes of $\beta_1(t)$ and $\beta_2(t)$ of the dynamic Nelson-Siegel be given by equations (3.12) and (3.13) where they are both influenced by a single Brownian motion $W(t)$. Then, for every $t > 0$, the probabilities of each yield curve shape to appear are given by:

1. $P_{\text{norm}} = P(\beta_1(t) < -|\beta_2(t)|)$
2. $P_{\text{inv}} = P(\beta_1(t) > |\beta_2(t)|)$
3. $P_{\text{hump}} = P(|\beta_1(t)| < \beta_2(t))$
4. $P_{\text{dip}} = P(-|\beta_1(t)| > \beta_2(t))$

Let us begin with the calculations for only P_{norm} below. The calculations for the probabilities of the other shapes follow similarly, and are detailed in Appendix Chapter A.

Calculations for Normal Curves P_{norm}

From (3.12), $\beta_1(t)$ is normally distributed as follows:

$$\beta_1(t) \sim N\Big(e^{-\mu t}\beta_1(0) + \mu \int_0^t e^{-\mu(t-u)} a_1(u)du, \quad \frac{b_1^2}{2\mu}(1 - e^{-2\mu t})\Big)$$

Similarly from (3.13), $\beta_2(t)$ is also normally distributed as follows:

$$\beta_2(t) \sim N\Big(e^{-\mu t}\beta_2(0) + \mu \int_0^t e^{-\mu(t-u)} a_2(u)du, \quad \frac{b_2^2}{2\mu}(1 - e^{-2\mu t})\Big)$$

For the case of a single Brownian motion $W(t)$, the solutions of the SDEs of $\beta_1(t)$ and $\beta_2(t)$ described in Theorem 3.3.1 are given as:

$$\beta_1(t) = \beta_1(0)e^{-\mu t} + \mu \int_0^t e^{-\mu(t-u)} a_1(u)du + b_1 \int_0^t e^{-\mu(t-u)} dW(u)$$
$$\beta_2(t) = \beta_2(0)e^{-\mu t} + \mu \int_0^t e^{-\mu(t-u)} a_2(u)du + b_2 \int_0^t e^{-\mu(t-u)} dW(u)$$

For normal curves, we have

$$
\begin{aligned}
P_{\text{norm}} &= P\Big(\beta_1(t) \leq -|\beta_2(t)|\Big) \\
&= P(\beta_1(t) \leq -\beta_2(t) \,\Big|\, \beta_2(t) > 0) \cdot P(\beta_2(t) > 0) \\
&\quad + P(\beta_1(t) \leq \beta_2(t) \,\Big|\, \beta_2(t) \leq 0) \cdot P(\beta_2(t) \leq 0) \quad (\because \text{law of total probability})
\end{aligned} \tag{3.29}
$$

If X is a normal random variable with parameters m and σ, then $Y = \frac{X-m}{\sigma}$ has a standard normal distribution with $Y \sim N(0,1)$. Therefore, we can calculate the following:

$$
\begin{aligned}
P(\beta_2(t) \leq 0) &= \Phi\left(\frac{0 - \left(\beta_2(0)e^{-\mu t} + \mu\int_0^t e^{-\mu(t-u)}a_2(u)du\right)}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \\
&= \Phi\left(\frac{-\beta_2(0)e^{-\mu t} - \mu\int_0^t e^{-\mu(t-u)}a_2(u)du}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \\
P(\beta_2(t) > 0) &= 1 - P(\beta_2(t) \leq 0) \\
&= 1 - \Phi\left(\frac{-\beta_2(0)e^{-\mu t} - \mu\int_0^t e^{-\mu(t-u)}a_2(u)du}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \\
&= \Phi\left(\frac{\beta_2(0)e^{-\mu t} + \mu\int_0^t e^{-\mu(t-u)}a_2(u)du}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \quad (\because 1 - \Phi(-x) = \Phi(x))
\end{aligned}
$$

where, $\Phi(\cdot)$ is the cumulative distribution function of a standard normal variable.

To calculate the remaining terms of (3.29), we use the SDE solutions of $\beta_1(t)$ and $\beta_2(t)$ as stated above.

Also, for the ease of calculation, let $a_1(u) = a_2(u) \equiv 0$.

For $P(\beta_1(t) \leq -\beta_2(t) \,\Big|\, \beta_2(t) > 0)$, we can obtain that

$$
\begin{aligned}
&\beta_2(t) > 0 \\
\Longrightarrow\ &\beta_2(0)e^{-\mu t} + b_2\int_0^t e^{-\mu(t-u)}dW(u) > 0 \\
\Longrightarrow\ &\int_0^t e^{-\mu(t-u)}dW(u) > \frac{-1}{b_2}\beta_2(0)e^{-\mu t} \quad (\text{assuming } b_2 > 0)
\end{aligned}
$$

and,

$$
\begin{aligned}
&\beta_1(t) \leq -\beta_2(t) \,\Big|\, \beta_2(t) > 0 \\
\Longrightarrow\ &\beta_1(t) + \beta_2(t) \leq 0 \,\Big|\, \beta_2(t) > 0 \\
\Longrightarrow\ &(\beta_1(0) + \beta_2(0))e^{-\mu t} + (b_1 + b_2)\int_0^t e^{-\mu(t-u)}dW(u) \leq 0
\end{aligned}
$$

$$\implies \int_0^t e^{-\mu(t-u)} dW(u) \leq \frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t} \quad \text{(assuming } (b_1+b_2) > 0)$$

Hence, we have

$$\underbrace{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}_{=:z_1} \leq \int_0^t e^{-\mu(t-u)} dW(u) \leq \underbrace{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}_{=:y_1}$$

The above holds, if and only if $z_1 < y_1$. Otherwise, the entire conditional probability $P(\beta_1(t) \leq \beta_2(t) \,\big|\, \beta_2(t) > 0) = 0$. Moreover, we can consider the middle term of the above inequality as follows

$$\underbrace{\int_0^t e^{-\mu(t-u)} dW(u)}_{=:N} \sim N\Big(0, \frac{1}{2\mu}(1-e^{-2\mu t})\Big)$$

Thus, the conditional probability $P(\beta_1(t) \leq \beta_2(t) \,\big|\, \beta_2(t) > 0)$ is calculated as:

$$\begin{aligned}
P(\beta_1(t) \leq -\beta_2(t) \,\big|\, \beta_2(t) > 0) &= \max\left(0, P\Big(z_1 \leq N \leq y_1\Big)\right) \\
&= \max\left(0, \Phi\left(\frac{y_1}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{z_1}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right) \\
&= \max\left(0, \Phi\left(\frac{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right)
\end{aligned}$$

Now for $P(\beta_1(t) \leq \beta_2(t) \,\big|\, \beta_2(t) \leq 0)$, we can similarly obtain

$$\begin{aligned}
&\beta_2(t) \leq 0 \\
\implies &\beta_2(0)e^{-\mu t} + b_2 \int_0^t e^{-\mu(t-u)} dW(u) \leq 0 \\
\implies &\int_0^t e^{-\mu(t-u)} dW(u) \leq \underbrace{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}_{\text{already defined as } z_1} \quad \text{(assuming } b_2 > 0)
\end{aligned}$$

and,

$$\begin{aligned}
&\beta_1(t) \leq \beta_2(t) \,\big|\, \beta_2(t) \leq 0 \\
\implies &\beta_1(t) - \beta_2(t) \leq 0 \,\big|\, \beta_2(t) \leq 0 \\
\implies &(\beta_1(0)-\beta_2(0))e^{-\mu t} + (b_1-b_2)\int_0^t e^{-\mu(t-u)} dW(u) \leq 0 \\
\implies &\int_0^t e^{-\mu(t-u)} dW(u) \leq \underbrace{\frac{-1}{b_1-b_2}(\beta_1(0)-\beta_2(0))e^{-\mu t}}_{=:y_2} \quad \text{(assuming } (b_1-b_2) > 0)
\end{aligned}$$

Since we do not know the exact relation between the parameters, we have

$$\underbrace{\int_0^t e^{-\mu(t-u)} dW(u)}_{\text{already defined as } \sim N\left(0, \frac{1}{2\mu}(1-e^{-2\mu t})\right)} \qquad \leq \min(z_1, y_2)$$

Hence, the conditional probability $P(\beta_1(t) \leq \beta_2(t) \,\Big|\, \beta_2(t) \leq 0)$ is:

$$\begin{aligned} P(\beta_1(t) < \beta_2(t) \,\Big|\, \beta_2(t) \leq 0) &= P\Big(N \leq \min(z_1, y_2)\Big) \\ &= \Phi\left(\frac{\min(z_1, y_2)}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) \end{aligned}$$

We substitute all of these values back in (3.29) to get the final P_{norm} value.

Formula for Normal Curves P_{norm}

Thus, the probability of obtaining normal curves P_{norm} is as follows:

$$\begin{aligned} P_{\text{norm}} &= P\Big(\beta_1(t) \leq -|\beta_2(t)|\Big) \\ &= P(\beta_1(t) \leq -\beta_2(t) \,\Big|\, \beta_2(t) > 0) \cdot P(\beta_2(t) > 0) \\ &\quad + P(\beta_1(t) \leq \beta_2(t) \,\Big|\, \beta_2(t) \leq 0) \cdot P(\beta_2(t) \leq 0) \\ &= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \cdot \Phi\left(\frac{\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \\ &+ \Phi\left(\frac{\min(z_1, y_2)}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) \cdot \Phi\left(\frac{-\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right) \end{aligned} \tag{3.30}$$

where, as stated previously,

$$\begin{aligned} z_1 &= \frac{-1}{b_2}\beta_2(0)e^{-\mu t} \\ y_2 &= \frac{-1}{b_1 - b_2}(\beta_1(0) - \beta_2(0))e^{-\mu t} \end{aligned}$$

In this method, our assumption of $a_1(u) = a_2(u) \equiv 0$ does not restrict our solution, as the general case of $a_1(u) \neq a_2(u) \neq 0$ can be dealt with in a similar way. Their impact will simply be the addition of some extra terms in the numerators of the cumulative distribution function Φ in (3.30).

Since (3.30) involves the parameter choices of b_1, b_2, $\beta_1(0)$, $\beta_2(0)$, and μ, the presentation of a closed form solution is quite cumbersome and offers no extra interpretation than the one already made in the paragraphs of this section. To compute exact probabilities, the reader can simply substitute their chosen parameters (provided they are appropriate and meet the conditions of our model in Theorem 3.3.1 and Section 3.3.2) and obtain the respective probability from the standard normal distribution table.

We can now present the respective probabilities of the other three shapes for this case, retaining our assumption of $a_1(t) = a_2(t) \equiv 0$ for simplicity. The proofs for these three probabilities, namely, P_{inv}, P_{hump}, P_{dip}, are presented in Appendix Chapter A.

Formula for Inverse Curves P_{inv}

$$\begin{aligned}
P_{\text{inv}} &= P(\beta_1(t) \geq |\beta_2(t)|) \\
&= P\left(\beta_1(t) > \beta_2(t) \,\middle|\, \beta_2(t) > 0\right) \cdot P(\beta_2(t) > 0) \\
&\quad + P\left(\beta_1(t) > -\beta_2(t) \,\middle|\, \beta_2(t) < 0\right) \cdot P(\beta_2(t) < 0) \\
&= \left(1 - \Phi\left(\frac{\max(z_1, y_2)}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)\right) \cdot \Phi\left(\frac{\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1 - e^{-2\mu t})}}\right) \\
&\quad + \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0) + \beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)\right]\right) \\
&\qquad \cdot \Phi\left(\frac{-\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1 - e^{-2\mu t})}}\right)
\end{aligned} \tag{3.31}$$

where,

$$\begin{aligned}
z_1 &= \frac{-1}{b_2}\beta_2(0)e^{-\mu t} \\
y_2 &= \frac{-1}{b_1 - b_2}(\beta_1(0) - \beta_2(0))e^{-\mu t}
\end{aligned}$$

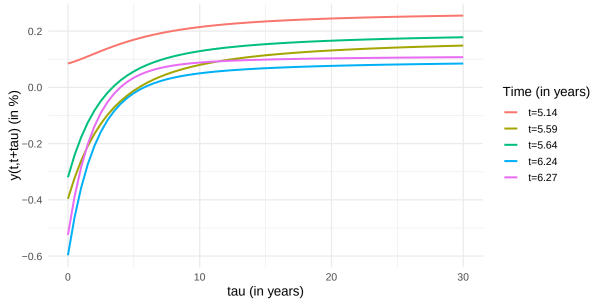
Formula for Humped Curves P_{hump}

$$\begin{aligned}
P_{\text{hump}} &= P(|\beta_1(t)| \leq \beta_2(t)) \\
&= P\left(\beta_1(t) < \beta_2(t) \,\middle|\, \beta_1(t) > 0\right) \cdot P(\beta_1(t) > 0) \\
&\quad + P\left(-\beta_1(t) < \beta_2(t) \,\middle|\, \beta_1(t) < 0\right) \cdot P(\beta_1(t) < 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1-b_2}(\beta_1(0) - \beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)\right]\right) \\
&\quad \cdot \Phi\left(\frac{\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1 - e^{-2\mu t})}}\right) \\
&\quad + \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0) + \beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)\right]\right) \cdot \Phi\left(\frac{-\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1 - e^{-2\mu t})}}\right)
\end{aligned} \tag{3.32}$$

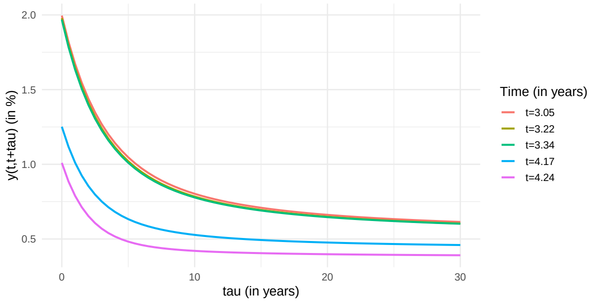
Formula for Dipped Curves P_{dip}

$$\begin{aligned}
P_{\text{dip}} &= P(-|\beta_1(t)| \geq \beta_2(t)) \\
&= P\left(-\beta_1(t) > \beta_2(t) \,\middle|\, \beta_1(t) > 0\right) \cdot P(\beta_1(t) > 0) \\
&\quad + P\left(\beta_1(t) > \beta_2(t) \,\middle|\, \beta_1(t) < 0\right) \cdot P(\beta_1(t) < 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \\
&\quad \cdot \Phi\left(\frac{\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right) \\
&\quad + \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{(b_1-b_2)}(\beta_1(0)-\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \\
&\quad \cdot \Phi\left(\frac{-\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right) \qquad (3.33)
\end{aligned}$$

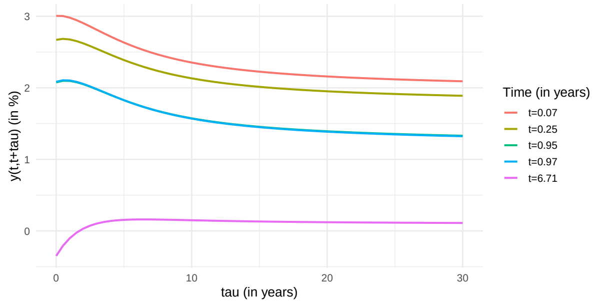
We present an example of how these probability conditions as explained in Remark 3.4.1 can be used to obtain instances of each of these four shapes. In each sub-figure of Figure 3.12, five random time instances at which that particular shape occurred are recorded and displayed. We choose the parameters as: $\lambda = 0.5374, \mu = 0.5$, $b_0 = 0.001, b_1 = 0.002, b_2 = 0.003$, $a_0 = 0.02, a_1(t) = 0.02\cos(t), a_2(t) = 0.02\sin(t)$. The initial values are $\beta_0(0) = 0.02, \beta_1(0) = 0.01, \beta_2(0) = 0.01$. The same algorithm used for Figure 3.7 is used to evolve $y(t, t+\tau)$ across the time-periods $\tau \in [0, 30]$ with $d\tau = 0.5$, and $t \in [0, 7]$ with $dt = 1/365$.



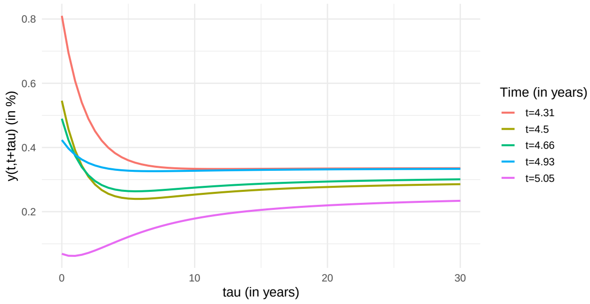
(a) Normal Curve



(b) Inverse Curve



(c) Humped Curve



(d) Dipped Curve

Figure 3.12: Single Brownian Motion: Generated Yield Curve Shapes

The choice of parameters play a crucial role in the exact probability of obtaining a particular curve shape. This coincides with our analysis of Equation (3.30).

3.4.2 Case: Multiple Brownian Motions

While the solutions in the case of a single Brownian motion are still comparatively obtainable, similar representations in the case of multiple Brownian motions (both uncorrelated and correlated) are much more complicated to obtain. Especially since the presence of several unknown parameters (such as $\beta_1(0), \beta_2(0), b_1, b_2$) and the correlation coefficients ρ_{ij} play a crucial role in determining the final probabilities.

However, for the sake of completeness, we present a suitable methodology to calculate the probabilities in such cases. Let us consider the case for uncorrelated Brownian motions.

Again for the ease of calculation, let $a_1(t) = a_2(t) \equiv 0$. We also know that $\beta_1(t)$ and $\beta_2(t)$ are normally distributed as:

$$\beta_1(t) \sim N\Big(e^{-\mu t}\beta_1(0), \quad \frac{b_1^2}{2\mu}(1-e^{-2\mu t})\Big)$$
$$\beta_2(t) \sim N\Big(e^{-\mu t}\beta_2(0), \quad \frac{b_2^2}{2\mu}(1-e^{-2\mu t})\Big)$$

which can be simplified to,

$$\beta_1(t) \sim N(\mu_1, \sigma_1^2)$$

$$\beta_2(t) \sim N(\mu_2, \sigma_2^2)$$

Now let $Z(t)$, $Y(t)$ be two independent standard normal random variables each with parameters $N(0,1)$. We can then write:

$$\begin{aligned}\beta_1(t) &= \mu_1(t) + \sigma_1(t)Z(t)\\ \beta_2(t) &= \mu_2(t) + \sigma_2(t)Y(t)\end{aligned}$$

For normal curves, we know from Remark 3.4.1 that

$$\begin{aligned}P_{\text{norm}} &= P\Big(\beta_1(t) \le -|\beta_2(t)|\Big)\\ &= P(\beta_1(t) \le -\beta_2(t) \;\Big|\; \beta_2(t) > 0) \cdot P(\beta_2(t) > 0)\\ &\quad + P(\beta_1(t) \le \beta_2(t) \;\Big|\; \beta_2(t) \le 0) \cdot P(\beta_2(t) \le 0)\end{aligned}$$

Expanding the above equation, for the first part $P(\beta_1(t) \le -\beta_2(t) \Big| \beta_2(t) > 0) \cdot P(\beta_2(t) > 0)$, we get

$$\begin{aligned}&P(\beta_1(t) \le -\beta_2(t) \;\Big|\; \beta_2(t) > 0)\\ &= P\Bigg(\mu_1(t) + \sigma_1(t)Z(t) + \mu_2(t) + \sigma_2(t)Y(t) < 0 \;\Bigg|\; \mu_2(t) + \sigma_2(t)Y(t) > 0\Bigg)\\ &= \frac{P(\mu_1(t) + \sigma_1(t)Z(t) + \mu_2(t) + \sigma_2(t)Y(t) < 0 \quad \cap \quad \mu_2(t) + \sigma_2(t)Y(t) > 0)}{P(\mu_2(t) + \sigma_2(t)Y(t) > 0)}\\ &\qquad\qquad \left(\because \text{ law of conditional probability } P(A|B) = \frac{P(A \cap B)}{P(B)}\right)\\ &= \frac{P\Bigg(Y > \frac{-\mu_2}{\sigma_2}, Z < \frac{-\mu_2 - \mu_1 - \sigma_2 Y}{\sigma_1}\Bigg)}{P\Bigg(Y > \frac{-\mu_2}{\sigma_2}\Bigg)}\end{aligned} \tag{3.34}$$

This is because

$$\begin{aligned}&\mu_1(t) + \sigma_1(t)Z(t) + \mu_2(t) + \sigma_2(t)Y(t) < 0\\ \implies &\sigma_1(t)Z(t) < -\mu_1(t) - \mu_2(t) - \sigma_2(t)Y(t)\\ \implies &Z(t) < \frac{-\mu_1(t) - \mu_2(t) - \sigma_2(t)Y(t)}{\sigma_1(t)}\end{aligned}$$

and,

$$\mu_2(t) + \sigma_2(t)Y > 0 \implies Y > \frac{-\mu_2(t)}{\sigma_2(t)}$$

Similarly for the second part of P_{norm}, $P(\beta_1(t) \le \beta_2(t) \Big| \beta_2(t) < 0) \cdot P(\beta_2(t) < 0)$, we get the following using the law of conditional probability:

$$
\begin{aligned}
&P(\beta_1(t) \leq \beta_2(t) \,\Big|\, \beta_2(t) < 0) \\
&= P\left(\mu_1(t) + \sigma_1(t)Z(t) - \mu_2(t) - \sigma_2(t)Y(t) < 0 \,\Big|\, \mu_2(t) + \sigma_2(t)Y(t) < 0\right) \\
&= \frac{P(\mu_1(t) + \sigma_1(t)Z(t) - \mu_2(t) - \sigma_2(t)Y(t) < 0 \quad \cap \quad \mu_2(t) + \sigma_2(t)Y(t) < 0}{P(\mu_2(t) + \sigma_2(t)Y(t) < 0)} \\
&= \frac{P\left(Y < \frac{-\mu_2}{\sigma_2}, Z < \frac{-\mu_1 + \mu_2 + \sigma_2 Y}{\sigma_1}\right)}{P\left(Y < \frac{-\mu_2}{\sigma_2}\right)}
\end{aligned} \tag{3.35}
$$

The property of independent random variables states that the joint probability density $f_{Z,Y}(z,y)$ of two independent random variables Z and Y is given as

$$f_{Z,Y}(z,y) = f_Z(z) f_Y(y)$$

where, $f(z)$ and $f(y)$ are the probability density functions of the standard normal random variables Z and Y respectively.

Furthermore, Korn (2014, Lemma 4.2, p. 196) state that if there are two independent normal random variables defined as follows:

$$
\begin{aligned}
X_1 &\sim N(\mu, \sigma^2) \\
X_2 &\sim N(0, 1)
\end{aligned}
$$

then, the following holds:

$$\int_{\tilde{x_1}}^{\infty} f_{\mu,\sigma^2}(x_1) \cdot \Phi(\alpha x_1 + \beta)\, dx_1 = P(X_1 \geq \tilde{x_1}, X_2 \leq \alpha X_1 + \beta)$$

where, $\Phi(\cdot)$ is the cumulative distribution function.

Hence, we can use this result to simplify the terms of (3.34) and (3.35) as follows:

$$P\left(Y > \frac{-\mu_2}{\sigma_2}, Z < \frac{-\mu_2 - \mu_1 - \sigma_2 Y}{\sigma_1}\right) = \int_{\frac{-\mu_2}{\sigma_2}}^{\infty} f_{0,1}(y) \cdot \Phi\left(\frac{-\sigma_2}{\sigma_1} y + \frac{-\mu_2 - \mu_1}{\sigma_1}\right) dy$$

and

$$P\left(Y \leq \frac{-\mu_2}{\sigma_2}, Z < \frac{\mu_2 - \mu_1 + \sigma_2 Y}{\sigma_1}\right) = \int_{-\infty}^{\frac{-\mu_2}{\sigma_2}} f_{0,1}(y) \cdot \Phi\left(\frac{\sigma_2}{\sigma_1} y + \frac{\mu_2 - \mu_1}{\sigma_1}\right) dy$$

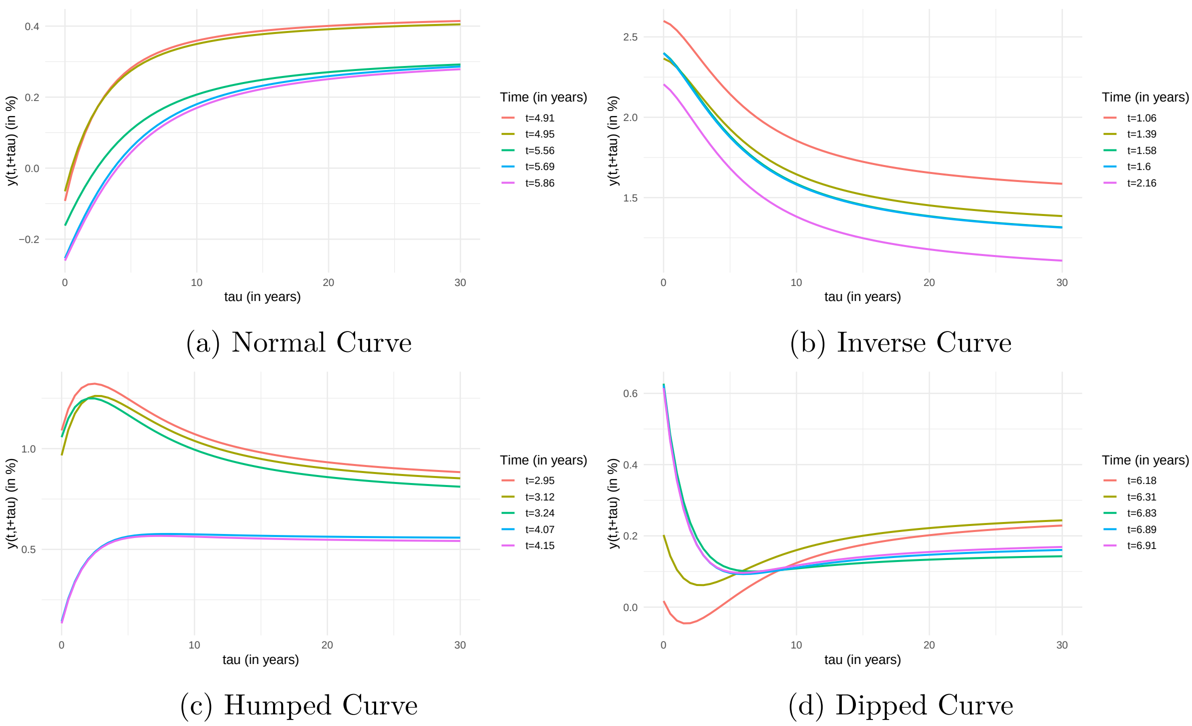
As we can observe, the above cannot be further explicitly integrated without any further information. However, one can substitute appropriate parameters values and employ intricate numerical approximation methods to calculate the required probabilities.

Furthermore, in the case of correlated Brownian motions, the integration limits become more challenging. With these justifications, we have decided to only present the aforementioned methodology and calculation steps required to solve such probabilities in this chapter. Since the methodology of obtaining the representations of P_{inv}, P_{hump}, and P_{dip} remains the same as P_{norm} above, our interpretations on the challenging nature of such computations also do not differ. As always, such calculations can be approximated using suitable numerical methods.

In the following figures, we present some graphical visualizations of these different yield curve shapes attainable by our Case 2: Uncorrelated and Case 3: Correlated models (as described in Section 3.3.4 and Section 3.3.5, respectively).

We choose the parameters as: $\lambda = 0.5374, \mu = 0.25, b_0 = 0.001, b_1 = 0.002, b_2 = 0.003, a_0 = 0.02, a_1(t) = 0.02\cos(t), a_2(t) = 0.02\sin(t)$. The initial values are $\beta_0(0) = 0.02, \beta_1(0) = 0.01, \beta_2(0) = 0.01$. For the correlated case, we choose the pairwise correlation coefficients as: $\rho_{01} = 0.4, \rho_{02} = 0.4, \rho_{12} = 0.4$. The same algorithm used for Figure 3.9 and Figure 3.11 is used to evolve $y(t, t+\tau)$ across the time-periods $\tau \in [0, 30]$ with $d\tau = 0.5$, and $t \in [0, 7]$ with $dt = 1/365$.

As we observe in both Figure 3.13 and Figure 3.14, all four possible shapes are attainable by the model specifications. Furthermore, the choice of parameters impact their magnitudes as well as their probabilities (as we have already concluded from this section).



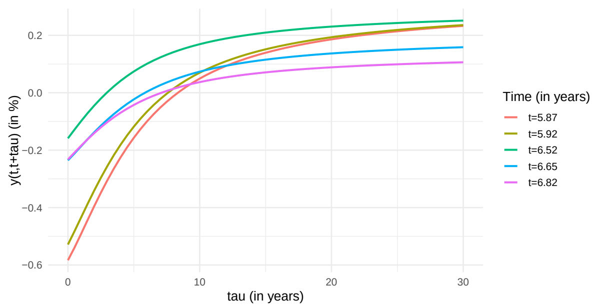
(a) Normal Curve

(b) Inverse Curve

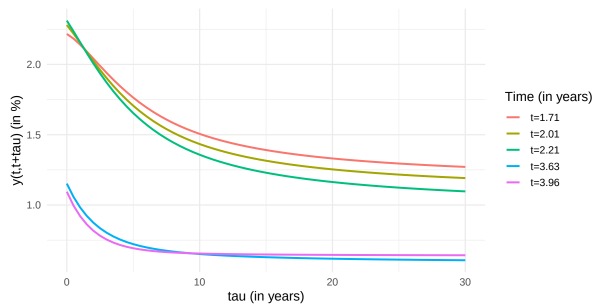
(c) Humped Curve

(d) Dipped Curve

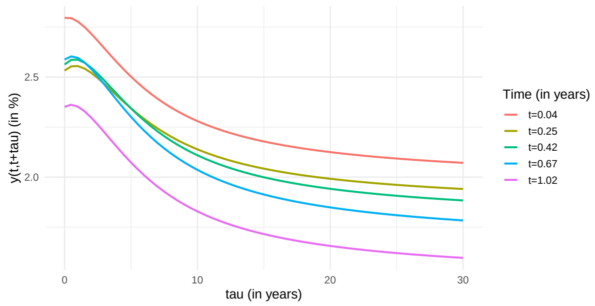
Figure 3.13: Uncorrelated Brownian Motions: Generated Yield Curve Shapes



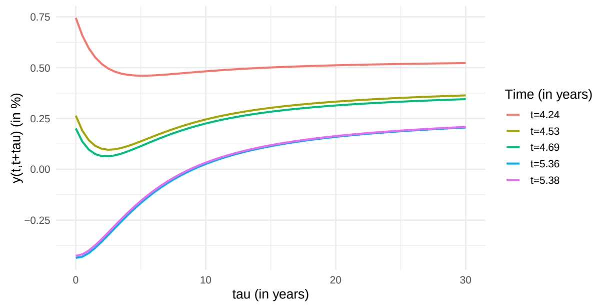
(a) Normal Curve



(b) Inverse Curve



(c) Humped Curve



(d) Dipped Curve

Figure 3.14: Correlated Brownian Motions: Generated Yield Curve Shapes

3.5 Concluding Remarks

In this chapter, we have analyzed the evolution of the dynamic Nelson-Siegel yield curve under a stochastic framework. Additionally, we have also used elementary mathematical concepts to provide a concrete proof of all possible shapes attainable by the original dynamic Nelson-Siegel model. This proof allowed us to determine exact conditions placed on $\beta_1(t)$ and $\beta_2(t)$ in order to obtain these four attainable shapes. On combining this contribution to our three different stochastic DNS models, we calculated and presented the methodology behind obtaining the probabilities of these different yield curve shapes. We make the following remark summarizing key inferences from our models:

Remark 3.5.1. By considering three different models of stochastic dynamics of the $\beta_0(t)$, $\beta_1(t)$, and $\beta_2(t)$, we observe the following:

1. **Comprehensive Specifications of the Three $\beta(t)$ Factors** The baseline stochastic model framework described in Section 3.3.2 presents a well-structured model with underlying Ornstein-Uhlenbeck processes driving the three $\beta(t)$ factors. Our framework can reproduce the mean-reverting tendencies often exhibited by yield curves in practice, and enable a parallel shift of the yield curve. Our specifications allow these $\beta(t)$ factors to retain their interpretations as the level, slope, and curvature of the yield curve, as originally stated in Diebold and Li (2006). Furthermore, this also allows for analytically tractable bond price processes that can be further utilized to answer pertinent research problems in the field of portfolio optimization. By considering three different models of the underlying Brownian

motions, we can provide a comprehensive framework that is consistent with the original DNS model.

2. **Capturing Realistic Yields.** Each of the three stochastic models is able to capture negative yields as well as all four different possible yield curve shapes. These features are often observed in the real-world bank data. The models for the single Brownian motion and uncorrelated Brownian motions are more consistent with regards to well-established empirical facts of the yield curve that are also cited in Diebold and Li (2006). In particular, both of these models allow for long-term rates to have higher average mean and to be less volatile than short-term rates, and for the $\beta_0(t)$ factor to be have the least variation with respect to its mean, thereby implying it is the most persistent of all three factors. We also observed how the presence correlation between the three stochastic factors does not serve as an ideal assumption for the yield curve framework. This behaviour was also explored by Diebold and Li (2006), wherein they asserted through empirical and statistical observations that the three $\beta(t)$ factors of the DNS model did not exhibit high pairwise correlations between each other, and thereby displayed little cross-factor interaction.

3. **Formulae for Probabilities of Yield Curve Shapes.** We present explicit representations for obtaining the probability of normal, inverse, humped, and dipped curves in the case of a single Brownian motion. For the case of multiple Brownian motions, we explain the methodology required to calculate these probabilities. Although exact computations are challenging due to the presence of multiple Brownian motions and several unknown parameters, our analysis shows that it is possible to obtain general framework for the corresponding probabilities and to subsequently calculate them using numerical methods.

Thus, we conclude this chapter on stochastic dynamic Nelson-Siegel models, and postulate that it offers significant research use-cases for practitioners and academicians alike.

Chapter 4

Optimal Portfolios with Stochastic Dynamic Nelson-Siegel

In this chapter, we consider the stochastic Nelson-Siegel framework described in Chapter 3 to solve the continuous-time portfolio selection problem studied by Merton (1969). Due to the presence of stochastic interest rates, we must employ the stochastic control methodology described by Korn and Kraft (2002). In comparison with the single-factor interest rate models used in their approach, the presence of the multi-factor stochastic DNS model requires a different corresponding verification theorem. We must construct a suitable step that ensures that the obtained optimal positions and the corresponding final wealth process satisfy a suitable Lipschitz condition. By doing so, we observe how the presence of a multi-factor stochastic yield curve impacts the final optimal positions of assets in a portfolio.

As an excursion, we discuss whether the lack of arbitrage-freeness in our stochastic Nelson-Siegel model impacts such portfolio problems significantly. We introduce the reason why Filipović (1999) considers the Nelson-Siegel family to never satisfy the no-arbitrage condition. Then, we consider the class of affine arbitrage-free Nelson-Siegel models (AFNS) developed by Christensen et al. (2011) and use them to solve the same portfolio problems. Our objective is to determine whether the optimal results obtained from these models exhibit behaviours that are different from our obtained optimal results. We find that on taking equal parameters, although there exists a constant time-independent difference in the optimal positions, the general representation and behaviour of the optimal portfolios are the same under both frameworks. Moreover, we also prove that the optimal positions obtained from our stochastic models that do not satisfy the so-called arbitrage-freeness have finite upper bounds.

The chapter is structured as follows: First, in Section 4.1, we construct a suitable theoretical foundation for our stochastic DNS framework that blends important results from stochastic control theory (see Fleming and Rishel (2012), Krylov (1980), Gikhman and Skorokhod (2004)), with their application to portfolio optimization problems (see Merton (1969) and Korn and Kraft (2002)). Then, in Section 4.2, we perform an application of this theory, and explicitly obtain valid and optimal solutions under the stochastic DNS framework driven by a single Brownian motion. Lastly, in Section 4.3 we present a comprehensive comparison of our results under the stochastic DNS framework driven by multiple Brownian motions, with those obtained under the corresponding AFNS framework. We conclude the chapter with significant remarks on our key findings.

Our main contributions in this chapter are:

1. The extension of the stochastic control methodology of Korn and Kraft (2002) to the stochastic dynamic Nelson-Siegel framework.
2. The formulation of an appropriate verification theorem for controlled diffusion processes represented by the stochastic DNS framework. This establishes the theoretical existence and uniqueness of such optimal solutions.
3. A justification of the validity of optimal portfolios under the stochastic DNS framework despite its theoretical lack of so-called consistent bond prices.

4.1 The Stochastic Control Approach

In finance, portfolio optimization has been one of the most significant areas of research, with each new development tackling an interesting subsection of the problem. With this thesis chapter, we aim to make a meaningful contribution to this area, both theoretically (with Section 4.1) and application-wise (with Section 4.2 and Section 4.3).

In particular, the continuous-time portfolio optimization problems solvable by the stochastic control methodology is of great interest to us. It was pioneered by Merton in his works to calculate the optimal investment strategy of an investor (see Merton (1969) and Merton (1975)). By modeling the asset prices as stochastic differential equations, Merton could solve the resulting control problem. However, a major drawback of his approach was the use of deterministic interest rates in the problem. Korn and Kraft (2002) overcome this restriction by incorporating stochastic interest rates into the continuous-time problem, thereby solving the optimization problem using stochastic control methodology. Their suggestion to extend their research to include other popular yield models motivates the application of stochastic DNS models in similar portfolio problems.

The martingale method is another major approach that is often used to solve continuous-time portfolio problems (see Karatzas et al. (1987)). However, since this approach is heavily reliant on the assumption of complete markets, we will exclude this methodology from our discussions.

This does not prove to be a drawback, as we will observe in this chapter how the underlying dynamic programming principle of the stochastic control approach allows optimization problems in such "incomplete" markets to be solved with adequate existence and uniqueness theorems. The ability to obtain optimal solutions using this approach accounts for real-world scenarios where the dynamic Nelson-Siegel model is arguably the most prominent underlying yield curve model.

Let us formulate the portfolio problem as well as explain the underlying theory to solve the problem. This will build the foundation necessary to solve such problems in the stochastic DNS setting and prove existence and uniqueness of the optimal solutions later in Section 4.2.

4.1.1 Formulation of the Bond Portfolio

In this chapter, by the phrase "continuous-time portfolio problem", we refer to a maximization problem of the total expected utility of the terminal wealth $X(T)$ of an investor in a trading time horizon $[0,T]$. The investor is allowed to make investment decisions continuous in time $t \in [0,T]$, and has an initial wealth of $X(0) = x_0 > 0$.

The utility of the terminal wealth of the investor will be measured by a suitable utility function. Korn (1997) defines a utility function as follows:

Definition 4.1.1 (Utility function (Korn (1997)))**.** A function $U : (0,\infty) \mapsto \mathbb{R}$ such that $U \in C^1$ is strictly concave and satisfies the following:

$$U'(0) = \lim_{x\to 0} U'(x) = +\infty$$
$$U'(\infty) = \lim_{x\to\infty} U'(x) = 0$$

is called a utility function.

For our problems in this chapter, we will consider the following utility function governing the investor's terminal wealth:

$$U(x) = \frac{1}{\gamma}x^{\gamma} \quad ,\gamma \in (0,1) \tag{4.1}$$

Although there are other utility functions (such as, the so-called hyperbolic absolute risk aversion (HARA) functions) often explored in similar literature, our choice allows for simplicity and ease of calculations (which will be useful to us when we solve for explicit solutions of the control problem). Moreover, it satisfies economic interpretations of utility as (4.1) implies that, when given the choice, the investor will always prefer a higher amount of wealth to a lower one. Additionally, (4.1) ensures that Alfred Marshall's law of diminishing marginal utility holds for the investor. This utility function is also comparable to the utility function of Korn and Kraft (2002) which will enable us to draw comparisons in later sections of this chapter.

Now that we have defined a suitable function that measures the utility that the investor obtains from his wealth, we can define exactly how the wealth is calculated. First, we introduce some assets that can be suitable in the portfolio of the investor in the stochastic DNS setting.

Definition 4.1.2 (Stochastic DNS Bond)**.** A discount bond, also known as a zero-coupon bond, maturing at time $T < \infty$ is called a stochastic DNS bond if its price $P(t,T)$ satisfies the following process at time t:

$$dP(t,T) = P(t,T)\Big(M(t)dt + \sigma(t)dW(t)\Big) \tag{4.2}$$

Here, the exact form of (4.2) depends on the type of underlying Brownian motions, as discussed in the previous chapter. For the case of a single Brownian motion, Theorem 3.3.1 applies, and the respective price process is followed. Similarly, for the case of uncorrelated Brownian motions, Theorem 3.3.3, and for the case of correlated Brownian motions, Theorem 3.3.4 applies. The drift process $M(t)$ takes the relevant form accordingly.

Remark 4.1.3 (Drift Functions of Stochastic DNS Bond)**.** It can be observed from corresponding equations of stochastic DNS bond prices under the three considered cases (*Case 1: single Brownian motion*,*Case 2: uncorrelated Brownian motions*, and *Case 3: correlated Brownian motions*), that the drift function $M(t)$ of (4.2) is of the following general form:

$$M_j(t) = \underbrace{f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)}_{\text{stochastic terms}} + \zeta_j(t) \tag{4.3}$$

where, $\zeta_j(t)$ is a deterministic function and $M_j(t)$ represents the drift function under j-th case of Brownian motions that influences the stochastic DNS framework, as described in Chapter 3. Namely,

1. $j = 1$ represents *Case 1: single Brownian motion* with the following drift function $M_1(t)$ taken directly from (3.18)

$$\begin{aligned} M_1(t) = {} & \beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big) \\ & + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ & - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ & + \frac{1}{2}(T-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2 \end{aligned}$$

2. $j = 2$ represents *Case 2: uncorrelated Brownian motions* with the following drift function $M_2(t)$ taken directly from (3.23)

$$\begin{aligned} M_2(t) = {} & \beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big) \\ & + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ & - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ & + \frac{1}{2}(T-t)^2(b_0^2 + g(t)^2b_1^2 + h(t)^2b_2^2) \end{aligned}$$

3. $j = 3$ represents *Case 3: correlated Brownian motions* with the following drift function $M_3(t)$ taken directly from (3.28)

$$\begin{aligned} M_3(t) = {} & \beta_0(t)(1 + \mu(T-t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda})\Big) \\ & + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))\Big) \\ & - a_0(T-t)\mu - a_1(t)(T-t)g(t)\mu - a_2(t)(T-t)h(t)\mu \\ & + \frac{1}{2}(T-t)^2\Big((b_0 + b_1g(t)\rho_{01} + h(t)b_2\rho_{02})^2 \\ & + (b_1g(t)\tilde{\rho}_{11} + b_2h(t)\tilde{\rho}_{12})^2 + (b_2h(t)\tilde{\rho}_{22})^2\Big) \end{aligned}$$

As observable, each of the three cases satisfies (4.3), with the stochastic terms being the same in all cases, and the corresponding deterministic functions $\zeta_j(t)$ taking a form relevant to the considered case. Moreover, we can collect the terms in the f functions as follows:

$$
\begin{aligned}
f_0(t) &:= 1 + \mu(T-t) \\
f_1(t) &:= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda}) \\
f_2(t) &:= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t))
\end{aligned} \tag{4.4}
$$

Definition 4.1.4 (Benchmark DNS Asset)**.** A benchmark DNS asset $B(t)$, or, simply, a benchmark asset, is an asset whose price process under the stochastic DNS framework is given by the following dynamics:

$$dB(t) = B(t)M_r(t)dt \tag{4.5}$$

where,

$$M_r(t) := M_j(t) - \zeta_j(t) \tag{4.6}$$

Here, j corresponds to the particular case of the stochastic DNS framework, as described in Remark 4.1.3.

Remark 4.1.5. Since $M_j(t)$ represents the drift of the stochastic DNS bond in the j-th case of the underlying Brownian motions (Case 1: single, Case 2: uncorrelated, and Case 3: correlated), as represented in Remark 4.1.3, we can observe that

$$M_r(t) = M_j(t) - \zeta_j(t) = f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t) \tag{4.7}$$

Hence,

$$dB(t) = B(t)\Big(f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)\Big)dt$$

Due to the functional forms of $f_0(t), f_1(t), f_2(t)$ from (4.4), it appears that the benchmark asset $B(t)$ is dependent on the time to maturity $T - t =: \tau$ of the stochastic DNS bond $P(t, T)$ (or, $P(t, t + \tau)$).

In the following paragraphs, we present a mild justification of the existence of such a benchmark asset in a market where the underlying yield curve is the dynamic Nelson-Siegel family.

The necessity of a benchmark asset $B(t)$, as defined in Definition 4.1.4, in our stochastic DNS framework is inspired by the treatment of the so-called rolling horizon bonds of Andersson and Lagerås (2013). Andersson and Lagerås (2013) show how an investment strategy where bonds of different times to maturity are repeatedly sold and bought such that the overall time to maturity of the strategy is fixed, can be used as a mathematical tool that simplifies the analysis of bond portfolios. They state that such an artificial portfolio of bonds which are continuously rolled over, also known as rolling horizon bonds, can be defined provided that an underlying affine term structure model exists. In particular, term structure models which are driven by Ornstein–Uhlenbeck (OU) processes and are represented as linear functions of these processes can be used to construct corresponding rolling horizon bonds.

The strategy adopted by Andersson and Lagerås (2013) is three-fold. First, they consider an affine forward rate curve to be represented as

$$F(t,t+\tau) := \kappa(\tau)'\mathbf{F}(t) \tag{4.8}$$

which is equivalent to assuming an affine relationship for the spot rates (yield curve) as

$$y(t,t+\tau) = \frac{1}{\tau}\overline{\kappa}(\tau)'\mathbf{F}(t)$$

where, $\overline{\kappa}(\tau) := \int_0^\tau \kappa(u)du$. Here, $\mathbf{F}(t)$ is the set of factors in the yield curve model, and $\kappa(t)$ is the corresponding set of functions determined by the dynamics of $\mathbf{F}(t)$.

Andersson and Lagerås (2013) showed that an example of such an affine forward curve is the dynamic Nelson-Siegel forward rate curve, represented as follows:

$$F(t,t+\tau) = \beta_0(t) + e^{-\lambda\tau}\beta_1(t) + \lambda\tau e^{-\lambda\tau}\beta_2(t)$$

Since we know that the underlying factors of the DNS model are $\beta_0(t), \beta_1(t), \beta_2(t)$, a correspondence with (4.8) implies,

$$\mathbf{F}(t) = (\beta_0(t), \beta_1(t), \beta_2(t))'$$

A significant remark here is that Andersson and Lagerås (2013) do not restrict their analysis to arbitrage-free term structure models. Instead, they simply state that for arbitrage-free models, the $\kappa(\tau)$ functions must be set appropriately under the risk-neutral measure. Furthermore, the authors justify that it is more fruitful to perform their analysis on the return and variance of portfolios containing rolling horizon bonds under the real-world measure $\mathbb{P}$. This also helps us to build a justification behind the use of a benchmark asset, as we have considered all of our yield curve and bond price dynamics under the $\mathbb{P}$-measure.

The next step after defining the forward curve as (4.8) in the strategy discussed by Andersson and Lagerås (2013) is to define a rolling horizon bond with the following representation:

$$R(t,t+\tau) := \exp\Big\{-\overline{\kappa}(\tau)'(\mathbf{F}(t) - \mathbf{F}(0)) + \kappa(\tau)'\overline{\mathbf{F}}(t)\Big\} \tag{4.9}$$

where, $\overline{\mathbf{F}}(t) := \int_0^t \mathbf{F}(s)ds$. This representation in their analysis is a direct consequence of (4.8) and the following representation of a rolling horizon bond given by Rutkowski (1999):

$$R(t,t+\tau) := \frac{P(t,t+\tau)}{P(0,\tau)}\exp\Big\{\int_0^\tau F(s,s+\tau)ds\Big\} \tag{4.10}$$

The final step in the analysis of Andersson and Lagerås (2013) is to construct a portfolio containing only rolling horizon bonds defined by (4.9) and thereby represent the return and variance of such portfolios using the cumulant generating functions of the underlying OU processes of the affine forward curve.

At the moment, for our definition of the benchmark asset $B(t)$, we do not require the final step. Hence, let us focus on the first two steps of their strategy and discuss the relevance to our benchmark asset representation.

We know that the value of the benchmark asset is

$$\begin{aligned} dB(t) &= B(t)M_r(t)dt \\ \text{or, } B(t) &= \exp\{f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)\} \end{aligned}$$

This is because from (4.7), we already know $M_r(t) = f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)$. Now from (4.4), we know that

$$\begin{aligned} f_0(t) &= 1 + \mu(T-t) \\ f_1(t) &:= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda}) \\ f_2(t) &:= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t)) \end{aligned} \tag{4.11}$$

As we know T is the maturity time of a stochastic DNS bond, and t is the current time, we can rewrite the above functions using the time to maturity $\tau := T - t$. Then,

$$\begin{aligned} f_0(\tau) &:= 1 + \mu\tau \\ f_1(\tau) &:= \frac{\mu}{\lambda} + e^{-\lambda\tau}(1 - \frac{\mu}{\lambda}) \\ f_2(\tau) &:= \frac{\mu}{\lambda} + e^{-\lambda\tau}(-\frac{\mu}{\lambda} - \mu\tau + \lambda\tau) \end{aligned} \tag{4.12}$$

Therefore, the value of the benchmark asset can be represented using τ as follows:

$$B(t) = B(t, t+\tau) = \exp\{f_0(\tau)\beta_0(t) + f_1(\tau)\beta_1(t) + f_2(\tau)\beta_2(t)\} \tag{4.13}$$

This trading strategy (4.13) resembles the rolling horizon bond of (4.9), with the following similarities:

$$\begin{aligned} \kappa(\tau) &= (f_0(\tau), f_1(\tau), f_2(\tau)) \\ \mathbf{F}(t) &= (\beta_0(t), \beta_1(t), \beta_2(t)) \end{aligned}$$

Hence, we justify our decision to include the benchmark asset $B(t)$ in a portfolio constructed under the stochastic DNS framework as follows: The benchmark asset $B(t)$ (or, $B(t, t+\tau)$) is an artificial trading strategy that has a fixed time to maturity τ, and offers a stable-horizon investment strategy. It is inspired by the conceptual framework of rolling horizon bonds of Andersson and Lagerås (2013), and can be thereby used by institutional investors as a hedging strategy against their liabilities. Furthermore, we can observe that for $\tau = 0$, the benchmark asset $B(t, t+\tau)$ behaves like a money market account which accumulates over the short rate.

In a way, we can consider the benchmark asset $B(t)$ to be a rolling horizon of short-term money market accounts (for example, US Treasury Bills). Although, it is an artificial asset (similar to the rolling horizon bonds of Rutkowski (1999) and of Andersson and Lagerås (2013)), it offers an economic advantage to practitioners, by offering a stable and relatively low-risk investment by allowing short-term assets (such as the money market account) to be continuously rebalanced for a fixed time to maturity.

Remark 4.1.6 (Economic Interpretation of the Benchmark Asset)**.** We interpret the benchmark asset $B(t)$, as defined in Definition 4.1.4, as a stable-horizon investment strategy, inspired by the treatment of rolling horizon bonds of Andersson and Lagerås (2013). This benchmark asset with its ability to continuously rollover short-term money market accounts can be regarded as a low-risk alternative to the stochastic DNS bond.

Remark 4.1.7. With Remark 4.1.6, we do not claim that the benchmark asset $B(t)$ is an exact replica of the rolling horizon bond defined by Andersson and Lagerås (2013). Rather, we imply that it is a relatively stable asset under the stochastic DNS framework that shares the economic intuition of a stable-horizon investment. Therefore, it is beneficial for enabling analytical tractability in the portfolio optimization problems in our stochastic DNS framework.

We have, thus, defined all kinds of assets that may be considered in our problem, thereby examining the sufficient conditions that ensures the admissibility of a portfolio within the stochastic DNS framework. We can now formally construct and represent such portfolios of the investor and the corresponding wealth process.

Definition 4.1.8 (Admissible Portfolio)**.** An investor is said to have an admissible portfolio if it consists of the following assets:

1. one benchmark asset $B(t)$ whose price process satisfies (4.5)
2. one stochastic DNS bond $P(t, T_i)$ whose price process satisfies (4.2) for some maturity $T_i < \infty$ for $i = 1, 2, \ldots$

Therefore, there are at least 2 assets in an admissible portfolio.

The portfolio process of the investor denotes the corresponding percentage of total wealth invested in each particular asset. Let the percentage of wealth invested in the stochastic DNS bond be denoted by $\pi(t)$, where $\pi(t)$ is a progressively measurable process with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. The percentage invested in the benchmark asset is then given by $1 - \pi$.

If the corresponding trading strategy of the portfolio is denoted by $\phi(t) = (\phi_1(t), \phi_2(t))$, then the wealth equation can be represented by the following:

$$X(t) = \phi_1(t) \cdot B(t) + \phi_2(t) \cdot P(t, T_i)$$

where, the respective trading strategies of the benchmark asset and of the bond asset are

$$\phi_1(t) := \frac{(1-\pi)X(t)}{B(t)}$$
$$\phi_2(t) := \frac{\pi(t)X(t)}{P(t, T_i)}$$

Here, we assume that the each maturity of the bond assets T_i is greater than the fixed investment horizon T.

Thus, the wealth dynamics of the investor is given by

$$dX(t) = X(t)\Big((\pi(t)(M_j(t) - M_r(t)) + M_r(t))dt + \pi(t)\sigma_j(t)dW(t)\Big) \quad , x_0 > 0 \qquad (4.14)$$

where, $M_j(t), \sigma_j(t)$ for $j = 1, 2, 3$ are the respective drift and volatility coefficients of the bond price processes represented by Cases 1,2,3 of Brownian motion, respectively, as described in Chapter 3. The exact specifications of (4.14) will depend on the chosen case of Brownian motions (*Case 1*, *Case 2*, or *Case 3*), and will be described in their respective sub-sections later in this chapter.

Remark 4.1.9. Our definition of a portfolio in Definition 4.1.8 also allows for the following extensions:

1. Other assets whose price processes differ from (4.2) can be additionally considered, assuming that they are also driven by a stochastic DNS framework established in the previous chapter.
2. Assume that an asset $S(t)$ is driven by both its own Brownian motion $W_s(t)$ along with the underlying Brownian motions of the stochastic DNS framework $W_j(t)$. Then, its price process can be of the form,

$$dS(t) = S(t)\Big(M_s(t)dt + \sigma_{sj}dW_j(t) + \sigma_s(t)dW_s(t)\Big)$$

   Here, the correlation between this new asset with the stochastic DNS bond must be taken into account. In Korn and Kraft (2002), such assets are called stocks.

With this setting, we may now proceed with the illustration of the control problem, and the adaptation of classical results of stochastic control required to solve the problem under the stochastic DNS framework.

4.1.2 Adaptation of Classical Results

In this chapter, the state of the system that we want to "control" consists of the following:

1. the wealth equation $X(t)$ and,
2. the stochastic DNS factors $\{\beta_0(t), \beta_1(t), \beta_2(t)\}$

This is because, in our problem setting, both the wealth equation and the underlying yield curve factors are stochastic. Classical results of stochastic control theory state that the solutions of a controlled process must always satisfy certain laws on existence and uniqueness. Therefore, an appropriate verification theorem must be levied on the obtained optimal solution.

In order to do so, we must first establish the theoretical foundation for our problem.

As we shall now deal with stochastic processes, let us recall the important concept of progressive measurability. From Karatzas and Shreve (2012, Definition 1.11), we define a progressively measurable function as follows:

Definition 4.1.10 (Progressively Measurable (Karatzas and Shreve (2012))). A stochastic process $\{X_t\}_{t\geq 0}$ is said to be progressively measurable with respect to the filtration $\mathcal{F}_T$ if for every $t \geq 0$, it is measurable relative to the product σ-algebra $\mathcal{B}_{[0,T]} \times \mathcal{F}_T$, when viewed as a function $X(t, \omega)$ on the product space $([0, T]) \times \Omega$.

Furthermore from Karatzas and Shreve (2012, Proposition 1.13), if $\{X_t\}_{t\in[0,T]}$ is an $\{\mathcal{F}_t\}$-adapted process with only left or right continuous paths, then it is progressively measurable.

Additionally, if $\int_{t\in[0,T]} \mathbb{E}|X_t|^\rho dt < \infty$ for some $\rho > 1$, then using Lalley (2012, Lemma 1)

$$Y_t := \int_{s\leq t} X_s ds$$

is continuous in $t \in [0, T]$ and progressively measurable.

Let the controlled diffusion process of the state of the system be represented as follows:

$$dY(t) = b(t, Y(t), u(t))dt + \sigma(t, Y(t), u(t))dW(t) \tag{4.15}$$

where, $Y(t)$ is the state of the system at time t, and $u(t)$ is a stochastic process that can be "controlled" by the investor to yield optimal results. The finite-dimensional Brownian motion is represented by $W(t)$.

Before we define the existence of the control process $u(t)$, let us recall how the existence of the solution to a stochastic diffusion process $Y(t)$ is proven. We take the help of Fleming and Rishel (2012, pp. 117 ff).

Let there be a stochastic integral equation as follows:

$$Y(t) = Y(0) + \int_0^t b(s, Y(s))ds + \int_0^t \sigma(s, Y(s))dW(s) \quad \text{, for } 0 \leq t \leq T_F \tag{4.16}$$

where, b and σ are Borel-measurable on $Q_0 := [T_0, T_F] \times \mathbb{R}^n$. Here, we take the initial time to be 0. It can also be taken as some t_0 with $T_0 \leq t_0 < T_F$.

For some constants C and K, if b and σ satisfy the following for all $(t, y) \in Q_0$:

$$\begin{aligned} |b(t, y)| &\leq C(1 + |y|) \\ |\sigma(t, y)| &\leq C(1 + |y|) \end{aligned} \tag{4.17}$$

and for any bounded $B \in \mathbb{R}^n$ and $T_0 \leq T' \leq T_F$, there exists a constant K which may depend on B and T' such that

$$\begin{aligned} |b(t, y) - b(t, y_1)| &\leq K|y - y_1| \\ |\sigma(t, y) - \sigma(t, y_1)| &\leq K|y - y_1| \end{aligned} \tag{4.18}$$

and if $Y(0)$ is independent of the Brownian motion $W(t)$ with $\mathbb{E}|Y(0)|^2 < \infty$, then a solution $Y(t)$ of (4.16) exists. This solution is unique in the sense that if $Y(t)'$ is also a solution of (4.16) with $Y(0)' = Y(0)$, then $Y(t)' = Y(t)$ with probability 1.

The existence and uniqueness of solutions to (4.16) naturally also hold under much stronger assumptions than (4.17) and (4.18). We present them here for completeness. In particular, (4.16) has a unique solution if $b(t, Y(t))$ and $\sigma(t, Y(t))$ satisfy the so-called Itô-conditions, as defined below:

Definition 4.1.11 (Itô conditions (Fleming and Rishel (2012)))**.** If the functions b and σ defined in (4.16) satisfy the following:

1. $b(t, Y(t))$ and $\sigma(t, X(t))$ are of class C^1 on Q_0, and
2. the first order partial derivatives of $b(t, Y(t))$ and $\sigma(t, Y(t))$ in the variables $Y(t) = (Y_i(t), \ldots, Y_n(t))$ are bounded

then, they are said to satisfy the so-called Itô conditions.

If the Ito conditions hold, then both b and σ satisfy (4.17) and (4.18) for a constant K not depending on B or T'.

The next step in our theoretical foundation would be to establish some classical results on the moments of these stochastic processes.

Proposition 4.1.12 (Estimates on moments)**.** Let us consider a stochastic process $Y(t)$ represented by (4.16). Then, we can state the following:

1. Let $b(t, Y(t))$ and $\sigma(t, Y(t))$ satisfy (4.18) and let the following hold:

$$\mathbb{E}\int_0^{T_F}\left(|Y(0)|^2 + |b(s,0)|^2 + \|\sigma(s,0)\|^2\right)ds < \infty$$

   Then, for any $t \leq T_F$, (4.16) has a unique solution with

$$\mathbb{E}\int_0^{T_F}|Y(s)|^2 ds < \infty$$

2. If we assume that the initial data $Y(0)$ has finite absolute moments of every order

$$\mathbb{E}|Y(0)|^k < \infty \quad , k = 1, 2, \ldots$$

   then, both the Ito conditions as well as the much weaker assumptions, (4.17) and (4.18), imply that for any $t_0 \leq t < T_F$

$$\mathbb{E}|Y(t)|^2 < \infty$$

   and

$$\mathbb{E}\int_{t_0}^{T_F}|b(s, Y(s))|ds < \infty$$
$$\mathbb{E}\int_{t_0}^{T_F}|\sigma(s, Y(s))|^2 ds < \infty$$

Proof. These assertions follow from results from classical literature on stochastic control theory.

1. For the first assertion, the proof is a direct result of Theorem 7 of Krylov (1980, p. 82) who state that if condition (4.18) is satisfied, and if $\mathbb{E}\int_0^T(|Y(0)|^2 + |b(t,0)|^2 + \|\sigma(t,0)\|^2)dt < \infty$, then for any $t \leq T$, (4.16) has a unique solution such that

$$\mathbb{E}\int_0^T |Y(t)|^2 dt < \infty$$

Moreover, if $Y_1(t)$ and $Y_2(t)$ are two solutions of (4.16), then

$$P\{\sup_{t\in[0,T]} |Y_1(t) - Y_2(t)| > 0\} = 0$$

2. For the second assertion, the proof is a direct result of Gikhman and Skorokhod (2004) who state that for each $k = 1, 2, \ldots$,

$$\mathbb{E}|Y(t)|^k \leq C_k(1 + \mathbb{E}|Y(0)|^k) \quad , t_0 \leq t \leq T_1$$

where, the constant C_k depends on k, $T_F - t_0$, and C defined in (4.17).

Hence, given the afore-stated assumptions, our assertions on the finiteness of the expectations of the solution to a stochastic integral equation are proved. For the interested reader, the proofs of the two original theorems are provided in Appendix Chapter B. □

Now that we have established important results of existence and uniqueness of solutions to a stochastic process represented by (4.16), we can proceed with establishing significant results on the controlled diffusion process represented by (4.15).

Repeating the equation, we know that the finite-dimensional state process of a control problem is represented by

$$dY(t) = \Lambda(t, Y(t), u(t))dt + \Sigma(t, Y(t), u(t))dW(t) \tag{4.19}$$

with initial condition $Y(t_0) = y_0$ for any $t_0 \leq t$. Here, $Y(t)$ is the state of the system to be controlled, $u(t)$ is the control applied at time t, and $W(t)$ is a d-dimensional Brownian motion. Let U be a set of all controls that can be applied at time t.

From Fleming and Rishel (2012, p. 156f), we can formulate the following definition of an admissible control:

Definition 4.1.13 (Admissible control (Fleming and Rishel (2012))). A control $u \in U$ is admissible if u is a Borel measurable function from $Q_0 := [T_0, T_F] \times \mathbb{R}^n \to U$ such that:

S.1 For each $(t_0, Y(t_0))$, $T_0 \leq t_0 < T_F$, there exists a Brownian motion $W(t)$ such that (4.19) with initial condition $Y(t_0) = y_0$ has a unique solution $Y(t)$ and,

S.2 for $k = 1, 2, \ldots$,

$$\mathbb{E}_{(t_0,y_0)}|Y(t)|^k < \infty \quad , \text{for } t_0 \leq t \leq T_F$$

and

$$\mathbb{E}_{(t_0,y_0)}\int_{t_0}^{T_F} |u(t)|^k dt < \infty$$

Here, $u(t) = u(t, Y(t))$. The bound in **S.2** may depend on (t_0, y_0).

Let the class of all admissible control u be denoted by $\mathcal{V}$. Any of the following three conditions is sufficient for the admissibility of $u \in \mathcal{V}$:

C.1 For some constant K_1,

$$u(t, y) \leq K_1(1 + |y|) \quad \forall (t, y) \in Q_0$$

Moreover, for any bounded $B \subset \mathbb{R}^n$ and $T_0 < t' < T_F$, there exists a constant K_2 such that

$$|u(t, y) - u(t, y_0)| \leq K_2 |y - y_0| \quad \forall y, y_0 \in B, \quad T_0 \leq t \leq t'$$

Here, K_2 may depend on B and t', and both K_1, K_2 may depend on u.

C.2 u satisfies a Lipschitz condition on Q_0.

C.3 Suppose that $\Lambda(t, y, u) = b(t, y) + \Sigma(t, y)\theta(t, y, u)$ where, b, Σ satisfy the Ito conditions and Σ is bounded, θ is of class $C^1(Q_0 \times U)$, and θ is bounded on $Q_0 \times U_1$ for a bounded $U_1 \subset U$. Then, any bounded Borel-measurable u is admissible.

If any of these conditions hold, then Λ^u and Σ^u satisfy hypotheses (4.17) and (4.18), thereby proving that the solution of (4.19) exists and is unique.

Now, for the class of admissible controls $\mathcal{V}$, we construct the maximization problem as follows:

Let L and Ψ be two continuous functions that satisfy the given polynomial growth conditions,

$$\begin{aligned} |L(t, y, u)| &\leq C(1 + |y| + |v|)^k \\ |\Psi(y)| &\leq C(1 + |y|)^k \end{aligned} \tag{4.20}$$

for some suitable constants C and k.

Our optimization problem is to now find a control u^* which maximizes the following:

$$J(t_0, y_0, u) := \mathbb{E}_{(t_0, y_0)} \left(\int_{t_0}^{\tau} L(t, Y(t), u(t))dt + \Psi(\tau, Y(\tau)) \right) \tag{4.21}$$

for $u \in \mathcal{V}$. Here, τ is the exit time from a given open set $Q \subset Q_0$ and the initial data $(t_0, y_0) \in Q$. If $Q = Q_0$, then $\tau = T_F$ and the respective function to be maximized becomes

$$J(t_0, y_0, u) := \mathbb{E}_{(t_0, y_0)} \left(\int_{t_0}^{T_F} L(t, Y(t), u(t))dt + \Psi(Y(T_F)) \right) \tag{4.22}$$

In our problem, we will consider the exit time of the process to be the final investment horizon (since the investor is allowed to hold the portfolio until the end of the investment horizon). Hence, (4.22) will be the relevant equation to be maximized for $u \in \mathcal{V}$. Condition **S.2** and Equation (4.20) ensure that $J(t_0, y_0, u)$ is finite for each $u \in \mathcal{V}$.

For the state system $Y(t)$ defined by (4.19), consider the diffusion operator $A^v(t_0)$ defined as the following:

$$A^v(t_0) = \frac{1}{2} \sum_{i,j=1}^{n} (\Sigma\Sigma')_{ij}(t_0, y_0, v) \frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^{n} \Lambda_i(t_0, y_0, v) \frac{\partial}{\partial y_i}$$

With this knowledge, we may now arrive at our verification theorem.

4.1.3 Verification Theorem for Controlled Diffusion Processes in the Stochastic DNS Framework

We consider the problem where an investor wants to maximize his expected utility of terminal wealth in a fixed time horizon. The underlying yield curve model influencing the market is the stochastic dynamic Nelson-Siegel framework of Chapter 3. The cases of single Brownian motion, uncorrelated Brownian motions, and correlated Brownian motions are considered in this framework. As we have observed in the previous chapter, our stochastic framework consists of mean-reverting diffusion processes.

Let the portfolio problem of the investor be represented by (4.22). In order to ensure an admissible and optimal solution to this problem, the following verification theorem must hold.

Theorem 4.1.14 (Verification Theorem)**.** Let the state of the process to be controlled be represented as follows:

$$dY(t) = \Lambda(t, Y(t), u(t))dt + \Sigma(t, Y(t), u(t))dW(t) \tag{4.23}$$

For initial conditions $(t_0, y_0) \in Q \subset Q_0$, let G be a solution of the following Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \sup_{v \in U}[A^v G(t, y) + L(t, y, v)] \tag{4.24}$$

with

$$G(t, y) = \Psi(t, y) \tag{4.25}$$

where, $(t, y) \in \partial^* Q$ and $\partial^* Q$ is a closed subset of ∂Q such that $(T_F, Y(T_F)) \in \partial^* Q$ with probability 1. Here, G is assumed to be $C^{1,2}(Q)$. If G is $C^{1,2}$ on a non-closed set Q, its continuity is required on the closure of it.

Then, for any admissible control u,

$$G(t, y) \geq J(t, y, u) \tag{4.26}$$

Furthermore, if u^* is an admissible control such that

$$A^{u^*}(t_0)G + L^{u^*}(t_0, y_0) = \max_{v \in U}[A^v(t_0)G + L(t_0, y_0, v)] \tag{4.27}$$

Then, $G(t_0, y_0) = J(t_0, y_0, u^*)$ for all $(t_0, y_0) \in Q$. Thus, u^* is optimal.

Proof. This verification theorem is a result of the corresponding theorem in Fleming and Rishel (2012, p. 159) (see §4, Theorem 4.1). We present the original theorem in Appendix Chapter B for the interested reader. Let us now prove how it can be modified to suit our relevant controlled diffusion problem.

Consider $Q_0 := [T_0, T_F] \times \mathbb{R}^n$, where T_F denotes the final investment time horizon, and a set of all controls $v \in U$. Also, let the initial values $(t_0, y_0) \in Q \subset Q_0$, for $T_0 \leq t_0 < T_F$, and $y_0 := Y(t_0)$.

We know that if Λ and Σ satisfy either (4.17) or the stronger Ito conditions, then the solution of (4.23) exists and is unique. Further, assume that G is a $C^{1,2}(Q)$ and solves the HJB equation (4.24). Then, using Theorem 5.1 of Fleming and Rishel (2012, pp. 124), we get

$$G(t,y) \geq \mathbb{E}_{t_0,y_0}\Big(\int_{t_0}^{T_F} L^u(t,Y(t))dt + G(T_F, Y(T_F))\Big)$$

Furthermore, (**S.2**) and (4.20) imply

$$\mathbb{E}_(t_0, y_0)\int_{t_0}^{T_F} |L^u(t,Y(t))|dt < \infty$$

Hence,

$$G(t,y) \geq \mathbb{E}_(t_0, y_0)\left(\int_{t_0}^{T_F} L^u(t,Y(t))dt + \Psi(Y(T_F))\right)$$

From (4.22), we know that the RHS of the above equation has already been defined as $J(t,y,u)$. Hence, we have proved that for any admissible control u,

$$G(t,y) \geq J(t,y,u)$$

holds true. Furthermore, for an optimal control u^*, we have equality in (4.24) for $u = u^*$. Then,

$$G(t,y) = J(t,y,u^*)$$

Thus, (4.27) holds true and u^* is the optimal control for the problem. □

Now that we have established the theoretical structure that quantifies and verifies our problem and its subsequent solution, we may apply it to the bond portfolio problem under the stochastic DNS framework, as described in the following section.

4.2 Application in the Stochastic DNS Framework

In this section, let us first consider the bond portfolio problem for the case when a single Brownian motion affects all three $\beta(t)$ factors of the stochastic DNS model. We use the model introduced in Section 3.3.3. The other two cases of the stochastic DNS framework will be discussed later in Section 4.3, as it will supplement our discussion of arbitrage in the Nelson-Siegel family.

4.2.1 Optimal Portfolio: Single Brownian Motion

Let there be an investor who has the following assets in his admissible portfolio:

1. one stochastic DNS bond whose dynamics are defined by $dP(t, T_1)$, as stated in *Case 1: single Brownian motion* of the previous chapter (see Theorem 3.3.1)
2. one benchmark asset whose dynamics are given by $dB(t)$.

We assume that the maturity time of the bond T_1 is greater than the investment horizon time T. This assumption prevents the assets from expiring before the final investment date.

We know from Theorem 3.3.1 (see page 28 ff.), that the bond price process is given by

$$dP(t,T_1) = P(t,T_1)\Big(M_1(t)dt + \sigma(t)dW(t)\Big)$$

where,

$$\begin{aligned} M_1(t) &= f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t) + \zeta_1(t) \\ \sigma(t) &= -(T_1 - t)(b_0 + g(t)b_1 + h(t)b_2) \end{aligned}$$

From (4.4) and (4.3), we also know that

$$\begin{aligned} f_0(t) &\coloneqq 1 + \mu(T_1 - t) \\ f_1(t) &\coloneqq \frac{\mu}{\lambda} + e^{-\lambda(T_1-t)}(1 - \frac{\mu}{\lambda}) \\ f_2(t) &\coloneqq \frac{\mu}{\lambda} + e^{-\lambda(T_1-t)}(-\frac{\mu}{\lambda} - \mu(T_1-t) + \lambda(T_1-t)) \\ \zeta_1(t) &\coloneqq -a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu \\ &\quad + \frac{1}{2}(T_1-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2 \end{aligned}$$

Here, $W(t)$ is a 1-D Brownian motion, since only one Brownian motion drives all three $\beta(t)$ factors.

From (4.5), we know that the benchmark asset is given by

$$dB(t) = B(t)M_r(t)dt$$

with $M_r(t) \coloneqq M_1(t) - \zeta_1(t) = f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)$

We can use (4.14) to formally obtain the wealth process $X(t)$ corresponding to the trading strategy represented by $(\phi_{P(t,T_1)}, \phi_{B(t)})$. Let $\pi(t)$ be the percentage of wealth invested in the bond. The percentage of wealth invested in the benchmark asset is $1 - \pi(t)$. The wealth equation (4.14) becomes

$$\begin{aligned} dX(t) &= X(t)\Bigg[\Bigg(\pi(t)M_1(t) + (1-\pi(t))M_r(t)\Bigg)dt + \pi(t)\sigma(t)dW(t)\Bigg] \\ &= X(t)\Bigg[\Bigg(\pi(t)(M_1(t) - M_r(t)) + M_r(t)\Bigg)dt + \pi(t)\sigma(t)dW(t)\Bigg] \\ &= X(t)\Bigg[\Bigg(\pi(t)\zeta_1(t) + f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t)\Bigg)dt + \pi(t)\sigma(t)dW(t)\Bigg] \end{aligned} \tag{4.28}$$

with $X(0) = x_0 \neq 0$.

Comparison of the Obtained Wealth Process with Korn and Kraft (2002):

Before we formulate the corresponding state process and solve the problem using the stochastic control methodology of Korn and Kraft (2002), let us take a look at the crucial differences between our wealth process and price process dynamics in comparison to theirs. This will not only help us establish the motivation behind solving such a problem for the stochastic DNS framework but also help us see the need to apply Theorem 4.1.14 on our obtained optimal portfolio process.

In Korn and Kraft (2002), the chosen stochastic interest rate models are the Ho-Lee model and the Vasicek model. Due to the presence of mean-reversion in our three stochastic $\beta(t)$ factors, let us only compare with their Vasicek approach which has a similar OU process.

They define a short rate modeled by the following SDE for their Vasicek approach:

$$dr(t) = \Big(\theta(t) - \alpha r(t) + b\zeta(t)\Big)dt + bdW(t)$$

where $\alpha > 0$ and $\zeta(t)$ is a risk premium assumed to be deterministic and continuous. Furthermore, their asset price processes are represented by

$$\begin{aligned} dB(t) &= B(t)r(t)dt \\ dP(t,T_1) &= P(t,T_1)\Big[\mu(t)dt + \sigma(t)dW(t)\Big] \end{aligned}$$

for $T_1 > T$. They choose the drift coefficient $\mu(t)$ and the volatility coefficient $\sigma(t)$ for the Vasieck model to be defined as follows

$$\begin{aligned} \mu(t) &:= r(t) + \zeta(t)\sigma(t) \\ \sigma(t) &:= \frac{-b}{a}(1 - \exp\{(-\alpha(T_1 - t))\}) \end{aligned}$$

which leads to the wealth equation to be defined as

$$\begin{aligned} dX(t) &= X(t)\left[\left(\pi(t)\mu(t) + (1-\pi(t))r(t)\right)dt + \pi(t)\sigma(t)dW(t)\right] \\ &= X(t)\left[\left(\pi(t)(\mu(t) - r(t)) + r(t)\right)dt + \pi(t)\sigma(t)dW(t)\right] \\ &= X(t)\left[\left(\pi(t)\zeta(t)\sigma(t) + r(t)\right)dt + \pi(t)\sigma(t)dW(t)\right] \end{aligned} \tag{4.29}$$

The drift coefficient in (4.29) has a different representation to the one in (4.28). First, by essentially defining $\mu(t) := r(t) + \zeta(t)\sigma(t)$ such that it contains the explicit risk premium $\zeta(t)$, the final drift coefficient of the wealth equation (4.29) only contains one additional stochastic term $r(t)$. Moreover, $\mu(t) - r(t)$ yields the risk premium $\zeta(t)$ multiplied with the volatility of the bond price $\sigma(t)$ (see Korn and Kraft (2002, p. 1255)).

On the other hand, for our stochastic portfolio problem, the representation of the drift coefficient in the wealth equation of (4.28) contains the three stochastic DNS factors, namely, $\beta_0(t)$, $\beta_1(t)$ and $\beta_2(t)$. This increases the dimensionality of the HJB problem and the corresponding optimal control solution. Furthermore, the deterministic term $\zeta_1(t)$ that arises from $M_1(t) - M_r(t)$ is the implicit risk premium of the stochastic DNS bond. Hence, its representation also confirms that it already implicitly accounts for the volatility of the stochastic DNS bond price.

We summarize these in the following remark.

Remark 4.2.1. The preliminary difference between the obtained wealth processes in the two works are highlighted as follows:

1. In Korn and Kraft (2002), the continuous-time portfolio problem was solved under the influence of a stochastic short rate. As a result, the savings account was driven by a stochastic short rate $r(t)$, while the bond asset was driven by an underlying drift function which is a sum of this stochastic short rate $r(t)$ and a suitable risk premium $\zeta(t)$.

2. On the other hand, we attempt to solve the continuous-time portfolio problem driven by an underlying stochastic DNS framework that possesses three stochastic factors $\beta_0(t), \beta_1(t), \beta_2(t)$. As a result of the analysis in Section 4.1.1, we must change the order in which we define the assets. The stochastic DNS bond asset must be defined first under a relevant underlying DNS setting with a stochastic drift function $M_1(t)$. Subsequently, we are allowed to define a benchmark asset whose underlying rate behaves like a continuously-rolled over short-term rates under the stochastic DNS framework. The risk premium $\zeta_1(t)$ is implied from the specifications of the stochastic DNS framework. Furthermore, we already examine in Remark 4.1.6 how this benchmark asset is a relatively stable investment strategy, similar to the classical savings bank account.

This allows us to obtain an interesting perspective on such stochastic portfolio problems. By changing the order in which the assets are defined, one can calculate valid and optimal portfolios. Essentially, by specifying sufficient conditions on assets defined in a stochastic DNS framework, the corresponding portfolio problem results in a deterministic solvable solution.

The Control Problem and its Solution:

Now we can finally solve the portfolio problem using the necessary stochastic control methodology. In our notations, we neglect the dependencies $t, x, \beta_0(t), \beta_1(t), \beta_2(t)$ in the brackets of the functions for simplicity and ease of reading. We reinstate them wherever necessary.

The corresponding state space formulation is:

$$
\begin{aligned}
Y(t) &= (X(t), \beta_0(t), \beta_1(t), \beta_2(t))' \\
\Lambda(t, x, \beta_0, \beta_1, \beta_2, \pi) &= (x(\pi\zeta_1 + f_0\beta_0 + f_1\beta_1 + f_2\beta_2), \mu(a_0 - \beta_0), \mu(a_1 - \beta_1), \mu(a_2 - \beta_2))' \\
\Sigma(t, x, \beta_0, \beta_1, \beta_2, \pi) &= (x\pi\sigma, b_0, b_1, b_2)' \\
\Sigma^*(t, x, \beta_0, \beta_1, \beta_2, \pi) &= \begin{bmatrix} x^2\pi^2\sigma^2 & b_0x\pi\sigma & b_1x\pi\sigma & b_2x\pi\sigma \\ b_0x\pi\sigma & b_0^2 & b_1b_0 & b_2b_0 \\ b_1x\pi\sigma & b_1b_0 & b_1^2 & b_1b_2 \\ b_2x\pi\sigma & b_2b_0 & b_1b_2 & b_2^2 \end{bmatrix} \\
A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi) &= G_t + 0.5(x^2\pi^2\sigma^2 G_{xx} + 2b_0x\pi\sigma G_{x\beta_0} + 2b_1x\pi\sigma G_{x\beta_1} + 2b_2x\pi\sigma G_{x\beta_2} \\
&\quad + b_0^2 G_{\beta_0\beta_0} + b_1^2 G_{\beta_1\beta_1} + b_2^2 G_{\beta_2\beta_2} + 2b_0b_1 G_{\beta_0\beta_1} + 2b_0b_2 G_{\beta_0\beta_2} + 2b_1b_2 G_{\beta_1\beta_2}) \\
&\quad + x(\pi\zeta_1 + f_0\beta_0 + f_1\beta_1 + f_2\beta_2)G_x + \mu(a_0 - \beta_0)G_{\beta_0} \\
&\quad + \mu(a_1 - \beta_1)G_{\beta_1} + \mu(a_2 - \beta_2)G_{\beta_2}
\end{aligned} \tag{4.30}
$$

The following HJB equation needs to be solved:

$$
\begin{aligned}
\sup_{\pi \le \delta} A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi) &= 0 \\
G(T, x, \beta_0, \beta_1, \beta_2) &= \frac{1}{\gamma} x^\gamma
\end{aligned}
$$

where $\delta > 0$ and $\gamma \in (0, 1)$. The chosen utility function is $U(x) = \frac{1}{\gamma}x^\gamma$.

The goal is to solve

$$
\max_{\pi(t) \in \mathcal{A}^*(0, x_0)} \mathbb{E}(\frac{1}{\gamma} X^\pi(T))^\gamma
$$

where, $X(t)$ is the wealth equation represented by (4.28) and

$$
\mathcal{A}^*(0, x_0) := \{\pi(t) \in \mathcal{A}^*(0, x_0) : X^\pi(s) \ge 0 \quad , P - \text{a.s for } s \in [0, T]\}
$$

is the set of admissible controls.

To solve the portfolio problem, we perform three crucial steps mentioned in Korn and Kraft (2002), as follows:

Step 1:

To formally get a candidate for the optimal portfolio, we differentiate $A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi)$ of (4.30) under the max-operator with respect to π and then set the gradient to zero. Then, by additionally assuming $G_{xx} < 0$, we obtain the following:

$$
\begin{aligned}
0 &= x^2\pi\sigma^2 G_{xx} + b_0x\sigma G_{x\beta_0} + b_1x\sigma G_{x\beta_1} + b_2x\sigma G_{x\beta_2} + G_x x\zeta_1 \\
\text{or, } \pi^* &= -\frac{b_0}{x\sigma}\frac{G_{x\beta_0}}{G_{xx}} - \frac{b_1}{x\sigma}\frac{G_{x\beta_1}}{G_{xx}} - \frac{b_2}{x\sigma}\frac{G_{x\beta_2}}{G_{xx}} - \frac{\zeta_1}{x\sigma^2}\frac{G_x}{G_{xx}}
\end{aligned} \tag{4.31}
$$

For calculation simplicity, let us define

$$
\eta(t) := \frac{\zeta_1(t)}{\sigma(t)}
$$

Step 2:

After carefully plugging (4.31) into the HJB equation and dropping the sup − operator, we get the following:

$$\begin{aligned}0 &= G_t G_{xx} - 0.5b_0^2 G_{x\beta_0}^2 - 0.5b_1^2 G_{x\beta_1}^2 - 0.5b_2^2 G_{x\beta_2}^2 - 0.5\eta^2 G_x^2 \\ &- b_0 b_1 G_{x\beta_0} G_{x\beta_1} - b_0 b_2 G_{x\beta_0} G_{x\beta_2} - b_1 b_2 G_{x\beta_1} G_{x\beta_2} \\ &+ 0.5b_0^2 G_{\beta_0\beta_0} G_{xx} + 0.5b_1^2 G_{\beta_1\beta_1} G_{xx} + 0.5b_2^2 G_{\beta_2\beta_2} G_{xx} \\ &+ b_0 b_1 G_{xx} G_{\beta_0\beta_1} + b_0 b_2 G_{xx} G_{\beta_0\beta_2} + b_1 b_2 G_{xx} G_{\beta_1\beta_2} \\ &+ x(f_0\beta_0 + f_1\beta_1 + f_2\beta_2) G_x G_{xx} - b_0\eta G_x G_{x\beta_0} - b_1\eta G_x G_{x\beta_1} - b_2\eta G_x G_{x\beta_2} \\ &+ \mu(a_0 - \beta_0) G_{\beta_0} G_{xx} + \mu(a_1 - \beta_1) G_{\beta_1} G_{xx} + \mu(a_2 - \beta_2) G_{\beta_2} G_{xx}\end{aligned} \tag{4.32}$$

We know that $G(T, x, \beta_0, \beta_1, \beta_2) = \frac{1}{\gamma}x^\gamma$. This motivates the following separation of G: $G(t, x, \beta_0, \beta_1, \beta_2) = f(t, \beta_0, \beta_1, \beta_2) \cdot \frac{1}{\gamma}x^\gamma$, where $f(T, \beta_0, \beta_1, \beta_2) = 1$ for all $\{\beta_0, \beta_1, \beta_2\}$.

Then, we can derive the respective first and second partial derivatives of G with respect to each variable $t, \beta_0, \beta_1, \beta_2$, as follows:

$$\begin{aligned}G_t &= f_t \frac{1}{\gamma} x^\gamma \\ G_x &= f x^{\gamma-1} \\ G_{\beta_0} &= f_{\beta_0} \frac{1}{\gamma} x^\gamma \\ G_{\beta_1} &= f_{\beta_1} \frac{1}{\gamma} x^\gamma \\ G_{\beta_2} &= f_{\beta_2} \frac{1}{\gamma} x^\gamma \\ G_{xx} &= (\gamma-1) f x^{\gamma-2} \\ G_{x\beta_0} &= f_{\beta_0} x^{\gamma-1} \\ G_{x\beta_1} &= f_{\beta_1} x^{\gamma-1}\end{aligned}$$

$$\begin{aligned}G_{x\beta_2} &= f_{\beta_2} x^{\gamma-1} \\ G_{\beta_0\beta_0} &= f_{\beta_0\beta_0} \frac{1}{\gamma} x^\gamma \\ G_{\beta_1\beta_1} &= f_{\beta_1\beta_1} \frac{1}{\gamma} x^\gamma \\ G_{\beta_2\beta_2} &= f_{\beta_2\beta_2} \frac{1}{\gamma} x^\gamma \\ G_{\beta_0\beta_1} &= f_{\beta_0\beta_1} \frac{1}{\gamma} x^\gamma \\ G_{\beta_0\beta_2} &= f_{\beta_0\beta_2} \frac{1}{\gamma} x^\gamma \\ G_{\beta_1\beta_2} &= f_{\beta_1\beta_2} \frac{1}{\gamma} x^\gamma\end{aligned}$$

We substitute these partial derivatives in (4.32) and also observe that all of the $x^{2\gamma-2}$ terms cancel out. Hence, we get

$$\begin{aligned}0 &= f_t f(\gamma-1) - 0.5\gamma b_0^2 f_{\beta_0}^2 - 0.5\gamma b_1^2 f_{\beta_1}^2 - 0.5\gamma b_2^2 f_{\beta_2}^2 - 0.5\gamma\eta^2 f^2 \\ &- b_0 b_1 \gamma f_{\beta_0} f_{\beta_1} - b_0 b_2 \gamma f_{\beta_0} f_{\beta_2} - b_1 b_2 \gamma f_{\beta_1} f_{\beta_2} \\ &+ 0.5b_0^2 f_{\beta_0\beta_0} f(\gamma-1) + 0.5b_1^2 f_{\beta_1\beta_1} f(\gamma-1) + 0.5b_2^2 f_{\beta_2\beta_2} f(\gamma-1) \\ &+ b_0 b_1 (\gamma-1) f f_{\beta_0\beta_1} + b_0 b_2 (\gamma-1) f f_{\beta_0\beta_2} + b_1 b_2 (\gamma-1) f f_{\beta_1\beta_2} \\ &+ (f_0\beta_0 + f_1\beta_1 + f_2\beta_2)(\gamma-1)\gamma f^2 - b_0\eta\gamma f f_{\beta_0} - b_1\eta\gamma f f_{\beta_1} - b_2\eta\gamma f f_{\beta_2} \\ &+ \mu(a_0 - \beta_0) f f_{\beta_0}(\gamma-1) + \mu(a_1 - \beta_1) f f_{\beta_1}(\gamma-1) + \mu(a_2 - \beta_2) f f_{\beta_2}(\gamma-1)\end{aligned} \tag{4.33}$$

Again, we are able to separate the function $f(t, \beta_0, \beta_2, \beta_2)$ into its components as follows: $f(t, \beta_0, \beta_2, \beta_2) = k(t)\exp\{(l_0(t)\beta_0 + l_1(t)\beta_1 + l_2(t)\beta_2)\}$ with terminal conditions $l_0(T) =$

$l_1(T) = l_2(T) = 0$ and $k(T) = 1$. Once again, we can derive the first and second partial derivatives of function f as follows:

$$
\begin{aligned}
&f_t = k' e^{l_0\beta_0 + l_1\beta_1 + l_2\beta_2} + k(l_0'\beta_0 + l_1'\beta_1 + l_2'\beta_2)e^{l_0\beta_0 + l_1\beta_1 + l_2\beta_2}\\
&f_{\beta_0} = fl_0, \quad f_{\beta_1} = fl_1, \quad f_{\beta_2} = fl_2, \quad f_{\beta_0\beta_1} = fl_0l_1, \quad f_{\beta_0\beta_2} = fl_0l_2, \quad f_{\beta_1\beta_2} = fl_1l_2,\\
&f_{\beta_0\beta_0} = fl_0^2, \quad f_{\beta_1\beta_1} = fl_1^2, \quad f_{\beta_2\beta_2} = fl_2^2
\end{aligned}
$$

We substitute these in (4.33) and also observe that all f^2 terms cancel out. Then, we get

$$
\begin{aligned}
0 = {} & k'(\gamma-1) + k(\gamma-1)(l_0'\beta_0 + l_1'\beta_1 + l_2'\beta_2)\\
& - k(\mu\beta_0 l_0(\gamma-1) + \mu\beta_1 l_1(\gamma-1) + \mu\beta_2 l_2(\gamma-1))\\
& + k\Bigg(-0.5\gamma b_0^2 l_0^2 - 0.5\gamma b_1^2 l_1^2 - 0.5\gamma b_2^2 l_2^2 - 0.5\eta^2\gamma - b_0b_1\gamma l_0l_1 - b_0b_2\gamma l_0l_2 - b_1b_2\gamma l_1l_2\\
& \quad + 0.5b_0^2l_0^2(\gamma-1) + 0.5b_1^2l_1^2(\gamma-1) + 0.5b_2^2l_2^2(\gamma-1) + b_0b_1(\gamma-1)l_0l_1\\
& \quad + b_0b_2(\gamma-1)l_0l_2 + b_1b_2(\gamma-1)l_1l_2 - b_0\eta\gamma l_0 - b_1\eta\gamma l_1 - b_2\eta\gamma l_2\\
& \quad + \mu a_0(\gamma-1)l_0 + \mu a_1(\gamma-1)l_1 + \mu a_2(\gamma-1)l_2 + (f_0\beta_0 + f_1\beta_1 + f_2\beta_2)(\gamma-1)\gamma\Bigg)
\end{aligned}
$$

Here f_0, f_1, f_2 are as defined in (4.4) and the stochastic term $k(f_0\beta_0 + f_1\beta_1 + f_2\beta_2)(\gamma-1)\gamma$ can be subsequently separated out of the large bracket.

Then, we get

$$
\begin{aligned}
0 = {} & k'(\gamma-1) + k(\gamma-1)(l_0'\beta_0 + l_1'\beta_1 + l_2'\beta_2)\\
& - k(\mu\beta_0 l_0(\gamma-1) + \mu\beta_1 l_1(\gamma-1) + \mu\beta_2 l_2(\gamma-1)) + k(f_0\beta_0 + f_1\beta_1 + f_2\beta_2)(\gamma-1)\gamma\\
& + k\underbrace{\Bigg(-0.5\gamma b_0^2 l_0^2 - 0.5\gamma b_1^2 l_1^2 - 0.5\gamma b_2^2 l_2^2 - 0.5\eta^2\gamma - b_0b_1\gamma l_0l_1 - b_0b_2\gamma l_0l_2 - b_1b_2\gamma l_1l_2\\
& \quad + 0.5b_0^2l_0^2(\gamma-1) + 0.5b_1^2l_1^2(\gamma-1) + 0.5b_2^2l_2^2(\gamma-1) + b_0b_1(\gamma-1)l_0l_1\\
& \quad + b_0b_2(\gamma-1)l_0l_2 + b_1b_2(\gamma-1)l_1l_2 - b_0\eta\gamma l_0 - b_1\eta\gamma l_1 - b_2\eta\gamma l_2\\
& \quad + \mu a_0(\gamma-1)l_0 + \mu a_1(\gamma-1)l_1 + \mu a_2(\gamma-1)l_2\Bigg)}_{=:c_1(t)}
\end{aligned}
\tag{4.34}
$$

Here, we can define all of the deterministic terms in the large bracket of (4.34) to be a deterministic function $c_1(t)$.

Our ansatz for f as made above will only be meaningful if we get an ODE for k which does not include any of the stochastic factors $\beta_0, \beta_1, \beta_2$. This is because it will eliminate the second, third, and fourth terms on the RHS of (4.34), leaving us with only the first term (which is $k'(\gamma-1)$) and the fifth term (which is $kc_1(t)$).

In order to do so, we must obtain suitable representations of $l_0(t), l_1(t), l_2(t)$. By looking at (4.34) carefully, we can extract and solve the following first-order non-homogeneous ODEs:

$$
i. \quad k(\gamma-1)l_0'\beta_0 - k\mu\beta_0 l_0(\gamma-1) + k\beta_0 f_0(\gamma-1)\gamma = 0
$$

$$\begin{aligned} ii. \quad & k(\gamma-1)l_1'\beta_1 - k\mu\beta_1 l_1(\gamma-1) + k\beta_1 f_1(\gamma-1)\gamma = 0 \\ iii. \quad & k(\gamma-1)l_2'\beta_2 - k\mu\beta_2 l_2(\gamma-1) + k\beta_2 f_2(\gamma-1)\gamma = 0 \end{aligned}$$

where we already know the explicit forms of f_0, f_1, f_2 from (4.4) as,

$$\begin{aligned} f_0 &= 1 + \mu(T-t) \\ f_1 &= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(1 - \frac{\mu}{\lambda}) \\ f_2 &= \frac{\mu}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu}{\lambda} - \mu(T-t) + \lambda(T-t)) \end{aligned}$$

Hence, the solutions of these ODEs are:

$$\begin{aligned} i. \quad & l_0(t) = \gamma(T-t) \\ ii. \quad & l_1(t) = \frac{\gamma}{\lambda}(1 - e^{-\lambda(T-t)}) \\ iii. \quad & l_2(t) = \frac{\gamma}{\lambda}(1 - e^{-\lambda(T-t)}(1 + \lambda(T-t))) \end{aligned}$$

Thus, (4.34) becomes the following ODE, which is subsequently solved using the method of integrating factors:

$$\begin{aligned} 0 &= k'(\gamma-1) + kc_1(t) \\ \text{or, } k(t) &= \exp\left\{\frac{1}{1-\gamma}(C_1(t) - C_1(T))\right\} \end{aligned}$$

where, $C_1(t)$ is the primitive of $c_1(t)$, which means $C_1(t) = \int c_1(t)dt$. Hence, we obtain the candidate of the value function as

$$\begin{aligned} G(t, x, \beta_0, \beta_1, \beta_2) &= \frac{1}{\gamma}x^\gamma k(t)\exp\{(l_0(t)\beta_0 + l_1(t)\beta_1 + l_2(t)\beta_2)\} \\ &= \frac{1}{\gamma}x^\gamma \exp\left\{\frac{1}{1-\gamma}(C_1(t) - C_1(T))\right\} \\ &\cdot \exp\left\{\gamma\Big((T-t)\beta_0(t) + \frac{1-e^{-\lambda(T-t)}}{\lambda}\beta_1(t) + \frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\beta_2(t)\Big)\right\} \end{aligned} \tag{4.35}$$

From this we can obtain the following partial derivatives,

$$\begin{aligned} G_x &= \frac{\gamma G}{x} \\ G_{xx} &= \frac{\gamma(\gamma-1)G}{x^2} \\ G_{x\beta_0} &= \frac{\gamma^2 G}{x}(T-t) \\ G_{x\beta_1} &= \frac{\gamma^2 G}{x}\frac{(1-e^{-\lambda(T-t)})}{\lambda} \\ G_{x\beta_2} &= \frac{\gamma^2 G}{x}\frac{(1-e^{-\lambda(T-t)}(1+\lambda(T-t)))}{\lambda} \end{aligned}$$

Inserting this into our candidate for the optimal bond position that we derived in Step 1, we finally get **the optimal bond position $\pi^*(t)$** to be:

$$\begin{aligned}\pi^*(t) &= -\frac{b_0}{x\sigma}\frac{G_{x\beta_0}}{G_{xx}} - \frac{b_1}{x\sigma}\frac{G_{x\beta_1}}{G_{xx}} - \frac{b_2}{x\sigma}\frac{G_{x\beta_2}}{G_{xx}} - \frac{\zeta_1(t)}{x\sigma^2}\frac{G_x}{G_{xx}} \\ &= \frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2} \\ &- \frac{\gamma}{1-\gamma}\cdot\left(b_0(T-t) + b_1\frac{1-e^{-\lambda(T-t)}}{\lambda} + b_2\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\right) \\ &\cdot\left(\frac{1}{(T_1-t)(b_0+g(t)b_1+h(t)b_2)}\right)\end{aligned}$$

We can succinctly represent this optimal bond position as follows:

$$\pi^*(t) = \frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2} - \frac{\gamma}{1-\gamma}\kappa_1(t) \tag{4.36}$$

where,

$$\begin{aligned}\kappa_1(t) \coloneqq &\left(b_0(T-t) + b_1\frac{1-e^{-\lambda(T-t)}}{\lambda} + b_2\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\right) \\ &\cdot\left(\frac{1}{(T_1-t)(b_0+g(t)b_1+h(t)b_2)}\right)\end{aligned}$$

Note that,

$$\begin{aligned}\zeta_1(t) &= -a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu \\ &+ \frac{1}{2}(T_1-t)^2(b_0+g(t)b_1+h(t)b_2)^2 \\ \sigma(t) &= -(T_1-t)(b_0+g(t)b_1+h(t)b_2) \\ b_0, b_1, b_2 &= \text{constant volatilities of } \beta_0(t), \beta_1(t) \text{ and } \beta_2(t) \text{ respectively} \\ \lambda &= \text{factor loading decay parameter of the DNS yield curve}\end{aligned}$$

Since the final optimal position contains purely deterministic terms, the corresponding verification step is simple to prove. This final step is performed to ensure that the optimal control $\pi^*(t)$ satisfies the verification theorem for controlled problems within the stochastic DNS framework.

Step 3:

In order to prove that our obtained $\pi^*(t)$ is the optimal and valid solution to the control problem defined in (4.30), we must be able to apply Theorem 4.1.14. We can successfully apply the verification theorem, as long as we can ensure the following three assumptions:

A1. the state process $dY(t)$ has a solution which is unique,

A2. $\pi^*(t)$ is an admissible control, and

A3. G is a $C^{1,2}$ solution of the HJB equation

Hence, the following assertions ensure that the three afore-mentioned assumptions of the theorem are met.

1. Each $\beta(t)$ factor of the stochastic DNS framework must satisfy a bound on their $k = 1, 2$ moments of their supremum in $[0, T]$. This means:

$$\mathbb{E}\left(\sup_{t\in[0,T]} |\beta_0(t)|^k\right) < \infty, \mathbb{E}\left(\sup_{t\in[0,T]} |\beta_1(t)|^k\right) < \infty, \mathbb{E}\left(\sup_{t\in[0,T]} |\beta_2(t)|^k\right) < \infty$$

2. $\pi^*(t)$ satisfies the following in $[0, T]$:

$$\mathbb{E}\left(\sup_{t\in[0,T]} |\pi^*(t)|^k\right) < \infty \quad \text{for } k = 1, 2$$

3. The state process $Y(t) = (X(t), \beta_0(t), \beta_1(t), \beta_2(t))'$ has a unique solution with initial condition $Y(t_0) = y_0$ in $[0, T]$.

4. The state process $Y(t) = (X(t), \beta_0(t), \beta_1(t), \beta_2(t))'$ satisfies the following on $[0, T]$:

$$\mathbb{E}\left(\int_0^T \left|Y^{\pi^*}(s)\right|\right) ds < \infty$$

5. G is a $C^{1,2}(Q)$ solution of the HJB equation on $Q \subset [0, T] \times \mathbb{R}^n$ with initial conditions $(t_0, y_0) \in Q$.

Assertions 1, 3, and 4 imply **A1**. Assertions 2 implies **A2**. Assertion 5 implies **A3**. Let us apply the theoretical results derived in Section 4.1 to prove the six assertions made above.

Proof of 1: From the model specifications made in Section 3.3.3, we know that all of the three $\beta(t) = \{\beta_0(t), \beta_1(t), \beta_2(t)\}$ factors of the stochastic DNS framework can be represented by the following stochastic differential equations:

$$\begin{aligned} d\beta_i(t) &= \mu(a_i(t) - \beta_i(t))dt + b_i dW(t) \\ \implies \beta_i(t) &= \beta_i(0) + \int_0^t \underbrace{\mu e^{-\mu(t-u)} a_i(u)}_{=:A_i(t)} du + \int_0^t \underbrace{b_i e^{-\mu(t-u)}}_{=:B_i(t)} dW(u) \end{aligned}$$

Here each $i = 0, 1, 2$ represents each respective $\beta_i(t)$ factor. Thus, from these stochastic integral equations, we can observe that $A_i(\cdot)$ and $B_i(\cdot)$ satisfy the assumptions (4.17) and (4.18).

Moreover, each initial data $\beta_i(0) \neq 0$ and has finite absolute moments of every order

$$\mathbb{E}(|\beta_i(0)|^k) < \infty \quad , k = 1, 2, \ldots$$

which allows us to apply the second assertion of Proposition 4.1.12. As a result,

$$\mathbb{E}(|\beta_i(t)|^k) < \infty \quad , k = 1, 2, \ldots$$

Moreover, since $\beta_i(0)$ is deterministic, we get

$$\mathbb{E}\Big(\max_{t\in[0,T]} |\beta_i(0)|^k\Big) < \infty$$

All of the afore-mentioned statements allow us to apply Krylov (1980, p. 83 f, Theorem 9) and its corollary Krylov (1980, p. 85, Corollary 10) which state that, given the afore-mentioned assumptions on a stochastic process, there exists a constant N such that for all $k \geq 1, t \in [0, T]$, the following holds true:

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,T]} |\beta_i(t)|^{2k}\Big) &\leq N\mathbb{E}\Big(\sup_{t\in[0,T]} |\beta_i(0)|^{2k}\Big) \\ &\quad + Nt^{k-1}e^{Nt}\mathbb{E}\Big(\int_0^t (|\beta_i(0)|^{2k} + |A_i(s)|^{2k} + \|B_i(s)\|^{2k})ds\Big)\end{aligned}$$

In our case, since $\beta_i(0)$ is always a constant and the other terms are all continuous functions in t, we are able to obtain a finite supremum $\in [0, T]$. Thus, the following is satisfied for all three factors, $\beta_0(t), \beta_1(t), \beta_2(t)$:

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\beta_i(t)|^k\Big) < \infty$$

Furthermore, note that since these three factors are independent of the control $\pi(t)$, we can treat the subsequent state process $Y(t)$ as if it only depends on the wealth equation $X(t)$.

□

Proof of 2: Upon solving the bond portfolio problem, our obtained optimal position $\pi^*(t)$ is given by (4.36). We simplify it further as follows:

$$\begin{aligned}\pi^*(t) &= \frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2} \\ &\quad - \frac{\gamma}{1-\gamma}\left(\frac{1}{(T_1-t)((b_0 + g(t)b_1 + h(t)b_2))}\right) \\ &\quad \cdot \left(b_0(T-t) + b_1\frac{1-e^{-\lambda(T-t)}}{\lambda} + b_2\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\right)\end{aligned}$$

where, $b_0, b_1, b_2, \mu, \lambda$ are constants and,

$$\begin{aligned}\zeta_1(t) &= -a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu \\ &\quad + \frac{1}{2}(T_1-t)^2(b_0 + g(t)b_1 + h(t)b_2)^2 \\ \sigma(t) &= -(T_1-t)(b_0 + g(t)b_1 + h(t)b_2)\end{aligned}$$

Since there are two terms in $\pi^*(t)$, we may refer to them as $f_3(t)$ and $f_4(t)$, respectively, for ease of discussion. Here, $f_3(t) := \frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2}$ and $f_4(t)$ is the other part containing $-\frac{\gamma}{1-\gamma}$.

We can observe that both $f_3(t)$ and $f_4(t)$ are deterministic functions of t, continuous, and bounded in $[0, T]$. This is because all of their inner terms are also strictly deterministic and continuous functions of t.

Thus, using the Cauchy-Schwarz inequality, we get

$$\begin{aligned} \|\pi^*(t)\|^2 &= \|f_3(t) + f_4(t)\|^2 \\ &\leq 2\|f_3(t)\|^2 + 2\|f_4(t)\|^2 \end{aligned}$$

Taking the expectation, we get

$$\begin{aligned} \mathbb{E}\|\pi^*(t)\|^2 &\leq 2\mathbb{E}\|f_3(s)\|^2 + 2\mathbb{E}\|f_4(s)\|^2 \\ &< \infty \end{aligned} \tag{4.37}$$

Hence, $\pi^*(t)$ is differentiable in t with bounded derivatives, thereby satisfying the Lipschitz condition on $[0, T]$ for some $T < T_1$ (in particular, the **C.2** condition for an admissible control). Thus,

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\pi^*(t)|^k\Big) < \infty \quad \text{for } k = 1, 2, \ldots$$

□

Proof of 3: As we stated previously in *Proof of 1*, the state process $Y(t)$ can be treated as if it consists of only $X(t)$, due to three $\beta_i(t)$ factors being independent of the control.

From (4.28),

$$dX(t) = X(t)\Bigg[\Big(\pi(t)\zeta_1(t) + M_r(t)\Big)dt + \pi(t)\sigma(t)dW(t)\Bigg]$$

whose solution is

$$\begin{aligned} X(t) = x_0 \exp&\Big\{\int_0^t \Big(\pi^*(s)\zeta_1(s) + M_r(s) - 0.5\pi^*(s)^2\sigma(s)^2\Big)ds\Big\} \\ &\cdot \exp\Big\{\int_0^t \pi^*(s)\sigma(s)dW(s)\Big\} \end{aligned}$$

Thus, with initial condition $X(0) = x_0$, we can apply Proposition 4.1.12 to obtain a unique solution with bounded moments. This is possible since both *Proof of 1* and *Proof of 2* show that the coefficients of $X(t)$ satisfy the Itô conditions specified in Definition 4.1.11. Therefore, the corresponding state process $Y^{\pi^*}(t)$ has a unique solution in $[0, T]$. □

Proof of 4: If a solution to the wealth equation (4.28) exists and is unique (as already proved in *Proof of 3*), then applying both of the assertions of Proposition 4.1.12, we can state that the expectation of the corresponding state process $Y(t)$ is bounded.

$$\mathbb{E}\Big(\int_0^T \Big|Y^{\pi^*}(s)\Big|\Big)ds < \infty$$

□

Proof of 5: In Step 1, we assumed $G_{xx} < 0$ and subsequently obtain a suitable candidate for the optimal bond position, with initial conditions as $X(0) = x_0$. Let us check if it is true.

From (4.35), we know that

$$G = \frac{1}{\gamma} x^{\gamma} \exp\left\{ \frac{1}{1-\gamma} (C_1(t) - C_1(T)) \right\}$$
$$\cdot \exp\left\{ \gamma \Big((T-t)\beta_0(t) + \frac{1 - e^{-\lambda(T-t)}}{\lambda} \beta_1(t) + \frac{1 - e^{-\lambda(T-t)}(1 + \lambda(T-t))}{\lambda} \beta_2(t) \Big) \right\}$$

On differentiating twice with respect to x, we obtain

$$G_{xx} = \frac{\gamma(\gamma-1)G}{x^2}$$

Due to our assumption of $\gamma \in (0,1)$ and the positivity of both G and x, the above representation yields $G_{xx} < 0$. Hence, G is a $C^{1,2}$ solution of the HJB equation on $Q \subset [0,T] \times \mathbb{R}^n$.

Thus, our candidate for the optimal portfolio, as obtained in Step 1, is indeed a maximum (and not a minimum). □

Optimal Portfolio Results: Single Brownian Motion

Hence, as a result of Steps 1,2, and 3, we obtain the optimal bond position of an investor's admissible portfolio under the influence of a single Brownian motion as follows:

$$\pi^*(t) = \frac{1}{1-\gamma} \frac{\zeta_1(t)}{\sigma(t)^2} - \frac{\gamma}{1-\gamma} \kappa_1(t) \tag{4.38}$$

where,

$$\kappa_1(t) := \left(b_0(T-t) + b_1 \frac{1 - e^{-\lambda(T-t)}}{\lambda} + b_2 \frac{1 - e^{-\lambda(T-t)}(1 + \lambda(T-t))}{\lambda} \right)$$
$$\cdot \left(\frac{1}{(T_1 - t)(b_0 + g(t)b_1 + h(t)b_2)} \right) \tag{4.39}$$

Analysis of Results:

Thus, we have proven that we can obtain an optimal bond position $\pi^*(t)$ of the portfolio problem represented by (4.30) in the framework of a single Brownian motion driving all three stochastic DNS $\beta_0(t), \beta_1(t), \beta_2(t)$ factors. We also proved that this optimal solution satisfies existence and uniqueness results of classical control theory. Let us now present some analysis of this result.

As we know, the optimal bond position is obtained to be,

$$\pi^*(t) = \frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2} - \frac{\gamma}{1-\gamma}\kappa_1(t)$$

The first term,

$$\frac{1}{1-\gamma}\frac{\zeta_1(t)}{\sigma(t)^2}$$

is a deterministic function of time. It corresponds to the **classical result** obtained in the portfolio optimization problem of Merton (1969) when the underlying term structure model is deterministic. As a result, our original interpretation of $\zeta_1(t)$ to be the implicit risk premium of the stochastic DNS bond is consistent with literature, as we examine below.

The optimal proportion of the risky asset calculated by Merton (1969) was found to be

$$\frac{1}{1-\gamma}\frac{a-r}{\sigma^2}$$

where, a is the expected rate of return of the risky asset and r is the deterministic return of the sure asset. Similarly, in Korn and Kraft (2002), the optimal investment percentage in the risky bond was calculated to be a function containing two parts. The first part was

$$\frac{1}{1-\gamma}\frac{\zeta(t)}{\sigma(t)}$$

where, as explained in Remark 4.2.1, $\zeta(t) := \frac{\mu-r}{\sigma}$ is taken to be the risk premium of the bond.

The second term of our $\pi^*(t)$,

$$\frac{\gamma}{1-\gamma}\kappa_1(t)$$

can be thought of as the **correction term** due to the three stochastic DNS factors. It is also a deterministic function of time, and has some interesting characteristics as observed in literature. Specifically, in Korn and Kraft (2002), their optimal investment percentage in the risky asset also includes the following correction term:

$$\frac{\gamma}{1-\gamma}\kappa(t)$$

where, $\kappa(t)$ takes an appropriate form according to the underlying single-factor stochastic short rate. This correction term was found to satisfy economically-intuitive features: First, the term is also positive and decreases monotonically to zero as the time approaches the end of the investment horizon. Second, the correction term increases with increasing risk-aversion of the investor. We would now like to verify if our correction term also satisfies these features.

On expanding our correction term, we get

$$\begin{aligned}\text{correction term} = \frac{\gamma}{1-\gamma}&\left(b_0(T-t) + b_1\frac{1-e^{-\lambda(T-t)}}{\lambda} + b_2\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\right)\\ &\cdot\left(\frac{1}{(T_1-t)(b_0+g(t)b_1+h(t)b_2)}\right)\end{aligned} \tag{4.40}$$

Here, we know that b_0, b_1, b_2 are the constant volatilities of the $\beta_0(t), \beta_1(t), \beta_2(t)$ factors, respectively. $g(t) := \frac{1-e^{-\lambda(T_1-t)}}{\lambda(T_1-t)}$ and $h(t) := \frac{1-e^{-\lambda(T_1-t)}}{\lambda(T_1-t)} - e^{-\lambda(T_1-t)}$ are the factor loadings of $\beta_1(t)$ and $\beta_2(t)$, respectively (see Section 3.3.3). The bond maturity time is T_1 which is greater than the investment termination time T. The coefficient of relative risk aversion of the investor is given by $1-\gamma$ with $\gamma \in (0,1)$. Hence, a more risk-averse investor will have a lower γ and vice-versa.

Since all of these values must be non-negative, it can be observed that the correction term is always strictly positive. Furthermore, as $t \to T$, due to the terms in the second large bracket,

$$\lim_{t\to T} \text{correction term} = 0$$

The derivative of this term with respect to t can be calculated as

$$\begin{aligned}&\frac{\gamma}{1-\gamma}\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\left(\frac{1}{(T_1-t)(b_0+g(t)b_1+h(t)b_2)}\right)}_{=:\text{part}_1}\\ &\cdot\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\left(b_0(T-t) + b_1\frac{1-e^{-\lambda(T-t)}}{\lambda} + b_2\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\right)}_{=:\text{part}_2}\end{aligned}$$

Then,

$$\begin{aligned}\frac{\mathrm{d}\text{part}_2}{\mathrm{d}t} &= -\left(b_0 + b_1 e^{-\lambda(T-t)} + b_2 e^{-\lambda(T-t)}\lambda(T-t)\right)\\ &< 0\end{aligned}$$

and

$$\frac{\mathrm{d}\text{part}_1}{\mathrm{d}t} = \frac{b_0 + b_1 e^{-\lambda(T_1-t)} + b_2 e^{-\lambda(T_1-t)}\lambda(T_1-t)}{(T_1-t)^2(b_0+g(t)b_1+h(t)b_2)^2} > 0$$

Hence, the derivative of the correction term with respect to t is non-positive. Thus, the correction term is monotonically decreasing to zero as $t \to T$.

Moreover, with our analysis on both part_1 and part_2, we see that a higher γ value results in a higher value of the correction term as a whole. This is because the coefficient $\frac{\gamma}{1-\gamma}$ in the correction term is an increasing function in γ. As we know from our utility function, the coefficient of relative risk aversion is given by $1-\gamma$. This implies with increasing risk-aversion, the investor's final position in the risky asset $\pi^*(t)$ decreases in comparison to his position in the low-risk asset. This is reasonable for a risk-averse investor, as the randomness of the three stochastic factors will prompt a highly risk-averse investor to shift his wealth to a safer strategy.

Thus, we arrive at the following remark.

Remark 4.2.2. When the underlying stochastic DNS yield curve is driven by a single Brownian motion, the continuous-time portfolio problem is solvable and produces a unique optimal position of the stochastic DNS bond as $\pi^*(t)$ whose representation is given in (4.36). Additionally, this optimal $\pi^*(t)$ satisfies intuitive economic interpretations that validates the behaviour of the investor with regards to his risk appetite.

Numerical Example:

Let us consider an example to visualize the remark made above. Consider an investor who is allowed to invest in a portfolio specified in this subsection (page 77 ff.).

Let the initial parameters of the three stochastic $\beta_0(t), \beta_1(t), \beta_2(t)$ factors be as follows: $a_0 = 0.05, a_1(t) = 0.02, a_2(t) = 0.02, b_0 = 0.005, b_1 = 0.011, b_2 = 0.025$. Further let the mean-reversion speed of these factors be $\mu = 0.5$ and the factor loading parameter be $\lambda = 0.5975$. Let the final investment horizon of the investor be $T = 25$ years, while the maturity time of the stochastic DNS bond be $T_1 = 30$ years. Finally, we take the range $\gamma = \{0.25, 0.45, 0.65\}$.

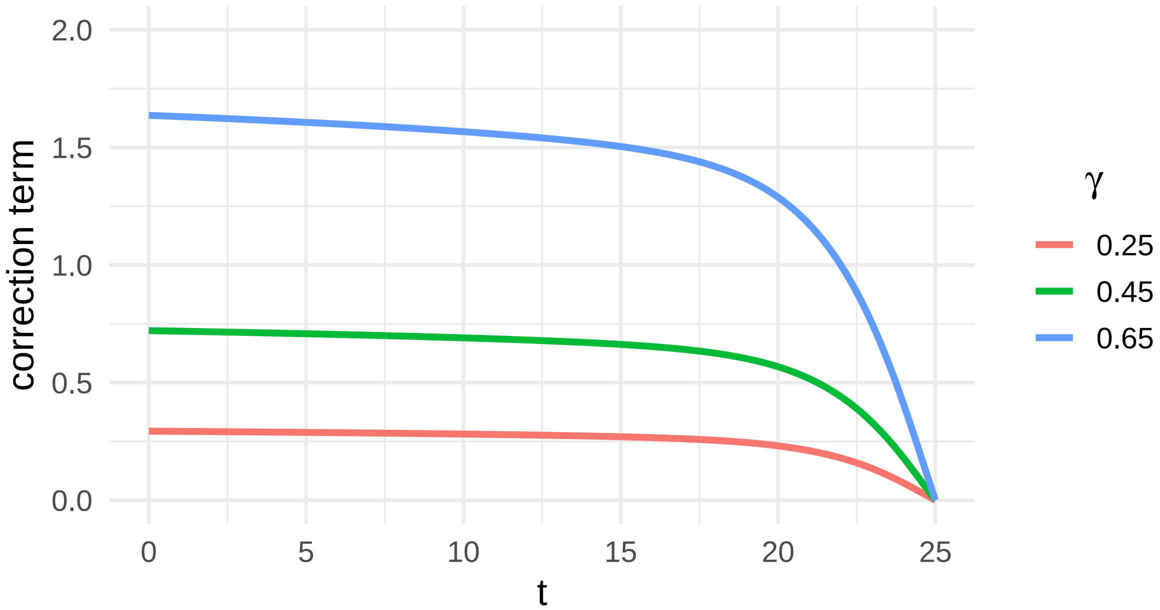


Figure 4.1: Relation between the Correction Term and the Risk-Aversion of the Investor

As we can observe in Figure 4.1, with increasing γ, we obtain an increasing term represented by $\frac{\gamma}{1-\gamma}\kappa_1(t)$. Now, let us visualize its impact on the final optimal portfolio $\pi^*(t)$ in the stochastic DNS bond in Figure 4.2. We can observe that a more risk-averse investor has a lower γ and thereby has a lower position in the risky asset, as compared to an investor who is more risk-taking.

Figure 4.2 also shows us that the optimal position in a stochastic DNS bond that is driven by a single Brownian motion is the highest at time 0 and gradually decreases towards the end of the investment horizon. Due to the absence of restrictions on the ability to short-sell or borrow the asset, the resultant $\pi^*(t)$ values are allowed to go beyond 100%. Similarly, these values can also be negative (not displayed in our numerical example).

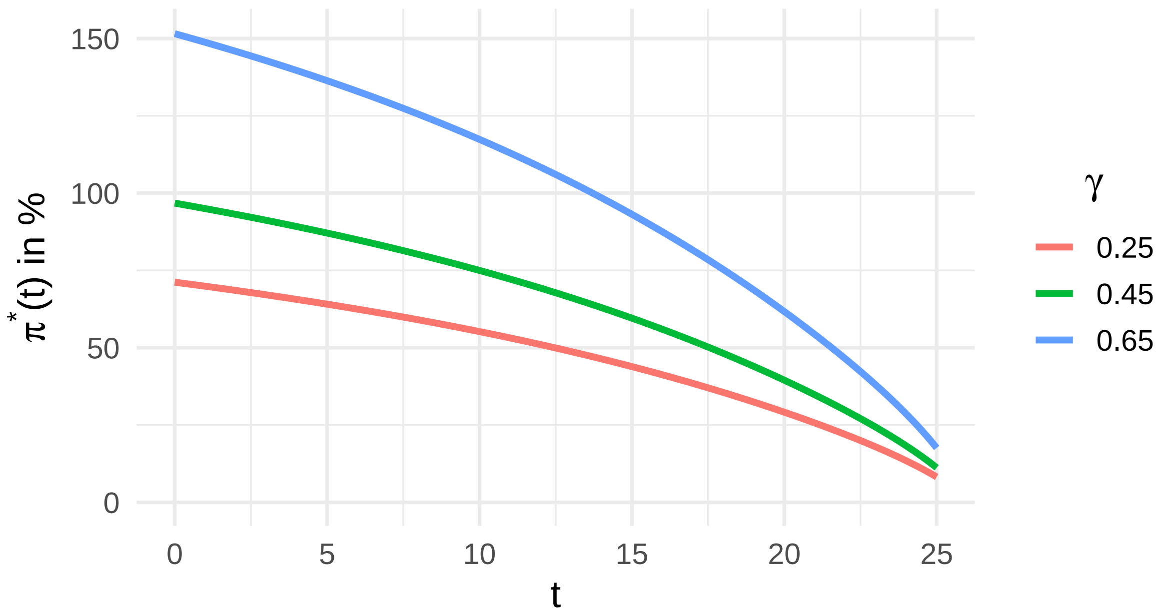


Figure 4.2: Optimal bond position $\pi^*(t)$: Single Brownian Motion

4.3 Comparison with Arbitrage-Free Nelson-Siegel Models

In the previous subsection, we obtained the optimal portfolio when the underlying yield curve market is influenced by a stochastic dynamic Nelson-Siegel framework, with a focus on the simplest of the three possible cases, *Case 1: single Brownian motion*. With this bird's eye overview on the topic, we now perform a relevant excursion that discusses how the concept of arbitrage affects the obtained optimal portfolios under such stochastic dynamics of the Nelson-Siegel family.

For this, we perform a three-fold approach: First, we discuss the concept of arbitrage, and subsequently delve into the reason behind why, in literature, the Nelson-Siegel family is considered to not satisfy the arbitrage-free condition. Second, we discuss an affine class of arbitrage-free yield curve (AFNS) models whose parameters match the Nelson-Siegel family, as introduced by Christensen et al. (2011). Finally, we compare the optimal portfolio results of the stochastic control problem obtained by our stochastic DNS framework with those obtained by the corresponding AFNS model.

4.3.1 Arbitrage in the Nelson-Siegel Family

In literature, an arbitrage opportunity can be commonly defined as follows:

Definition 4.3.1 (Arbitrage opportunity)**.** An arbitrage opportunity is the phenomenon that occurs when there exists an admissible self-financing trading strategy $\phi(t)$ with initial wealth $x_0 = 0$ such that, the final wealth corresponding to the strategy satisfies $X_T(\phi) \geq 0$ $P - a.s$ with $P(X_T(\phi) > 0) > 0$.

This definition implies that there exists a risk-less strategy in the market that produces a non-negative final capital with a strictly positive probability for even being positive, despite zero initial wealth.

However, independent of this assumption on initial wealth or the construction of the control problem in general, the Nelson-Siegel family of yield curves has been considered to fail to impose theoretical restrictions on its parameters that ensure the absence of arbitrage opportunities. This finding was first highlighted by Björk and Christensen (1999) within the Heath-Jarrow-Morton (HJM) framework, and was generalized and detailed by Filipović (1999). In particular, Filipović (1999) introduces the concept of "consistent Itô processes" to define the parameters of the Nelson-Siegel family in order to prove that the resultant discounted bond prices can never be $\mathbb{P}$-martingale. We discuss the results briefly below.

Consider the original Nelson-Siegel family of yield curves of Nelson and Siegel (1987), described in (3.2) repeated as follows:

$$y(t, t+\tau) = \beta_0(t) + (\beta_1(t) + \beta_2(t)) \cdot \frac{\left(1 - \exp\left(\frac{-\tau}{\lambda}\right)\right)}{\frac{\tau}{\lambda}} - \beta_2(t) \cdot \exp\left(-\frac{\tau}{\lambda}\right) \tag{4.41}$$

The above equation is derived from the following relation between yield curves and forward rate curves:

$$y(t, t+\tau) = \frac{1}{\tau} \int_0^\tau f(s, s+\tau) ds \tag{4.42}$$

Thus, the corresponding equation representing the family for forward rates is

$$F(t, t+\tau) = \beta_0(t) + \beta_1(t) \exp\{-\lambda\tau\} + \beta_2(t)\tau \exp\{-\lambda\tau\} \tag{4.43}$$

Filipović (1999) expresses (4.43) as a state process $F(\tau, Z(t))$, where

$$Z(t) = (\beta_0(t), \beta_1(t), \beta_2(t), \lambda(t)) \tag{4.44}$$

is the parameter space of the Nelson-Siegel family. Here, $\lambda(t)$ is also considered to be a time-dependent parameter.

Then, each element of $Z(t)$ is assumed to be an Itô process of the form

$$Z_i(t) = Z_i(0) + \int_0^t b_i(s) ds + \sum_{j=1}^{d} \int_0^t \sigma_{ij}(s) dW_j(s) \quad, t \in [0, \infty) \tag{4.45}$$

where, i represents each element of the space $Z(t)$, and $Z(0)$ is $\mathcal{F}(0)$-measurable. Here, b and σ are progressively measurable functions, and W is a d-dimensional $(\mathcal{F}_t)$ Brownian motion.

Thus, this implies that all three $\beta(t)$ factors of the DNS model as well as the $\lambda(t)$ parameter follow an Itô process given by (4.45). Filipović (1999) hypothesizes that they can commonly be solutions of stochastic differential equations, and the results can also be extended to a wider class of state processes, including time homogeneous Markov processes and Itô processes with jumps.

Then, the price process for a zero-coupon bond maturing at time T is represented by

$$P(t,T) := \exp\left\{-\int_0^{T-t} F(s, Z(t))ds\right\} \tag{4.46}$$

and the corresponding money market account is given by

$$B(t) := \exp\left\{\int_0^t F(0, Z(s))ds\right\} \tag{4.47}$$

We now explain how the concept of "consistent Itô process" introduced by Filipović (1999) can be used to ensure the lack of arbitrage opportunities.

Definition 4.3.2 (Consistent state space process (Filipović (1999))). A state space process $Z(t)$, represented by (4.44), is consistent with the Nelson-Siegel family, if and only if

$$\left(\frac{P(t,T)}{B(t)}\right) \quad , t \in [0,T]$$

is a $\mathbb{P}$-martingale for all $T < \infty$.

If such a parameter state space process exists for $Z(t) = (\beta_0(t), \beta_1(t), \beta_2(t), \lambda(t))$, then the discounted bond price processes is a $\mathbb{P}$-martingale, thereby indicating that there exists an equivalent martingale measure $\mathbb{Q}$. It is known that the presence of an equivalent martingale measure implies the absence of arbitrage opportunities (see Delbaen and Schachermayer (2006) and Delbaen (2006)).

To find whether such a consistent state process exists, the primary question posed by Filipović (1999) is as follows: Are there suitable coefficients b and σ for which the Itô processes of $Z(t)$ represented by (4.45) are consistent?

The answer to this question was found to be: no. It turns out that all Itô processes consistent with the Nelson-Siegel family have essentially deterministic dynamics, as illustrated in the results below:

Remark 4.3.3. Filipović (1999, Theorem 4.1) states that for all $t \in [0, \infty)$, if $Z(t) = (\beta_0(t), \beta_1(t), \beta_2(t), \lambda(t))$ is a consistent Itô process, then it is of the form:

$$\begin{aligned}
\beta_0(t) &= \beta_0(0) \\
\beta_1(t) &= \beta_1(0)e^{-\lambda(0)\cdot t} + \beta_2(0) \cdot te^{-\lambda(0)\cdot t} \\
\beta_2(t) &= \beta_2(0)e^{-\lambda(0)\cdot t}
\end{aligned}$$

$$\lambda(t) = \lambda(0) + \Big(\int_0^t b_\lambda(s)ds + \sum_{j=1}^{d} \int_0^t \sigma_{\lambda j}(s)dW_j(s)\Big)\mathbb{1}_{\Omega_0}$$

where, $\Omega_0 \coloneqq \{\beta_1(0) = \beta_2(0) = 0\}$.

The above remark thereby implies that for the Nelson-Siegel family, the only corresponding bond price process that ensures the lack of arbitrage opportunities is purely deterministic. Thus, there can be no stochastic evolution of the Nelson-Siegel factors that ensures $P(t,T)$, as defined in (4.46), satisfies the no-arbitrage condition.

Remark 4.3.4. With this interpretation of arbitrage, it can be inferred that the primary criticism towards the theoretical structure of bond prices resulting from a stochastic Nelson-Siegel framework lies in the fact that a bond price derived within a stochastic DNS framework at time t no longer stays within the same Nelson-Siegel family at time $t+1$.

In order to remedy this theoretical deficit, Christensen et al. (2011) proposed an affine class of arbitrage-free models that imposes the absence of arbitrage on the Nelson-Siegel family of bond prices.

4.3.2 The Arbitrage-Free Nelson-Siegel (AFNS) Model

Christensen et al. (2009) and Christensen et al. (2011) introduced a term structure model whose factor loadings exactly matched the dynamic Nelson-Siegel (DNS) factor loading structure. These models, called the arbitrage-free Nelson-Siegel (AFNS) models, attempted to combine the most favorable feature of affine term structure models (namely, their ability to satisfy the theoretical arbitrage-free condition) and of the DNS yield curve (namely, its high empirical success in term structure estimation and forecasting).

The strategy behind the AFNS model construction is based on the framework of Duffie and Kan (1996) who state that arbitrage-free zero-coupon bond prices are exponential affine functions of the state variables. In the DNS framework, the state variables are the three factors, $\beta_0(t), \beta_1(t)$ and $\beta_2(t)$. As a result, the zero-coupon bond price under the absence of arbitrage becomes:

$$P(t,T) = \exp\{B(t,T)'\beta(t) + A(t,T)\} \tag{4.48}$$

where,

$$\beta(t) = (\beta_0(t), \beta_1(t), \beta_2(t))$$

Here, the term $B(t,T)$ is chosen such that it matches the factor loadings of the DNS yield curve. Thus,

$$B(t,T) = \left(-(T-t), -\frac{1-e^{\lambda(T-t)}}{\lambda}, -\Big(\frac{1-e^{\lambda(T-t)}}{\lambda} - (T-t)e^{-\lambda(T-t)}\Big)\right)$$

Since we already know that the factor loadings of the DNS yield curve is $\Big(1, \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}, \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}\Big)$, it can be observed that the structure of $B(t,T)$ is a result of a relevant system of ODEs.

The final term $A(t,T)$ is the key difference between the DNS and AFNS models. It is called the yield-adjustment term, and is responsible for the affine arbitrage-free structure of the model. The analytical form of the yield-adjustment term is stated by Christensen et al. (2011) as follows:

$$\begin{aligned}\frac{A(t,T)}{T-t} &= \frac{1}{2}\frac{1}{T-t}\sum_{j=1}^{3}\int_t^T \left(\Sigma' B(s,T)B(s,T)'\Sigma\right)_{j,j} ds \\ &= \bar{A}\frac{(T-t)^2}{6} + \bar{B}\left[\frac{1}{2\lambda^2} - \frac{1}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{1}{4\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right] \\ &+ \bar{C}\left[\frac{1}{2\lambda^2} + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{4\lambda}(T-t)e^{-2\lambda(T-t)} - \frac{3}{4\lambda^2}e^{-2\lambda(T-t)}\right. \\ &\left. - \frac{2}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{5}{8\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right] \\ &+ \bar{D}\left[\frac{1}{2\lambda}(T-t) + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t}\right] \\ &+ \bar{E}\left[\frac{3}{\lambda^2}e^{-\lambda(T-t)} + \frac{1}{2\lambda}(T-t) + \frac{1}{\lambda}(T-t)e^{-\lambda(T-t)} - \frac{3}{\lambda^2}\frac{1-e^{-\lambda(T-t)}}{T-t}\right] \\ &+ \bar{F}\left[\frac{1}{\lambda^2} + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{2\lambda^2}e^{-2\lambda(T-t)} - \frac{3}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t}\right. \\ &\left. + \frac{3}{4\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right] \end{aligned} \tag{4.49}$$

where,

$$\begin{aligned}\bar{A} &= \sigma_{11}^2 + \sigma_{12}^2 + \sigma_{13}^2 \\ \bar{B} &= \sigma_{21}^2 + \sigma_{22}^2 + \sigma_{23}^2 \\ \bar{C} &= \sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2 \\ \bar{D} &= \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} + \sigma_{13}\sigma_{23} \\ \bar{E} &= \sigma_{11}\sigma_{31} + \sigma_{12}\sigma_{32} + \sigma_{13}\sigma_{33} \\ \bar{F} &= \sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32} + \sigma_{23}\sigma_{33}\end{aligned}$$

As we can see, it has a complicated analytical form, where the σ_{ij} values are obtained from the following triangular matrix:

$$\Sigma = \begin{pmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{pmatrix}$$

This triangular matrix Σ represents the constant volatility structure of the three DNS $\beta(t)$ factors. The complete derivations of the AFNS model specifications and the yield-adjustment term $\frac{A(t,T)}{(T-t)}$ can be found in Christensen et al. (2011). We do not repeat them in this thesis.

Now it can be seen from (4.49) that the yield-adjustment term essentially has the following form:

$$A(t,T) = \frac{1}{2}\sum_{j=1}^{3}\int_t^T \left(\Sigma' B(s,T)B(s,T)'\Sigma\right)_{j,j} ds \tag{4.50}$$

Christensen et al. (2011) obtain this form by setting the mean-reversion levels of the $\beta(t)$ factors to zero under the $\mathbb{Q}$-measure. On one hand, the analytical formula of $A(t,T)$ allows for empirical tractability, thereby ensuring that the implementation of AFNS models is feasible. On the other hand, with this additional term, it removes the intuitive interpretation that the original DNS model possessed, that allowed the yield curve formula to be suitably condensed such that it enabled the three $\beta_0(t), \beta_1(t), \beta_2(t)$ factors to represent the level, slope, and curvature of the yield curve, respectively.

Let us now observe the resultant yield curve dynamics. The pricing function (4.48) implies that the zero-coupon yields under the AFNS framework are given by

$$\begin{aligned} y(t,T) &= -\frac{1}{T-t}\log P(t,T) \\ &= -\frac{B(t,T)}{(T-t)}\beta(t) - \frac{A(t,T)}{(T-t)} \end{aligned} \tag{4.51}$$

We can condense (4.51) as follows:

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t) - \frac{A(t,T)}{(T-t)} \tag{4.52}$$

where we already defined (in Section 3.3.3),

$$\begin{aligned} g(t) &= \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} \\ h(t) &= \frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)} \end{aligned}$$

The first three terms of (4.52) are exactly similar to the original DNS yield curve, as given by

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t)$$

The difference lies in the additional term $\frac{-A(t,T)}{(T-t)}$, which Christensen et al. (2011) claims is only dependent on the bond maturity and not on time.

Let us now discuss the specifications of the three $\beta(t)$ factors under the AFNS model proposed by Christensen et al. (2011). Under the P-measure, the $\beta(t)$ factors of the AFNS models are specified to be as follows:

1. the **independent-factor AFNS model**, where all three $\beta_0(t), \beta_1(t), \beta_2(t)$ factors are independent and are defined by the following stochastic differential equations:

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu_0 & 0 & 0 \\ 0 & \mu_1 & 0 \\ 0 & 0 & \mu_2 \end{pmatrix} \left[\begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} \tilde{b}_0 & 0 & 0 \\ 0 & \tilde{b}_1 & 0 \\ 0 & 0 & \tilde{b}_2 \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{4.53}$$

2. the **correlated-factor AFNS model**, where all three $\beta_0(t), \beta_1(t), \beta_2(t)$ factors interact dynamically and are defined by the following stochastic differential equations:

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu_{00} & \mu_{01} & \mu_{02} \\ \mu_{10} & \mu_{11} & \mu_{12} \\ \mu_{20} & \mu_{21} & \mu_{22} \end{pmatrix} \left[\begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} \tilde{b}_{00} & 0 & 0 \\ \tilde{b}_{10} & \tilde{b}_{11} & 0 \\ \tilde{b}_{20} & \tilde{b}_{21} & \tilde{b}_{22} \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{4.54}$$

Bond Price Dynamics

The bond price dynamics can be calculated from (4.48) and (4.52), as follows:

$$dP(t,T) = P(t,T)\Big(- dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \quad \text{(from (3.16))}$$

For the independent-factor AFNS model,

$$\begin{aligned} dP^{\text{AFNS}}_{\text{indp}}(t,T) &= P^{\text{AFNS}}_{\text{indp}}(t,T)\Big(- dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \\ &= P^{\text{AFNS}}_{\text{indp}}(t,T)\bigg(\Big(\beta_0(t)(1+\mu_0(T-t)) + \beta_1(t)\Big(\frac{\mu_1}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu_1}{\lambda})\Big) \\ &\qquad + \beta_2(t)\Big(\frac{\mu_2}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu_2}{\lambda} - \mu_2(T-t) + \lambda(T-t))\Big) \\ &\qquad - \theta_0(T-t)\mu_0 - \theta_1(T-t)g(t)\mu_1 - \theta_2(T-t)h(t)\mu_2 \\ &\qquad + A'(t,T) + \frac{1}{2}(T-t)^2(\tilde{b}_0^2 + g(t)^2\tilde{b}_1^2 + h(t)^2\tilde{b}_2^2)\Big)dt \\ &\qquad - (T-t)\tilde{b}_0 dW_0(t) - (T-t)g(t)\tilde{b}_1 dW_1(t) - (T-t)h(t)\tilde{b}_2 dW_2(t)\bigg) \end{aligned} \tag{4.55}$$

and for the correlated-factor AFNS model,

$$\begin{aligned} dP^{\text{AFNS}}_{\text{corr}}(t,T) &= P^{\text{AFNS}}_{\text{corr}}(t,T)\Big(- dy(t,T)(T-t) + y(t,T)dt + \frac{1}{2}(T-t)^2 d < y(t,T) > \Big) \\ &= P^{\text{AFNS}}_{\text{corr}}(t,T)\bigg(\Big[\beta_0(t)(1+\mu_0(T-t)) + \beta_1(t)\Big(\frac{\mu_1}{\lambda} + e^{-\lambda(T-t)}(1-\frac{\mu_1}{\lambda})\Big) \\ &\qquad + \beta_2(t)\Big(\frac{\mu_2}{\lambda} + e^{-\lambda(T-t)}(-\frac{\mu_2}{\lambda} - \mu_2(T-t) + \lambda(T-t))\Big) \end{aligned}$$

$$
\begin{aligned}
&- \theta_0(T-t)\mu_0 - \theta_1(T-t)g(t)\mu_1 - \theta_2(T-t)h(t)\mu_2 \\
&+ A'(t,T) + \frac{1}{2}(T-t)^2\Big((\tilde{b}_{00} + \tilde{b}_{10}g(t) + h(t)\tilde{b}_{20})^2 \\
&+ (\tilde{b}_{11}g(t) + \tilde{b}_{21}h(t))^2 + (\tilde{b}_{22}h(t))^2\Big)\Big]dt \\
&- (T-t)(\tilde{b}_{00} + \tilde{b}_{10}g(t) + h(t)\tilde{b}_{20})d\tilde{W}_0(t) \\
&- (T-t)(\tilde{b}_{11}g(t) + \tilde{b}_{21}h(t))d\tilde{W}_1(t) - (T-t)h(t)\tilde{b}_{22}d\tilde{W}_2(t)\Big)
\end{aligned}
\tag{4.56}
$$

Except for the time-derivative of the yield-adjustment term $A'(t,T)$, both of the AFNS bond price representations, (4.55) and (4.56), appear similar to the corresponding representation of the stochastic DNS models presented in (3.23) and (3.28). However, the non-vanishing analytical form of $A(t,T)$ makes this difference non-negligible.

In contrast to the AFNS model specifications of (4.53) and (4.54), let us now recall our corresponding stochastic DNS model specifications from Section 3.3.4 and Section 3.3.5 below:

1. the **uncorrelated stochastic DNS model** of Theorem 3.3.3, where all three $\beta_0(t), \beta_1(t), \beta_2(t)$ factors are driven by uncorrelated Brownian motions (from (3.19)):

$$
\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & b_2 \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix}
\tag{4.57}
$$

2. the **correlated stochastic DNS model** of Theorem 3.3.4, where all three $\beta_0(t), \beta_1(t), \beta_2(t)$ factors are driven by correlated Brownian motions (from (3.24)):

$$
\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ b_1\rho_{01} & b_1\tilde{\rho}_{11} & 0 \\ b_2\rho_{02} & b_2\tilde{\rho}_{21} & b_2\tilde{\rho}_{22} \end{pmatrix} \begin{pmatrix} d\tilde{W}_0(t) \\ d\tilde{W}_1(t) \\ d\tilde{W}_2(t) \end{pmatrix}
\tag{4.58}
$$

Differences between AFNS and DNS

We summarize the key differences between the AFNS models and our proposed stochastic DNS models of Chapter 3 in the following remark:

Remark 4.3.5 (Differences between the Models)**.** The primary similarity between the stochastic DNS models proposed in Chapter 3 and the AFNS models proposed by Christensen et al. (2011) is the overall structure of the stochastic factors $(\beta_0(t), \beta_1(t), \beta_2(t))$. In both of these models, the three DNS factors are considered to be mean-reverting Itô processes. On the other hand, the differences between the two models are as follows:

1. The mean-reversion speeds of the factors in our stochastic DNS models are the same. Each $\beta(t)$ factor has a constant mean-reversion speed μ. On the other hand, the mean-reversion speeds of each of the three $\beta(t)$ factors in the AFNS models are different $\mu_0 \neq \mu_1 \neq \mu_2$.
2. Only the $\beta_0(t)$ factor of our stochastic DNS models has a constant mean-reversion level a_0. The other two factors, $\beta_1(t)$ and $\beta_2(t)$, have time-dependent mean-reversion levels $a_1(t)$ and $a_2(t)$, respectively. On the other hand, all three mean-reversion levels of the $\beta(t)$ factors in the AFNS models are constants θ_0, θ_1 and θ_2, respectively.
3. In our stochastic DNS framework, the yield curve is calculated as follows:

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t)$$

   The resultant bond price dynamics $dP(t,T)$ is calculated accordingly. However, the yield curve formula for the AFNS framework contains an extra yield-adjustment term, as observable below:

$$y(t,T) = \beta_0(t) + \beta_1(t)g(t) + \beta_2(t)h(t) - \frac{A(t,T)}{(T-t)}$$

   which also changes the resultant bond price dynamics $dP^{\text{AFNS}}(t,T)$. The primary difference is the additional term $A'(t,T)$, the time-derivative of the yield-adjustment term.

The first two observations of Remark 4.3.5 are not highly significant since the parameters can be easily changed in either models. However, the third statement discussing the yield-adjustment term is likely to impact the final optimal result. Thus, let us now compute the optimal portfolio solutions of all four models and compare them correspondingly: the uncorrelated stochastic DNS model with the independent AFNS model, and the correlated stochastic DNS model with the correlated AFNS model. With this, we can determine if the arbitrage-free models exhibit a significantly different behaviour in such portfolio optimization problems.

4.3.3 Comparison of Optimal Portfolios: Multiple Brownian Motions

In this section, we will solve the portfolio problem described in Section 4.2 for the case of multiple Brownian motions. With this, we can thereby analyze and compare the optimal results of the stochastic DNS models (*Case 2: uncorrelated Brownian motions* and *Case 3: correlated Brownian motions*) represented by (4.57) and (4.58), respectively, with that of the independent-factor and correlated-factor AFNS models represented by (4.53) and (4.54), respectively.

Since the treatment and methodology of the control problem will be the same for both the uncorrelated as well as the correlated Brownian motions, we combine them together to avoid redundancy. Thus, in this section, we shall focus on the control problem where the stochastic DNS $\beta(t)$ factors are described by (3.24) of Section 3.3.5, and repeated as follows:

$$\begin{pmatrix} d\beta_0(t) \\ d\beta_1(t) \\ d\beta_2(t) \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \left[\begin{pmatrix} a_0 \\ a_1(t) \\ a_2(t) \end{pmatrix} - \begin{pmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \end{pmatrix} \right] dt + \begin{pmatrix} b_0 & 0 & 0 \\ b_1\rho_{01} & b_1\tilde{\rho}_{11} & 0 \\ b_2\rho_{02} & b_2\tilde{\rho}_{21} & b_2\tilde{\rho}_{22} \end{pmatrix} \begin{pmatrix} dW_0(t) \\ dW_1(t) \\ dW_2(t) \end{pmatrix}$$

The influence of the three uncorrelated Brownian motions that drive the three DNS factors $\beta_0(t), \beta_1(t), \beta_2(t)$ are $W_0(t), W_1(t), W_2(t)$, respectively. In the above, we can observe that the case of uncorrelated models can be extracted when $\rho_{01} = \rho_{02} = \rho_{12} = 0$, thereby producing (3.19).

The portfolio problem in consideration is the same as described in Section 4.2. Let the investor be allowed to allocate shares in his portfolio that consists of the following assets:

1. one stochastic DNS bond whose price process is either $dP(t, T_1)$ (for our stochastic DNS framework) or $dP^{\text{AFNS}}(t, T_1)$ (for the AFNS framework)
2. one benchmark asset $dB((t)$ whose dynamics is defined by (4.5) and is the relatively less risky asset

Optimal Portfolios: Stochastic DNS Models

We begin with the optimal portfolio positions within our stochastic DNS framework. Let us assume that the investment horizon is T. The asset price processes in this case are as follows:

The bond price dynamics for a bond maturing at $T_1 > T$ is given by

$$dP(t, T_1) = P(t, T_1)\Bigg(M_P(t)dt + \sigma_0(t)dW_0(t) + \sigma_1(t)dW_1(t) + \sigma_2(t)dW_2(t) \Bigg) \tag{4.59}$$

where,

$$\begin{aligned} M_P(t)dt &= \beta_0(t)(1 + \mu(T_1 - t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T_1 - t)}(1 - \frac{\mu}{\lambda})\Big) \\ &\quad + \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T_1 - t)}(-\frac{\mu}{\lambda} - \mu(T_1 - t) + \lambda(T_1 - t))\Big) \\ &\quad - a_0(T_1 - t)\mu - a_1(t)(T_1 - t)g(t)\mu - a_2(t)(T_1 - t)h(t)\mu \\ &\quad + \frac{1}{2}(T_1 - t)^2\Big((b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})^2 + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})^2 \\ &\qquad\qquad + (h(t)b_2\tilde{\rho}_{22})^2\Big) \\ &=: f_0(t)\beta_0(t) + f_1(t)\beta_1(t) + f_2(t)\beta_2(t) + \zeta_P(t) \end{aligned}$$

where,

$$\begin{aligned} \zeta_P(t) &:= -a_0(T_1 - t)\mu - a_1(t)(T_1 - t)g(t)\mu - a_2(t)(T_1 - t)h(t)\mu \\ &\quad + \frac{1}{2}(T_1 - t)^2\Big((b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})^2 + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})^2 \end{aligned}$$

$$+ (h(t)b_2\tilde{\rho}_{22})^2\Big)$$

and

$$\begin{aligned}\sigma_0(t) &:= -(T_1 - t)(b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})\\ \sigma_1(t) &:= -(T_1 - t)(g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})\\ \sigma_2(t) &:= -(T_1 - t)(h(t)b_2\tilde{\rho}_{22})\end{aligned}$$

We also know that the benchmark asset is represented by

$$dB(t) = B(t)M_r(t)dt \tag{4.60}$$

In this case,

$$\begin{aligned}M_r(t) := M_P(t) - \zeta_P(t) &= f_0\beta_0(t) + f_1\beta_1(t) + f_2\beta_2(t)\\ &= \beta_0(t)(1 + \mu(T_1 - t)) + \beta_1(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T_1-t)}(1 - \frac{\mu}{\lambda})\Big)\\ &+ \beta_2(t)\Big(\frac{\mu}{\lambda} + e^{-\lambda(T_1-t)}(-\frac{\mu}{\lambda} - \mu(T_1 - t) + \lambda(T_1 - t))\Big)\end{aligned}$$

As we can observe, the $M_r(t)$ value is the same as the one obtained for the portfolio optimization under the single Brownian motion setting. The only difference arises from the terms inside the implicit risk premium of the bond $\zeta_P(t)$ in this case, as opposed to the $\zeta_1(t)$ of the previous case.

Now, using (4.14) to formulate the wealth process $X(t)$ of the investor, we obtain

$$\begin{aligned}dX(t) = X(t)\Bigg[&\Bigg(\pi(t)(M_P(t) - M_r(t)) + M_r(t)\Bigg)dt\\ &+ \pi(t)\Bigg(\sigma_0(t)dW_0(t) + \sigma_1(t)dW_1(t) + \sigma_2(t)dW_2(t)\Bigg)\Bigg]\\ = X(t)\Bigg[&\Bigg(\pi(t)\zeta_P(t) + f_0\beta_0(t) + f_1\beta_1(t) + f_2\beta_2(t)\Bigg)dt\\ &+ \pi(t)\Bigg(\sigma_0(t)dW_0(t) + \sigma_1(t)dW_1(t) + \sigma_2(t)dW_2(t)\Bigg)\Bigg]\end{aligned} \tag{4.61}$$

with $X(0) = x_0 \neq 0$.

The Control Problem and its Solution:

We may now apply the stochastic control methodology adapted to the stochastic DNS framework, as illustrated in Section 4.2. The corresponding state space formulation is:

$$
\begin{aligned}
Y(t) &= (X(t), \beta_0(t), \beta_1(t), \beta_2(t))' \\
\Lambda(t, x, \beta_0, \beta_1, \beta_2, \pi) &= (x(\pi\zeta_P + f_0\beta_0 + f_1\beta_1 + f_2\beta_2), \mu(a_0 - \beta_0), \mu(a_1 - \beta_1), \mu(a_2 - \beta_2))' \\
\Sigma(t, x, \beta_0, \beta_1, \beta_2, \pi) &= \begin{bmatrix} x\pi\sigma_0 & x\pi\sigma_1 & x\pi\sigma_2 \\ b_0 & 0 & 0 \\ b_1\rho_{01} & b_1\tilde{\rho}_{11} & 0 \\ b_2\rho_{02} & b_2\tilde{\rho}_{21} & b_2\tilde{\rho}_{22} \end{bmatrix} \\
\Sigma^*(t, x, \beta_0, \beta_1, \beta_2, \pi) &= \Sigma\Sigma' \\
&= \begin{bmatrix} x^2\pi^2(\sigma_0^2 + \sigma_1^2 + \sigma_2^2) & b_0x\pi\sigma_0 & b_1x\pi(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11}) & b_2x\pi\begin{matrix}(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} \\ +\sigma_2\tilde{\rho}_{22})\end{matrix} \\ b_0x\pi\sigma_0 & b_0^2 & b_0b_1\rho_{01} & b_0b_2\rho_{02} \\ b_1x\pi(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11}) & b_0b_1\rho_{01} & b_1^2(\rho_{01}^2 + \tilde{\rho}_{11}^2) & b_1b_2(\rho_{01}\rho_{02} + \tilde{\rho}_{11}\tilde{\rho}_{21}) \\ b_2x\pi\begin{matrix}(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} \\ +\sigma_2\tilde{\rho}_{22})\end{matrix} & b_0b_2\rho_{02} & b_1b_2(\rho_{01}\rho_{02} + \tilde{\rho}_{11}\tilde{\rho}_{21}) & b_2^2(\rho_{02}^2 + \tilde{\rho}_{21}^2 + \tilde{\rho}_{22}^2) \end{bmatrix} \\
A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi) &= G_t + 0.5\Big(x^2\pi^2(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)G_{xx} + 2b_0x\pi\sigma_0G_{x\beta_0} \\
&\quad + 2b_1x\pi(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11})G_{x\beta_1} + 2b_2x\pi(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} + \sigma_2\tilde{\rho}_{22})G_{x\beta_2} \\
&\quad + b_0^2G_{\beta_0\beta_0} + b_1^2(\rho_{01}^2 + \tilde{\rho}_{11}^2)G_{\beta_1\beta_1} + b_2^2(\rho_{02}^2 + \tilde{\rho}_{21}^2 + \tilde{\rho}_{22}^2)G_{\beta_2\beta_2} \\
&\quad + 2b_0b_1\rho_{01}G_{\beta_0\beta_1} + 2b_0b_2\rho_{02}G_{\beta_0\beta_2} + 2b_1b_2(\rho_{01}\rho_{02} + \tilde{\rho}_{11}\tilde{\rho}_{21})G_{\beta_1\beta_2}\Big) \\
&\quad + x(\pi\zeta_P + f_0\beta_0 + f_1\beta_1 + f_2\beta_2)G_x \\
&\quad + \mu(a_0 - \beta_0)G_{\beta_0} + \mu(a_1 - \beta_1)G_{\beta_1} + \mu(a_2 - \beta_2)G_{\beta_2}
\end{aligned} \tag{4.62}
$$

Now, the following HJB equation needs to be solved:

$$
\sup_{\pi \leq \delta} A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi) = 0
$$

$$
G(T, x, \beta_0, \beta_1, \beta_2) = \frac{1}{\gamma}x^\gamma
$$

where $\delta > 0$ and $\gamma \in (0, 1)$. The chosen utility function is $U(x) = \frac{1}{\gamma}x^\gamma$.

In comparison with the state space formulation of *Case 1: single Brownian motion*, we observe that the volatility matrix Σ of the control problem in (4.62) is a (4×3) matrix, as opposed to the single dimensional matrix (4×1) of (4.30). We conjecture that it will affect the final optimal process $\pi^*(t)$, especially in its corresponding correction term. Let us now solve the problem and check if this claim is true.

Here, we once again perform the three crucial steps in order to obtain the optimal portfolio $\pi^*(t)$, as detailed in Korn and Kraft (2002) and as solved previously in Section 4.2.

Step 1:

To obtain a suitable candidate for the optimal portfolio process, we differentiate $A^\pi G(t, x, \beta_0, \beta_1, \beta_2, \pi)$ of (4.62) under the max-operator with respect to π and set the gradient to zero. Assuming $G_{xx} < 0$, we obtain the following:

$$
0 = x^2\pi(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)G_{xx} + b_0x\sigma_0G_{x\beta_0}
$$

$$+ b_1 x(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11})G_{x\beta_1} + b_2 x(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} + \sigma_2\tilde{\rho}_{22})G_{x\beta_2} + G_x x\zeta_P$$

$$\text{or, } \pi^* = -\frac{b_0\sigma_0}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_0}}{G_{xx}} - \frac{b_1(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_1}}{G_{xx}} - \frac{b_2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_2}}{G_{xx}} - \frac{\zeta_P}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_x}{G_{xx}} \tag{4.63}$$

Let $\sigma_{\text{total}} \coloneqq \sigma_0^2 + \sigma_1^2 + \sigma_2^2$.

Step 2:

We plug (4.63) into the HJB equation above and drop its sup-operator. Then, we obtain

$$\begin{aligned}
0 = {} & G_t G_{xx} - 0.5\frac{b_0^2\sigma_0^2}{\sigma_{\text{total}}}G_{x\beta_0}^2 - 0.5\frac{b_1^2(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})^2}{\sigma_{\text{total}}}G_{x\beta_1}^2 - 0.5\frac{b_2^2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})^2}{\sigma_{\text{total}}}G_{x\beta_2}^2 \\
& - 0.5\frac{\zeta_P^2}{\sigma_{\text{total}}}G_x^2 - \frac{b_0 b_1\sigma_0(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})}{\sigma_{\text{total}}}G_{x\beta_0}G_{x\beta_1} - \frac{b_0 b_2\sigma_0(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})}{\sigma_{\text{total}}}G_{x\beta_0}G_{x\beta_2} \\
& - \frac{b_1 b_2(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})}{\sigma_{\text{total}}}G_{x\beta_1}G_{x\beta_2} \\
& + 0.5b_0^2 G_{\beta_0\beta_0}G_{xx} + 0.5b_1^2(\rho_{01}^2+\tilde{\rho}_{11}^2)G_{\beta_1\beta_1}G_{xx} + 0.5b_2^2(\rho_{02}^2+\tilde{\rho}_{21}^2+\tilde{\rho}_{22}^2)G_{\beta_2\beta_2}G_{xx} \\
& - \frac{b_0\sigma_0}{\sigma_{\text{total}}}\zeta_2 G_x G_{x\beta_0} - \frac{b_1(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})}{\sigma_{\text{total}}}\zeta_2 G_x G_{x\beta_1} - \frac{b_2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})}{\sigma_{\text{total}}}\zeta_2 G_x G_{x\beta_2} \\
& + b_0 b_1\rho_{01}G_{\beta_0\beta_1}G_{xx} + b_0 b_2\rho_{02}G_{\beta_0\beta_2}G_{xx} + b_1 b_2(\rho_{01}\rho_{02}+\tilde{\rho}_{11}\tilde{\rho}_{21})G_{\beta_1\beta_2}G_{xx} \\
& + x(f_0\beta_0 + f_1\beta_1 + f_2\beta_2)G_x G_{xx} \\
& + \mu(a_0-\beta_0)G_{\beta_0}G_{xx} + \mu(a_1-\beta_1)G_{\beta_1}G_{xx} + \mu(a_2-\beta_2)G_{\beta_2}G_{xx}
\end{aligned} \tag{4.64}$$

with $G(T, x, \beta_0, \beta_1, \beta_2) = \frac{1}{\gamma}x^\gamma$. Similar to the calculations made for the single Brownian motion, we can also separate the function G into its components as follows: $G(t, x, \beta_0, \beta_1, \beta_2) = f(t, \beta_0, \beta_1, \beta_2) \cdot \frac{1}{\gamma}x^\gamma$ where, $f(T, \beta_0, \beta_1, \beta_2) = 1$.

Furthermore, the function $f(t, \beta_0, \beta_2, \beta_2)$ can be separated into its components as follows: $f(t, \beta_0, \beta_2, \beta_2) = k(t)\exp\{(l_0(t)\beta_0 + l_1(t)\beta_1 + l_2(t)\beta_2)\}$ with terminal conditions $l_0(T) = l_1(T) = l_2(T) = 0$ and $k(T) = 1$.

By separating these functions into such sub-functions, we are able to substitute the partial derivatives and obtain simpler representations of the final differential equation. This methodology is the same as the one performed for the case of a single Brownian motion. For the sake of conciseness, we only show the final equation that contains all of the relevant partial derivatives in this case. This further allows us to obtain the form of the value function G. The equation is as follows, and is comparable to that of the case of a single Brownian motion (see (4.34)):

$$\begin{aligned}
0 = {} & k'(\gamma-1) + k(\gamma-1)(l_0'\beta_0 + l_1'\beta_1 + l_2'\beta_2) \\
& - k\Big(\mu\beta_0 l_0(\gamma-1) + \mu\beta_1 l_1(\gamma-1) + \mu\beta_2 l_2(\gamma-1)\Big) + k(f_0\beta_0 + f_1\beta_1 + f_2\beta_2)(\gamma-1)\gamma \\
& + k\Bigg(- 0.5\gamma\frac{b_0^2\sigma_0^2}{\sigma_{\text{total}}}l_0^2 - 0.5\gamma\frac{b_1^2(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})^2}{\sigma_{\text{total}}}l_1^2 - 0.5\gamma\frac{b_2^2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})^2}{\sigma_{\text{total}}}l_2^2
\end{aligned}$$

$$
\begin{aligned}
&- 0.5\frac{\zeta_P^2}{\sigma_{\text{total}}}\gamma - \frac{b_0 b_1 \sigma_0(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11})}{\sigma_{\text{total}}}\gamma l_0 l_1 - \frac{b_0 b_2 \sigma_0(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} + \sigma_2\tilde{\rho}_{22})}{\sigma_{\text{total}}}\gamma l_0 l_2 \\
&- \frac{b_1 b_2(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11})(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} + \sigma_2\tilde{\rho}_{22})}{\sigma_{\text{total}}}\gamma l_1 l_2 \\
&+ 0.5 b_0^2 l_0^2(\gamma - 1) + 0.5 b_1^2 l_1^2(\rho_{01}^2 + \tilde{\rho}_{11}^2)(\gamma - 1) + 0.5 b_2^2 l_2^2(\rho_{02}^2 + \tilde{\rho}_{21}^2 + \tilde{\rho}_{22}^2)(\gamma - 1) \\
&+ b_0 b_1 \rho_{01}(\gamma - 1) l_0 l_1 + b_0 b_2 \rho_{02}(\gamma - 1) l_0 l_2 + b_1 b_2(\rho_{01}\rho_{02} + \tilde{\rho}_{11} + \tilde{\rho}_{21})(\gamma - 1) l_1 l_2 \\
&- \frac{b_0\sigma_0\zeta_P}{\sigma_{\text{total}}}\gamma l_0 - \frac{b_1(\sigma_0\rho_{01} + \sigma_1\tilde{\rho}_{11})\zeta_P}{\sigma_{\text{total}}}\gamma l_1 - \frac{b_2(\sigma_0\rho_{02} + \sigma_1\tilde{\rho}_{21} + \sigma_2\tilde{\rho}_{22})\zeta_P}{\sigma_{\text{total}}}\gamma l_2 \\
&+ \mu a_0(\gamma - 1) l_0 + \mu a_1(\gamma - 1) l_1 + \mu a_2(\gamma - 1) l_2 \Bigg)_{=:c_2(t)}
\end{aligned}
\tag{4.65}
$$

In the above equation, the deterministic terms are represented by $c_2(t)$. Similar to the methodology made in Section 4.2, the second, third and fourth terms in (4.65) is zero after solving the ODEs for the following l_0, l_1, l_2, thereby ensuring that our ansatz for f is valid.

$$
\begin{aligned}
i.\quad & l_0(t) = \gamma(T - t) \\
ii.\quad & l_1(t) = \frac{\gamma}{\lambda}(1 - e^{-\lambda(T-t)}) \\
iii.\quad & l_2(t) = \frac{\gamma}{\lambda}(1 - e^{-\lambda(T-t)}(1 + \lambda(T - t)))
\end{aligned}
$$

Then, (4.65) becomes the following ODE:

$$
0 = k'(\gamma - 1) + k c_2(t) \tag{4.66}
$$

and can be solved by the method of integrating factors to yield

$$
k(t) = \exp\left\{\frac{1}{1-\gamma}(C_2(t) - C_2(T))\right\}
$$

where, $C_2(t)$ is the primitive of $c_2(t)$, defined as $C_2(t) := \int c_2(t)dt$. Hence, our candidate for the value function G is

$$
\begin{aligned}
G(t, x, \beta_0, \beta_1, \beta_2) &= \frac{1}{\gamma} x^\gamma k(t) \exp\{l_0(t)\beta_0(t) + l_1(t)\beta_1(t) + l_2(t)\beta_2(t)\} \\
&= \frac{1}{\gamma} x^\gamma \exp\left\{\frac{1}{1-\gamma}(C_2(t) - C_2(T))\right\} \\
&\cdot \exp\left\{\gamma\left((T - t)\beta_0(t) + \frac{1 - e^{-\lambda(T-t)}}{\lambda}\beta_1(t) + \frac{1 - e^{-\lambda(T-t)}(1 + \lambda(T - t))}{\lambda}\beta_2(t)\right)\right\}
\end{aligned}
\tag{4.67}
$$

From this we obtain the following partial derivatives,

$$
\begin{aligned}
G_x &= \frac{\gamma G}{x} \\
G_{xx} &= \frac{\gamma(\gamma - 1)G}{x^2}
\end{aligned}
$$

$$
\begin{aligned}
G_{x\beta_0} &= \frac{\gamma^2 G}{x}(T-t) \\
G_{x\beta_1} &= \frac{\gamma^2 G}{x}\frac{(1-e^{-\lambda(T-t)})}{\lambda} \\
G_{x\beta_2} &= \frac{\gamma^2 G}{x}\frac{(1-e^{-\lambda(T-t)}(1+\lambda(T-t)))}{\lambda}
\end{aligned}
$$

Inserting this into our candidate of optimal bond position that we derived in Step 1, we get **the optimal bond position $\pi^*(t)$ for stochastic DNS framework of multiple Brownian motions** to be:

$$
\begin{aligned}
\pi^*(t) &= -\frac{b_0\sigma_0}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_0}}{G_{xx}} - \frac{b_1(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_1}}{G_{xx}} \\
&\quad - \frac{b_2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_{x\beta_2}}{G_{xx}} - \frac{\zeta_P}{x(\sigma_0^2+\sigma_1^2+\sigma_2^2)}\frac{G_x}{G_{xx}} \\
&= \frac{1}{1-\gamma}\frac{\zeta_P(t)}{(\sigma_0^2+\sigma_1^2+\sigma_2^2)} \\
&\quad + \frac{\gamma}{1-\gamma}\cdot\Bigg(b_0\sigma_0(T-t) + b_1(\sigma_0\rho_{01}+\sigma_1\tilde{\rho}_{11})\frac{1-e^{-\lambda(T-t)}}{\lambda} \\
&\quad + b_2(\sigma_0\rho_{02}+\sigma_1\tilde{\rho}_{21}+\sigma_2\tilde{\rho}_{22})\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t)}{\lambda}\Bigg)\cdot\frac{1}{(\sigma_0^2+\sigma_1^2+\sigma_2^2)}
\end{aligned}
$$

On expanding $(\sigma_0^2+\sigma_1^2+\sigma_2^2)$ in the second term of $\pi^*(t)$, we get

$$
\begin{aligned}
\pi^*(t) &= \frac{1}{1-\gamma}\frac{\zeta_P(t)}{(\sigma_0^2+\sigma_1^2+\sigma_2^2)} \\
&\quad - \frac{\gamma}{1-\gamma}\cdot\Bigg(b_0(b_0+g(t)b_1\rho_{01}+h(t)b_2\rho_{02})(T-t) \\
&\quad + b_1((b_0+g(t)b_1\rho_{01}+h(t)b_2\rho_{02})\rho_{01} + (g(t)b_1\tilde{\rho}_{11}+h(t)b_2\tilde{\rho}_{21})\tilde{\rho}_{11})\frac{1-e^{-\lambda(T-t)}}{\lambda} \\
&\quad + b_2((b_0+g(t)b_1\rho_{01}+h(t)b_2\rho_{02})\rho_{02} + (g(t)b_1\tilde{\rho}_{11}+h(t)b_2\tilde{\rho}_{21})\tilde{\rho}_{21} + (h(t)b_2\tilde{\rho}_{22})\tilde{\rho}_{22}) \\
&\qquad \cdot\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t))}{\lambda}\Bigg) \\
&\quad \cdot\left(\frac{1}{(T_1-t)(b_0+g(t)b_1\rho_{01}+h(t)b_2\rho_{02})^2+(g(t)b_1\tilde{\rho}_{11}+h(t)b_2\tilde{\rho}_{21})^2+(h(t)b_2\tilde{\rho}_{22})^2}\right)
\end{aligned}
$$

Thus, we represent the optimal position $\pi^*(t)$ succinctly below:

$$
\pi^*(t) = \frac{1}{1-\gamma}\frac{\zeta_P(t)}{(\sigma_0^2+\sigma_1^2+\sigma_2^2)} - \frac{\gamma}{1-\gamma}\kappa_P(t) \tag{4.68}
$$

where,

$$\kappa_P(t) = \Bigg(b_0(b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})(T-t)$$
$$+ b_1((b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})\rho_{01} + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})\tilde{\rho}_{11})\frac{1-e^{-\lambda(T-t)}}{\lambda}$$
$$+ b_2((b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})\rho_{02} + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})\tilde{\rho}_{21} + (h(t)b_2\tilde{\rho}_{22})\tilde{\rho}_{22})$$
$$\cdot \frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t)}{\lambda}\Bigg)$$
$$\cdot \Bigg(\frac{1}{(T_1-t)(b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})^2 + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})^2 + (h(t)b_2\tilde{\rho}_{22})^2}\Bigg)$$

and,

$$\zeta_P(t) = -a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu$$
$$+ \frac{1}{2}(T_1-t)^2\Big((b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})^2 + (g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})^2$$
$$(h(t)b_2\tilde{\rho}_{22})^2\Big)$$
$$\sigma_0(t) = -(T_1-t)(b_0 + g(t)b_1\rho_{01} + h(t)b_2\rho_{02})$$
$$\sigma_1(t) = -(T_1-t)(g(t)b_1\tilde{\rho}_{11} + h(t)b_2\tilde{\rho}_{21})$$
$$\sigma_2(t) = -(T_1-t)(h(t)b_2\tilde{\rho}_{22})$$

Step 3:

Since the optimal bond position obtained here in (4.68) contains purely deterministic terms, the application of the verification theorem is exactly similar to the one made for the case of a single Brownian motion (see 85 ff.). The only difference is the fact that the current case considers three-dimensional Brownian motion, while the previous considered a single-dimensional one. However, since the proofs of Step 3 rely on a finite-dimensional Brownian motion, the results can be exactly applied here as well. Hence, we do not repeat it to avoid redundancy.

Optimal Portfolios: AFNS Models

Now, we solve the same portfolio problem using the AFNS models. We consider both the independent-factor and the correlated-factor AFNS models introduced by Christensen et al. (2011), and as represented in (4.53) and (4.54), respectively. Note that we can obtain the independent-factor model from the correlated-factor, by simply putting the correlation coefficients to zero (as done before).

The stochastic DNS bond has the following price process under the arbitrage-free framework:

$$dP^{\text{AFNS}}(t,T_1) = P^{\text{AFNS}}(t,T_1)\Big(M_P^{\text{AFNS}}(t)dt + \tilde{\sigma}_0(t)dW_0(t) + \tilde{\sigma}_1(t)dW_1(t) + \tilde{\sigma}_2(t)dW_2(t)\Big) \tag{4.69}$$

where,

$$
\begin{aligned}
M_P^{\text{AFNS}}(t) &= \beta_0(t)(1+\mu_0(T_1-t)) + \beta_1(t)\Big(\frac{\mu_1}{\lambda} + e^{-\lambda(T_1-t)}(1-\frac{\mu_1}{\lambda})\Big) \\
&\quad + \beta_2(t)\Big(\frac{\mu_2}{\lambda} + e^{-\lambda(T_1-t)}(-\frac{\mu_2}{\lambda} - \mu_2(T_1-t) + \lambda(T_1-t))\Big) \\
&\qquad - a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu \\
&\quad + A'(t,T_1) + \frac{1}{2}(T_1-t)^2\Big((\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})^2 + (g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})^2 \\
&\qquad\qquad (h(t)\tilde{b}_{22})^2\Big) \\
&=: \tilde{f}_0(t)\beta_0(t) + \tilde{f}_1(t)\beta_1(t) + \tilde{f}_2(t)\beta_2(t) + \zeta_P^{\text{AFNS}}(t)
\end{aligned}
$$

where,

$$
\begin{aligned}
\zeta_P^{\text{AFNS}}(t) &:= -a_0(T_1-t)\mu - a_1(t)(T_1-t)g(t)\mu - a_2(t)(T_1-t)h(t)\mu \\
&\quad + A'(t,T_1) + \frac{1}{2}(T_1-t)^2\Big((\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})^2 + (g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})^2 + (h(t)\tilde{b}_{22})^2\Big)
\end{aligned}
$$

From (4.50), we can calculate the time-derivative of the yield-adjustment term to be as follows:

$$
\begin{aligned}
A(t,T_1) &= \frac{1}{2}\sum_{j=1}^{3}\int_t^{T_1}\Big(\Sigma' B(s,T_1)B(s,T_1)'\Sigma\Big)_{j,j} ds \\
\implies A'(t,T_1) &= -\frac{1}{2}\sum_{j=1}^{3}\Big(\Sigma' B(s,T_1)B(s,T_1)'\Sigma\Big)_{j,j} \\
&= -\frac{1}{2}(T_1-t)^2\Big((\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})^2 + (g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})^2 + (h(t)\tilde{b}_{22})^2\Big)
\end{aligned}
\tag{4.70}
$$

Moreover, we notice that in comparison with the stochastic DNS model, there are slight differences with the corresponding functions of the AFNS model. Namely,

$$
\begin{aligned}
f_0(t) &\neq \tilde{f}_0(t) \\
f_1(t) &\neq \tilde{f}_1(t) \\
f_2(t) &\neq \tilde{f}_2(t) \\
\zeta_P(t) &\neq \zeta_P^{\text{AFNS}}(t)
\end{aligned}
$$

Here, the left-hand sides of the inequalities represent the functions observed in the stochastic DNS model and the right-hand sides represent the corresponding functions of the independent-factor AFNS model. The differences are highlighted in Remark 4.3.5. For the f_0, f_1, f_2 functions, the differences are due to the different mean-reversion levels and speeds. However, for the ζ_P function, the corresponding function for the AFNS model contains an extra term $A'(t,T_1)$ which we have calculated in (4.70).

Furthermore, for simplicity let,

$$
\tilde{\sigma}_0(t) = -(T_1-t)(\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})
$$

$$\tilde{\sigma}_1(t) = -(T_1 - t)(g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})$$
$$\tilde{\sigma}_2(t) = -(T_1 - t)(h(t)\tilde{b}_{22})$$

The benchmark asset has the following process:

$$dB(t) = B(t)M_r^{\text{AFNS}}(t)dt \tag{4.71}$$

In this case,

$$M_r^{\text{AFNS}}(t) := M_P^{\text{AFNS}}(t) - \zeta_P^{\text{AFNS}}(t) = \tilde{f}_0\beta_0(t) + \tilde{f}_1\beta_1(t) + \tilde{f}_2\beta_2(t) \tag{4.72}$$

Using (4.14) to formulate the wealth process $X(t)$ of the investor, we obtain

$$\begin{aligned} dX(t) = X(t)\Bigg[&\Bigg(\pi(t)\zeta_P^{\text{AFNS}}(t) + \tilde{f}_0\beta_0(t) + \tilde{f}_1\beta_1(t) + \tilde{f}_2\beta_2(t)\Bigg)dt \\ &+ \pi(t)\Bigg(\tilde{\sigma}_0(t)dW_0(t) + \tilde{\sigma}_1(t)dW_1(t) + \tilde{\sigma}_2(t)dW_2(t)\Bigg)\Bigg] \end{aligned} \tag{4.73}$$

with $X(0) = x_0 > 0$.

Remark 4.3.6. It can be observed, that on comparing (4.73) with (4.61), the overall structure of the wealth process remains the same. Hence, the methodology to obtain the optimal portfolio result is the same for both models.

The Controlled Solution:

Now that we have established the asset price and investor's wealth processes for the independent-factor AFNS model, we simply make the necessary substitutions of the corresponding functions in the control problem, as described in (4.62).

In order to avoid redundancy, we simply present the solution of the corresponding control problem for the case of the AFNS model.

Thus, on performing Steps 1, 2 and 3 as described in page 103 ff., we arrive at the final result. The optimal bond position in the case of the AFNS model is as follows:

$$\begin{aligned} &\pi^{*\text{AFNS}}(t) \\ &= \frac{1}{1-\gamma}\frac{\zeta_P^{\text{AFNS}}(t)}{(\tilde{\sigma}_0^2 + \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)} \\ &+ \frac{\gamma}{1-\gamma}\frac{\left(\tilde{b}_{00}\tilde{\sigma}_0(T-t) + (\tilde{b}_{10}\tilde{\sigma}_0 + \tilde{b}_{11}\tilde{\sigma}_1)\left(\frac{1-e^{-\lambda(T-t)}}{\lambda}\right) + (\tilde{b}_{20}\tilde{\sigma}_0 + \tilde{b}_{21}\tilde{\sigma}_1 + \tilde{b}_{22}\tilde{\sigma}_2)\left(\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t)}{\lambda}\right)\right)}{(\tilde{\sigma}_0^2 + \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)} \end{aligned}$$

which, after expanding the $(\tilde{\sigma}_0^2 + \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)$ part in the second term, can be restructured as follows:

$$\pi^{*\text{AFNS}}(t) = \frac{1}{1-\gamma}\frac{\zeta_P^{\text{AFNS}}(t)}{(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)} - \frac{\gamma}{1-\gamma}\kappa_P^{\text{AFNS}}(t) \tag{4.74}$$

where, the $\kappa_P^{\text{AFNS}}(t)$ term is

$$\kappa_P^{\text{AFNS}}(t) = \Bigg(\tilde{b}_{00}(\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})(T-t)$$
$$+ (\tilde{b}_{10}(\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20}) + \tilde{b}_{11}(g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})\frac{1-e^{-\lambda(T-t)}}{\lambda}$$
$$+ (\tilde{b}_{20}(\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20}) + \tilde{b}_{21}(g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21}) + \tilde{b}_{22}(h(t)\tilde{b}_{22}))\frac{1-e^{-\lambda(T-t)}(1+\lambda(T-t)}{\lambda}\Bigg)$$
$$\cdot \left(\frac{1}{(T_1-t)((\tilde{b}_{00} + g(t)\tilde{b}_{10} + h(t)\tilde{b}_{20})^2 + (g(t)\tilde{b}_{11} + h(t)\tilde{b}_{21})^2 + (h(t)\tilde{b}_{22})^2)}\right)$$

Remark 4.3.7 (Comparison Between the Optimal Positions)**.** The following observations can be made by comparing the formulae of the optimal portfolio position of the risky DNS bond in the case of our stochastic DNS framework, represented in (4.68), and that in the case of the arbitrage-free DNS framework represented in (4.74):

1. Both of the optimal results have the same correction term. It can be observed from the expansion of $\kappa_P(t)$ in (4.68) and $\kappa_P^{\text{AFNS}}(t)$ in (4.74) that

$$\kappa_P(t) = \kappa_P^{\text{AFNS}}(t)$$

2. Thus, the primary difference between the two optimal results, $\pi^*(t)$ and $\pi^{*\text{AFNS}}(t)$, lies within their first terms, due to their respective $\zeta_P(t)$ form. As we recall, the $\zeta_P(t)$ arises from the drift function of the respective DNS bond price process.

   Under the arbitrage-free framework, the $\zeta_P^{\text{AFNS}}(t)$ term contains an additional yield-adjustment term whose effect is visualized in the final optimal result. Hence, since

$$\zeta_P(t) \neq \zeta_P^{\text{AFNS}}(t)$$

   thus,

$$\pi^*(t) \neq \pi^{*\text{AFNS}}(t)$$

3. With the above two assertions, and given the fact that the general representation of the optimal position is similar in both cases, it is possible to infer that the behaviour of the optimal positions of the assets are similar in both the stochastic DNS framework as well as in the arbitrage-free DNS framework.

We visualize the third assertion made in Remark 4.3.7 in the following numerical example.

Numerical Example:

Let us now visualize the difference between $\pi^*(t)$ and $\pi^{*\text{AFNS}}(t)$ through a numerical example. We can then concretely determine whether the optimal portfolios obtained in each case exhibit different behaviours.

Consider the following parameters:

Let the bond maturity time be $T_1 = 30$ years, and the investment time horizon be $T = 25$ years. Let the factor loading decay parameter of the yield curve be $\lambda = 0.45$. Let the risk-aversion of the investor be $\gamma = 0.15$.

In order to make a comparison, we take as many similar parameters as we can for the two models. Hence, we consider the mean-reversion speed of the stochastic DNS model to be equal to those of the AFNS model, $\mu = \mu_0 = \mu_1 = \mu_2 = 0.45$. We also consider the mean-reversion level of each $\beta(t)$ factor to be equal in both models. So, $a_0 = \theta_0 = 0.023$, $a_1(t) = \theta_1 = 0.042$, $a_2(t) = \theta_2 = 0.028$. For the volatilities of the three $\beta(t)$ factors: $b_0 = 0.084$, $b_1 = 0.011$, $b_2 = 0.025$.

In Figure 4.3, we plot both $\pi^*(t)$ of the uncorrelated-factor stochastic DNS model and $\pi^{*\text{AFNS}}(t)$ of the independent-factor AFNS model. Note that the blue line represents $\pi^*(t)$ and the red line represents $\pi^{*\text{AFNS}}(t)$.

In the second figure, Figure 4.4, we plot both $\pi^*(t)$ of the correlated-factor stochastic DNS model and $\pi^{*\text{AFNS}}(t)$ of the correlated-factor AFNS model. We take the correlations between the three DNS $\beta(t)$ factors as follows: $\rho_{01} = 0.4$, $\rho_{02} = 0.4$, $\rho_{12} = 0.3$. Note that the corresponding volatility matrix for the AFNS model can be obtained by multiplying the correlation coefficients with the volatility coefficients of the stochastic DNS model (b_0, b_1, b_2).

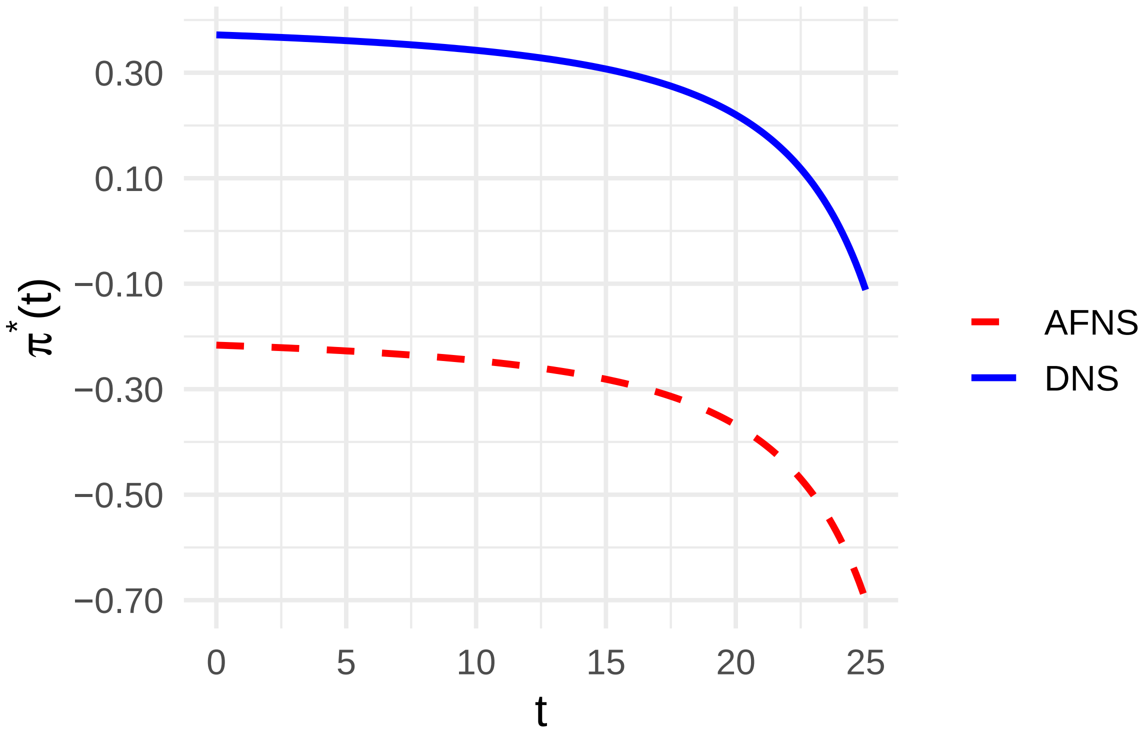


Figure 4.3: Comparison of optimal bond positions: Uncorrelated Brownian Motions

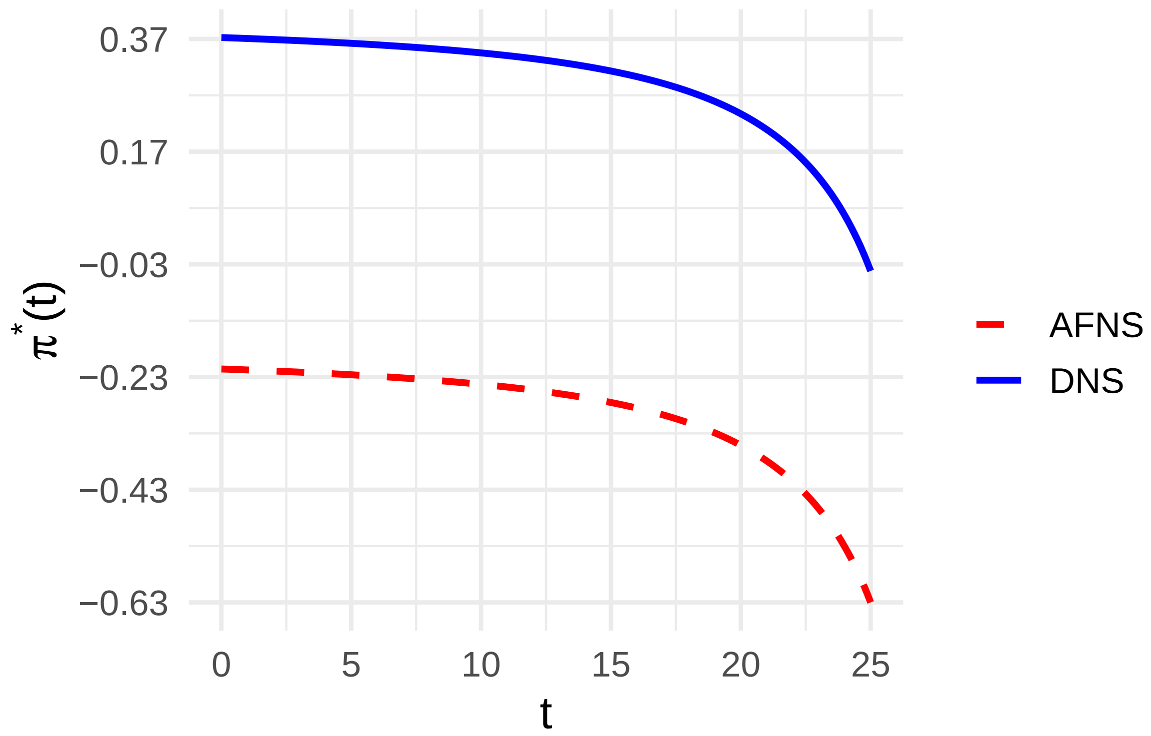


Figure 4.4: Comparison of optimal bond positions: Correlated Brownian Motions

As we can observe from both Figure 4.3 and Figure 4.4, the optimal position $\pi^*(t)$ under our stochastic DNS framework and the optimal position $\pi^{*\text{AFNS}}(t)$ under the arbitrage-free AFNS framework exhibit a similar behaviour. A reasonable risk-averse investor assumes the position of $\pi^*(t)$ for a risky asset, and as he approaches the end of the investment horizon, this position decreases as it becomes safer to invest in the less risky asset. This behaviour is consistent with reason because a risk-averse investor always tries to hedge against the additional risk induced by the three underlying stochastic DNS factors $\beta_0(t), \beta_1(t), \beta_2(t)$.

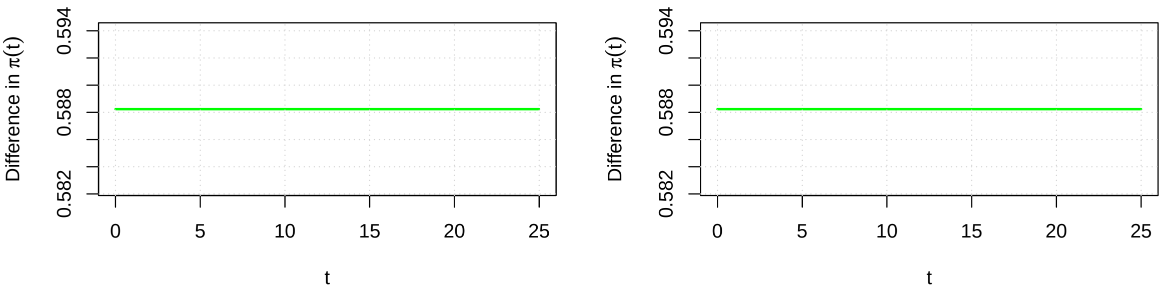
Furthermore, we observe that in the case of correlated Brownian motions, the optimal positions in both frameworks, namely the stochastic DNS and the AFNS frameworks, are higher in magnitude compared to their counterparts in the case of uncorrelated Brownian motions. All of the optimal positions, as observable, have finite upper bounds.

From Figure 4.3 and Figure 4.4, we can also observe that the only difference between the optimal $\pi^*(t)$ values is the respective magnitude of their values, with the stochastic DNS model exhibiting a higher optimal strategy as compared to that of the AFNS model. This difference can be calculated and plotted, as displayed in Figure 4.5. As we can observe, the difference in the optimal positions is constant and independent of time. By calculating the difference between $\pi^*(t)$ and $\pi^{*\text{AFNS}}(t)$, and using our observations of Remark 4.3.7, we obtain the difference to be

$$\frac{1}{1-\gamma} \cdot \frac{1}{2}$$

Hence, this constant difference is only dependent on the risk-aversion of the investor.

If the investor is more risk-averse, thereby implying a lower γ, the difference between the optimal positions under both frameworks decreases. On the other hand, for a more risk-taking investor, the difference increases as he is more willing to take his chance with our framework that has no additional arbitrage restrictions.



(a) Uncorrelated Brownian Motions (b) Correlated Brownian Motions

Figure 4.5: Difference in Optimal Positions Between Stochastic DNS and AFNS Frameworks

We summarize the key observations in the following remark:

Remark 4.3.8 (Interpretations of the Optimal $\pi^*(t)$ Under Both Frameworks)**.** In this section, we have constructed a comparative analysis of the optimal trading strategy $\pi^*(t)$ obtained under our stochastic DNS framework with that obtained under the so-called arbitrage-free Nelson-Siegel framework. On taking similar parameters for the analysis, we infer the following:

1. The behaviour of the optimal position under both frameworks is similar to each other. Furthermore, they are similar to the behaviour observed in the case of a single Brownian motion under our stochastic DNS framework (see Figure 4.2). This implies that there is no significant deviation in the mathematical as well as the economic interpretations of the optimal position throughout the investment horizon. Thus, our stochastic DNS framework is comparable with the results of the AFNS framework as well as with that of the existing analysis of similar portfolio problems discussed in Merton (1969) and Korn and Kraft (2002).

2. The economic behaviour of the optimal positions $\pi^*(t)$ can be interpreted as follows: An investor is more likely to hold a risky bond whose price process is governed by the stochastic DNS framework (or the AFNS framework) at the beginning of the investment duration in order to maximize his utility of the total portfolio wealth. As time approaches the end of the investment horizon T, the less risky asset (known in our analysis as the benchmark asset) whose position is given by $1 - \pi^*(t)$, offers a better solution. This is reasonable as a traditional risk-averse investor seeks to invest more heavily into the relatively less risky asset for short-term investing for a more stable wealth.

Since both frameworks depict the same economically-reasonable behaviour, we conclude that on keeping similar parameters, the investor does not suffer from choosing the stochastic DNS framework over the so-called arbitrage-free Nelson-Siegel framework.

4.4 Concluding Remarks

In this chapter, we extended the continuous-time portfolio problem of Merton (1969) by essentially considering three underlying stochastic factors in the yield curve that drive asset price processes and portfolio wealth dynamics. Using the methodology performed for one-factor short-rate models in Korn and Kraft (2002) and constructing an appropriate verification theorem in Theorem 4.1.14, we have showed that it possible to obtain valid optimal portfolio positions under the stochastic DNS yield curve framework. We thereby ensured that the results are consistent with theoretical existence and uniqueness results established in classical stochastic control literature. Furthermore, we demonstrated through an analysis of the optimal results (see Section 4.2.1, Remark 4.2.2, and Remark 4.3.8) how these optimal positions also satisfy previously-established economic interpretations of optimal investment behaviour for similar portfolio problems.

We can now construct the following theorem summarizing our key results, whose proof can be observed in all of the calculations made in Section 4.2 and Section 4.3.3:

Theorem 4.4.1 (Optimal Portfolios under the Stochastic DNS Framework)**.**
Let there be an investor who operates under our stochastic dynamic Nelson-Siegel framework. The three stochastic $\beta_0(t), \beta_1(t), \beta_2(t)$ factors are driven by either a single or multiple underlying Brownian motions. Assume that the investor holds an admissible portfolio of assets constructed under this framework, as defined in Definition 4.1.8, and wishes to maximize his utility of the final portfolio wealth. Let the position in the stochastic risky bond and in the relatively less-risky benchmark asset be $\pi(t)$ and $1 - \pi(t)$, respectively. Then, the optimal positions in the risky bond $\pi^*(t)$ are calculated to be as follows:

If the stochastic DNS framework is driven by a single Brownian motion (see Section 4.2.1), then

$$\pi^*(t) = \frac{1}{1-\gamma}\frac{\zeta(t)}{\sigma(t)^2} - \frac{\gamma}{1-\gamma}\kappa(t)$$

and if the stochastic DNS framework is driven by three (either uncorrelated or correlated) Brownian motions (see Section 4.3.3), then

$$\pi^*(t) = \frac{1}{1-\gamma}\frac{\zeta(t)}{(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)} - \frac{\gamma}{1-\gamma}\kappa(t)$$

The respective $\zeta(t)$ and $\kappa(t)$ representations depend on the considered case of Brownian motions, and can be found in (4.36) and (4.68), respectively. Furthermore, the existence and uniqueness of the optimal $\pi^*(t)$ is verified by Theorem 4.1.14.

Remark 4.4.2. In this remark, we state how the above Theorem 4.4.1 and the results obtained in this chapter enable us to justify the significance of solving such portfolio problems under the stochastic DNS framework.

1. **Mathematical Tractability.** We show that by establishing a suitable theoretical foundation that uses the classical results of stochastic control theory and adapts them to the portfolio optimization problem, we are able to construct a valid control problem that produces a unique and valid optimal solution. We show that by

imposing sufficient conditions on assets defined in this framework, it is possible to obtain valid portfolio solutions. With the help of our verification theorem, Theorem 4.1.14, we are able to prove the existence and uniqueness of the bounded and optimal portfolio positions of Theorem 4.4.1.

2. **Economic-Consistent Interpretations.** Additionally, with the mathematical analysis of the optimal results, as well as with the help of graphical numerical examples, we demonstrate how these optimal portfolio positions in the risky bond corroborates with the investment strategy of an investor with a measurable risk-aversion and wishes to maximize his utility of the total portfolio wealth.

3. **Validity despite the so-called Lack of Consistency.** We demonstrated how the long-standing criticism of the lack of consistent bond prices in the dynamic Nelson-Siegel family (often referred to in literature as the lack of arbitrage in the Nelson-Siegel family) does not create any negative repercussions on our portfolio problem. In fact, we prove that our optimal positions have a finite upper bound, thereby removing any consequent arbitrage opportunity. Moreover, the behaviour of our optimal results throughout the investment horizon is consistent with those obtained under the so-called arbitrage-free framework of the AFNS models proposed by Christensen et al. (2011).

Thus, we conclude this chapter on the analysis of portfolio optimization problems under the stochastic dynamic Nelson-Siegel framework, and state that the results hold significance in both literature and in practice.

Chapter 5

Summary and Outlook

In this final chapter, we shall summarize the key analyses performed and the main results obtained in this thesis.

Summary of Key Findings

In Chapter 2, we established a preliminary base to the concept of term structure models. With this backdrop, we were able to introduce the significance of the parameterized dynamic Nelson-Siegel (DNS) model as arguably one of the most popular term structure models used across all domains in the world of finance and economics.

The numerous technical strengths of the DNS model including its high empirical success in yield forecasting and intuitive economic interpretations of its model parameters have been well-established in literature. On the other hand, the theoretical examination of the dynamic Nelson-Siegel model has not been adequately discussed in literature pertinent to yield curve models. This acted as the prime motivation to dissect the model mathematically in both Chapter 3 and Chapter 4.

In Chapter 3, we first provided a detailed and elementary proof of existence of all possible yield curve shapes (normal, inverse, humped, and dipped) that can be attained by the Nelson-Siegel family. We also proved the necessary conditions required to obtain each shape. This is a significant finding as it provides a theoretical justification to an empirical fact of the model that has not only been observed in practice but has also been stated (but not proved) in the original papers of Nelson and Siegel (1987) and Diebold and Li (2006). With our elementary proof, we hope to substantiate their assertions mathematically.

Our second contribution in Chapter 3 was developing and thoroughly analyzing realistic stochastic models for the DNS family. By doing so, we offer an alternative to the deterministic $\beta(t)$ factors often provided by central banks around the world. Our stochastic models consider that these three $\beta(t)$ factors of the DNS are driven by individual stochastic OU processes, which then result in stochastic yield curve dynamics. We considered three cases elaborately: the influence of a single Brownian motion, of three uncorrelated Brownian motions, and of three pairwise correlated Brownian motions. By treating three separate cases individually, we were able to obtain analytically tractable yield curve dynamics that accounted for the most common behaviours of yield curves, and thereby captured realistic scenarios. These realistic behaviours include demonstrating the long-term normal shape of yield curves, the higher volatility of the yield curve at the short-end compared to the long-end, the persistence of the level factor $\beta_0(t)$ when compared to the other two factors, the presence of negative interest rates, and the parallel shift of the yield curve. Our framework also allows the $\beta(t)$ factors to retain their interpretations as the level, slope, and curvature factors of the yield curve.

Furthermore, with our stochastic DNS framework, we were able to reproduce the original four attainable shapes of the DNS family. We demonstrated a suitable methodology and calculated explicit formulae of the probabilities of attaining each of these four shapes. Our formulae showed that the probabilities highly depend on the chosen parameters of the stochastic framework, and explicit values can then be obtained through common numerical methods.

In Chapter 4, we demonstrated how our established stochastic DNS framework can be applied to one of the most common problems in financial mathematics: the continuous-time portfolio optimization problem. Our first contribution in this chapter was to build a solid theoretical foundation of assets prices which are modeled on the stochastic DNS framework. We constructed a stochastic DNS bond whose price processes were proved in Chapter 3 to satisfy $P(T,T) = 1$. Then, we constructed a low-risk alternative to this asset and called it a benchmark asset whose price process can be interpreted as a rolling-horizon of short-term maturity bonds. In this manner, we demonstrated how such assets modeled on the stochastic DNS framework offer investors the option to construct an admissible portfolio of a mixture of risky stochastic bonds and more-stable counterparts.

With this, we then employed the stochastic control methodology established in Korn and Kraft (2002) in order to solve portfolio problems under the stochastic DNS framework. Thus, for our second contribution in this chapter, we constructed a suitable theoretical framework that adapted classical results of stochastic control theory as well as their well-established applications in portfolio optimization literature. We constructed an appropriate verification theorem that proved the existence and uniqueness of results obtained within the stochastic DNS framework. We demonstrated the validity of both the portfolio control problem as well as the optimal obtained position for the case of a single Brownian motion through a three-step approach. Furthermore, with an analysis of our obtained results, we demonstrated how the optimal portfolio investment decision resembles the realistic behaviour of an investor with risk-aversion.

Given the mathematical and economic validity of our optimal portfolio results under the influence of a single Brownian motion, we proceeded to perform an excursion to investigate the impact of a long-standing criticism of the Nelson-Siegel family on such optimal investment problems, namely the lack of consistent bond prices. It is often discussed how the dynamic Nelson-Siegel model fails to produce bond prices that are consistent under a martingale measure. The primary criticism then boils down to how the price process of a DNS bond constructed under a martingale measure at a time t no longer belongs to the same DNS family at a time $t + 1$. This is commonly referred to as the lack of arbitrage in the Nelson-Siegel family. To address this, a class of affine arbitrage-free Nelson-Siegel (AFNS) models were developed by Christensen et al. (2011). Thus, we decided to test the impact of the lack of the arbitrage within our stochastic framework (that does not use the AFNS specifications) on optimal investment decision behaviours.

This formed the basis of our final contribution in this chapter and thesis. We performed the stochastic control methodology as discussed earlier and solved for the optimal portfolio positions of an investor investing within our stochastic DNS framework. We assumed that both uncorrelated and correlated multiple Brownian motions could influence the underlying $\beta(t)$ factors, and obtained optimal results that also satisfied our verification theorem. Then, we performed the same methodology for the AFNS frameworks, and compared the two optimal results. Our comparative analysis demonstrated that not only did our obtained optimal portfolio exhibit a finite upper bound throughout the investment horizon (a feature that was also seen in the case of a single Brownian motion), but also the general functions representing the optimal results in both frameworks were similar. Coupled with the fact that the difference in magnitudes of the optimal positions (with ours having the higher value) was constant and time-independent, we were able to infer that the investor does not suffer from unfair unbounded opportunities in our stochastic DNS framework. As Diebold and Li (2006) mentioned how their model, which is not arbitrage-free, can produce good forecasts, our comprehensive analysis is able to show that the stochastic DNS framework, which is not arbitrage-free, can be good representatives of such portfolio problems.

Outlook

With our extensive analysis of this stochastic framework of the dynamic Nelson-Siegel model, we hope to impart a productive contribution to the relevant literature as well as to practitioners of yield curve models. We also realize the scope of further possible extensions to this work, and hope to discuss them in the near future.

Due to the highly parameterized framework that dictates the behaviour of the yield curves and of the subsequent asset price processes modeled on it, we believe that a proper calibration of parameters with existing central bank data may enable us to obtain a realistic range of parameter values. This shall further solidify our probability formulae. Moreover, the analyses performed in this thesis can be extended to include the term structure model of Svensson (1995). This yield curve model extends the existing DNS formula with the addition of a second curvature factor $\beta_3(t)$ which is influenced by an extra parameter λ_2. This discussion may prove how well our interpretations and calculations hold up in the presence of generalized multi-factor parametric yield curve models.

Appendix A

Appendix to Chapter 3

Descriptive Statistics of Generated Yield Curves: Single Brownian Motion

	Mean	SD	CV	Min	Max
Yield (6 months)	1.8219	3.1000	1.7015	-4.2664	8.2819
Yield (10 years)	2.1326	1.5058	0.7061	-1.0102	5.3457
Yield (20 years)	2.1651	1.0368	0.4789	-0.0532	4.3902
$\beta_0(t)$	0.0220	0.0056	0.2548	0.0095	0.0339
$\beta_1(t)$	-0.0044	0.0265	-6.0456	-0.0533	0.0492
$\beta_2(t)$	0.0012	0.0236	20.3471	-0.0469	0.0511

Table A.1: For $\mu = 0.5$ Descriptive statistics of generated yield curves: Single Brownian Motion

	Mean	SD	CV	Min	Max
Yield (6 months)	1.9632	2.8073	1.4299	-3.2879	7.7831
Yield (10 years)	2.1277	1.3723	0.6450	-0.6241	5.0418
Yield (20 years)	2.1347	0.9492	0.4447	0.1786	4.1597
$\beta_0(t)$	0.0214	0.0052	0.2432	0.0102	0.0324
$\beta_1(t)$	-0.0022	0.0239	-10.8901	-0.0439	0.0458
$\beta_2(t)$	0.0016	0.0215	13.4530	-0.0402	0.0468

Table A.2: For $\mu = 0.75$ Descriptive statistics of generated yield curves: Single Brownian Motion

Descriptive Statistics of Generated Yield Curves: Uncorrelated Brownian Motions

	Mean	SD	CV	Min	Max
Yield (6 months)	2.1076	1.0051	0.4769	-0.8722	4.4528
Yield (10 years)	2.6471	0.5518	0.2085	0.9442	3.8691
Yield (20 years)	2.6525	0.4809	0.1813	1.1826	3.5476
$\beta_0(t)$	0.0265	0.0046	0.1717	0.0133	0.0338
$\beta_1(t)$	-0.0070	0.0083	-1.1870	-0.0276	0.0169
$\beta_2(t)$	0.0069	0.0123	1.7741	-0.0213	0.0373

Table A.3: For $\mu = 0.5$ Descriptive statistics of generated yield curves: Uncorrelated Brownian Motions

	Mean	SD	CV	Min	Max
Yield (6 months)	2.0513	1.0183	0.4964	-0.8721	4.5525
Yield (10 years)	2.4517	0.5391	0.2199	0.8303	3.5967
Yield (20 years)	2.4544	0.4743	0.1932	1.0599	3.3174
$\beta_0(t)$	0.0245	0.0045	0.1835	0.0119	0.0324
$\beta_1(t)$	-0.0052	0.0087	-1.6898	-0.0277	0.0202
$\beta_2(t)$	0.0053	0.0108	2.0572	-0.0211	0.0301

Table A.4: For $\mu = 0.75$ Descriptive statistics of generated yield curves: Uncorrelated Brownian Motions

Additional Proofs of Probability Distribution of Yield Curve Shapes

We present the proofs of the probabilities of obtaining inverse, humped, and dipped yield curves in the case of a single Brownian motion influencing all three dynamic Nelson-Siegel factors, $\beta_0(t)$, $\beta_1(t)$, and $\beta_2(t)$.

Inverse Curves

For P_{inv} (see (3.31)):

$$\begin{aligned} P_{\text{inv}} &= P(\beta_1(t) \geq |\beta_2(t)|) \\ &= P\left(\beta_1(t) > \beta_2(t) \,\middle|\, \beta_2(t) > 0\right) \cdot P(\beta_2(t) > 0) \\ &\quad + P\left(\beta_1(t) > -\beta_2(t) \,\middle|\, \beta_2(t) < 0\right) \cdot P(\beta_2(t) < 0) \end{aligned}$$

For $P\left(\beta_1(t) > \beta_2(t) \,\middle|\, \beta_2(t) > 0\right) \cdot P(\beta_2(t) > 0)$, we know that

$$\beta_2(t) > 0 \implies \int_0^t e^{-\mu(t-u)} dW(u) > \underbrace{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}_{=:z_1}$$

and

$$\begin{aligned}
&\beta_1(t) > \beta_2(t) \,\Big|\, \beta_2(t) > 0 \\
&\implies \beta_1(t) - \beta_2(t) > 0 \,\Big|\, \beta_2(t) > 0 \\
&\implies \int_0^t e^{-\mu(t-u)} dW(u) > \underbrace{\frac{-1}{b_1 - b_2}(\beta_1(0) - \beta_2(0))e^{-\mu t}}_{=:y_3}
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int_0^t e^{-\mu(t-u)} dW(u) > \max(z_1, y_3) \\
\implies P(\beta_1(t) > \beta_2(t) \,\Big|\, \beta_2(t) > 0) &= 1 - \Phi\left(\frac{\max(z_1, y_3)}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)
\end{aligned}$$

And for $P(\beta_1(t) > -\beta_2(t) \,\Big|\, \beta_2(t) < 0) \cdot P(\beta_2(t) < 0)$, we know that

$$\beta_2(t) < 0 \implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}_{z_1}$$

and

$$\begin{aligned}
&\beta_1(t) > -\beta_2(t) \,\Big|\, \beta_2(t) < 0 \\
&\implies \beta_1(t) + \beta_2(t) > 0 \,\Big|\, \beta_2(t) < 0 \\
&\implies \int_0^t e^{-\mu(t-u)} dW(u) > \underbrace{\frac{1}{b_1 + b_2}(\beta_1(0) + \beta_2(0))e^{-\mu t}}_{*}
\end{aligned}$$

Hence,

$$\begin{aligned}
&* < \int_0^t e^{-\mu(t-u)} dW(u) < z_1 \quad (\text{iff, } * < z_1) \\
&\implies P(\beta_1(t) > -\beta_2(t) \,\Big|\, \beta_2(t) < 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0) + \beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1 - e^{-2\mu t})}}\right)\right]\right)
\end{aligned}$$

Hence, the total probability is

$$\begin{aligned}
P_{\text{inv}} &= P(\beta_1(t) \geq |\beta_2(t)|) \\
&= P(\beta_1(t) > \beta_2(t) \,\Big|\, \beta_2(t) > 0) \cdot P(\beta_2(t) > 0) \\
&\quad + P(\beta_1(t) > -\beta_2(t) \,\Big|\, \beta_2(t) < 0) \cdot P(\beta_2(t) < 0)
\end{aligned}$$

$$= \left(1 - \Phi\left(\frac{\max(z_1, y_3)}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right) \cdot \Phi\left(\frac{\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right)$$
$$+ \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_2}\beta_2(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \cdot \Phi\left(\frac{-\beta_2(0)e^{-\mu t}}{\sqrt{\frac{b_2^2}{2\mu}(1-e^{-2\mu t})}}\right)$$

Humped Curves

For P_{hump} (see (3.32)):

$$\begin{aligned} P_{\text{hump}} &= P(|\beta_1(t)| \le \beta_2(t)) \\ &= P(\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) > 0) \cdot P(\beta_1(t) > 0) \\ &\quad + P(-\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) < 0) \cdot P(\beta_1(t) < 0) \end{aligned}$$

For $P(\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) > 0) \cdot P(\beta_1(t) > 0)$, we know that

$$\beta_1(t) > 0 \implies \int_0^t e^{-\mu(t-u)} dW(u) > \underbrace{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}_{=:z_2}$$

and

$$\begin{aligned} &\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) > 0 \\ &\implies \beta_1(t) - \beta_2(t) < 0 \\ &\implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_1 - b_2}(\beta_1(0) - \beta_2(0))e^{-\mu t}}_{y_3} \end{aligned}$$

Hence, iff $z_2 < y_3$, the above conditional holds, otherwise we get zero. Thus,

$$\begin{aligned} &z_2 < \int_0^t e^{-\mu(t-u)} dW(u) < y_3 \\ &\implies P(\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) > 0) \\ &= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{(b_1-b_2)}(\beta_1(0)-\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \end{aligned}$$

And for $P(-\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) < 0) \cdot P(\beta_1(t) < 0)$, we know that

$$\beta_1(t) < 0 \implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}_{z_2}$$

and

$$-\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) < 0$$

$$\begin{aligned}&\implies \beta_1(t)+\beta_2(t)>0\\&\implies \int_0^t e^{-\mu(t-u)}dW(u) > \underbrace{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}_{**}\end{aligned}$$

The above conditional holds iff $** < z_2$, otherwise we get zero. Thus,

$$\begin{aligned}&P(-\beta_1(t)<\beta_2(t)\,\Big|\,\beta_1(t)<0)\\&\implies \max\left(0,\left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)-\Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right)\end{aligned}$$

Hence, the total probability is

$$\begin{aligned}P_{\text{hump}} &= P(|\beta_1(t)|\le\beta_2(t))\\&= P(\beta_1(t)<\beta_2(t)\,\Big|\,\beta_1(t)>0)\cdot P(\beta_1(t)>0)\\&\quad + P(-\beta_1(t)<\beta_2(t)\,\Big|\,\beta_1(t)<0)\cdot P(\beta_1(t)<0)\\&= \max\left(0,\left[\Phi\left(\frac{\frac{-1}{(b_1-b_2)}(\beta_1(0)-\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)-\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right)\cdot\Phi\left(\frac{\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right)\\&\quad+\max\left(0,\left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)-\Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right)\cdot\Phi\left(\frac{-\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right)\end{aligned}$$

Dipped Curves

For P_{dip} (see (3.33)):

$$\begin{aligned}P_{\text{dip}} &= P(-|\beta_1(t)|\ge\beta_2(t))\\&= P(-\beta_1(t)>\beta_2(t)\,\Big|\,\beta_1(t)>0)\cdot P(\beta_1(t)>0)\\&\quad + P(\beta_1(t)>\beta_2(t)\,\Big|\,\beta_1(t)<0)\cdot P(\beta_1(t)<0)\end{aligned}$$

For $P(-\beta_1(t)>\beta_2(t)\,\Big|\,\beta_1(t)>0)\cdot P(\beta_1(t)>0)$, we know that

$$\beta_1(t)>0 \implies \int_0^t e^{-\mu(t-u)}dW(u) > \underbrace{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}_{z_2}$$

and

$$\begin{aligned}&-\beta_1(t)>\beta_2(t)\,\Big|\,\beta_1(t)>0\\&\implies \beta_1(t)+\beta_2(t)<0\end{aligned}$$

$$\implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_1+b_2}(\beta_1(0)+\beta_2(0))e^{-\mu t}}_{**}$$

Hence, iff $z_2 < **$, the above conditional holds, otherwise we get zero. Thus,

$$\begin{aligned}
&z_2 < \int_0^t e^{-\mu(t-u)} dW(u) < ** \\
&\implies P(-\beta_1(t) < \beta_2(t) \,\Big|\, \beta_1(t) > 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right)
\end{aligned}$$

And for $P(\beta_1(t) > \beta_2(t) \,\Big|\, \beta_1(t) < 0) \cdot P(\beta_1(t) < 0)$, we know that

$$\beta_1(t) < 0 \implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}_{z_2}$$

and

$$\begin{aligned}
&\beta_1(t) > \beta_2(t) \,\Big|\, \beta_1(t) < 0 \\
&\implies \beta_1(t) - \beta_2(t) > 0 \\
&\implies \int_0^t e^{-\mu(t-u)} dW(u) < \underbrace{\frac{-1}{b_1-b_2}(\beta_1(0)-\beta_2(0))e^{-\mu t}}_{y_3}
\end{aligned}$$

The above conditional holds iff $y_3 < z_2$, otherwise it is zero. Hence,

$$\begin{aligned}
&y_3 < \int_0^t e^{-\mu(t-u)} dW(u) < z_2 \\
&\implies P(\beta_1(t) > \beta_2(t) \,\Big|\, \beta_1(t) < 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1-b_2}(\beta_1(0)-\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right)
\end{aligned}$$

Hence, the total probability is

$$\begin{aligned}
P_{\text{dip}} &= P(-|\beta_1(t)| \geq \beta_2(t)) \\
&= P(-\beta_1(t) > \beta_2(t) \,\Big|\, \beta_1(t) > 0) \cdot P(\beta_1(t) > 0) \\
&\quad + P(\beta_1(t) > \beta_2(t) \,\Big|\, \beta_1(t) < 0) \cdot P(\beta_1(t) < 0) \\
&= \max\left(0, \left[\Phi\left(\frac{\frac{-1}{(b_1+b_2)}(\beta_1(0)+\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \cdot \Phi\left(\frac{\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right) \\
&\quad + \max\left(0, \left[\Phi\left(\frac{\frac{-1}{b_1}\beta_1(0)e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right) - \Phi\left(\frac{\frac{-1}{b_1-b_2}(\beta_1(0)-\beta_2(0))e^{-\mu t}}{\sqrt{\frac{1}{2\mu}(1-e^{-2\mu t})}}\right)\right]\right) \cdot \Phi\left(\frac{-\beta_1(0)e^{-\mu t}}{\sqrt{\frac{b_1^2}{2\mu}(1-e^{-2\mu t})}}\right)
\end{aligned}$$

Appendix B

Appendix to Chapter 4

Krylov (1980, Theorem 7)

The following result from Krylov (1980, Theorem 7, p. 82) is used to justify the assertions made in Proposition 4.1.12 of Chapter 4 (see page 73).

Theorem B.1. Let there be a stochastic integral equation as follows:

$$x_t = x_0 + \int_0^t b(s, x_s)ds + \int_0^t \sigma(s, x_s)dW(s)$$

for any $t \in [0, T]$, $T > 0$. Let

$$\|\sigma(t, x) - \sigma(t, y)\| \leq K|x - y|$$
$$|b(t, x) - b(t, y)| \leq K^2|x - y|$$

for $K > 0$. Further, let

$$\mathbb{E}\int_0^T \left(|x_0|^2 + |b(s, 0)|^2 + \|\sigma(s, 0)\|^2\right)ds < \infty$$

Then, for any $t \leq T$, there exists a solution x_t such that $\mathbb{E}\int_0^T |x_t|^2 dt < \infty$. If x_t and y_t are two solutions of the stochastic integral equation, then $\mathbb{P}\{\sup_{t\in[0,T]} |x_t - y_t| > 0\} = 0$.

Proof. Let there be an operator I defined as

$$Ix_t = \int_0^t b(s, x(s))ds + \int_0^t \sigma(s, x(s))dW(s)$$

Let V be a space of progressively measurable processes x_t with values in $\mathbb{R}^d$ such that

$$\|x_t\| = \left(\mathbb{E}\int_0^T |x_t|^2 dt\right)^{\frac{1}{2}} < \infty$$

Then, operator I maps V into V. Furthermore, for any $\alpha := 2K^2(1 + TK^2)$,

$$\mathbb{E}|Ix_t - Iy_t|^2 \leq \alpha\mathbb{E}\int_0^t |x_s - y_s|^2 ds$$

Let $x_t^{(0)} \equiv 0$ and $x_t^{(n+1)} = x_0 + Ix_t^{(n)}$ for any $n = 0, 1, 2, \ldots$, then

$$\mathbb{E}\left|x_t^{(n+1)} - x_t^{(n)}\right|^2 \leq \alpha\mathbb{E}\int_0^t \left|x_s^{(n)} - x_s^{(n-1)}\right|^2 ds$$

Iterating the above inequality, we get

$$\left\|x_t^{(n+1)} - x_t^{(n)}\right\|^2 \leq \frac{T^n\alpha^n}{n!}\left\|x_t^{(1)}\right\|^2$$

Since the RHS of the above inequality converges, the functions $x_t^{(n+1)}$ converge in V. There exists a process $\tilde{x}_t \in V$ such that $\left\|x_t^{(n)} - \tilde{x}_t\right\| \to 0$ as $n \to \infty$. Further,

$$\|Ix_t - Iy_t\| \leq \alpha T\|x_t - y_t\|$$

Thus, the operator I is continuous in V. It follows thta

$$I\tilde{x}_t = I(x_0 + I\tilde{x}_t)_t$$

for almost all $t, \omega \in W$.Since both sides of this equality are continuous with respect to t, they coincide for all t (a.s). Taking $x_t = x_0 + I\tilde{x}_t$, we have $x_t = x_0 + Ix_t$ for all t (a.s).

Thus, x_t is a solution to the given stochastic integral equation.

Moreover, $\mathbb{E}|x_t - y_t|^2 = 0$ for each t, and the process $x_t - y_t$ can be represented as the sum of stochastic integrals and ordinary integrals. Hence, it is continuous almost surely. The equality $x_t = y_t$ (a.s.) for each t implies that it holds true for all t (a.s.). Thus, $\mathbb{P}\{\sup_{t\in[0,T]} |x_t - y_t| > 0\} = 0$ holds true. □

Gikhman and Skorohod (1972, Theorem 3)

The following result from Gikhman and Skorohod (1972, Theorem 3, pp. 136 ff.) and is used to justify the assertions of Proposition 4.1.12 in Chapter 4 (see page 73).

Theorem B.2. A stochastic integral equation satisfying

$$x(t) = x(0) + \int_0^t b(s, x(s))ds + \int_0^t \sigma(s, x(s))dW(s)$$

possesses a unique solution in $Q \subset Q_0$ defined for all $t \in [0, T]$. This solution has the properties

$$\mathbb{E}(\sup_{s\in[0,T]} |x(s)|^2) \leq A(1 + \|\varphi\|^2)$$

and

$$\mathbb{E}(\sup_{s\in[t,t+h]} |x(s) - x(t)|^2) \leq B(1 + \|\varphi\|^2 + \sup_{u\in[t,a]} |x(u)|^2)h$$

where, A and B are constants that only depend on C, T, K.

Proof. The first part of the theorem follows a similar proof as the one made by Krylov (1980, Theorem 7, p. 82). In order to demonstrate the second part of the theorem, namely the proof of the two inequalities, set $z(t) = \mathbb{E}(\sup_{s\in[0,t]} |x(s)|^2)$. Using Gikhman and Skorohod (1972, Lemma 9, p. 134),

$$\begin{aligned} z(t) &\le 2|\varphi(0)|^2 + 2\mathbb{E}\sup|I(s,x)|^2 \\ &\le C_2(\|\varphi\|^2 + t + \int_0^t z(s)ds) \end{aligned}$$

where C_2 is some constant that depends on only C, T, K. The last inequality implies

$$z(t) \le A(\|\varphi\|^2 + 1)e^{C_2 t}$$

Analogously for

$$z_1(t) = \mathbb{E}(\sup_{s\in[a,t]} |x(s) - x(a)|^2)$$

we have

$$z_1(t) \le C_3(t(1 + \|\varphi\|^2) + \int_0^t z_1(s)ds)$$

where C_3 is some constant that depends on only C, T, K. This implies that

$$z_1(t) \le (1 + \|\varphi\|^2 + 2a \sup_{u\in[0,a]} |x(u)|^2)(e^{C_3(t-a)} - 1)$$

The theorem is proved. □

Fleming and Rishel (2012, Theorem 4.1)

The following result is taken directly from Chapter VI of Fleming and Rishel (2012, Theorem 4.1, pp. 159 ff.) wherein a suitable verification theorem for a system whose state at time t is represented by a controlled diffusion process. All of the assertions made prior to Section 4.1.3 before hold true. Then, the following theorem can be applied:

Theorem B.3. Let $W(s,y)$ be a solution of the dynamic programming equation

$$0 = W(s) + \min_{v\in U}[A^v(s)G + L(s,x,v)]$$

for $(s,y) \in Q$ with boundary data

$$W(s,y) = \Psi(s,y), \quad (s,y) \in \partial^* Q$$

such that W is in $C_P^{1,2}(Q)$ and continuous on the closure $\overline{Q}$. Then:

1. $W(s,y) \le J(s,y,\mathbf{u})$ for any admissible feedback control $\mathbf{u}$ and any initial data $(s,y) \in Q$.
2. If $\mathbf{u}^*$ is an admissible feedback control such that

   $$A^{\mathbf{u}^*}(s)W + L^{\mathbf{u}^*}(s,y) = \min_{v\in U}[A^v(s)W + L(s,y,v)]$$

   for all $(s,y) \in Q$, then $W(s,y) = J(s,y,\mathbf{u}^*)$ for all $(s,y) \in Q$. Thus $\mathbf{u}^*$ is optimal.

Bibliography

Patrik Andersson and Andreas N Lagerås. Optimal bond portfolios with fixed time to maturity. *Insurance: Mathematics and Economics*, 53(2):429–438, 2013.

Bank for International Settlements. *Zero-coupon yield curves: Technical documentation.* SSRN, 2014.

Tomas Björk and Bent Jesper Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348, 1999.

Damiano Brigo and Fabio Mercurio. *Interest rate models—Theory and practice: With smile, inflation and credit.* Springer, 2006.

Jens HE Christensen, Francis X Diebold, and Glenn D Rudebusch. An arbitrage-free generalized nelson–siegel term structure model, 2009.

Jens HE Christensen, Francis X Diebold, and Glenn D Rudebusch. The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, 164(1):4–20, 2011.

John C Cox, Jonathan E Ingersoll, Stephen A Ross, et al. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

Freddy Delbaen. *The mathematics of arbitrage.* Springer, 2006.

Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing (1994). In *The Mathematics of Arbitrage*, pages 149–205. Springer, 2006.

Sascha Desmettre and Ralf Korn. *Moderne Finanzmathematik–Theorie und Praktische Anwendung Band 2.* Springer, 2018.

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

Francis X Diebold and Glenn D Rudebusch. *Yield curve modeling and forecasting: the dynamic Nelson-Siegel approach.* Princeton University Press, 2013.

Francis X Diebold, Glenn D Rudebusch, and S Boragan Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of econometrics*, 131(1-2): 309–338, 2006.

Francis X Diebold, Canlin Li, and Vivian Z Yue. Global yield curve dynamics and interactions: a dynamic nelson–siegel approach. *Journal of Econometrics*, 146(2): 351–363, 2008.

Franziska Diez. *Yield Curves and Chance-Risk Classification: Modeling, Forecasting, and Pension Product Portfolios.* Fraunhofer Verlag, 2022.

Franziska Diez and Ralf Korn. Yield curve shapes of vasicek interest rate models, measure transformations and an application for the simulation of pension products. *European Actuarial Journal*, 10:91–120, 2020.

Darrell Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical finance*, 6(4):379–406, 1996.

Eugene F Fama and Robert R Bliss. The information in long-maturity forward rates. *The American Economic Review*, pages 680–692, 1987.

Damir Filipović. A note on the nelson–siegel family. *Mathematical finance*, 9(4):349–359, 1999.

Damir Filipovic. *Term-structure models: A graduate course.* Springer Science & Business Media, 2009.

Wendell H Fleming and Raymond W Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

Jeffrey A Frankel and Cara S Lown. An indicator of future inflation extracted from the steepness of the interest rate yield curve along its entire length. *The Quarterly Journal of Economics*, 109(2):517–530, 1994.

Iosif I. Gikhman and Anatolií Vladimirovich Skorohod. Stochastic differential equations. In *The theory of stochastic processes III*, pages 113–219. Springer, 1972.

Iosif I. Gikhman and Anatolií Vladimirovich Skorokhod. *The theory of stochastic processes II.* Springer Science & Business Media, 2004.

Refet S Gürkaynak, Brian Sack, and Jonathan H Wright. The us treasury yield curve: 1961 to the present. *Journal of monetary Economics*, 54(8):2291–2304, 2007.

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

Thomas SY Ho and Sang-Bin Lee. Term structure movements and pricing interest rate contingent claims. *the Journal of Finance*, 41(5):1011–1029, 1986.

John Hull and Alan White. Pricing interest-rate-derivative securities. *The review of financial studies*, 3(4):573–592, 1990.

Kiyosi Itô. On a formula concerning stochastic differentials. *Nagoya Mathematical Journal*, 3:55–65, 1951.

Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

Ioannis Karatzas, John P Lehoczky, and Steven E Shreve. Optimal portfolio and consumption decisions for a “small investor” on a finite horizon. *SIAM journal on control and optimization*, 25(6):1557–1586, 1987.

Martin Keller-Ressel. The classification of term structure shapes in the two-factor vasicek model—a total positivity approach. *International Journal of Theoretical and Applied Finance*, 24(05):2150027, 2021.

Martin Keller-Ressel and Felix Sachse. Term structure shapes and their consistent dynamics in the svensson family. *arXiv preprint arXiv:2410.08808*, 2024.

Siem Jan Koopman, Max IP Mallee, and Michel Van der Wel. Analyzing the term structure of interest rates using the dynamic nelson–siegel model with time-varying parameters. *Journal of Business & Economic Statistics*, 28(3):329–343, 2010.

Ralf Korn. *Optimal portfolios: stochastic models for optimal investment and risk management in continuous time.* World scientific, 1997.

Ralf Korn. *Moderne Finanzmathematik–Theorie und praktische Anwendung Band 1*. Springer, 2014.

Ralf Korn. Interest rate theory manuscript for the lectures ws 2021/22, 2022.

Ralf Korn and Holger Kraft. A stochastic control approach to portfolio problems with stochastic interest rates. *SIAM Journal on Control and Optimization*, 40(4):1250–1269, 2002.

Ralf Korn, Elke Korn, and Gerald Kroisandt. *Monte Carlo methods and models in finance and insurance.* CRC press, 2010.

Leo Krippner. A theoretical foundation for the nelson–siegel class of yield curve models. *Journal of Applied Econometrics*, 30(1):97–118, 2015.

Nicolai V Krylov. *Controlled diffusion processes.* Springer, 1980.

Steven P. Lalley. Notes on the ito calculus, 2012. URL `https://www.stat.uchicago.edu/~lalley/Courses/385/Old/ItoIntegral-2012.pdf`.

Márcio Poletti Laurini and Armênio Dias Westin Neto. Arbitrage in the term structure of interest rates: a bayesian approach. *International Econometric Review*, 6(2):77–99, 2014.

Jared Levant and Jun Ma. A dynamic nelson-siegel yield curve model with markov switching. *Economic Modelling*, 67:73–87, 2017.

Yuh-Dauh Lyuu. *Financial Engineering and Computation: Principles, Mathematics, Algorithms.* Cambridge University Press, 2001.

Robert C Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257, 1969.

Robert C Merton. Optimum consumption and portfolio rules in a continuous-time model. In *Stochastic optimization models in finance*, pages 621–661. Elsevier, 1975.

Charles R Nelson and Andrew F Siegel. Parsimonious modeling of yield curves. *Journal of business*, pages 473–489, 1987.

Ken Nyholm and Rositsa Vidova-Koleva. Nelson-siegel affine and quadratic yield curve specifications: which one is better at forecasting?, 2010. URL `http://ssrn.com/abstract_id=1610230`.

Michael Debabrata Patra, Harendra Behera, and Joice John. A Macroeconomic View of the Shape of India's Sovereign Yield Curve, 2021. URL https://rbi.org.in/scripts/BS_ViewBulletin.aspx?Id=20321.

Marek Rutkowski. Self-financing trading strategies for sliding, rolling-horizon, and consol bonds. *Mathematical Finance*, 9(4):361–385, 1999.

Sebastian T. Schich. Estimating the german term structure, 1997. URL https://www.bundesbank.de/resource/blob/622394/ab816caf466eeea7dd2fef26d0fbfc30/472B63F073F071307366337C94F8C870/1997-10-01-dkp-04-data.pdf.

Lars EO Svensson. Estimating forward interest rates with the extended nelson & siegel method. *Sveriges Riksbank Quarterly Review*, 3(1):13–26, 1995.

Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188, 1977.

Arnold Zellner. Statistics, science and public policy. *Journal of the American Statistical Association*, 87(417):1–6, 1992.

Academic Career

10/2022 - 03/2026	PhD student of Prof. Dr. Ralf Korn, Financial Mathematics Research Group *RPTU Kaiserslautern-Landau, Kaiserslautern, Germany*
06/2018 - 10/2020	Master of Science in Financial Mathematics, *The Maharaja Sayajirao University of Baroda, Vadodara, India*
06/2015 - 05/2018	Bachelor of Science (Honors) in Physics, *The Maharaja Sayajirao University of Baroda, Vadodara, India*
01/2003 - 05/2015	Grades 1 to 12, *Schools in Brunei and India*

Wissenschaftlicher Werdegang

10/2022 - 03/2026	Doktorandin bei Prof. Dr. Ralf Korn, Arbeitsgruppe Finanzmathematik *RPTU Kaiserslautern-Landau, Kaiserslautern, Deutschland*
06/2018 - 10/2020	Master of Science in Finanzmathematik, *The Maharaja Sayajirao University of Baroda, Vadodara, Indien*
06/2015 - 05/2018	Bachelor of Science (Honors) in Physik, *The Maharaja Sayajirao University of Baroda, Vadodara, Indien*
01/2003 - 05/2015	Schuljahre 1 bis 12, *Schulen in Brunei und Indien*