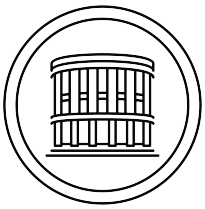
COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS



Quadratic hedging with path-dependent volatility

MASTER'S THESIS

2024 Bc. Jakub ČIŽNÁR

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND
INFORMATICS

Quadratic hedging with path-dependent volatility

MASTER'S THESIS

Study Programme:	Mathematics of Economy, Finance and Modeling
Field of Study:	1114 Applied Mathematics
Department:	Department of Applied Mathematics and Statistics
Supervisor:	doc. Ing. Aleš Černý, PhD.

Bratislava 2024 **Bc. Jakub ČIŽNÁR**

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Kvadratické zaistenie s cenovou volatilitou závislou od cesty

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomicko-finačná matematika a modelovanie
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. Ing. Aleš Černý, PhD.

Bratislava 2024 **Bc. Jakub ČIŽNÁR**





Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Bc. Jakub Čižnár
Study programme:	Mathematical Economics, Finance and Modelling (Single degree study, master II. deg., full time form)
Field of Study:	Mathematics
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Quadratic hedging with path-dependent volatility

Annotation: A particular type of stochastic volatility model that has been shown to capture well the empirical behavior of stock prices wil be studied. The standard quadratic hedging theory to (a Markovian version of) this model will be applied to obtain prices, hedge ratios, and hedging error for a contingent claim on the stock and simple volatility derivatives.

Supervisor:	doc. Ing. Aleš Černý, PhD.
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Head of department:	doc. Mgr. Igor Melicherčík, PhD.

Assigned: 10.01.2023

Approved: 11.01.2023 prof. RNDr. Daniel Ševčovič, DrSc.
Guarantor of Study Programme

.....
Student

.....
Supervisor





Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Jakub Čižnár
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Quadratic hedging with path-dependent volatility
Kvadratické zaistenie s cenovou volatilitou závislou od cesty

Anotácia: Úlohou bude študovať konkrétny typ modelu stochastickej volatility, o ktorom sa ukázalo, že dobre zachytáva empirické správanie cien akcií. Na tento model sa použije štandardná teória kvadratického zaistenia (Markovove reťazce), aby sa získala cena, zaisťovacia stratégia a chyba zaistenia pre derivát akcie a jednoduché deriváty volatility.

Vedúci: doc. Ing. Aleš Černý, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 10.01.2023

Dátum schválenia: 11.01.2023

prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Acknowledgements: I would like to sincerely thank my supervisor, doc. Ing. Aleš Černý, PhD., for his numerous helpful tips, guidance, patience, and words of encouragement. I would also like to thank my family and friends for all their support during my studies.

Abstract

Bc. Jakub Čižnár: *Quadratic hedging with path-dependent volatility* [Master's thesis] Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, supervisor: doc. Ing. Aleš Černý, PhD., Bratislava, 2024, 75p.

This thesis focuses on the comprehensive application of quadratic hedging theory to the 2-factor Markovian path-dependent volatility model. We begin with a description of the model and then briefly review the development of hedging strategies. Then, the application of quadratic hedging theory to the model leads us to derive two crucial processes: the opportunity process and the adjustment process. Hence, we compute the optimal hedging coefficient and the expected squared hedging error. We show that the variance-optimal martingale measure (VOMM) is equal to the opportunity measure and the objective measure for the model, which simplifies the computation of the contingent claim price. We solve the problem numerically using Monte Carlo simulations and price the American and European put options. Completing the picture, we backtest the optimal hedging strategy against the model delta hedging strategy and the Black-Scholes delta hedging strategy from the perspective of a European put option writer. It shows no significant difference in terms of the squared hedging error between each strategy on the sample of 394 non-paying dividend stocks from the New York Stock Exchange.

Keywords: quadratic hedging, 2-factor Markovian path-dependent volatility model, Monte Carlo simulations, optimal hedging strategy

Abstrakt

Bc. Jakub Čižnár: *Kvadratické zaistenie s cenovou volatilitou závislou od cesty* [Diplomová práca] Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. Ing. Aleš Černý, PhD., Bratislava, 2024, 75s.

Táto diplomová práca sa zameriava na komplexnú aplikáciu teórie kvadratického zaistenia na dvojfaktorový markovovský model s cenovou volatilitou závislou od cesty. Úvod práce sa venuje opisu modelu a prehľadu vývoja stratégií zaistenia. Aplikácia kvadratickej teórie zaistenia pre náš zvolený model vedie k odvodeniu dvoch kľúčových procesov pomocou, ktorých následne odvodíme optimálny koeficient zaistenia a očakávanú kvadratickú chybu zaistenia. Ukázalo sa, že miera variance-optimal martingale measure (VOMM) sa rovná takzvanej miere opportunity measure, ktorá sa rovná objektívnej miere pre zvolený model, čo zjednodušuje výpočet ceny finančného derivátu. Problém riešime numericky pomocou simulácií Monte Carlo a oceňujeme americké a európske put opcie. Na záver spätne testujeme optimálnu zaisťovaciu stratégiu oproti modelovej delta zaisťovacej stratégii a Black-Scholesovej delta zaisťovacej stratégii z pohľadu vypisovateľa európskej put opcie. Výsledky neukázali žiadny významný rozdiel z hľadiska kvadratickej chyby zaistenia medzi jednotlivými stratégiami na určenej vzorke 394 akcií nevyplácajúcich dividendy z Newyorskej burzy cenných papierov.

Kľúčové slová: kvadratické zaistenie, dvojfaktorový markovovský model s cenovou volatilitou závislou od cesty, simulácie Monte Carlo, optimálna stratégia zaistenia

Contents

Introduction **11**

1 Two-factor Markovian PDV model **13**

1.1 Introduction to PDV models 13

1.2 Defining the 2-factor Markovian PDV model 14

1.2.1 Volatility constrains 17

1.2.2 Factors description 18

1.2.3 Kernels 19

2 Preliminaries **20**

2.1 Essential stochastic calculus theory 20

2.2 Martingale theory 22

2.3 Notation for Itô formula 24

2.4 Other processes 25

3 Brief history of hedging strategies **26**

3.1 Locally risk-minimising strategy 28

3.2 Mean-variance optimal strategy 29

4 Quadratic hedging **31**

4.1 Derivation of processes for quadratic hedging 31

4.1.1 Price process is semimartingale 31

4.1.2 Derivation of opportunity process 32

4.1.3 Opportunity neutral measure 34

4.1.4 Modified mean-variance trade-off process 34

4.1.5 Derivation of adjustment process 35

4.1.6 Variance-optimal logarithm process 35

4.1.7 Variance-optimal martingale measure 35

4.2 Option pricing 36

4.3 Optimal hedging coefficient 37

4.4 Expected squared hedging error 38

5 Monte Carlo numerical approach **41**

5.1 Valuing put option 41

5.1.1 Simulating paths 42

5.1.2 Computing put option value 45

5.1.3 Mean value process analysis 46

5.2 Computing first derivatives 46

5.3 Computing optimal hedging coefficient 48

5.4 Computing expected squared hedging error 50

5.5 Numerical and practical implementation issues 52

6 Real data application **57**

6.1 Inputs to the model 58

6.2 Calibration 58

6.2.1 VIX index 59

6.3 Backtesting 59

6.3.1 Single stock backtesting: case study of Alphabet Inc. . . 60

6.3.2 Comprehensive backtesting of 394 stocks 63

6.4 Analysis and discussion of backtesting results 65

Conclusion **67**

Bibliography **70**

Appendix **73**

A Mean value process analysis 73

Introduction

Hedging strategies to minimise risk have been misprized but, more importantly, have not been systematically incorporated into pricing throughout history. That's why it's still crucial to understand the actual level of risk in a realistic model of price formation, which could potentially prevent crises like the one in 2008. Consequently, this master's thesis directs its focus toward hedging strategies in incomplete markets within a model where volatility is primarily derived from the historical price data of the asset itself. The aim is to apply the theory of quadratic hedging to a realistic model of price formation and to analyse the optimal hedging strategy compared to the classical Black-Scholes framework.

We adopt the mean-variance hedging methodology outlined in [6]. We approach the problem as in [5], where the optimal hedge, the derivative value, and the expected squared hedging error resulting from imperfect replication are ascertained from the Heston model. We extend this approach to the 2-factor Markovian path-dependent volatility model proposed in [12], and then solve the problem numerically using the Monte Carlo method. Afterwards, we backtest the optimal hedging strategy on the European options historical data from LSEG Datastream.

The master's thesis is structured into six chapters. In the initial chapter, we introduce the 2-factor Markovian path-dependent volatility model, outlining its configuration, analyzing its properties, providing the historical context of path-dependent models, and justifying their application in the real world. Chapter 2 sets forth the preliminaries for subsequent chapters in the thesis, including essential definitions, propositions for deriving the variance-optimal martingale measure, and the notation of the Itô formula employed for the simpler derivation of the optimal hedging coefficient and the expected squared hedging error. Chapter 3 provides a brief history of hedging strategies in incomplete markets, explaining the core ideas and rationale behind mean-variance hedging and underscoring its importance in realistically capturing world dynamics.

Moving forward, Chapter 4 begins with deriving the opportunity process,

the adjustment process, the opportunity-neutral measure, and other processes to attain the variance-optimal martingale measure (VOMM) for the 2-factor Markovian path-dependent volatility model. We subsequently compute the optimal hedging coefficient, the expected squared hedging error, and the derivative value, given by a bounded continuous function. In Chapter 5, we apply the Monte Carlo simulation method to estimate the values of European and American put options, determining the optimal hedging coefficient and computing the mean squared hedging error.

Chapter 6 scrutinises the applied numerical results on historical data from LSEG Datastream, backtesting the optimal hedging strategy from the perspective of someone who is selling or writing an option against the model delta hedging strategy and the classic Black-Scholes delta hedging strategy on European put options for 394 public companies.

1 Two-factor Markovian PDV model

This chapter dissects the *Path Dependent Volatility* (PDV) model which tries to explain volatility by capturing the underlying behaviour of asset prices. Specifically, the 2-factor Markovian PDV model, which was proposed by [12]. The chapter also explains important properties of the model, potential advantages and disadvantages, and also explains the motivation for creating such a PDV model. Definition and description of the model is obtained from [11] and [12]. The theory of stochastic calculus, such as the Itô lemma or the Itô integral, is adapted from [20]. The notation used is combined from both literature [12], [20] in order to follow the notation from [5] since later in Chapter 4 we apply the model to the theory of the quadratic hedging. Throughout the paper and in subsequent chapters, we use the term model or our model to refer to the 2-factor Markovian PDV model unless otherwise stated.

1.1 Introduction to PDV models

Capturing volatility with PDV models arose from *Stochastic Volatility* (SV) and *Local Volatility* (LV) models. Motivation came from several directions. As presented in [12] from philosophical, intuitive, financial, theoretical, and empirical.

Philosophical is that time goes only in one direction. This means that the future isn't solely determined by the past through the present moment. Researchers often use a simplified assumption called the Markovian assumption to make things easier to understand and calculate, even though it's not always a fundamental rule. We are employing the same rationale.

Intuition can be enlightened through an illustrative scenario wherein one is tasked with estimating the volatility of an underlying asset that was valued at 1000 units one month prior and currently stands at 2000 units. Subsequent to this initial valuation, further information is introduced, indicating that the asset was valued at 3000 units a week ago. The iterative incorporation of additional information results in a dynamic revision of estimations. Suggesting

usages of PDV as a great modeling approach.

Financial motivation describes the fact that in volatility modeling, the two common quantities that possess a natural scale are the volatility levels and the asset returns. Therefore a model should represent these two quantities. LV explains volatility levels on asset levels. SV explains volatility returns on asset returns. PDV model explains volatility levels on asset returns, which fits the financial requirement.

Theoretical motivation comes from the fact that PDV models explain volatility in a purely endogenous way. Implies that PDV models are complete models in which derivatives have a unique, unambiguous price, independent of any preferences or utility functions.

Lastly, several pieces of literature have already reported the empirical motivation of path dependence in volatility. Efforts to combine SV and LV are presented in [11]. The evidence that volatility is dependent on past prices is presented in [27]. Introduction to PDV models on a specific model is explored in [7]. For a deeper understanding of the development of PDV models, we are also referring to the second chapter of [12]. The above literature should give the reader a perspective to better comprehend the idea of PDV modelling.

1.2 Defining the 2-factor Markovian PDV model

At the very beginning, we assume a stochastic differential equation that models movement in the underlying asset price:

$$dS_t = S_t \, g(R_{1,t}, R_{2,t}) \, dW_t \ . \tag{1}$$

Where S_t is a stock price at time t. W_t is the Wiener process, defined as in the Definition 2.3. Finally, function $g : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is a volatility function of two Markovian factors, defined as:

$$g(R_1, R_2) = \beta_0 + \beta_1 R_1 + \beta_2 \sqrt{R_2} \ . \tag{2}$$

While the volatility g drives the dynamics of the asset price S, there is a feedback loop from past prices to volatility. PDV models are thus pure feedback

models in which the volatility is fully endogenous. Parameters β in (2) are in range of $\beta_0 > 0$, $\beta_1 < 0$, and $\beta_2 \in (0, 1)$. Negative β_1 refers to the so-called trend feature and β_2 refers to the so-called activity feature described later in Section 1.2.2. For a more detailed description, development, and interpretation of parameters, β we refer to the third chapter of [12]. To clarify the time-specific application of the function g we denote it by g_t, where the subscript t indicates the evaluation of the function with R_1 and R_2 at time t.

According to [12] volatility is mostly path-dependent and estimating volatility with a simple linear function (2), which largely captures market volatility, is an acceptable approach to volatility modelling. The function g represents instantaneous volatility which, as we have shown, is well determined by two factors, R_1 and R_2, which are constructed as a continuous weighted sum of past returns and past squared returns, with the functions $K_1(t)$ and $K_2(t)$ representing kernels, loosely given that they weight past returns. According to [12] using single normalized exponential kernels such as:

$$\begin{aligned} K_1(t) &:= \lambda_1 e^{-\lambda_1 t} \qquad \text{for } t \geq 0 \\ K_2(t) &:= \lambda_2 e^{-\lambda_2 t} \end{aligned} \tag{3}$$

with positive constants $\lambda_1, \lambda_2 > 0$ yields a Markovian model which is very fast and easy to simulate. Also derivation $K_1' = -\lambda_1 K_1$ of normalized exponential kernels will be further used. So the factors R_1 and R_2 are defined as:

$$\begin{aligned} R_{1,t} &:= \int_{-\infty}^{t} K_1(t-u)\,\frac{dS_u}{S_u} = \int_{-\infty}^{t} K_1(t-u) g_u\, dW_u \\ R_{2,t} &:= \int_{-\infty}^{t} K_2(t-u)\left(\frac{dS_u}{S_u}\right)^2 := \int_{-\infty}^{t} K_2(t-u) g_u^2\, du\ . \end{aligned} \tag{4}$$

We will further derive dynamics dR_1 and dR_2 which we will need to fully and exactly define our 2-factor Markovian PDV model. We use the lower index t to indicate the points in time at which the factors are calculated. In the Section 1.2.2 we will take a closer look at the two factors R_1, R_2 and how they define volatility. As mentioned in [12] single normalized exponential kernels such as (3) in the 2-factor Markovian PDV model fail to describe volatility in the data and we will need at least 4 factors to fit the real volatility data. It will

suffice as an illustration and proof of concept, together with our assumption that the model is good for modelling asset prices for our purposes.

Furthermore, we will derive dR_1 and dR_2 from (4) using the Itô calculus theory from [20]. Firstly, we will look at R_1 and subsequently, we will apply the same approach for R_2. From (4) we have well defined Itô process constructed as in section 3.1 in [20]:

$$R_{1,t} = \int_{-\infty}^{t} K_1(t-u) g_u \, dW_u \ . \tag{5}$$

After moving t-dependent to the left-hand side we receive:

$$R_{1,t} e^{\lambda_1 t} = \lambda_1 \int_{-\infty}^{t} e^{\lambda_1 u} g_u \, dW_u \ . \tag{6}$$

Now we derive the equation (6) and on the right-hand side we have a derivation of integral with upper bound or, to put it differently, the Itô process as is often written in a differential form:

$$dR_{1,t} e^{\lambda_1 t} + R_{1,t} \lambda_1 e^{\lambda_1 t} dt = \lambda_1 e^{\lambda_1 t} g_t dW_t \ . \tag{7}$$

We multiply it with $e^{-\lambda_1 t}$ and modify the equation (7) to receive dynamics of R_1:

$$dR_{1,t} = \lambda_1 (-R_{1,t} dt + g_t dW_t) \ . \tag{8}$$

The same straightforward approach can be used to derive $dR_{2,t}$ assuming an Itô process with the assumption that squared increments equal to increments of time as described in Section 2.2, formally $(dW_t)^2 = dt$ then we have well defined Itô process:

$$R_{2,t} = \int_{-\infty}^{t} K_2(t-u) g_u^2 \, du \ . \tag{9}$$

Moving t-dependent to the left-hand side and subsequently derive the formula we receive:

$$dR_{2,t} e^{\lambda_2 t} + R_{2,t} \lambda_2 e^{\lambda_2 t} dt = \lambda_2 e^{\lambda_2 t} g_t^2 dt \ . \tag{10}$$

We multiply it with $e^{-\lambda_2 t}$ and after modifying the equation we receive dynamics of R_2:

$$dR_{2,t} = \lambda_2 \Big(g_t^2 - R_{2,t} \Big) \, dt \ . \tag{11}$$

As you can see R_2 is a pure deterministic process. After we derived dynamics of the R_1 and R_2 we have a complete 2-factor PDV model with Markovian properties:

$$\begin{aligned} \frac{dS_t}{S_t} &= g_t\, dW_t \qquad \text{where} \quad g(R_{1,t}, R_{2,t}) = \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}} \\ dR_{1,t} &= \lambda_1 \Big(-R_{1,t}\, dt + g_t\, dW_t \Big) \\ dR_{2,t} &= \lambda_2 \Big(g_t^2 - R_{2,t} \Big)\, dt \ . \end{aligned} \tag{12}$$

Instantaneous volatility is a deterministic function of the two Markovian factors $(R_{1,t}, R_{2,t})$. The model is fully described by the four-dimensional Markovian vector $(T-t, R_{1,t}, R_{2,t}, S_t)$ where T is a fixed time in the future, R_1 and R_2 are autonomous, and S is dependent on R_1, R_2, and itself.

1.2.1 Volatility constrains

One might think that volatility could turn negative. We can easily prove that this is not the case for the one-dimensional diffusion process g. First, we look at the dynamics of volatility. After deriving the g, we obtain the process with the dynamics dR_1 and dR_2, which we already have:

$$\begin{aligned} dg_t &= \beta_1 dR_{1,t} + \frac{\beta_2}{2} \frac{1}{\sqrt{R_{2,t}}} dR_{2,t} \\ dg_t &= \left(-\beta_1 \lambda_1 R_{1,t} + \frac{\beta_2 \lambda_2}{2} \frac{(g_t^2 - R_{2,t})}{\sqrt{R_{2,t}}} \right) dt + \beta_1 \lambda_1 g_t dW_t \ . \end{aligned} \tag{13}$$

Then, while we are in a continuous world when the volatility reaches zero, the stochastic part disappears and the volatility becomes purely deterministic with the already mentioned properties of $\lambda_1, \lambda_2, \beta_0, \beta_2 > 0$, and $\beta_1 < 0$ and from (2) we separate $\sqrt{R_{2,t}}$, we get the drift:

$$-\beta_1 \lambda_1 R_{1,t} - \frac{\beta_2 \lambda_2}{2} \sqrt{R_{2,t}} \ = \ \frac{\beta_0 \lambda_2}{2} + \beta_1 R_{1,t} \left(\frac{\lambda_2}{2} - \lambda_1 \right) . \tag{14}$$

Now volatility equals zero and $\beta_0, \beta_2 > 0$, and $\beta_1 < 0$ then from (2) we have $R_{1,t} > 0$. When condition

$$\lambda_2 < 2\lambda_1, \tag{15}$$

holds, the drift (14) would push the volatility away from zero, and we have shown that we have non-negative volatility. This sufficient condition holds from the data shown in [12]. Later in the thesis, we will test the sufficiency condition during the asset price simulation.

1.2.2 Factors description

When we defined the 2-factor Markovian PDV model, we didn't describe exactly what is behind the factors and why we use them to define volatility. This subsection attempts to answer these questions.

The first factor, or R_1, attempts to capture trend features from asset price history to learn the leverage effect, the stylised fact about the negative relationship between asset price and volatility. From the construction, it is easy to see that R_1 could go to negative values, meaning that the recent price is falling, and therefore from (2) and the property that the coefficient β_1 is less than zero means that volatility is going up. A more significant price decline would cause higher volatility. This relationship also holds vice versa, with price increases leading to lower volatility. It is important to note that, as constructed (4), the factor gives less weight to further history and more weight to the present, while it is intuitive that the price movement one year ago would be less significant in explaining volatility than yesterday's price. This property is defined by kernels for both R_1 and R_2, which is explained in more detail in Section 1.2.3.

The second factor or R_2 is trying to capture activity features or volatility features and focus on recent movements in the asset price, irrespective of the direction of the price. The aim is to capture the concept of volatility clustering, the stylised fact that periods of high volatility are often followed by periods of high volatility and periods of lower volatility are followed by periods of lower volatility. From the volatility function g, we can see that the use of $\sqrt{R_2}$ with the property that β_2 is greater than zero implies that the second factor pushes the price up or down more when the price movement is more significant. This can also be observed from the construction (4) with an asset return to the power of two.

1.2.3 Kernels

In defining our model, we only mentioned that using single normalized exponential kernels results in a Markovian model. But we didn't mention whether there are other ways to define kernels and how they weigh past returns or give different weight to different price movements in history.

Kernels mix short and long memories of the asset price, which can be easily seen from exponential kernels defined as (3). Considering kernels with power-law decay is because the data shows that according to [12] where it is well supported by many studies therefore for a deeper understanding of the topic we refer the reader to section 3.2 in [12]. When we talk about a power law, we can observe that after a certain time, the weight of price in very far history is diminishing or is close to zero. One can choose between time-shifted power law (TSPL) kernels, exponential kernels, or a mix of exponential kernels. Our choice is based on [12], where TSPL kernels are approximated by two exponential kernels. Using 2 or more exponential kernels in a 4-factor PDV model can capture all the stylised facts, which is a desirable feature of the PDV model [12]. Practical engineering purposes led to the creation of a model with Markovian properties, which simplifies the simulation of asset prices.

2 Preliminaries

We devote this chapter to all definitions, remarks and other characterisations needed for further derivations. We will also try to explain the necessary basic terminology and preliminaries for the derivations in the master's thesis.

The chapter is structured as follows: first, we define the terminology and all the definitions as the basis of the stochastic calculus theory. Then we select definitions from martingale and semimartingale theory. After that we introduce the Itô lemma in notation adapted from [5]. Finally, we show necessary remarks from the quadratic hedging theory related to the semimartingale theory. Each definition, characterisation or remark is adapted from the literature as follows: He, Wang and Yan [14], Jacod and Shiryaev [16], Øksendal [20], Protter [22], or Shreve [24]. Unexplained notation is generally used, as in [16].

2.1 Essential stochastic calculus theory

We start with the definition of a *filtration* and a *stochastic* process, the fundamental concepts of stochastic calculus.

Definition 2.1. (Filtration). *A filtration on the space* $(\Omega, \mathscr{F})$ *is a family* $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$ *of nested* σ*-algebras* $\mathscr{F}_t \subseteq \mathscr{F}$ *that is*

$$0 \leq s \leq t \Rightarrow \mathscr{F}_s \subseteq \mathscr{F}_t$$

Each σ-algebra with the index t, which represents the time from $\mathbb{R}_+$, represents the information available at that moment. Information is not lost over time due to the non-decreasing property. We work on a filtered probability space $(\Omega, \mathscr{F}_T, \mathbb{F} = \{\mathscr{F}_t\}_{t\in[0,T]}, P)$ where T is a fixed time horizon.

Definition 2.2. (Stochastic process). *A stochastic process is a parameterised collection of random variables* $\{X_t\}_{t\geq 0}$ *defined on a probability space* $(\Omega, \mathscr{F}, P)$ *with values in* $\mathbb{R}$.

Throughout the paper, when we use the word process, we refer to the Definition 2.2. We can also define the process X_t as a function of two variables

(t, ω) from $\mathbb{R}_+ \times \Omega$ into $\mathbb{R}$. For fixed $\omega \in \Omega$ we have a function from $\mathbb{R}_+$ into $\mathbb{R}$

$$t \mapsto X_t(\omega)$$

called a *path* of the process X_t.

We say that a process $\{X_t\}_{t\geq 0}$ is *adapted* to the filtration $\{\mathscr{F}_t\}$ if for all t the random variable X_t is $\mathscr{F}_t$-measurable. In other words, a process X_t is determined based on the information available up to time t.

Definition 2.3. (standard Brownian motion). *Standard Brownian motion is a continuous, adapted stochastic process* $\{W_t\}_{t\geq 0}$ *with properties:*

(i) with a probability of 1 it applies $W_0 = 0$

(ii) for $0 \leq s < t$*, the increment* $W_t - W_s$ *is independent of* $\mathscr{F}_s$ *and is normally distributed with mean zero and variance* $t - s$*.*

Standard Brownian motion is named after a 19th-century Scottish botanist. The motion is also known as the Wiener process, after the American mathematician who formulated a mathematical model to accurately describe this random process. In this thesis, we will use the terms Brownian motion and Wiener process interchangeably.

We define *geometric Brownian motion* as

$$S_t = S_0 e^{X_t} = S_0 e^{\mu t + \sigma W_t} \tag{16}$$

for *Brownian motion* defined as a process X of the form $X_t = \mu t + \sigma W_t$, whose distribution is $X \sim \mathcal{N}(\mu t, \sigma^2 t)$. The geometric Brownian motion is said to have a *lognormal distribution* because the logarithm of S is a Brownian motion that has a normal distribution. We call

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{17}$$

a *stochastic differential equation* (SDE) with $\mu(t, X_t)$ called the *drift* and $\sigma(t, X_t)$ called the *diffusion*. Therefore as shown in section 6.2 in [16] the Brownian motion from (16) with $\mu(t, X_t) = \mu$ and $\sigma(t, X_t) = \sigma$ being constant can be written as SDE

$$dX_t = \mu dt + \sigma dW_t \ . \tag{18}$$

The process X_t is called *càdlàg* if all its paths are right-continuous with left limits. Right-continuity means that for every $t \geq 0$ you have $\lim_{s \to t^+} X_s = X_t$. Left continuity means that the limit $X_{t-} := \lim_{u \to t^-} X_u$ exists for every $t > 0$.

A process X is said to be *increasing* if all its paths are non-negative, increasing, right-continuous functions. A process is said to have *finite variation* if it is the difference between two increasing processes.

2.2 Martingale theory

Definition 2.4. (Martingale). *A process $\{M_t\}_{t \geq 0}$ on a probability space $(\Omega, \mathscr{F}, P)$ is called martingale with respect to a filtration $\{\mathscr{F}_t\}_{t \geq 0}$ if satisfies properties:*

(i) M_t *is adapted to* $\{\mathscr{F}\}_{t \geq 0}$,

(ii) $E(|M_t|) < \infty$ *for all* t,

(iii) $E(M_t \mid \mathscr{F}_s) = M_s$ *for all* $s \leq t$.

The term "martingale" comes from a French acronym for the gambling strategy of doubling your bets until you win. The martingale property means that the player's average wealth in the next game is only her current wealth.

A *stopping time* is a random variable τ for which the event $\tau < t$ is $\mathscr{F}_t$-measurable for every t. This means that, given the information available at time t, we can tell whether τ has already happened or not. For a stopping time τ and an adapted process X, we define the *stopped process* X_t^τ as

$$X_t^\tau := X_{\min\{t,\tau\}}$$

A process $\{X_t\}_{t \geq 0}$ is *uniformly integrable* if, for every $\varepsilon > 0$, there exists a constant $K \in \mathbb{R}$ such that for all $t \geq 0$ we have

$$E[X_t | 1_{|X_t| > K}] < \varepsilon \ .$$

The function $\mathbf{1}_{|X_t| > K}$ is the indicator function that returns 1 if $|X_t| > K$ and 0 otherwise. This function essentially "turns on" where the condition $|X_t| > K$ is true.

Definition 2.5. (Local martingale). *An adapted, càdlàg process* $\{X_t\}_{t\geq 0}$ *is a local martingale if there is an increasing sequence of stopping times* $\{\tau_n\}_{n=1}^{\infty}$ *such that*

$$\lim_{n\to\infty} \tau_n = \infty$$

and for each n *the stopped process* X^{τ_n} *is a uniformly integrable martingale.*

Remark 2.6. *Every martingale is also a local martingale, but not every local martingale is also a martingale, as shown in section 6 in [22].*

Definition 2.7. (Semimartingale). *A process* $\{X_t\}_{t\geq 0}$ *is called a semimartingale if it can be decomposed as*

$$X_t = M_t + A_t, \tag{19}$$

where M_t *is a local martingale and* A_t *is a càdlàg adapted process with a finite variation.*

The next definitions will be inspired by section 3.4.2 from [16].

Definition 2.8. (Quadratic variation of semimartingale). *Let* X *be a semimartingale. The quadratic variation process of* X*, denoted* $[X]_t = [X,X]_t$ *is defined by*

$$[X]_t = X_t^2 + X_0^2 - 2\int_0^t X_u\, du.$$

Definition 2.9. (Quadratic covariation of semimartingales). *Let* X *and* Y *be semimartingales. The quadratic covariation process of* X *and* Y*, is denoted as* $[X,Y]_t$ *defined by*

$$[X,Y]_t = X_tY_t - X_0Y_0 - \int_0^t X_u\, du - \int_0^t Y_u\, du.$$

We highlight the special case when $d[W]_t = (dW_t)^2 = dt$ or in other words the increment of Brownian motion is an increment of time, as shown in theorem 3.4.3. in [24]. Throughout the thesis we use rules

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0$$

as presented in section 3.1 in [20].

2.3 Notation for Itô formula

Similarly as in [5], where Černý and Kallsen use the basis of semimartingale theory from [16], where the characteristics of semimartingale processes are defined. We will use the same approach in this paper, so the following notation is taken from [5].

According to Jacod and Shiryaev in section II. proposition 2.9 [16] there exists an increasing random process A, a $\mathbb{R}^n$-valued random process b^X and a $\mathbb{R}^{n\times n}$-valued random process c^X whose values are symmetric, non-negative definite matrices. In this paper, the activity process A can be chosen such that $A_t := t$ and we will adopt this convention. Thus, in this paper, b^X refers to the drift and $\sqrt{c^X}$ to the volatility when X is a stochastic process.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be in C^1 and we denote first and second partial derivatives of function f as $f_i = \frac{\partial f}{\partial x_i}$ and $f_{ij} = \frac{\partial^2 f}{\partial x_i x_j}$ where x_i represents i-th variable. Then assume X as an $\mathbb{R}^n$-valued semimartingale then $f(X)$ is a semimartingale and the Itô formula in our notation reads

$$\begin{aligned} b^{f(X)} &= \sum_{i=1}^{n} f_i(X) b_i^X + \frac{1}{2}\sum_{i,j=1}^{n} f_{ij}(X) c_{ij}^X, \\ c^{f(X)} &= \sum_{i,j=1}^{n} f_i(X) f_j(X) c_{ij}^X, \\ c^{f(X)X_i} &= \sum_{j=1}^{n} f_j(X) c_{ij}^X. \end{aligned} \tag{20}$$

Where lower index $b_{ii}^X = b_i^X$ means the drift of the i-th variable. The same logic is applied to c_i^X where c_{ij}^X means quadratic covariance of i-th and j-th

variables. In our case for $X = (T-t, S, R_1, R_2)$ in (12) we have,

$$\begin{bmatrix} b^{(T-t)} \\ b^{R_1} \\ b^{R_2} \\ b^{S} \end{bmatrix} = \begin{bmatrix} -1 \\ \lambda_1 R_1 \\ \lambda_2(g^2 - R_2) \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} c^{(T-t)} & c^{(T-t)R_1} & c^{(T-t)R_2} & c^{(T-t)S} \\ & c^{R_1} & c^{R_1R_2} & c^{R_1S} \\ & & c^{R_2} & c^{R_2S} \\ & & & c^{S} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & \lambda_1^2 g^2 & 0 & \lambda_1 g^2 S \\ & & 0 & 0 \\ & & & g^2 S^2 \end{bmatrix} \tag{21}$$

where the drift of the processes is derived from (12). The matrix in (21) is symmetric. One can easily see that the deterministic processes have zero volatility and zero covariance with other processes. We have c^{R_1} and c^S as the volatility of the R_1 and S processes. And lastly, we have a quadratic covariation of R_1 and S as c^{R_1S} derived using $d[W]_t = dt$.

2.4 Other processes

With the defined notation, a *stochastic exponential* for a semimartingale X is given by $\mathscr{E}(X) := \exp(X - \frac{1}{2}[X, X])$.

Remark 2.10. *The stochastic exponential for the semimartingale $X = 0$ is a trivial one which is equal to $\mathscr{E}(0) = 1$.*

Definition 2.11. (Stochastic logarithm). *For a semimartingale X with values in $\mathbb{R}$, such that the two processes X_t and X_{t-} do not vanish, the stochastic logarithm $\mathcal{L}(X)$ is defined as $\mathcal{L}(X)_t = \frac{X_t}{X_{t-}}$ with the dynamics $d\mathcal{L}(X)_t = \frac{dX_t}{X_{t-}}$.*

The dynamics of processes can be expressed with the stochastic logarithm, such as the geometric Brownian motion written in SDE, $dS_t = \mu S_t dt + \sigma S_t dW_t$ as in (18), can be written as $d\mathcal{L}(S)_t = \frac{dS_t}{S_t} = \mu dt + \sigma dW_t$.

3 Brief history of hedging strategies

This chapter presents selected results from the theory of mean-variance hedging, mainly based on [21] and [25]. Knowing and understanding motivation and current research developments in this field is important. We aim to provide only a brief overview of the concepts and evolution of hedging strategies and therefore do not provide detailed and rigorous mathematical formulations.

In finance, there is the problem of correctly pricing contingent claims and then hedging them against market risk by trading the underlying assets. The classic model assumes a risk-free asset and n risky assets; in our paper, we use only one risky asset, denoted S to simplify the interpretation of our results. Asset S is usually modelled by semimartingales. A risk-free asset is usually used as a numeraire also called the reference asset. This means that there is a risk-free asset which has a discounted price and other assets whose prices are discounted by numeraire. The question may arise: what happens if there is no risk-free asset? This scenario is explored in the paper [3], which examines quadratic hedging in which the asset is represented by a semimartingale with and without a numeraire. In our paper, we assume the existence of nice numeraire defined as definition 3.8 in [3].

We have also a basic assumption of a no-arbitrage condition. In the standard theory of complete markets, the assumption is satisfied by the equivalent martingale measure Q, which is a probability measure equivalent to the objective probability measure P, such that the price S is a local martingale under Q. Additionally, the contingent claim can be replicated in a preference-independent way with a so-called self-financing portfolio strategy based on risk-free and risky assets. We assume a value of the replicating portfolio in time t defined as

$$V_t = \vartheta_t S_t + \eta_t \tag{22}$$

where ϑ is a predictable process and η is an adapted process. The predictability of ϑ is mathematically formulated as an informational constraint, which states that ϑ cannot predict the movement of S. The process ϑ_t describes the number of units of the risky asset S held at time t, and η_t is the amount invested in a

risk-free asset at time t. Therefore, the cumulative profit from trading a risky asset at time t is

$$G_t(\vartheta) = \int_0^t \vartheta_u \, dS_u \tag{23}$$

where S is, as we said, a semimartingale and $G(\vartheta)$ is a stochastic integral of ϑ with respect to S. Cumulative cost in time t is defined as:

$$C_t = V_t - G_t(\vartheta) \ . \tag{24}$$

Strategy is self-financing when the cumulative cost process is constant over time. We say that a contingent claim H is attainable if there exists a self-financing strategy with the value $V_T = H$ which means that H can be written as

$$H = x + \int_0^T \vartheta_t \, dS_t \tag{25}$$

with respect to the objective probability measure P. We call the constant x the initial endowment, or, to put it another way, the initial risk.

Therefore, with an appropriate dynamic hedging strategy, one can reduce the risk of the portfolio to zero. This was the basic idea that led to the creation of the famous Black-Scholes [1] and Merton [19] formulas for option pricing. The issue has been addressed in the scenario where S represents a one-dimensional geometric Brownian motion, focusing on the pricing and hedging of both European call options and Warrants. J. M. Harrison and D. M. Kreps later developed and clarified the mathematical structure of the problem and its connections with the martingale theory. Following their terminology, we refer to a contingent claim as H. We also refer to a complete market when every contingent claim is attainable. Although our approach lacks mathematical precision and offers only a brief summary, it's important to note that several conditions need to be met for it to be attainable.

But in reality, creating a perfectly replicating portfolio is far from the truth. In incomplete markets, reducing risk to zero is no longer the case and we need to measure or more intuitively minimize the error of the dynamic hedging strategy. The concept of risk minimisation was introduced by Föllmer and Sondermann in 1986. Schweizer then introduced the case of the semimartingale

price process in local risk minimisation. In other words, given a non-attainable contingent claim H, it is by definition impossible to find a strategy with a final value $V_T = H$ that is also self-financing. There are several concepts of pricing and hedging of contingent claims in incomplete markets. We would like to present two of the most important and frequently cited theories: the *locally risk-minimising strategy* and the *mean-variance optimal strategy.* We'll be applying the mean-variance optimal strategy to the model outlined in Chapter 1, so we'll focus more on describing this strategy in detail.

3.1 Locally risk-minimising strategy

A locally risk-minimising strategy can be characterised by two properties: its cost process C must be a martingale, and the strategy is no longer self-financing but remains mean self-financing if C is a martingale under P. The martingale C must be orthogonal to the martingale part of the price process S. So there exists locally risk-minimising strategy when H admits a so-called Föllmer-Schweizer decomposition defined as

$$H = x + \int_0^T \vartheta_t dS_t + M_T \tag{26}$$

under the objective probability measure P. Where M_T is the martingale orthogonal to the martingale part of the price process S, while S is a semi-martingale that can be decomposed into a local martingale and a process of finite variation, see Definition 2.7. For a finite discrete time scenario, one can compute ϑ and M recursively, moving backwards in time. In 2007, Černý and Kallsen revisited this problem using sequential regression [4]. In the context of a continuous S, the Föllmer-Schweizer decomposition under P is achievable by the Galtchouk-Kunita-Watanabe decomposition, which is computed under the minimal martingale measure, denoted as $\hat{P}$. The fact that you have to work with strategies that are not self-financing is a disadvantage of the current approach.

3.2 Mean-variance optimal strategy

On the other hand, if you want to insist on the self-financing constraint then the possible outcomes of such strategies are of the form $x + G_T(\vartheta)$ for initial capital $x \in \mathbb{R}$ and ϑ from all allowed integrands of V_t.

The idea is to minimise the residual risk between the contingent claim and the final portfolio value. It's a L^2 projection into the space of stochastic integrals. So we are looking for the best approximation of H by the final value $x + G_T(\vartheta)$ of any pair (x, ϑ). The use of a quadratic criterion to measure the quality of this approximation was proposed by Bouleau/Lamberton in 1989. This was done when S is both a martingale and a function of a Markov process. In more general cases, it has been the subject of derivations by Duffie and Richardson in 1991 and by Schweizer in 1994, among others. Again, in the case where S is the martingale, the optimal strategy can be constructed using the Kunita-Watanabe projection technique. In the case of the semimartingale, the problem is more complex. In general, we aim to minimise:

$$E((x + \vartheta \cdot S_T - H)^2) \tag{27}$$

where, for simplicity, we use the notation of [5], where $\vartheta \cdot S_T = \int_0^T \vartheta_t dS_t$ is a stochastic integral, so dot represents stochastic integration. Minimizing expected squared hedging error over all reasonable hedging strategies $\vartheta \in \Theta$ and possibly all initial endowments $x \in \mathbb{R}$. The set of all reasonable trading strategies Θ is defined as in chapter 2 in [6]. The solution to the minimisation problem is $\varphi^* = (x^*, \vartheta^*)$. The value x^*, if it exists, is called the approximate price of H. Using a quadratic criterion results in equal penalisation for both positive and negative values of the hedging error. While this may initially appear to be a disadvantage, it actually makes the strategy suitable for both buyers and sellers of options, favouring close replication over profit-making.

The strategy that minimises the mean squared hedging error among all admissible strategies is referred to as the *globally optimal strategy*, or known as quadratic hedging or mean-variance hedging. These terms will be used interchangeably throughout our master's thesis to refer to the same concept.

In Chapter 4 we will define a rigorous and mathematically tenable quadratic hedging theory, which we will apply to the model from Chapter 1. We will search for an optimal hedging strategy that minimises the (27). We will then look for the expected quadratic hedging error of the contingent claim at time $t = 0$ up to the maturity T.

4 Quadratic hedging

In this chapter, we will derive the pure hedge coefficient and hence the optimal hedge, which leads to obtaining the excepted squared hedging error, and we will also derive the value of the derivative given by the PDE. The optimal hedge is the solution to the minimisation problem (27). For the derivation of the optimal hedge we will use the following procedure, which is shown by Černý and Kallsen in [6]. The same approach is used to solve the mean-variance hedging problem in the Heston model by Černý and Kallsen in [5]. The notation is the same as the one used in the [5] and [6]. The definitions and remarks used are also given in the preliminaries, Section 2, which we refer to if necessary. The standard assumption in the literature [5] of zero interest rates, which does not entail a loss of generality in the context of models with deterministic interest rates, is also adopted in this chapter.

4.1 Derivation of processes for quadratic hedging

To derive the optimal hedging coefficient, we first show that the price process S is a semimartingale, then we determine the *opportunity process* L and the *adjustment process* $\tilde{a}$. For the determination of the two processes, we are using theorem 3.25 from [6]. The above processes L and $\tilde{a}$ result in the modified *mean-variance trade-off process* denoted K, the *opportunity-neutral measure* represented by P^*, the *variance-optimal logarithm process* denoted N, and the *variance-optimal martingale measure* (VOMM) denoted as Q^*. Lastly, the mean-value process V leads to the pure hedge coefficient and hence to the optimal hedge. Finally, we show the connection between option pricing with semimartingale theory and PDE and then derive a formula for the expected squared hedging error.

4.1.1 Price process is semimartingale

To show that the price process is a semimartingale, we must first show that the price process S is a local martingale and then prove from the definition

that it is a semimartingale.

We have a price process defined with a stochastic differential equation (SDE) as (1). One can see that we are working with a drift-less process. From the setup in [12] the model has Markovian property and the model is set up as a local martingale defined as in Definition 2.6.

Now let's define $A = 0$, an adapted process with a finite variation. Then we define a new process X as

$$X_t = S_t + A_t \tag{28}$$

from Definition 2.7 we have that X is a semimartingale or as we desire S is a semimartingale. In the following sections we will use the knowledge that the price process S is a semimartingale.

4.1.2 Derivation of opportunity process

Opportunity process denoted as L, in general, defines the so-called *opportunity-neutral* measure denoted as P^*, which we further describe in Section 4.1.3. The opportunity process has also a connection to the famous Sharpe ratio, which we also further describe for a better understanding of the process. From [5] we will introduce the definition of *candidate opportunity process* and *candidate adjustment process*:

Definition 4.1. *We say that L is a candidate opportunity process if*

(i) $L \in (0, 1]$ *is a valued continuous semimartingale,*

(ii) $L_T = 1$,

(iii) For $K := \mathcal{L}(L)$ *we have*

$$b^K = (b^S + c^{KS})^2 / c^S. \tag{29}$$

In such case, we call $\tilde{a} = (b^S + c^{KS})/c^S$ *the candidate adjustment process corresponding to* L.

Alternatively, from chapter 3 in [6] the opportunity process defined in each time t can be seen as,

$$L_t = E(1 - \lambda^{(\tau)} \cdot S_T | \mathscr{F}_t) \tag{30}$$

where the expected value of $1 - \lambda^{(\tau)} \cdot S_T$ where $\lambda^{(\tau)}$ represents admissible strategy, the definition of admissible strategy is given by definition 2.2 from [6]. The importance of admissible strategies lies in the idea that one can create arbitrage profits in models that are considered arbitrage-free. These kinds of strange strategies that can create something out of nothing are called *doubling strategies*. Harrison and Kreps were the first to identify this problem in 1979 in Merton's theory of optimal portfolio [13]. In equation (30) term $\lambda^{(\tau)} \cdot S_T$ stands for Itô integral or $\int_0^T \lambda_t \, dS_t$.

The opportunity process is linked to optimal investment opportunity with a relationship to the Sharpe ratio introduced as proposition 3.6 from [6],

$$\varrho = \sqrt{\frac{1}{L} - 1} \tag{31}$$

or more intuitively,

$$L = \frac{1}{1 + \varrho^2} \tag{32}$$

where ϱ^2 represents the squared maximal Sharp ratio. The proof is also provided under proposition 3.6 in [6]. Intuitively, we can see that (32) sets boundaries for the process L, and also we can see that maximizing of Sharpe ratio is pushing the value of L closer to zero.

We have shown that the price process S is a local martingale, whereas in the setup of Föllmer and Sondermann [8] we need S to be a martingale, but it helped us to come up with the candidate for the opportunity process:

$$L_t = 1 \qquad \text{for all } \; t \in [0, T] \; . \tag{33}$$

Now we apply proposition 4.1 from [3], which gives sufficient conditions for the true opportunity process L, and since the process is bounded away from zero, and provided other conditions are satisfied, we can claim that (33) is the true opportunity process.

4.1.3 Opportunity neutral measure

The explanation of the *opportunity-neutral measure* denoted as P^* can be found from [4] in the context of local risk minimization in a sequence of one-period least squares regressions, which are running recursively backward in time: If under original probability measure P some states offer a higher Sharpe ratio in the future than others, then the one-period realized hedging error in those states can be higher because better investment opportunities in the future will allow them to make up for the higher error today. Whereas the least squares under P fail to incorporate the changing investment opportunities, P^* modifies the probability weights so that the least squares regression without intercept taken under P^* yields the globally optimal hedging strategy. For this reason, we call P^* the *opportunity-neutral measure*. A more detailed description and explanation can be found in chapter 4 in [4].

To derive P^* we will use proposition 3.28 from [6]. We can easily see that opportunity process L is a predictable process of finite variation and L_0 is deterministic, which implies:

$$P = P^*, \tag{34}$$

original measure P equals opportunity measure P^*. Now with P^* in hand, we can further derive an adjustment process as the myopic mean-variance stock portfolio weight under P^* as it is done in [5].

4.1.4 Modified mean-variance trade-off process

We specified the process L, and now we can specify the so-called *modified mean-variance tradeoff process* (MMVT), denoted $K := \mathcal{L}(L)$ from Definition 4.1. Operator $\mathcal{L}$ stands for stochastic logarithm and if a stochastic process $L = 1$ is constant with a value of one over time, it means that there is no change or variability in the process as time progresses. In such a case, both the drift and diffusion coefficients of the process are zero. Therefore, the stochastic logarithm of a constant process would be trivial and have a value of zero. This can also be easily seen from definition 3.11 in [6] from which the definition of

K is adopted. Therefore we have

$$K = 0. \tag{35}$$

4.1.5 Derivation of adjustment process

We can understand the *adjustment process* as an optimal number of shares per unit of wealth. Following the literature, it is denoted by $\tilde{a}$. To derive the desired adjustment process, we use the relation $b^{KS} = 0$ while (35). The candidate adjustment process from Definition 4.1 is,

$$\tilde{a} = \frac{b^S}{c^S} = \frac{b^{S^*}}{c^S} \tag{36}$$

where b^{S^*} is a drift of the S under the measure P^* and b^S is a drift of S under original measure P, which equals because of (34). Moreover, $\sqrt{c^S}$ represents a volatility of the S under P. The price process S is a drift-less process or formally ($b^S = 0$), so the candidate for the adjustment process at time t is:

$$\tilde{a}_t = 0 \qquad \text{for all } \ t \in [0, T]. \tag{37}$$

We have only given the simplified definition of $\tilde{a}$, whereas the detailed description and derivation of $\tilde{a}$ is given in chapter 3 in [6].

4.1.6 Variance-optimal logarithm process

Variance-optimal logarithm process denoted as N is defined as in definition 3.12 in [6], from which we can easily see

$$N := K - \tilde{a} - [\tilde{a} \cdot S, K] = 0 \ , \tag{38}$$

while $K = 0$ and $\tilde{a} = 0$. Now with L, $\tilde{a}$, K, and N in hand, we can further derive the so-called *variance-optimal martingale measure*, which we describe in Section 4.1.7.

4.1.7 Variance-optimal martingale measure

Variance-optimal martingale measure (VOMM) is minimal martingale measure relative to P^*, denoted as Q^*. From definition 3.12 in [6] we have,

$$\frac{dQ^*}{dP} := \frac{L_0}{E[L_0]} \mathscr{E}(N)_T = 1 \tag{39}$$

Where $\mathscr{E}(N)_T$ is a stochastic exponential of the process N at fixed time T, given as in Section 2. Using (33) and Remark 2.10, so $L_0 = 1$, meaning that the change of measure is equal to one, which means,

$$P = P^* = Q^* \tag{40}$$

that difficult part of the change of measures falls away and applying all the results we have verified we have, $L = 1$, $\tilde{a} = 0$, $K = 0$, $N = 0$, and $P = P^* = Q^*$. We have the same results as in the case where S is a martingale and thus we are in the setup with Föllmer and Sondermann [8].

4.2 Option pricing

The formula for the *mean value process* as in [8] is reduced to the equation:

$$V_t = E_{Q^*}[H|\mathscr{F}_t] = E[H|\mathscr{F}_t] \qquad \text{for all } t \in [0,T] \tag{41}$$

while (40) holds. We showed that the mean value process value is the expected value of H on filtration at time t under the original probability measure, written as (41).

Proposition 4.2. *If the contingent claim H is given by $h(R_{1,T}, R_{2,T}, S_T)$, where h is a bounded continuous function, then $V_t = f(T-t, R_{1,t}, R_{2,t}, S_t)$ for $f \in C^{1,2,2,2}$ and also f is the unique classical solution of the PDE*

$$\begin{aligned} 0 &= -f_1 - f_2\lambda_1 R_1 + f_3\lambda_2(g^2 - R_2) + \frac{g^2}{2}(f_{22}\lambda_1^2 + 2f_{24}\lambda_1 S + f_{44}S^2) \\ f(0, R_1, R_2, S) &= h(R_1, R_2, S) \end{aligned} \tag{42}$$

with the first and second partial derivatives $f_i = \frac{\partial f}{\partial x_i}$, $f_{ij} = \frac{\partial^2 f}{\partial x_i x_j}$, $i = 1,2,3,4$ where x_i is the i-th variable of the function f.

Proof. The proof is adapted from [15] where the following procedure is applied to the Heston model. If certain 3 conditions hold, then f satisfies a PDE from theorem 1 in [15] or as we have presented it in (42). The price process S follows the stochastic differential equation (SDE) (12). We price a European put on S, so the put has a bounded payoff function h.

Following the proof for the Heston model in section 2 in [15], we have a drift and a matrix of quadratic variance and covariance (20) which do not depend on time t and are obviously in C^1 for each variable on $(-\infty, \infty)^2$. Then we have satisfied the first condition, denoted $(A1)$ in [15].

We have already shown in Section 1.2.1 that $g \in (0, \infty)^2$ and the same holds for S because we have the condition that price is always bounded away from zero. The factors R_1 and R_2, defined as (5), (9) respectively from the definition, also hold in $(-\infty, \infty)^2$. Hence we have also satisfied the second condition, $(A2)$ in [15].

Finally, the third condition $(A3)$ from [15] also holds because the drift and volatility (20) are in C^1.

□

4.3 Optimal hedging coefficient

To derive the *optimal hedging coefficient* or the solution to the minimisation problem (27), we use theorem 4.10 from [6], where the optimal hedging strategy, denoted as φ, coincides with the *pure hedge*, denoted as ξ,

$$\varphi_t(x, H) = \xi_t + \tilde{a}_t(V_t - x - \varphi(x, H) \cdot S_t) \tag{43}$$

where x represents initial capital, coefficient ξ_t represents so-called *pure hedge coefficient* at time t formulated as in definition 4.6 in [6],

$$\xi = \frac{c^{VS}}{c^S} \tag{44}$$

where c^{VS} represents instantaneous covariance between V and S, and c^S represents instantaneous variance of S described as in Section 2. While the so-called *feedback element* of (43), defined as $\tilde{a}(V - x - \varphi(x, H) \cdot S)$, is zero, because of (37) our problem is reduced to deriving the pure hedge ξ. To clarify when we refer to pure hedge ξ_t with lower index t we are referring to hedging at time t, therefore also drift and volatility would be presented with a lower index t. To simplify, we will use the notation without t, switching to t notation at the end.

We have already shown in (21) that

$$c^S = g^2 S^2 \tag{45}$$

so we need only to derive c^{VS}. In Section 2, in order to solve this problem, we have introduced an instrument or rewritten the Itô formula in our adapted notation. In our case we have $c^{f(X)X_4} = c^{VS}$ where a contingent claim V is a function f of 4 variables $V = f(T-t, R_1, R_2, S)$ defined as in Proposition 4.2 and X_4 represents 4-th variable of the function f as S. Following (20) and (21) we have

$$\begin{aligned} c^{VS} =& f_1(T-t, R_1, R_2, S)c^{T-t} + f_2(T-t, R_1, R_2, S)c^{R_1} \\ &+ f_3(T-t, R_1, R_2, S)c^{R_2} + f_4(T-t, R_1, R_2, S)c^S \\ c^{VS} =& f_2(T-t, R_1, R_2, S)\lambda_1 g^2 S + f_4(T-t, R_1, R_2, S)g^2 S^2 \end{aligned} \tag{46}$$

So the optimal hedging coefficient in the notation at time t equals:

$$\varphi_t = \xi_t = f_2(T-t, R_{1,t}, R_{2,t}, S_t)\frac{\lambda_1}{S_t} + f_4(T-t, R_{1,t}, R_{2,t}, S_t). \tag{47}$$

The optimal hedge corresponds to a pure hedge and is decomposed into the standard delta hedge $f_4(T-t, R_{1,t}, R_{2,t}, S_t)$ using the representative agent price V_t and a component utilising the influence of the volatility process $f_2(T-t, R_{1,t}, R_{2,t}, S_t)$ to the asset price S through factor R_1 defined as in (12). We can observe a similar decomposition if we compare the results with [5], where the optimal hedging coefficient is derived for the Heston model. In the Heston model, the optimal hedge coefficient is decomposed into a standard delta hedge and a leverage component that exploits the correlation of V_t with the volatility process.

4.4 Expected squared hedging error

After deriving the optimal hedging coefficient at each time t, we can derive the expected squared hedging error at time $t = 0$. We define the expected squared hedging error of the optimal hedge as in theorem 4.12 in [6]. A more intuitive description would be that we are applying an optimal hedging strategy

and comparing it with the derivative value, defined in Theorem 4.2. While we have used the optimal hedge coefficient, which minimises the expected squared hedge error, we obtain *the expected minimal squared hedging error*. Therefore, at each time t we calculate

$$\gamma_t := (\varphi_t dS_t - dV_t)^2 = \varphi_t^2 d[S]_t - 2\varphi_t d[S,V]_t + d[V]_t$$

difference between replicating portfolio and derivative value. The quadratic criterion results in an equal penalty for both positive and negative hedge error values, as described in Section 3.2. After applying the quadratic criterion, we obtain the quadratic variance and covariance of the processes dS_t and dV_t. In our notation from Section 2 and using (47) we have

$$\gamma_t = \left(\frac{c_t^{VS}}{c_t^S}\right)^2 c_t^S - 2\frac{(c_t^{VS})^2}{c_t^S} + c_t^V = c_t^V - \frac{(c_t^{VS})^2}{c_t^S} \tag{48}$$

the squared error of the replicating strategy at time t. We now need to compute the minimal expected squared hedging error to maturity T, as in theorem 4.12 in [6]. Therefore we have

$$\varepsilon_0^2 = E\left[\int_0^T \gamma_t L_t dt\right]. \tag{49}$$

the minimal expected squared hedging error until maturity. The lower index ε_0^2 represents ε^2 at time $t = 0$ or the lower integral limit. In our case, we have $L_t = 1$ for $t \in [0, T]$ defined as (30), so we just compute γ_t. To simplify, we will use the notation without t, switching to t notation at the end of the section as in Section 4.3. First we compute each component c^V, c^S and $(c^{SV})^2$ using the (20) we get:

$$\begin{aligned}
c^V &= f_2^2(T-t,R_1,R_2,S)c^{R_1} + f_4^2(T-t,R_1,R_2,S)c^S \\
&\quad + 2f_2(T-t,R_1,R_2,S)f_4(T-t,R_1,R_2,S)c^{R_1 S} \\
c^V &= f_2^2(T-t,R_1,R_2,S)\lambda_1^2 g^2 + f_4^2(T-t,R_1,R_2,S)g^2 S^2 \\
&\quad + 2f_2(T-t,R_1,R_2,S)f_4(T-t,R_1,R_2,S)\lambda_1 g^2 S \\
c^S &= g^2 S^2 \\
(c^{VS})^2 &= f_2^2(T-t,R_1,R_2,S)\lambda_1^4 g^4 + f_4^2(T-t,R_1,R_2,S)g^4 S^4 \\
&\quad + 2f_2(T-t,R_1,R_2,S)f_4(T-t,R_1,R_2,S)\lambda_1^2 g^4 S^2.
\end{aligned} \tag{50}$$

Therefore we have

$$\begin{aligned}\gamma =&\lambda_1 g^2(f_2^2(T-t,R_1,R_2,S)\lambda_1(1-\frac{\lambda_1^2}{S^2})\\&+2f_2(T-t,R_1,R_2,S)f_4(T-t,R_1,R_2,S)(S-\lambda_1))\end{aligned} \tag{51}$$

and the expected minimal squared hedging error in t notation reads:

$$\varepsilon_0^2=\lambda_1 E\left[\int_0^T g_t^2\left(f_2^2(T-t,R_{1,t},R_{2,t},S_t)\lambda_1\left(1-\frac{\lambda_1^2}{S_t^2}\right)+2f_2(T-t,R_{1,t},R_{2,t},S_t)f_4(T-t,R_{1,t},R_{2,t},S_t)(S_t-\lambda_1)\right)dt\right]. \tag{52}$$

Now that we have computed the expected minimal squared hedging error, we can compute it numerically in Chapter 5.

5 Monte Carlo numerical approach

There are two ways to calculate a contingent claim value: first, by using Monte Carlo simulation to generate paths of the underlying security under the risk-neutral measure and using these paths to estimate the risk-neutral expected discounted payoff; second, by numerically solving a partial differential equation in our case defined as in Proposition 4.2, which is the time-dependent PDE in 3 dimensions.

This chapter discusses the first of these methods, starting by valuing the European and American put options at time t and then calculating the optimal hedge and the mean squared hedging error. Specifically, we will use the Monte Carlo approach of Longstaff and Schwartz [17]. At the end of the chapter, we will discuss the potential complications and trade-offs associated with the Monte Carlo method and how we approach it. The source for understanding the numerical drawbacks of applying the Monte Carlo method to option pricing is the book by Glasserman [10]. All calculations in this and the following chapter are carried out in the Matlab version 2015b software environment.

5.1 Valuing put option

The main idea behind the Monte Carlo method of option pricing, as outlined earlier, is to simulate a sufficient number of paths. Then calculate the payoffs or cash flows from each path, starting at the option's maturity, and then calculate the conditional cash flows to each discretized time before that. Then sum all the discounted cash flows and take the average, which is the American option value at time zero. The European option value is easier to derive. We simply discount all cash flows from each path at expiration and take an average.

The conditional expectation function is derived as a regression of constant plus some degree of Laguerre polynomials as in [17]. The dependent variable is the discounted payoff from the previous time, remember we are pricing the option backwards, and the independent variable is the value of the stock at the time the conditional expectation is calculated. The conditional expectation

calculation is called the Least Squares Method (LSM) algorithm. The method is efficient in the sense that it calculates the conditional expectation only in paths where the cash flow is positive, thus saving on computing power and reducing the computing time. For a better and more detailed understanding of the Monte Carlo method used in our research, we refer the reader to the numerical example in the first chapter in [17].

5.1.1 Simulating paths

Our model (12) is easy to simulate, so it's a good candidate for the chosen Monte Carlo method. The parameters $\beta_0, \beta_1, \beta_2, \lambda_1, \lambda_2$ are taken from [12]. They are calibrated to the S&P 500 index or SPX index, the 500 most capitalised companies in the US. For a description and understanding of the S&P 500 index, we refer to [23]. We take this into account so that for practical application the model needs to be calibrated to the precise security. Our assumption is that these parameters are good for our explanatory purposes. We chose $S_0 = 40$ as the initial price. We also needed to set $R_{1,0}, R_{2,0}$ as the starting point for the simulation, so we generated a random path ending in the value S_0 and calculated initial $R_{1,0}, R_{2,0}$ from this historical path.

In Figure 1 we show the simulation of the asset price for with the historical path following a geometric Brownian motion in the upper left graph, the calculated volatility from the simulated R_1, R_2 in the upper right graph, the simulation of the factor R_1 in the lower left graph, and the simulation of the factor R_2 in the lower right graph. We generated 50,000 paths plus 50,000 antithetic paths to reduce the standard error of the simulation. We are referring to chapter 2 in [10] for a more detailed explanation of reducing standard error in simulations. The option can be exercised 50 times per year. Asset values are simulated for exactly one year. The red dotted line represents the average value of simulation at every point in the year. We summarize all parameters and initial values in Table 1 for the simulation in Figure 1. We note that the condition for non-negative volatility (15) holds.

We also perform a simulation of a simple model with constant volatility

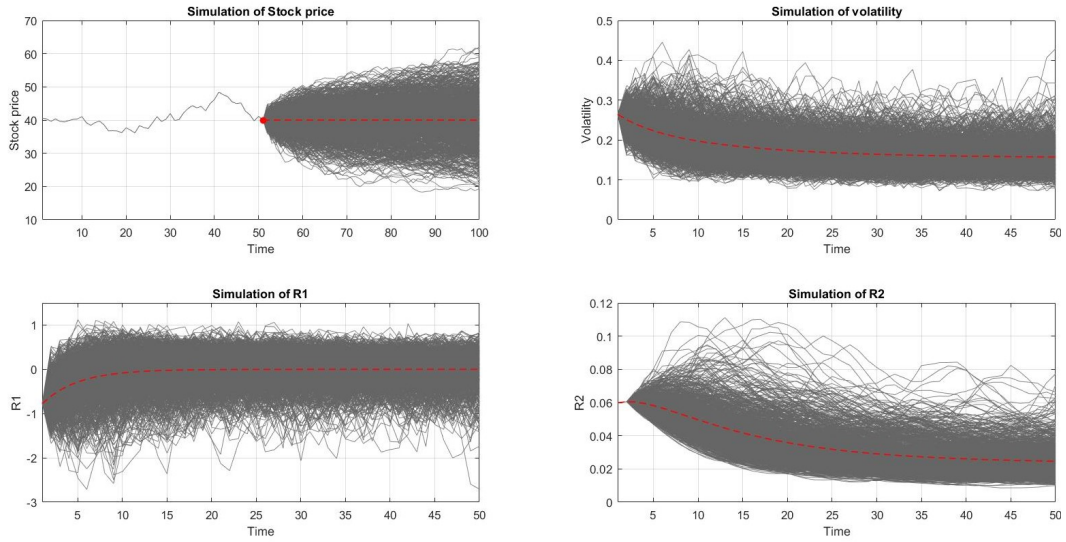


Figure 1: Simulation of 2-factor Markovian PDV model

Table 1: Parameters and initial values for the simulation of 2-factor Markovian PDV model in Figure 1

β_0	β_1	λ_1	β_2	λ_2	S_0	$R_{1,0}$	$R_{2,0}$
0.08	-0.08	12.500	0.5	7.937	40	-0.773	0.0599

and zero drift,

$$dS_t = \sigma dW_t \tag{53}$$

to visually compare simulated prices. Both asset prices follow geometric Brownian motion defined as (16) and the initial value is set to $S_0 = 40$. In the simple model, we set the constant volatility to $\sigma = 0.178$ as an average of the averages over paths and time from the 2-factor Markovian PDV simulation in Figure 1. The comparison can be seen in Figure 2, where on the left side is a simulation with our model (12) and on the right is the simple model (53). The seed is set to the same value for both simulations. Parameters and initial values are set as in Table 1.

Figure 1 can explain differences in Figure 2 while it's the same simulation of the 2-factor PDV Markovian model. In Figure 2 at time zero we can observe higher volatility for the simulation on the left side. This can be described with higher volatility in the past prices, which the Simple model does not take into account. As a result, we can observe different distributions at the end of the

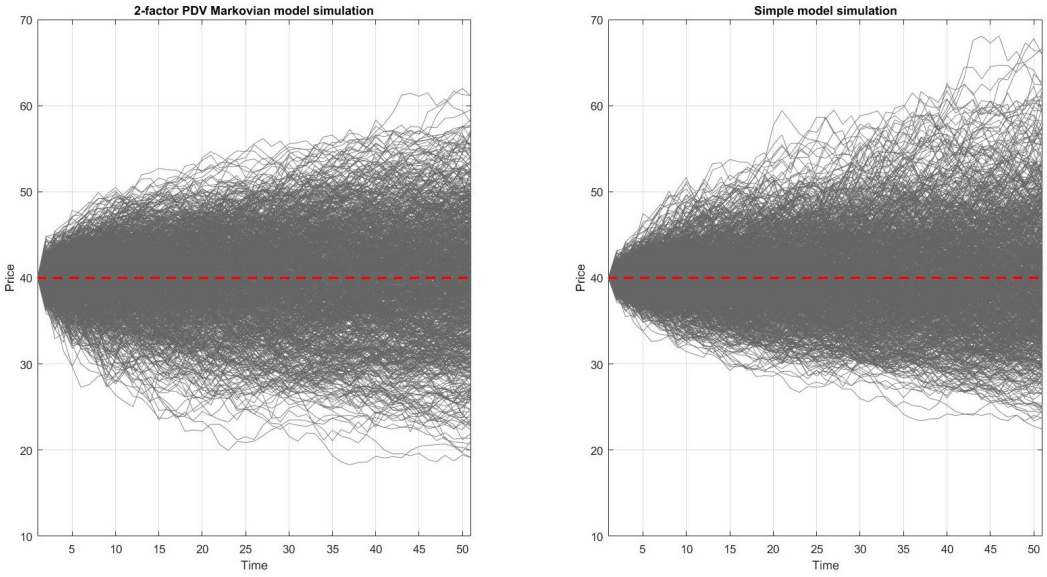


Figure 2: Comparison of asset price paths with variable volatility input to the 2-factor Markovian PDV model on the left and with constant volatility to the simple model on the right. The red dotted line represents the average of all simulations.

simulation shown in Figure 3.

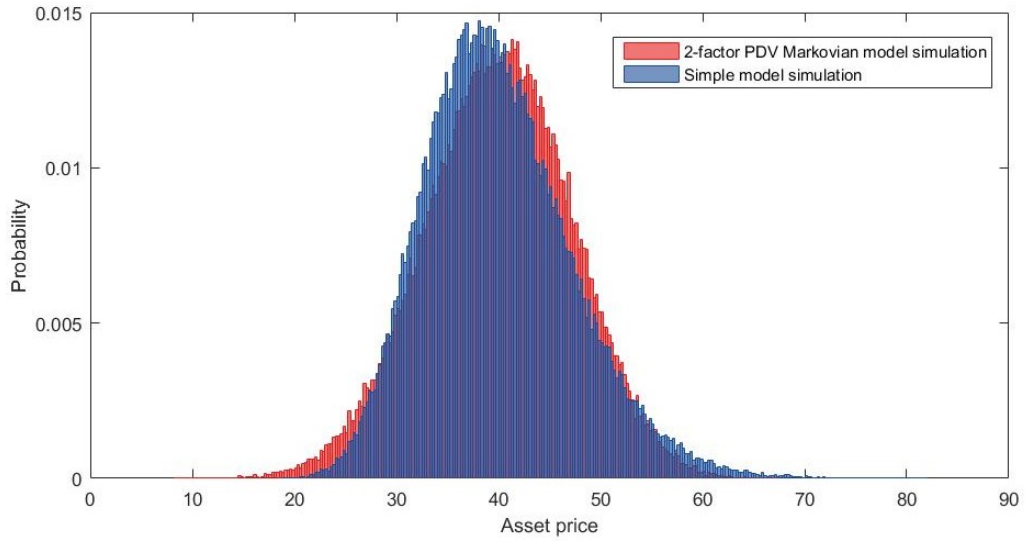


Figure 3: Distribution of the simulations of the asset prices for two different models: the 2-factor PDV model and the Simple model at the end of the period at the time $t = 50$ with the initial price $S_0 = 40$.

At first glance, we have two different distributions. The statistical characteristics of distributions at the end of the simulation period produced by two different models: the 2-factor PDV Markovian model and the Simple model are summarised in Table 2. For the 2-factor PDV Markovian model, the median

is 44.338, slightly higher than the mean of 43.703, indicating a slight left skew in the distribution, which is confirmed by a negative skewness value of -0.221. The model has a relatively low standard deviation of 2.992 and a kurtosis of 1.487, indicating a flatter distribution compared to a normal distribution. In contrast, the Simple model has a higher median of 45.108 and a mean of 44.360 and also shows a slight left skew with a skewness of -0.118. The Simple model has a higher standard deviation of 3.077 due to constant volatility compared to the different volatility modelling by the 2-factor PDV model. The kurtosis is almost similar with a value of 1.470, suggesting a similar tail behaviour to the 2-factor PDV model.

Table 2: Key statistical measures for two asset price distributions from Figure 3 at time $t = 50$, such as central tendencies, variability and distribution shapes, providing a quick overview of their statistical characteristics.

	Median	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**
2-factor PDV Markovian model	44.338	43.703	2.992	-0.221	1.487
Simple model	45.108	44.360	3.077	-0.118	1.470

5.1.2 Computing put option value

For simulated paths, we can now compute the value of the European and American put options. We set different initial prices S_0, maturity T, and initial path (different $R_{1,0}, R_{2,0}$ values). Results are summarised in Table 3.

Options prices show a stylised example with decreasing prices when we choose a higher initial price S_0 with a fixed strike price K and also American options have some premium over European options, the so-called early exercise value. We can also observe differences when we set different $R_{1,0}, R_{2,0}$, which represents a different path of an asset price. Positive $R_{1,0}$ represents positive returns in near history and negative $R_{1,0}$ represents negative returns in near history. Higher $R_{1,0}$ values represent lower volatility in near history and vice versa. While higher $R_{2,0}$ represents higher volatility in the long run. In Table 3 we can see that higher long-term volatility pushes up option prices as one would expect. A more detailed description of R_1 and R_2 is provided in Section 1.2.2.

Table 3: Option values for different initial prices S_0, $R_{1,0}, R_{2,0}$, and maturity T in years, with strike price $K = 40$. The discount rate is set at $r = 0.06$ per annum, with continuous discounting. An option can be exercised 50 times per year. The simulation is based on 100,000 (50,000 plus 50,000 antithetic) paths for the asset price process.

S_0	T = 1				T = 2			
	$R_{1,0} = 0.556,\ R_{2,0} = 0.062$				$R_{1,0} = 0.556,\ R_{2,0} = 0.062$			
	American 2PDV	American Simple	European 2PDV	European Simple	American 2PDV	American Simple	European 2PDV	European Simple
38	3.706	3.725	3.611	3.665	4.434	4.611	4.198	4.431
39	3.205	3.193	3.131	3.148	3.973	4.123	3.770	3.979
40	2.766	2.715	2.707	2.684	3.554	3.681	3.380	3.564
41	2.386	2.296	2.334	2.273	3.169	3.280	3.027	3.182
42	2.048	1.928	2.009	1.911	2.827	2.915	2.708	2.834
	$R_{1,0} = -0.733,\ R_{2,0} = 0.059$				$R_{1,0} = -0.733,\ R_{2,0} = 0.059$			
38	3.431	3.489	3.362	3.438	4.219	4.287	4.021	4.124
39	2.914	2.946	2.869	2.909	3.747	3.791	3.585	3.661
40	2.473	2.464	2.438	2.440	3.325	3.348	3.191	3.238
41	2.091	2.044	2.065	2.028	2.945	2.936	2.837	2.854
42	1.762	1.684	1.744	1.671	2.608	2.569	2.519	2.507

5.1.3 Mean value process analysis

We conduct a brief mean value process (41) analysis, which shows the difference between prices produced by our model and the Black-Scholes pricing approach from [1] for European put options. Constant volatility for the Black-Scholes pricing approach is chosen as initial volatility produced by $R_{1,0}$ and $R_{2,0}$ as in Section 5.1.2. We summarise the results in Appendix A. We observe little skewness towards the higher mean value process prices produced by the Black-Scholes pricing approach for out-of-the-money (OTM) options and skewness towards higher mean value process prices produced by our model for in-the-money (ITM) options. The difference in results is due to the different volatilities of the two pricing approaches.

5.2 Computing first derivatives

The objective of our study is to determine the optimal hedge at a specific instance, time t. To accomplish this, it is necessary to first calculate the first derivatives of the function f at the same time point, focusing on the second and fourth variables, which are R_1 and S. There are two methods

for calculating the derivative when using the Monte Carlo approach: Finite-Difference Approximations and The Likelihood Ratio Method. All definitions and information in this section are taken from the book [10].

First, we consider the finite-difference approximation method. To follow notation from [10] it's so-called *pathwise* (PW) method when considering the assumption that derivative exists with probability one. The computation of the function f is computationally relatively large, so the main problem in applying our method is the trade-off between bias, variance, and computation time. To reduce the convergence rate of the bias, we chose a central-difference estimator of the derivative at the cost of greater computational time, while having to compute an additional point of the function. The computation of the central-difference estimator follows the approach from section 7.1 in [10], where the input to the central difference estimator is the mean of all simulations of the function f. It is important to note that in order to compute the means as input, they must be i.i.d., which is represented by different seeds set up in the Matlab environment. We also had to reduce the number of repetitions for the mean computation due to the scale of our problem. An additional repetition would not only increase the calculation of the derivative but would also later affect the calculation of the hedging error. Our aim is therefore twofold: firstly, to minimize the step for the calculation of the central difference, but too low a value could lead to greater variability in the results, and the next aim is to maximize the number of repetitions in order to calculate the mean as the value of the function at the chosen point, at the cost of greater computational time.

We will have a look at this difficulty in an example later in Section 5.5. We also refer a reader to sections 7.1 and 7.2 in [10] for a more detailed explanation of approximating derivatives with the PW method in the Monte Carlo simulation.

Secondly, we can assume the *likelihood ratio* (LR) method to compute the first derivatives of the function f representing option value in time. The likelihood ratio method provides an alternative way of estimating derivatives besides the finite-difference approximation method. In our exploration of the pathwise

method, we assumed the existence of a functional relationship between a variable of the function f (and subsequently a value of the function f) and a point where we want to compute the derivative. Conversely, in the LR method, we assume that a variable has a probability density function, with a point of derivation acting as a parameter within that density. Research in section 7.3 in [10] suggests the use of LR in the calculation of second derivatives and not the first. Using the LR method in combination with the PW method shows superior performance in terms of variance of the estimator when calculating second derivatives. Application of the LR method to a first derivative has many times greater variance of the estimator and therefore in our case use of the LR method would lead to less accurate results. We have chosen the PW method as the best approach to estimate the first derivative in the Monte Carlo simulation.

Once we have the derivatives at time t, we can simply apply them to the equation (47) and we have an optimal hedge at time t, which we can also use later to calculate the hedging error (52).

5.3 Computing optimal hedging coefficient

The numerical derivation of an optimal hedging coefficient is a straightforward application of the numerical calculation of the first derivatives to the equation (47).

To begin with, we choose a central difference estimator with a step of $h_2 = 0.5$ when calculating the derivative of the second variable, namely R_1, and step $h_4 = 1$ when computing the fourth variable, namely S. We perform five i.i.d. simulations to take an average of all simulations as input to the central difference estimator. Further explanation of our chosen setup as step and number of repetitions to compute the central difference estimator will be examined later in Section 5.5. With derivatives in hand, we simply apply the results to (47).

In Figure 4 we graphically show the change in the hedge coefficient over the option's life until maturity alongside a randomly generated path. The blue line

represents an asset price path in one year and the red dotted line represents the initial asset price value $S_0 = 40$, with the y-axis on the right. On the left-hand side of the y-axis, we can see the optimal hedge levels, represented by the grey bars at the bottom of the chart.

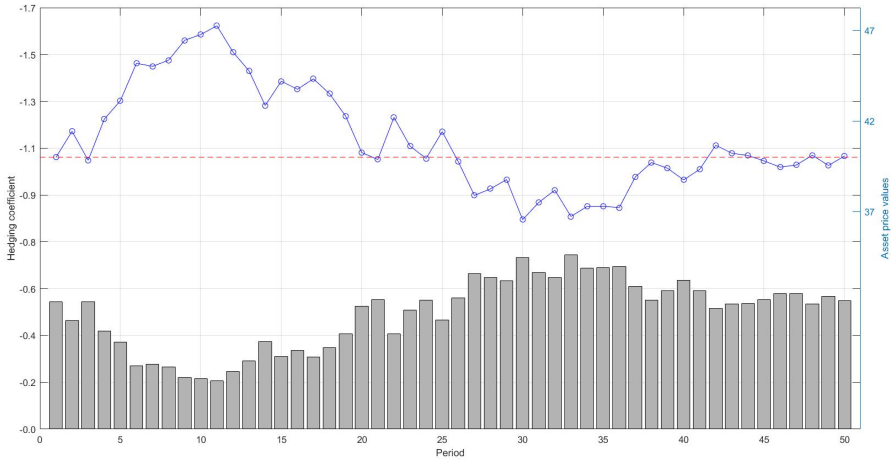


Figure 4: Generated stock price path with European put option maturity $T = 1$ in years and the option exercisable 50 times per year with put option strike price $K = 40$. The discount rate is set at $r = 0.06$ per annum, with continuous discounting. The left axis shows the values of the hedging coefficients as bars and the right axis shows the stock price as a blue line with circles when the option is exercisable.

Figure 4 shows a stylised example of a put option where the hedging coefficient moves closer towards zero as the price rises and vice versa. We present optimal hedging coefficient values in the first 10 iterations in Table 4 for the chosen path with the setup as in Table 1. We also present two different results in Table 4, one with a lower accuracy and one with a higher accuracy. The lower accuracy is based on 1,000 generated paths to estimate the derivatives of the function f. The higher accuracy is based on 100,000 generated paths. We also show the difference between these two estimates in the last row.

It can be seen that the difference between higher and lower accurate estimates is two or three decimal places after the decimal point. The cost of more accurate results is time, while it takes 20 times longer to compute the optimal hedge. In our case, this is not a problem because the lower estimate takes 2 seconds to calculate and the higher estimate takes 40 seconds. But if we want to calculate the optimal hedge in a portfolio of thousands of stocks,

Table 4: Each column represents the optimal hedging coefficient for a chosen path in the first 10 iterations. The put option strike price is $K = 40$. The discount rate is set at $r = 0.06$ per annum, with continuous discounting. An option can be exercised 50 times per year with maturity $T = 1$. The lower accuracy estimate is based on 1,000 simulations and the higher accuracy is based on 100,000 estimated paths for the asset price process. The last row represents the difference between the lower and higher estimates in absolute terms.

Accuracy	1	2	3	4	5	6	7	8	9	10
Lower	-0.5148	-0.4396	-0.5151	-0.3977	-0.3524	-0.2569	-0.2618	-0.2518	-0.2080	-0.2052
Higher	-0.5035	-0.4229	-0.5147	-0.3854	-0.3380	-0.2567	-0.2631	-0.2513	-0.2165	-0.2074
difference	*0.0113*	*0.0167*	*0.0004*	*0.0123*	*0.0144*	*0.0002*	*0.0013*	*0.0005*	*0.0085*	*0.0022*

it will make a huge difference. If we have thousands of options, we need to buy an equivalent number of assets. As a result, a difference of two or three decimal places in the calculation can lead to a variation of a few units. Again this issue becomes more pronounced when we are dealing with thousands of different assets. So it's important to decide whether you want to prioritise more accurate results or save time on calculations.

5.4 Computing expected squared hedging error

The process for calculating the expected squared hedging error is similar to that for finding the optimal hedging coefficient. At each time point t until maturity, we calculate derivations, then compute the hedging error using the formula in equation (52), and sum up these errors for the duration of the path. This procedure is repeated for a set number of generated paths, and then we take the average to determine the expected hedging error. The accuracy of this expected hedging error improves with an increased number of paths. For a more detailed explanation of how the expected hedging error is numerically computed, we refer the reader to chapter 13 in [2].

As mentioned before, we have to choose the trade-off between bias, variance, and computational time. In the case of the hedging error, we are computing derivatives to maturity, so our holding factors have a snowballing effect, producing a larger error in the end. To determine the hedging error at a specific point in time, we must calculate the derivative of the second and fourth variables. This needs to be done along the path, which multiplies the

time by the number of days available for exercising the option until maturity. As we get closer to maturity, we save on computation time since there are fewer days to simulate. Having computed one path, achieving precise results requires thousands more computations, thereby further multiplying the computational time. To present the idea of computational robustness, we compute the average hedging error of 100 paths. This takes about 13 minutes with the dimension to compute derivatives of 10,000 paths. We also run a simulation of 500 paths, which takes almost 50 minutes with a dimension of 1,000 to compute the derivatives. To illustrate the enormous computational burden, if we simulate 500 paths, we have to calculate the value of the American put option half a million times and run nearly 12 million regressions to calculate the option value. The results of the simulations are summarised in Table 5.

Table 5: Results from the computation of average hedging error with different dimensions. Simulation is the number of paths used to calculate the average error. Min and Max represent the minimum and maximum error from all simulations. Dimension represents the number of paths used in the calculation of the derivation of the second and fourth variables to calculate the hedging error.

Sim.	Min	Max	Max − Min	Average	Sd.	Total [min]	Avg. [s]	Dimension
5×10^2	15.4170	170.2640	154.8470	108.0077	6.1079	49.1	5.8951	10^3
10^2	24.2169	157.3152	133.0983	99.8576	6.0452	12.7	7.6153	10^4

A notable difference of around 8 units is observed in the average hedging errors when comparing the two simulations. The first, less precise but with five times more simulations, is used to compute the average hedging error, while the second, more precise with higher dimension in computing derivatives, involves fewer simulations for the same calculation. When these 8 units are considered as a percentage, there is an 8% increase or decrease from one error to the other, representing a significant difference that cannot be overlooked. It can be seen that the standard deviation of the more accurate calculation is lower, which supports our statement of higher accuracy. Therefore, selecting the appropriate dimension and number of simulations is crucial and must be precisely established. We will explore these and other implementation-related matters in the following section.

5.5 Numerical and practical implementation issues

We would like to discuss questions that may arise about the standard error of the simulation, why we chose Laguerre polynomials for the regression, the computational speed, the step for the central difference estimator, and the number of iterations for computing the average in the derivative approximation. All of the above problems will be discussed only briefly, and for a deeper understanding, we refer the reader to chapter 8 in [17].

First, we look at why we chose Laguerre polynomials for regression in the computation of conditional expectations in option valuation. The results of various computational experiments, as detailed in [17], show that the results obtained by the LSM algorithm exhibit considerable robustness with respect to the choice of basis functions. However, if the choice of basis functions produces a cross-moment matrix that is close to singular, this can lead to numerical errors in some least squares methods. These errors can change the functional form of the estimated conditional expectation function. To avoid these numerical problems, we use the QR algorithm to perform the least squares calculations in our regressions as suggested by [17]. To leverage the capabilities of the Matlab environment and save computational time, we use the recursive definition of Laguerre polynomials. Other types of basis functions such as the Hermite, Legendre, Chebyshev, Gegenbauer and Jacobi polynomials can also be used. We provide a numerical example where we compare the results using simple regression with constant, X, and X^2 and Laguerre polynomials of different degrees and how they affect the option price. The results are shown in Table 6, where we can see lower average values from the simple regression and a greater difference between the maximum and minimum values. We can also see from the simple regression a higher standard deviation in the choice of both degrees, which means that the results are more volatile. This implicitly supports our choice of Laguerre polynomials instead of using the simple ones. The choice of Laguerre polynomials is also based on a geometrical argument. For a more detailed explanation, we refer the reader to section 8.3 in [17].

Now we will focus on the degree of polynomials to choose. With fewer

Table 6: Comparison of simple regression with constant and some degree polynomials as X, X^2 and regression with second and third degree Laguerre polynomials. The difference between simple and Laguerre regression is shown in the given degree. The dimension is based on 1,000 generated paths calculated on 100 simulations with the setup as in Table 4.

Regression	Degree	Min	Max	Difference	Average	Standart deviation
Laguerre	3	2.7055	3.0678	0.3623	2.8671	0.2936
Simple		2.6464	3.0550	0.4086	2.8448	0.2991
difference		*0.0591*	*0.0128*	*-0.0463*	*0.0223*	*0.0018*
Laguerre	2	2.6554	3.0460	0.3906	2.8501	0.2918
Simple		2.6198	3.0256	0.4058	2.8112	0.2987
difference		*0.0265*	*0.0293*	*0.0028*	*0.0389*	*-0.0070*

degrees, we get faster computation but less accurate results and vice versa. We present in Table 7 why we have chosen a 3-degree Laguerre polynomial. One can observe almost no change in the standard deviation of the option value in 100 simulations after 3 degrees. Moreover, the difference between the mean when we use 3 degrees and 6 degrees is almost half a cent, which is well below the bid-ask spreads in liquid equity markets. Maximum or minimum values move within a few cents or less, as well as the difference between maximum and minimum values from all simulations. We also show the relative difference, which also does not change. This supports our choice of 3-degree Laguerre polynomials intending to achieve the most accurate and least computationally intensive result.

Table 7: Calculation of the value of American put option with different degrees of Laguerre polynomials on 100 simulations. The dimension is based on 100,000 generated paths with the setup as in Table 4.
*Relative difference is a difference divided by an average of the most accurate estimate, which is in the 6-degree row (last row).

Degree	Min	Max	Difference	Relative diff.*	Average	Standard Deviation
0	2.4907	2.5228	0.0321	0.0116	2.5070	0.0793
1	2.7207	2.7614	0.0407	0.0147	2.7437	0.0937
2	2.7283	2.7720	0.0437	0.0158	2.7528	0.0945
3	2.7374	2.7750	0.0376	0.0136	2.7575	0.0901
4	2.7369	2.7772	0.0403	0.0146	2.7593	0.0909
5	2.7381	2.7771	0.0390	0.0141	2.7607	0.0908
6	2.7412	2.7789	0.0377	0.0137	2.7616	0.0903

Next, we will look at choosing the optimal step h when computing derivatives of the function f or equivalently option value. Decreasing h can increase

the variance while decreasing the bias. Or to put it another way, if h is too small, inaccuracies in the estimates of the function f will make the derivatives more volatile, and if h is too large, the results will not be as accurate. Therefore, it's important to find a balance between these two factors in order to minimise the mean square error (MSE). The choice of the central difference estimator with random sampling when computing derivative values provides the smallest bias, variance and fastest convergence when compared to the forward difference estimator with independent or common sampling according to [10].

In Figure 5, we present the results of the simulation with different choices of the optimal step h and we are focusing on the MSE, standard deviation, and the difference between the option values from 100 simulations. First, we have to distinguish between the derivation of f with the second variable R_1 and the fourth variable S. We can observe different sensitivities between the two. On the left axis, we can observe the derivation with R_1 where we choose $h_2 = 0.5$ and on the right axis the derivation with S where we choose $h_4 = 1$. After the proposed thresholds, the MSE is almost zero and the average derivative values and standard deviations are almost unchanged.

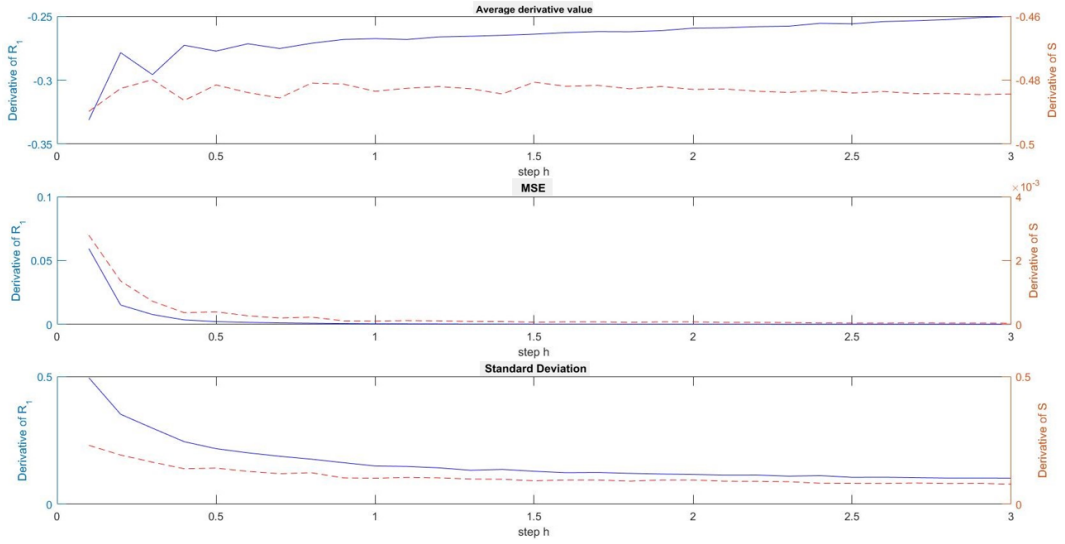


Figure 5: The graph shows the change in average derivative value, MSE and standard deviation as we increase the step h. The results are based on 100 simulations and the derivative is calculated as the average of 5 independent simulations. The derivative of R_1 is on the left axis with blue colour and the derivative of S is on the right axis with orange colour.

Another difficulty is choosing a sufficient number of simulations to calculate the derivative value from the central difference estimator. The choice of steps is adopted as suggested in Figure 5 we present the same approach as before to choose the threshold above which the values do not change. From Figure 6 we can observe that MSE and standard deviation have the same trend. However, the mean values of the derivation by R_1 move downwards and then increase and stabilise around the value -0.0276. For the derivation by S, the values are decreasing and also stabilising. Nevertheless, the values of the derivatives move around the third decimal place after the decimal point, so we suggested that calculating the average 5 times for both derivatives should be sufficient to save calculation time. Finally, to clarify our approach, we discuss the number of replicates but not the important i.i.d. property, so different seeds are used to compute the average, which is then used as input to the central difference estimator.

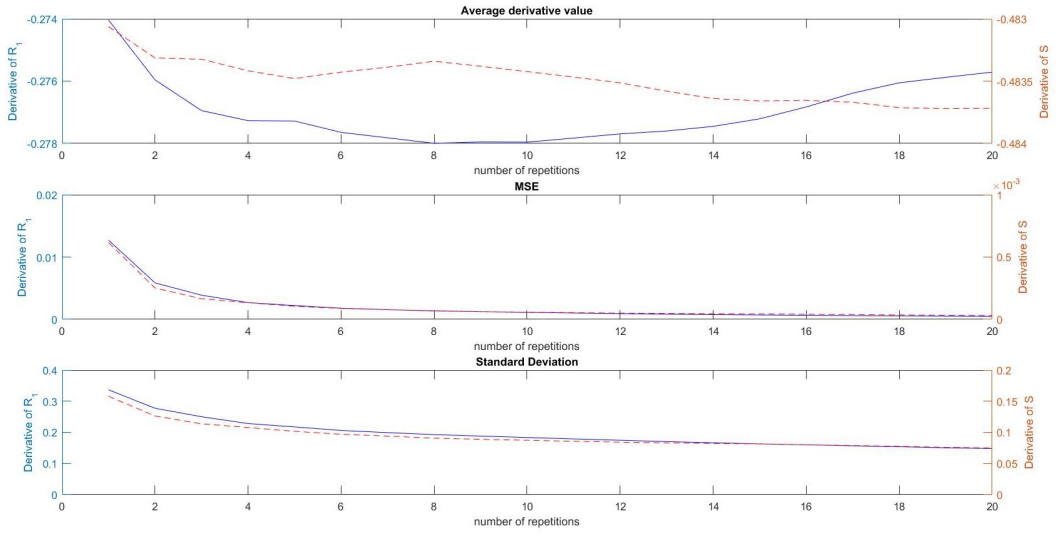


Figure 6: The graph shows the change in average derivative value, MSE and standard deviation as we increase the number of repetitions in the computation of the central difference estimator. The results are based on 100 simulations and steps are chosen as follows $h_2 = 0.5$ and $h_4 = 1$. The derivative of R_1 is on the left axis with blue colour and the derivative of S is on the right axis with orange colour.

At the end of this sub-section, we would like to mention that we have tried to implement time-saving improvements in the case of increasing computing speed suggested by [17]. We tried to simulate increments only once when

calculating option values and used it for further path generations. We have also tried to remember regression coefficients when valuing options and use them for all other valuations. However, none of these improvements gives us a significant computational advantage and in the case of a more precise value we do not recommend using it. Still, we don't rule out that it could be a time saver for large portfolios.

6 Real data application

We ask reasonable questions from the perspective of the option writer or, as is usually the case, the bank. Would our model (12) better minimise the squared hedging error of the equity portfolio when we hedge the European put option along the path with the optimal hedge (47) or delta hedge from the model than the classical simple Black-Scholes delta hedge approach? If so, what might be the reason, and if not, why shouldn't we stick with the classical approach?

However, limited resources for financial data in the industry allowed only data for all stocks from the New York Stock Exchange (NYSE). We selected from the universe of NYSE stocks that do not pay dividends and filtered out 394 public companies listed on the NYSE. Data are taken from the LSEG Datastream. We hedge the European put option contract that expires on 19 April. However, we wouldn't be able to see the price of any further expired options, while LSEG only provided us with one month's data from the last expired option, so our experiment wouldn't be as consequential as the illustrative one.

As we said at the beginning, we behave like a bank that is writing down a European put option, or rather selling a European put option, and trying to hedge its position with the aim of minimising the expected squared hedging error to maturity. We use daily data, so we hedge our position on a daily basis and assume no transaction costs. Hedging is done just before the market closes, so we assume it is possible to buy or sell shares at the closing price. We have one month to hedge before expiry in one realized path in history. This approach is usually referred to as backtesting the strategy, and we will refer to it likewise. The chapter will be structured as follows: at the beginning, we will describe the data as inputs to the model, then we will explain the detailed procedure related to the calibration of the model, and then we will show step by step the procedure of our analysis and evaluation, and at the end, we will present our results and try to answer the research questions.

6.1 Inputs to the model

Model (12) is entered with data from the LSEG Datastream. Table 8 presents all input data categories. Data are structured as follows: Stock price history as mid-price calculated as an average of bid and ask prices, European put option price also as mid-price, implied volatility, and finally option delta from the Black-Scholes approach. All prices are taken as closing prices. We have used a two-year history for stock price and implied volatility, and a 1-month history for option price and option delta. As mentioned above, we were unable

Table 8: Data categories used in backtesting.
*Implied volatility is calculated from the at-the-money option. We provide more detailed information in Section 6.2

Data categories	Description
Stock price	Closing mid-price as an average between the bid and ask prices
Implied volatility*	Volatility as the output of the Black-Scholes model when we insert option mid-price
Option price	Closing mid-price as the average between the bid and ask prices
Option delta	Black-Scholes option delta from [1]

to obtain data on expired options and some data was missing, which made our backtesting more difficult. We can't use data with missing values, so we reduced the sample size to 394 public companies, and a smaller sample size reduces the significance of the results.

6.2 Calibration

Approach to calibration of the model (12) was taken from [12]. The first step is to calculate the two factors R_1 and R_2 from the stock price history at each point in time, or in our case, daily. In the second step, run the simple linear regression (54) with the dependent variable being implied volatility and the independent variables being R_1, R_2, as it is done in chapter 3 in [12].

$$\text{Implied volatility}_t = \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}} \tag{54}$$

We used the first year of two years of stock price history to calculate the factors R_1 and R_2. The following year was used for the regression calculation (54), which implied the constants $\beta_0, \beta_1, \beta_2$. To clarify, we do not use a calibration

period in backtesting, or in other words, we calibrate the model out of the test sample.

While the calibration in [12] was done on the S&P 500 Index (SPX) with implied volatility defined as the Volatility Index (VIX), we try to apply a similar methodology in our approach. LSEG provided us with implied volatility from at-the-money option prices with a one-year history on selected stocks from the NYSE. We think it is important to explain how the VIX index is constructed and that our approach is similar to that used in [12], so we have dedicated a Section 6.2.1 to describing it. In our case, we are using implied volatility from LSEG Datastream with their applied methodology described only briefly in [18]. We requested a more detailed answer on how LSEG calculate implied volatility from option prices, but unfortunately, we didn't receive a reply.

6.2.1 VIX index

In 1993, Cboe Global Markets, Incorporated (Cboe) introduced the original version of the Cboe VIX Index, which was initially designed to measure the market's expectation of 30-day volatility implied by at-the-money option prices on the S&P 100 Index (OEX Index). Ten years later, in 2003, Cboe, along with Goldman Sachs, updated the VIX Index to reflect a new way of measuring expected volatility that continues to be widely used by financial theorists, risk managers and volatility traders alike. The new VIX Index is based on the SPX, the core US equity index and estimates expected volatility by aggregating the weighted prices of SPX puts and calls over a wide range of strike prices. The price history for the Cboe S&P 100 Volatility Index (VXO) matches the values for the original version of the VIX index, which was introduced in 1993. [26]

6.3 Backtesting

To illustrate the backtesting process, we will look at the example of one stock in more detail and then show how volatility is captured compared to implied or historical volatility. We then describe the whole process of backtesting and

how we try to measure the effectiveness of different hedging strategies and the reasoning behind them.

6.3.1 Single stock backtesting: case study of Alphabet Inc.

As an illustration, we choose *Alphabet Inc. Class C* share or the better-known company behind the famous internet search engine Google. In Figure 7 we plot 4 data series on the three graphs. In the first graph, starting at the top, we show the hedge coefficient from the model optimal hedge defined in (47), the delta hedge from the model as the second part of (47), and the delta hedge from the Black-Scholes option pricing method explained in LSEG Datastream [18]. The second graph shows two option prices, one historical from the stock market and the other calculated using the Monte Carlo method described in Section 5.1.2 on the left axis and the stock price and the strike price on the right axis. The last graph shows volatility levels from implied volatility, historical volatility as rolling 100-day standard deviation, and model volatility defined as (2).

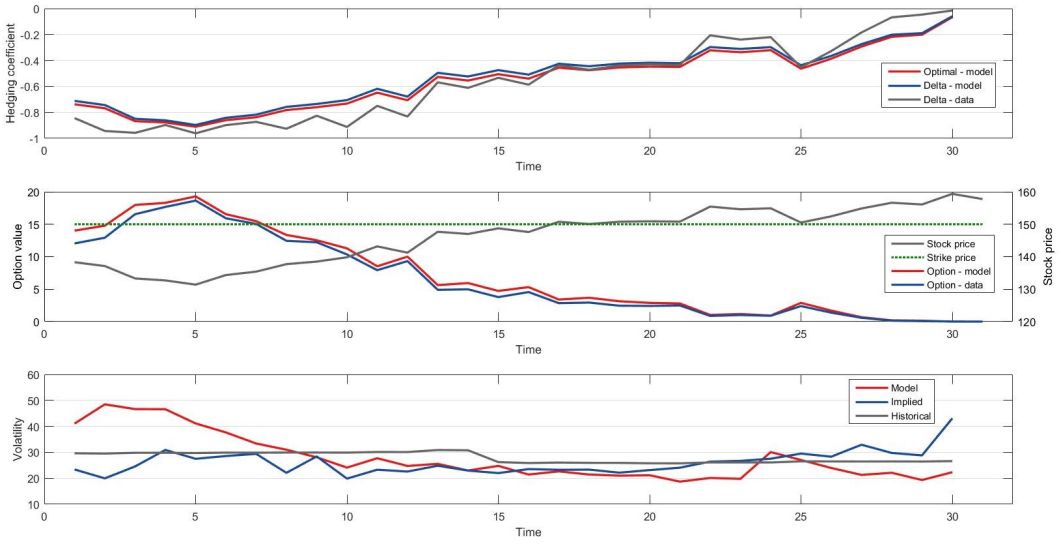


Figure 7: Alphabet Inc. Class C share: hedging coefficients, option price, share price and volatility from model and market data. The strike price is randomly chosen around the share price at the start of the valuation period ($K = 150$).

We can see the difference between implied and model volatility in Figure 7. The steep downward move in Alphabet's share price just before the option

was valued affected the first factor R_1 and captured the leverage effect, pushing volatility up, almost doubling it relative to implied volatility. Therefore, we can observe a premium in the option valuation at the beginning of the 30-day period. The option was valued at 14.02 dollars and can be bought on the NYSE at 12.05 dollars at the beginning of the period. The strike price is chosen randomly around the initial share price as a whole number to meet the requirement of the existence of the option with that strike price. The hedging coefficients have similar trajectories, depending on the different distributions of the simulated paths explained in Figure 3, and therefore different delta hedge and optimal hedge produced by the model.

All the results of hedging along the historical path are summarised in Table 9, where we show the quadratic hedging errors for different strategies generated from the hedging coefficients calculated as (27). We show the Mean Absolute Percentage Error (MAPE) between the market option price and the model option price, and between implied and model volatility. We set the initial endowment x as the option price in the market month before expiration, we will explain this choice in more detail later in Section 6.3.2. From now on, we will refer to daily hedging to maturity with a chosen hedging coefficient as a hedging strategy.

Table 9: Alphabet Inc. Class C shares: quadratic hedging error, the option price difference between model and market data using delta and optimal hedge, and the difference between implied and model volatility along a historical path. The MAPE is used to assess the difference between option pricing and volatility modelling.

Quadratic hedging error			**Option price**	**Volatility**
Optimal (model)	**Delta (model)**	**Delta (B-S)**	**MAPE**	**MAPE**
3.210	3.381	3.578	18.813	49.753

To explain the difference in option prices in Figure 7, we need to understand the model (12) that explains the price movement. We can observe a small premium at the beginning, determined by the higher underlying volatility. After 10 periods we can still see the small premium explained by the higher long-term volatility determined by R_2, while the simulated paths from which the option price is calculated as a discounted cash flow have different

distributions as illustrated in Figure 3. In the last 10 periods where the option is out-of-the-money (OTM), the premium disappears and the small difference could be explained by the calibration of the model, which is still 2-factorial and doesn't sufficiently explain the volatility in the data, as explained in [12].

For the illustration, we also simulate the 500 paths of the share price, with the initial price S_0 set as the Google share price at the time of the first option valuation, i.e. one month before expiration. We then calculate the squared hedging error on each path with the optimal hedging coefficient. We plot the distribution of the squared hedging errors in Figure 8. The median of around 11.2 and the average of around 14.9 shows that we are closer to zero on the distribution than the average or median. Not surprisingly, the distribution is concentrated around zero, and the probability of a higher squared hedging error is significantly lower. We haven't experienced any major market downturns or other market shocks that would distort our results during the backtesting period.

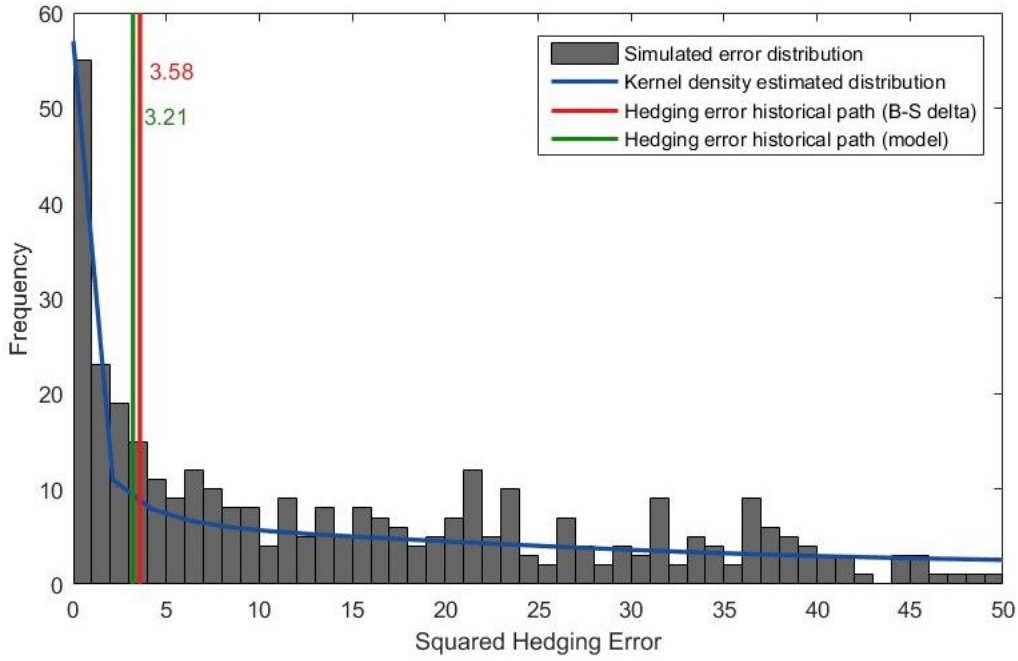


Figure 8: Squared hedging errors distribution of hedging with optimal hedging coefficient for Alphabet Inc. based on over 500 simulations, with hedging error on the historical path with optimal and Black-Scholes delta hedging coefficients.

6.3.2 Comprehensive backtesting of 394 stocks

Now we backtest 394 historical stock paths each as in Section 6.3.1 calibrated as explained in Section 6.2. Then, if we need to evaluate the effectiveness of the optimal hedge, the delta hedge from the model (12) and the Black-Scholes delta hedge, we need to set a performance measure for that.

As an indicator of the effectiveness of hedging coefficients, we look at the difference between the quadratic hedging errors produced by daily hedging until maturity. We buy or sell a fraction of the shares on a daily basis to hedge as determined by the optimal hedge, the model's delta hedge and the Black-Scholes delta hedge coefficients. We apply the squared hedging error minimisation problem (27) to the realised historical path. As we said, we are looking at the problem from the bank's perspective, so we are writing down the option and selling it at a price. The value for which we sell an option is our initial endowment or x as in (27).

We selected the initial endowment based on the price of an option on the NYSE. The rationale for this choice is simple: if our valuation of the option is above the market price, we would prefer to sell it at that higher price, reflecting our view of its value. However, potential buyers are unlikely to pay a premium if the option is available from other banks at a lower market price. Market dynamics therefore force us to sell at the market price. Conversely, if our valuation is below the market price, selling the option at the market price will result in a surplus as it will exceed our estimated value. We hedge our position using the optimal hedge, the model's delta hedge, and the Black-Scholes delta hedge. At expiration, we settle the option based on the predetermined strike price. We then calculate the quadratic hedging error and compare the results using different hedging coefficients.

We present the results in Figure 9, which shows a box plot of the differences across all 394 stocks on the NYSE, starting on the left, we plot the difference between the quadratic hedging errors with the optimal hedge and the Black-Scholes delta hedging strategies, then the optimal hedge and the delta hedge in the middle, and finally the difference between quadratic hedging errors of

the delta hedge from the model and the Black-Scholes delta hedging strategies.

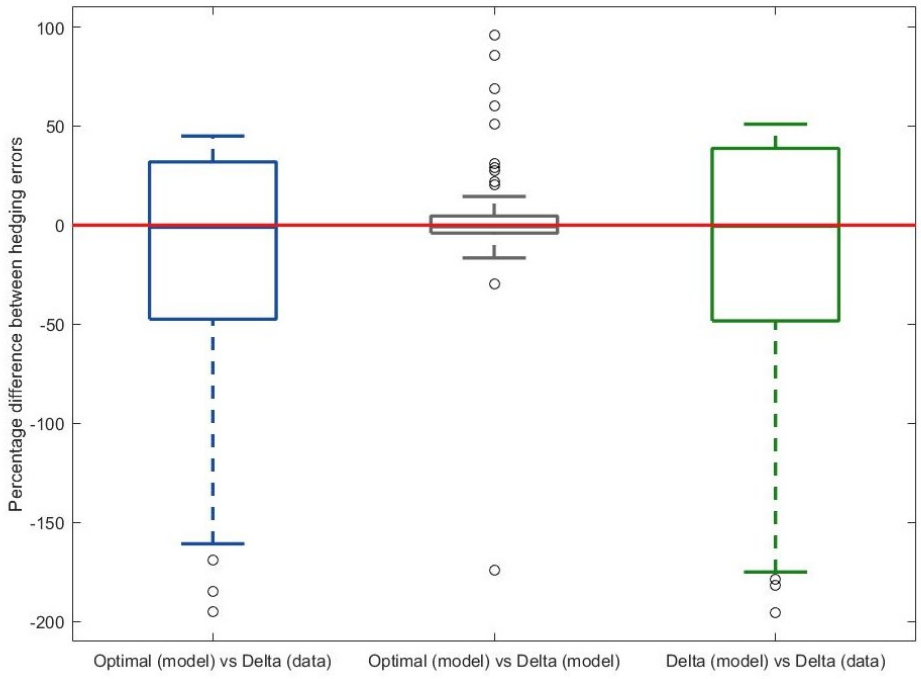


Figure 9: Difference in percentage between quadratic hedging errors of 394 stocks with different hedging strategies. The difference is calculated as the first value on the x-axis name minus the second value on the x-axis name. The zero level is highlighted in red.

At first glance, it is easy to see that the quadratic hedging error, produced by applying the optimal hedging strategy, does not perform significantly better than a simple Black-Scholes delta hedging strategy with the median of around -1.05%. The difference between quadratic hedging errors produced by the optimal hedging and the model delta hedging strategies is almost similar, with a median difference of around -0.20%. The difference in quadratic hedging errors between the model's delta hedging and the Black-Scholes delta hedging strategies is close to zero, with a median of -0.83%, and all free boxplots fit zero within the first and third quartiles, indicating no significant difference between the three hedging strategies. We choose the median instead of the average because of the higher number of outliers in the sample. We present in Table 10 average and median MAPE across all stocks. We wanted to highlight the difference between the market option price and the model option price around 11.2% in median, which we think is more accurate than average in

Table 10: Average and median for each of the stocks MAPE: The market option price and model option price, implied volatility and the model volatility.

	MAPE	
	Option value	**Volatility**
Average	13.247	28.789
Median	11.206	28.817

terms of the few outliers that appeared in our analysis. The difference between implied and model volatility is around 28.8% at the median. The MAPE measure is used to give a picture of the error that occurs between option prices and volatility. While both option prices and volatility sometimes over- or under-react to each other, we choose the MAPE measure while penalising error in absolute terms.

It is also worth noting that the average or median percentage return on stocks in the 10 days prior to the first option pricing was slightly skewed to the downside, with an average return of -0.53% and a median return of -0.47% for all stocks.

6.4 Analysis and discussion of backtesting results

In our analysis, we showed that quadratic hedging errors were not significantly lower than the classic Black-Scholes delta hedging approach. However, the data indicated that optimal hedging slightly outperformed model delta hedging, while model delta hedging results were relatively similar to those of the Black-Scholes delta hedging approach. It is important to recognise several limitations to our analysis, which make the results more illustrative than informative.

Firstly, our dataset contained only 394 data for different stocks. This sample size is considered small to hedge against expected squared hedging errors, where millions of paths are expected to be hedged. In addition, we restricted our analysis to a single month or instrument, choosing a period in which, as noted in the previous sections, the market experienced a mild downturn of only -0.47%, whereas typical SPX movements range from -1% to 1% [23]. As a result, we think that our sample was relatively limited, covering only one

hedging period or only one contract to maturity.

Unfortunately, we did not have access to historical option price data for a wider range of stocks. Two main dimensions could be further explored. First, it would be beneficial to conduct analyses on more contracts with a long historical record. Second, extending the analysis to a larger pool of stocks, ideally thousands, would potentially increase the robustness and relevance of the results. The limitations sets the stage for future research.

Second, our analysis is based entirely on the 2-factor PDV model for the formation of prices, which, as Guyon and Lekeufack note [12], may not sufficiently capture volatility. To more accurately model stock price movements, they recommend using at least a 4-factor Markovian PDV model. As a result, our option pricing and the consequent hedging coefficients may not be as accurate. Examining the effectiveness of a 4-factor Markovian PDV model through backtesting with a larger dataset could provide insightful results.

In addition, we used the Monte Carlo option pricing method with only 1,000 simulated paths, which limits the accuracy of our pricing and hedging calculations. Using greater computing power to simulate at least 100,000 paths, as suggested by Longstaff [17], would likely improve the accuracy of our results.

These improvements address potential drawbacks and could further verify whether the optimal hedging strategy derived from the PDV model outperforms the classic Black-Scholes delta hedging strategy and answer our first question, if the optimal hedging strategy produces lower quadratic hedging error than the classic Black-Scholes delta hedging strategy. In addition, our second question concerns the development and backtesting of optimal hedging strategies. In our view, these strategies are crucial for a more realistic assessment of market risks and could provide banks with more effective tools for minimising hedging errors, which could lead to better risk management and more realistic pricing in financial markets.

Conclusion

In this master's thesis, we examined the quadratic hedging problem in the 2-factor Markovian path-dependent volatility model, which derives the volatility from the historical price data of the asset itself. We diverged from the classic Black-Scholes framework to pricing the contingent claims [1]. Therefore, we derived the optimal hedging coefficient, the contingent claim value and the expected squared hedging error resulting from the replication error in the incomplete market. We then solved the problem numerically using a Monte Carlo simulation applied to European and American put options and applied it to the real data. Backtesting on historical data for European put options quoted on the New York Stock Exchange showed no significant difference in the optimal hedging strategy compared to the Black-Scholes delta hedging strategy and the delta hedging strategy generated by the 2-factor Markovian PDV model. A smaller data sample characterised the results.

In the first chapter, we introduced the 2-factor Markovian path-dependent volatility model from [12]. We briefly explained its construction and interpreted the motivation for PDV models in general. Then we showed constraint (15) for volatility to be bounded away from negative values for a one-dimensional diffusion process (13). We discussed the choice of using the single normalized exponential kernels (3) in the model and described two factors, R_1 and R_2. We also explained how R_1 captures the leverage effect and how R_2 captures the volatility clustering effect.

Chapter 2 was devoted to presenting the essential theory of stochastic calculus and martingale theory as a basis for further derivations in the thesis. In Chapter 3 we conducted a review of the brief history of hedging strategies from [21] and [25], namely local risk-minimising strategy and mean-variance optimal strategy.

We applied the quadratic hedging problem from [6] to the 2-factor Markovian path-dependent volatility model. Firstly we derived two crucial processes as suggested by [6]: the opportunity process (33) and the adjustment process (36). Afterwards, we derived the variance-optimal martingale measure

(VOMM). Our results demonstrate that the objective measure P, the opportunity measure P^*, and the VOMM as Q^* are equal (40), assigning identical probabilities to all events in the considered probability space. Accordingly, we have calculated the optimal hedging coefficient (47), the expected squared hedging error (52), and derived continent claim value given by bounded continuous function as a unique classical solution of a time-dependent 3-dimensional PDE (42). We numerically solved the problem in Chapter 5 by applying the Monte Carlo method from [17] to price the European and American put options. We presented practical implementation issues supported by [10] to show all the drawbacks associated with implementing the 2-factor Markovian PDV model to the option pricing.

In Chapter 6 we applied the theory of quadratic hedging with a combination of derived mathematical and numerical methods for optimal hedging and put option pricing from a bank's perspective. We backtested hedging of the European put option 30 days before expiry until maturity and measured the squared hedging error produced by the optimal hedging, the model delta hedging, and Black-Scholes delta hedging. The results for the smaller sample of 394 stocks showed no significant outperformance of any of these strategies in terms of squared hedging error, with the value of the zero differences for each comparison falling within the interquartile range, with median differences of -1.05%, -0.20% comparing the optimal hedging strategy with delta hedging and Black-Scholes delta hedging strategies respectively, and a median difference of -0.83% comparing model delta hedging and Black-Scholes delta hedging strategies. The option value had an average mean absolute percentage error of around 13.3 comparing the model and market prices. Implied volatility and the model volatility had an average mean absolute percentage error of around 28.8, supporting our assumption that the 2-factor Markovian PDV model cannot sufficiently explain the volatility in the data, as shown in [12].

The main contribution of this thesis lies in the comprehensive description and complex solution of the quadratic hedging problem. Starting with an explanation of the perspective type of the PDV model, the quadratic hedging

theory is applied to the model, followed by a numerical solution of the problem, and finally, the derived methods are used for back-testing on real historical data.

We see a number of possible future extensions of this work. First, we have applied the quadratic hedging theory only to the 2-factor Markovian PDV, and one can instead use the 4-factor Markovian PDV model, which captures the market volatility relatively well according to [12]. We would suggest using the quadratic hedging theory from the conceptual framework proposed by [6]. Secondly, our backtesting could be extended to the broader sample of stocks from two dimensions: one with a deeper history of expired options and the other with a greater number of stocks used from different markets.

Bibliography

[1] Black, F., Scholes, M.: The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, Vol. 8, 1973, pp. 637-654.

[2] Černý, A.: *Mathematical techniques in finance: Tools for incomplete markets*, Second edition, Princeton University Press, 2009, ISBN 978-0-691-14121-3.

[3] Černý, A., Czichowsky, C. and Kallsen, J.: Numeraire-Invariant Quadratic Hedging and Mean-Variance Portfolio Allocation, *Mathematics of Operations Research*, 2024, Website available at (10.05.2024): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3944947

[4] Černý, A., Kallsen, J.: Hedging by Sequential Regressions Revisited, *Mathematical Finance*, Vol. 19(4), 2009, pp. 591-617.

[5] Černý, A., Kallsen, J.: Mean-Variance hedging and optimal investment in Heston's model with correlation, *Mathematical Finance*, Vol. 18(3), 2008, pp. 473–492.

[6] Černý, A., Kallsen, J.: On the structure of general mean-variance hedging strategies, *The Annals of Probability*, Vol. 35(4), 2007, pp. 1479-1531.

[7] Foschi, P., Pascucci, A.: Path dependent volatility. *Decisions Econ Finan*, Vol. 31, 2008, pp. 13–32.

[8] Föllmer, H., Sondermann, D.: *Hedging of non-redundant contingent claims*, chapter 12, pp. 205–223, North Holland, Amsterdam, 1986.

[9] Föllmer, H., Schweizer, M.: Hedging of Contingent Claims under Incomplete Information, In Mark H.A. Davis and Robert J. Elliott (editors): *Applied Stochastic Analysis*, Vol. 5, 1990, pp. 389–414, Gordon and Breach, London/New York.

[10] Glasserman, P.: *Monte Carlo Methods in Financial Engineering*, Springer New York, 2003, ISBN 978-0-387-00451-8

[11] Guyon, J.: Path-Dependent Volatility, 2014, Website available at (10.05.2024): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2425048

[12] Guyon, J., Lekeufack, J.: Volatility Is (Mostly) Path-Dependent, *Quantitative Finance*, Vol. 23(9), 2022, pp. 1221-1258.

[13] Harrison, J. M., Kreps, D. M.: Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory*, Vol. 20, 1979, pp. 381-408.

[14] He, S., Wang, J., Yan, J.: *Semimartingale Theory and Stochastic Calculus*, Science Press and CRC Press, 1992, ISBN 0-8493-7715-3.

[15] Heath, D., Schweizer, M.: Martingales versus PDEs in Finance: An Equivalence Result with Examples, *Journal of Applied Probability*, Vol. 37(4), 2000, pp. 947–957.

[16] Jacod, J., Shiryaev, A. N.: *Limit Theorems for Stochastic Processes*, Springer-Verlag Berlin Heidelberg, 2003, ISBN 3-540-43932-3.

[17] Longstaff, F.A., Schwartz, E.S.: Valuing American Options by Simulation: A Simple Least-Squares Approach, *The review of financial studies*, Vol. 14(1), 2001, pp. 113-147.

[18] LSEG Datastream, Options and Derivatives Methodology, Website available at (10.05.2024): https://www.lseg.com/en/data-analytics/financial-data/pricing-and-market-data/fixed-income-pricing-data/futures-options-derivatives

[19] Merton, R. C.: Theory of Rational Option Pricing, *The Bell Journal of Economics and Management Science*, Vol. 4(1), 1973, pp. 141–183.

[20] Øksendal, B.: *Stochastic Differential Equations: An Introduction with Applications*, Fifth Edition, Springer-Verlag Berlin Heidelberg, 1998, ISBN 3-5400-4758-1.

[21] Pham, H.: On quadratic hedging in continuous time, *Mathematical Methods of Operations Research*, Vol. 51, 2000, pp. 315-339.

[22] Protter, P.: *Stochastic Integration and Differential Equations*, Second Edition, Springer-Verlag Berlin Heidelberg, 2005, ISBN 3-540-00313-4.

[23] S&P Dow Jones Index, Overview, Website available at (10.05.2024): https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview

[24] Shreve, S.: *Stochastic Calculus for Finance II*, Springer-Verlag New York, 2004, ISBN 0-387-40101-6.

[25] Schweizer, M.: A guided tour through quadratic hedging approaches, In E. Jouini, J. Cvitanić, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, 2001, pp. 538-574.

[26] VIX Index, Overview, Website available at (10.05.2024): https://www.cboe.com/tradable_products/vix/

[27] Zumbach, G.: Volatility conditional on price trends, *Quantitative Finance*, Vol. 10(4), 2010, pp. 431-442.

Appendix

A Mean value process analysis

We showed in (41) that the mean process value at time t is the expected value of H on filtration. We then solved the problem numerically in Chapter 5. We intend to analyse the mean value process in comparison to the classical Black-Scholes option pricing approach with constant volatility defined as the model volatility at the time the option is priced. We take a look at European put options with different maturities and plot the results. For the Black-Scholes option pricing approach from [1] we use the function *blsprice* in the Matlab environment.

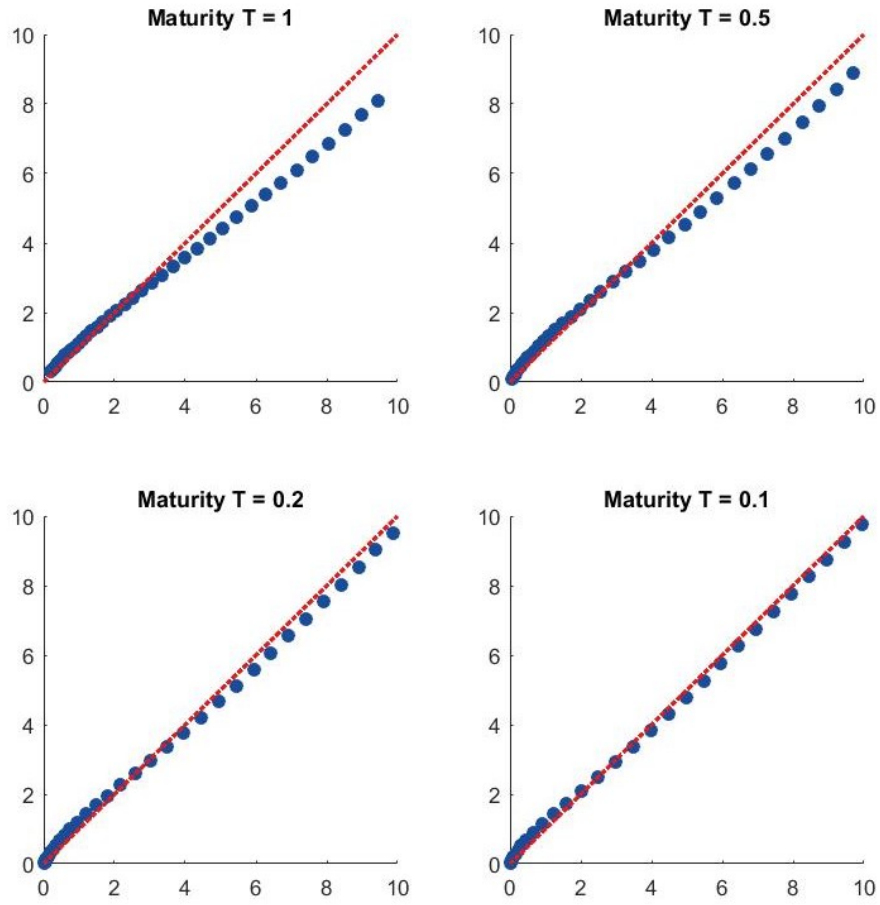


Figure 10: Mean value process comparison for the Black-Scholes option price on the y-axis and for the 2-factor Markovian PDV model option price on the x-axis exercisable **50 times** in a year. European put options with S_0 ranging from 20 to 60 with the step of 0.5 and different maturities in years are shown in the graphs. Risk-free rate $r = 6\%$ and strike price $K = 40$, and number of paths simulated for the model $N = 50,000$.

We plot two series of graphs: one where the option is exercisable 50 times per year in Figure 10 and one where it is exercisable 250 times per year in Figure 11 for the Monte Carlo pricing approach. We have chosen an initial price S_0 in the range 20 to 60 with a step of 0.5 and strike price $K = 40$. We are simply pointing out that out-of-the-money options are less valuable than in-the-money options, as one would expect. We plot the values of the options with different maturities. We can see a convergence of results as we approach maturity, seen as points closer to the red line. With a maturity of more than one month, represented as $T = 0.1$, we can observe a greater divergence for higher-priced options. The difference is skewed towards the lower value for the

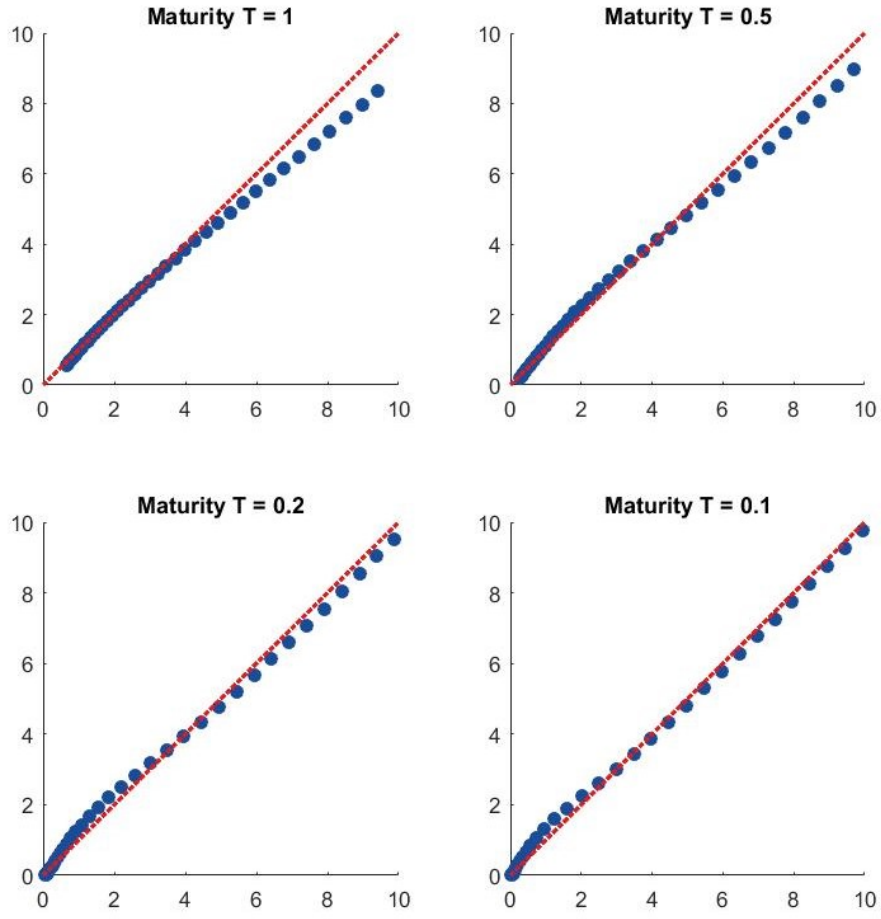


Figure 11: Mean value process comparison for the Black-Scholes option price on the y-axis and for the 2-factor Markovian PDV model option price on the x-axis exercisable **250 times** in a year. European put options with S_0 ranging from 20 to 60 with the step of 0.5 and different maturities in years are shown in the graphs. Risk-free rate $r = 6\%$ and strike price $K = 40$, and number of paths simulated for the model $N = 50,000$.

Black-Scholes pricing approach for ITM options. For OTM options, we see higher option prices for the Black-Scholes pricing approach.

We explain the skewness with the initial factor values $R_{1,0}$, $R_{2,0}$ and other coefficients set as in Table 1. As we can see in Figure 1, the average volatility of all paths is initially priced higher and then converges to a lower stable value. This is determined with the asset price drawdown close to the option pricing period. There is also the error in the price determined by the Monte Carlo simulation approach, which could not be overlooked.

We also see little difference in the results when we compare the cases of 50 or 250 times exercisable per year. However, we have to spend more computing time to get slightly more accurate results. So again, one needs to be aware of the trade-off between computing time and accuracy.