

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMETRICS AND OPERATIONS

Dynamic Hedging in Incomplete Markets

MASTER THESIS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of
Science in Quantitative Finance and Actuarial Science

Author: Julien Ottink

DEPARTMENT OF ECONOMETRICS AND OPERATIONS
RESEARCH:
QUANTITATIVE FINANCE AND ACTUARIAL SCIENCE

Principal Supervisor	DR. N. SCHWEIZER
Second Reader	PROF. DR. B. MELENBERG
Student number	1273799
Field of Study	Quantitative Finance and Actuarial Science
Contact Details	m.j.ottink@tilburguniversity.edu
Submission Date	25.09.2022

Abstract

This master thesis explores the hedging problem in a general incomplete market setting. The goal of the hedger is to reduce the basis risk of a non-tradable asset using imperfect hedging instruments. We follow a methodology developed in the context of dynamic minimum-variance hedging by Basak and Chabakauri (2011). We extend the class of dynamic hedges presented in this paper to the setting where the hedging instrument and the risk position are cointegrated. Furthermore, we constructed a numerical approach to extend the hedging strategy for an Asian option using a LSMC algorithm in a market setting with two Geometric Brownian Motions. Using Monte Carlo simulations, we evaluate the hedging performance of these two extensions. We find that a hedger can significantly improve his/her hedging effectiveness by accounting for incorporating the possibility of cointegration between assets. In a numerical experiment we show that the hedging effectiveness is approximately equal for Asian options of different strike prices, but that their hedging positions differ over the hedging horizon.

Contents

1	Introduction	1
1.1	Literature review	2
1.2	Research Objective	3
2	Theoretical framework	4
2.1	Stochastic processes	4
2.2	Incomplete market	5
2.3	Variance-minimize criterion	5
2.4	Stationary Basis	6
2.5	Trading cost	11
3	Hedging strategies	12
3.1	Time consistency	12
3.2	Dynamically time-consistent hedging strategy	13
3.3	Pre-commitment strategy	17
3.4	Myopic hedging strategy	19
3.5	Delta Tolerance strategy	19
4	Option Hedging strategies	21
4.1	Plain Vanilla Options	21
4.2	Asian Option	21
4.3	Least Squares Monte Carlo (LSMC) algorithm	22
5	Models	27
5.1	2GBM model	27
5.2	Stationary spread model	30
6	Simulation results	33
6.1	Model error	34
6.2	Option hedging	36
7	Conclusion	44
A	Appendix	45
A.1	Proof of Corollary 3.2.1	45
A.2	Derivation of Proposition 3	46

List of Figures

1	Simulated processes.	9
2	Ratio of the standard deviation of hedged $(X_T - M_T^*)$ and unhedged position X_T plotted for $\rho \in [-1, 1]$. Evaluated for $T = 2$, $X_0 = 10$, $S_0 = 10$, $r = 0.01$, $m = 0.2$, $\nu = 0.2$ and Sharpe ratio $S_a \in [0.45, 3.15]$...	28
3	The difference in hedging effectiveness for different correlation values and mean reversion parameters. Asset parameters: $\mu = 0.1$, $\sigma = 0.2$, $\bar{X} = 0.1$, $I_0 = 12$, $S_0 = 10$, $X_0 = 0.08$. Monte Carlo parameters: $T \in [0.2, 2]$, $m = 100000$ and $N \in [50, 500]$	35
4	Hedging position quantiles for $\rho = 0.9$, $\kappa = 3$ and $T = 1$	35
5	HE as function of moneyness for different basis function sets. Monte Carlo parameters: $N = 250$ and $m = 100000$. Basis functions parameters: $J = 22$, $L = 4$, $P = 32$; $3 \leq K_i \leq 25$; for the linear spline basis functions; $3 \leq K_i \leq 35$. for the bucket functions.	37
6	Measure of error $\mathcal{M}_1$ as function of moneyness for different basis function sets under the driftless $\mu = m = 0$ scenario. Monte Carlo parameters: $N = 250$ and $m = 100000$	38
7	Quantiles of optimal dynamic hedging strategies for different levels of moneyness and volatility. Monte Carlo parameters: $N = 250$ and $m = 100000$	40
8	Simulated paths for X_t and ATM and OTM hedging strategies.	41
9	Illustration of the cumulative transaction cost and percentage of trades for different values of bandwidth H	42

List of Tables

1	Statistics of Case 1 and Case 2	10
2	Functional form of basis functions. $u(x) = a(x) + (1 - \frac{t}{T})x$ and c is a constant	24
3	Functional form of basis functions; $g_j(x) = [u(x) - K_i]^+ - [u(x) - K_{i+1}]^+$; $k_p(x) = [x - K_i]^+ - [x - K_{i+1}]^+$	24
4	Hedging strategy sensitivity with respect to volatility parameters.	29
5	Parameter values for base case Asian option	36
6	Parameters: $T = 1, S_0 = 10, X_0 = 10, M_0 = 1, r = 0.02$ and $\sigma = \nu = 0.2$, $\mu^* = 0.1$ and $m^* = 0.1$. Simulated using $m = 100000$ paths and $N = 250$ time steps.....	39
7	Hedging effectiveness as function of bandwidth H for different transaction cost scenario's. Calculated for an OTM option. Asset parameter as given in Table 5.....	42

1 Introduction

Hedging is a strategy in finance that is used to reduce risk. It is used to offset the risk of any adverse price movement and is therefore an important part of risk management. In the case of dynamic trading and a frictionless complete market perfect hedging, constructed through no-arbitrage methods, could theoretically eliminate risk completely. However, in reality investors and intermediaries are not able to hedge perfectly, because of market frictions such as the inability to trade continuously, stochastic volatility, jumps in the underlying investment, and transaction cost. Therefore, perfect replication is not feasible and hedging is only an approximation of the final payoff function, and as such it causes a non-trivial error. In this thesis we will focus on the risk associated with hedging a non-tradable security using an imperfectly correlated tradable asset. The risk that arises due to the introduction of the aforementioned non-trivial error is called *residual risk* or *basis risk*. Prominent examples of parts in the financial sector where residual risk is present are:

- Futures and forward contracts that are used by energy market participants to buy or sell a specific volume of gas or electricity, at a certain price, for settlement on a specific date in the future. Futures and forwards can be used to hedge price risk of a future risk position. To guarantee liquidity in the energy market, futures are highly standardized. However, there could be a mismatch between the expiration date of the futures and the date on which the risk position is sold or received, therefore inducing residual risk.
- Stocks that have been grouped in an index are called index options. Index options are generally hedged by using a certain number of, but not all, the underlying assets. As a consequence, basis risk arises due to imperfect correlation between the underlying assets and the index.

So why do we care to minimize basis risk by hedging? Ankirchner and Imkeller (2011) and Ankirchner, Schneider, et al. (2014) illustrate that in commodity and complex life insurance hedging even a relatively large correlation between assets only reduces basis risk by half. Insurance companies generally use dynamic hedging in financial products to hedge the risk from their variable annuity business. Basis risk arises due to imperfect correlation between the underlying mutual fund and the assets used to cross hedge the financial obligations.

A static hedge minimizes the basis risk at the initial start of the hedging horizon and holds the number of assets used to hedge constant throughout the this horizon. While static hedges are easy to implement, they can be very inefficient. A static hedging strategy does not have the mechanism of updating itself as new information unfolds with time, and is therefore vulnerable to future market changes and generally less effective than a dynamic strategy. Dynamic hedging involves a strategy that rebalances hedging positions as market conditions change. A breakthrough in dynamic hedging is made by Basak and Chabakauri (2010) who obtain an explicit time-consistent optimal policy using a mean-variance framework. Basak and Chabakauri (2011) extended this framework to a minimum variance hedging error criterion. The paper by

Basak and Chabakauri (2011) will be the basis of the hedging strategies derived in this thesis.

1.1 Literature review

Schweizer (1994) and Duffie and Richardson (1991) solve the dynamic minimum-variance hedging problem using a quadratic theory framework. Duffie and Richardson (1991) derive an explicit solution for the minimum variance criterion subject to a predetermined expected hedging error under Geometric Brownian Motion (GBM) asset prices. The hedging strategy derived minimizes the hedging error variance at an initial date. These strategies are classified as "pre-commitment" strategies, as the strategy becomes sub optimal at future time point on the hedging horizon unless the hedger is able to pre-commit to the strategy. Several authors have formulated the problem of hedging basis risk in terms of the utility maximization approach. The utility maximization objective is another way of specifying the hedgers appetite for the amount of basis risk. Monoyios (2004) used the utility framework to derive analytic approximations for the price and hedging strategy of the claim using an exponential utility. Monoyios (2007) developed analytic expansions for the indifference price and hedging strategies using this utility framework. A key expansion is the incorporation of a filtering approach, where the hedger integrates observations of asset prices to leverage better estimates of the market price of the asset models.

Basak and Chabakauri (2010) developed in the context of dynamic mean variance portfolio optimization a recursive formulation for the hedger's objective function. Using the same methodology as Basak and Chabakauri (2010) in Basak and Chabakauri (2011) dynamic time consistent hedging strategies for the minimum-variance criterion are derived. The dynamic strategies are defined mathematically similar to complete market dynamic hedging strategies, which are constructed using no-arbitrage methods. The similarity lies in " 'Greeks' that quantify the sensitivities of the asset value under the unique risk-neutral measure to the pertinent stochastic variables in the economy" (Basak and Chabakauri, 2011). The difference is twofold: the addition of a parameter that accounts for the degree of market incompleteness. And a change in drift parameters, resulting in the consideration of hedging cost due to the incompleteness of the market.

Ankirchner, Pigorsch, et al. (2014) derive an approximation of the variance minimal hedging strategy based on decomposing the risk position into a hedgeable and an orthogonal non-hedgeable component. They construct an algorithm for calculating hedging error characteristics. The algorithm delivers tight upper and lower bounds on the minimal hedging variance and consequently contains an efficient built-in control for the quality of the approximation. Ankirchner, Schneider, et al. (2014) study the effect of a long-term relationship on optimal cross-hedging strategies; thereby accounting for a possible dependence of the correlation and time to maturity in the cross hedge framework. In their research the long-term relationship of asset prices is replicated modelling the spread between two assets as a Ornstein-Uhlenbeck process and by modeling the futures price as a GBM. Using the decomposition method

described by Foellmer et al. (1985) they derive a variance minimizing cross hedge strategy in an continuous martingale setting of an incomplete market.

1.2 Research Objective

The aim of the thesis is to solve the dynamic hedging problem by implementing a minimum-variance criterion and derive tractable optimal hedging strategies in a incomplete market. We consider a hedger that wishes to reduce the risk of a non-tradable or illiquid asset at some future date using a correlated tradable asset. Our research has both similarities and extensions compared to previous research. We adopt the dynamic hedging strategy derived in Basak and Chabakauri (2011) and extend their time-consistent strategy framework to a broader range of settings. Our research objectives can best be summarized as follows:

- Analyze the impact of different parameter settings on the performance of the Geometric Brownian Motion model.
- Propose an numerical approach to extend the time-consistent hedging strategy to Asian options.
- Study the effect of a long-term relationship on optimal strategies and the consequences of model misspecification.

We will evaluate the performance of the last two research objectives using Monte Carlo simulation techniques. The remainder of this thesis is structured as follows. Section 2 starts by presenting the general theoretical framework. After that, we introduce the concept of the time consistency property and the dynamic optimal hedging strategy. Section 4 addresses an implementation of hedging strategies on options. In Section 5 we present two models on which we apply our dynamic optimal hedging strategy. Section 6 outlines the simulation results for our option hedging strategies and analysis of model risk. Finally, Section 7 summarizes and concludes.

2 Theoretical framework

2.1 Stochastic processes

We consider a continuous-time incomplete market Markovian economy with finite horizon $[0, T]$. The price processes of the two risk assets are defined over the common probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Define the state space Ω as the set of all possible realization of the underlying processes between $[0, T]$. $\mathcal{F}$ represents a sigma field of distinguishable events at time T , and $\mathcal{P}$ is the probability measure that is defined on $\mathcal{F}$. Define $\mathbb{F} = \{\mathcal{F}_t : t \in [0, T]\}$ as the augmented filtration generated by the relevant underlying processes and assume $\mathcal{F} = \mathcal{F}_T$. Denote by $(\widehat{W}_1, \widehat{W}_2) = \{\widehat{W}_1(t), \widehat{W}_2(t) : t \in [0, T]\}$ a two-dimensional Brownian Motion on the common probability space. Given that $\widehat{W}_1 \perp \widehat{W}_2$ we use the Cholesky decomposition to generate two correlated Wiener processes. Let Σ denote the correlation matrix for two Brownian Motions, then there exists a lower-triangular matrix L , such that the following holds

$$\Sigma = LL^T$$

where

$$L = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix}$$

The vector of correlated Brownian Motions W is given by

$$W = L\widehat{W}$$

The two Brownian motions (W_1, W_2) have instantaneous correlation denoted by $\rho \in [-1, 1]$.

We will consider an hedger who is committed to hold the derivative of a non-tradable asset with payoff $h(X_T)$ at time T . The function h is determined by the derivative, e.g. for a call option with strike K we have $h(X_T) = \max(X_T - K, 0)$. The price of the non-tradable asset follows the dynamics.

$$\frac{dX_t}{X_t} = m(X_t, t)dt + \nu(X_t, t)dW_2(t) \quad (2.1)$$

The risk of holding the derivative of the non-tradable asset can be hedged by continuous trading in a risk-free asset earning at a constant rate of $r > 0$ and a tradable risky security. The dynamics for the price of the tradable security are given by

$$\frac{dS_t}{S_t} = \mu(S_t, t)dt + \sigma(S_t, t)dW_1(t) \quad (2.2)$$

We will denote $\mu_t, \sigma_t, m_t, \nu_t$ as shorthand for the coefficients in (2.1) and (2.2), and $W_{1,t}, W_{2,t}$ for the Brownian motions.

2.2 Incomplete market

In a complete financial market we can perfectly replicate the payoff of a derivative via continuous trading. Discreteness of stock price observations, jumps in the stock price, stochastic volatility, transaction cost and market liquidity render perfect replication impracticable.

We focus on liquidity as the source of incompleteness. Consider an investor that seeks to hedge a derivative where the underlying security is either non-tradable or not sufficiently liquid. Consequently, the investor uses another asset that is tradable, or highly liquid, and ideally highly correlated to the underlying asset of the derivative. We call this *cross hedging*. Let's assume we have a risk position equal to $N_x X_T$ and a hedge position $N_s S_T$ where N_x and N_s are the units of X_t and S_t respectively. The basis is defined as

$$N_x X_T - N_s S_T \quad (2.3)$$

for a risk position maturing at time T . *Basis risk* is often defined as the variance of the basis. The unconditional variance of the basis is equal to

$$\text{Var}[N_x X_T - N_s S_T] = N_x^2 \text{Var}[X_T] + N_s^2 \text{Var}[S_T] - 2N_x N_s \text{Cov}(X_T, S_T)$$

The basis risk is zero when the variances of X_t and S_t are equal. This only holds if the assets are perfectly correlated:

$$\rho = \frac{\text{Cov}(X_T, S_T)}{\sqrt{\text{Var}[X_T]}\sqrt{\text{Var}[S_T]}} = \frac{\text{Cov}(X_T, X_T)}{\text{Var}[X_T]} = \pm 1$$

However, in practice assets are rarely perfectly correlated and this imperfection creates basis risk. Under basis risk, hedging can only reduce the risk of a position to a limited extent. For the assets defined in Section 2.1 we observe that for $\rho \in (-1, 1)$ the randomness in S_t , indicated by $W_{1,t}$, and the randomness in X_t , indicated by $W_{2,t}$, are imperfectly correlated and thus entail basis risk in hedging.

A simple example of basis risk is given by Ankirchner et al., 2012. Consider the German gas forward market, which is not very liquid. In order for an energy company to hedge its commitment to sell gas in the German spot market, they must use forwards of other countries to hedge their position. The degree of correlation between different regions depends on the local demand, local production and local availability with respect to pipelines etc. The energy company can hedge its risk position by buying Dutch gas on the forward market, however because of the geographical spread in commodity prices the Dutch forward price will not exactly converge to the German spot price at maturity. This gives rise to basis risk.

2.3 Variance-minimize criterion

The agents wealth at time t is given by $M_t = N_x X_t + \theta_t$, where $\theta_t = \pi_t S_t$ denotes the amount of money invested in the tradable security S_t . Hence we have an amount of

money $(M_t - \theta_t)$ that is invested in the riskless bond. The dynamics of the bond are given by

$$dB_t = rB_t dt \quad (2.4)$$

The agents wealth, given initial wealth M_0 , follows the process given by

$$dM_t = [rM_t + \theta_t(\mu_t - r)]dt + \theta_t\sigma_t dW_{1,t} \quad (2.5)$$

The goal of the agent is to minimize the variance of the *hedging error* $h(X_T) - M_T$. Denote the conditional variance at time t by $\text{Var}_t[\cdot] = \text{Var}_t[\cdot \mid X_t = x, S_t = s] = \text{Var}_t[\cdot \mid \mathcal{F}_t]$. The variance-minimizing criterion can be defined as

$$\min_{\theta_t} = \text{Var}_t[h(X_T) - M_T] \quad (2.6)$$

The criterion above is subjected to the dynamic budget constraint of (2.5). The purpose of minimizing the variance of the hedging error is to reduce basis risk. Note that if the objective of the agent would be to minimize potential losses, the agent could adopt for example a VaR (Value at Risk) minimizing criterion. The mean-variance criterion would be suitable for agents looking to maximize the trade off between expected return and risk.

The goal of minimizing the variance replication error is adopted in a large part of the risk management literature (Hull, 2009). Instinctively, the motive behind a minimum-variance hedge is to construct a hedge similar to the delta neutral hedge in the complete market setting as a means to eliminate all risk. However, in incomplete markets not all risk can be eliminated and hence the objective is to eliminate as much as risk as possible by replicating the risk position as accurately as possible. The motive for the minimum variance objective originates from the nature of the hedger. Take market-makers for example, they provide liquidity by allowing buyers and sellers to trade in the market at any given time. At the same time they act as a hedger by trading in the underlying or cross correlated asset on which they write derivatives. Their hedging purpose is largely based on minimizing the risk that originates from writing of derivatives, and is therefore not mean return driven. The variance minimization criterion is used as well in other contexts in economics. Monetary economics is a prime example, in this environment the primary goal of a intuition such as the government is to achieve stability and decrease uncertainty in the economy.

2.4 Stationary Basis

In this section we will introduce *cointegrated* stochastic processes. Cointegration is referred to as the existence of a linear relationship between two stochastic processes

that has constant mean and standard deviation. Ankirchner et al. (2012) emphasize the sizeable dependence of correlation on the selected time measured interval. They give an example of weekly, monthly and yearly log returns of kerosene and crude oil, and show that short-term correlation is considerably lower than long-term correlation. This could be evidence of the existence of a long-term linear relationship with short-term deviations. In non mathematical terms, if this linear relationship is *stationary* then the two assets are cointegrated. If correlation is a function of time to maturity then we must account for the time dependence of the correlation parameter in our cross hedge. "Intuitively, for long-term hedges it is likely that the two assets are in their equilibrium relationship, whereas in the short-term the dynamics are dominated by noisy fluctuations due to shortage or oversupply of kerosene" (Ankirchner et al., 2012).

2.4.1 Stationarity

Stationarity means that the statistical properties of the process are roughly constant over longer periods; non-stationary thus implies that movements are not necessarily within a certain range, and that the process has no limits with respect to where it goes.

Definition 2.4.1. *Let X_t be a stochastic process. Then X_t is a weakly stationary process if*

$$\begin{aligned}\mu_X(t) &= \mu_X < \infty \quad \text{constant}, \\ \sigma_X(t) &= \sigma_X < \infty \quad \text{constant}, \\ \text{Cov}(X_t, X_s) &= \text{Cov}(X_{t+n}, X_{s+n}),\end{aligned}$$

for all integers t, s and n .

In order to test if the process X_t is weakly stationary, a unit root statistical test should be employed. Some of the most basic unit root tests are the Augmented Dickey-Fuller (ADF) test, the Phillips-Ouliaris test and the KPSS test. Consider the following discrete time AR(1) process

$$X_t = \alpha + \beta X_{t-1} + \epsilon_t \tag{2.7}$$

and take the first difference to obtain

$$\Delta X_t = \alpha + \phi X_{t-1} + \epsilon_t \tag{2.8}$$

where $\Delta X_t = X_t - X_{t-1}$ and $\phi = (\beta - 1)$. The Dickey-Fuller test has the following hypothesis

$$H_0 : \phi = 0 \quad H_1 : \phi < 0 \tag{2.9}$$

if we reject H_0 then the process X_t does not have a unit root (i.e. the process is stationary). Consider the discrete time AR(p) process

$$\Delta X_t = \alpha + \phi X_{t-1} + \gamma_1 \Delta X_{t-1} + \cdots + \gamma_p \Delta X_{t-p} + \epsilon_t \quad (2.10)$$

then the ADF-test tests whether ϕ is different from zero. The ADF-test statistic τ_{ADF} is defined as

$$\tau_{ADF} = \frac{\hat{\phi}}{s.e.(\hat{\phi})}$$

where $\hat{\phi}$ is the estimated coefficient of ϕ in (2.14). Note that this test statistic has its own Dickey-Fuller distribution which causes the hypothesis test to be asymmetrical. The number of lags p for which we test in the ADF-test is generally determined empirically.

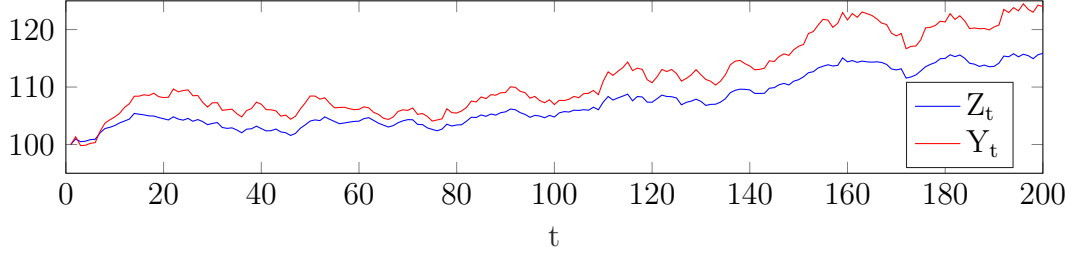
2.4.2 Cointegration

Consider two stochastic processes S_t and I_t . In order to test if these processes are cointegrated we use the Engle and Granger (1987) 2-step approach

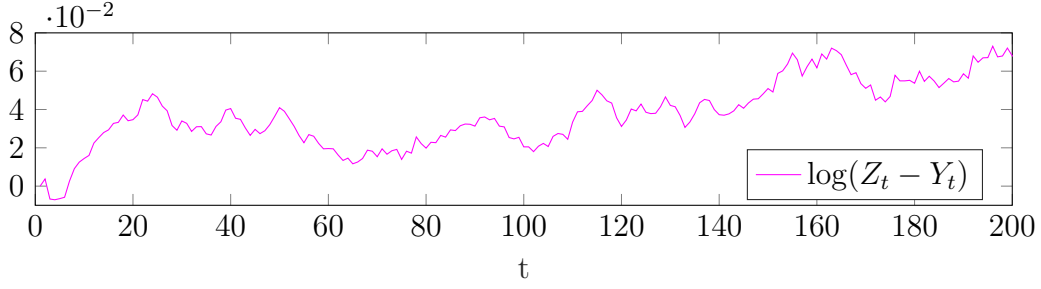
- Use a statistical unit root test to test if S_t and I_t are non-stationary.
- Run the regression equation $S_t = \alpha + \beta I_t + u_t$. Use a unit root test to determine if the estimated residuals $\hat{u}_t$ are stationary. Here β is called the cointegration coefficient.

The correlation between selected assets is a critical factor of the performance of a cross hedge. Correlation is a measure for how two processes move jointly over time. Cointegration is a measure of the similarity in behaviour of two processes in the long term. Cointegration is referred to a more stable and robust relationship than correlation. Next, we show with a basic illustration the difference between correlation and cointegration. Consider two different cases:

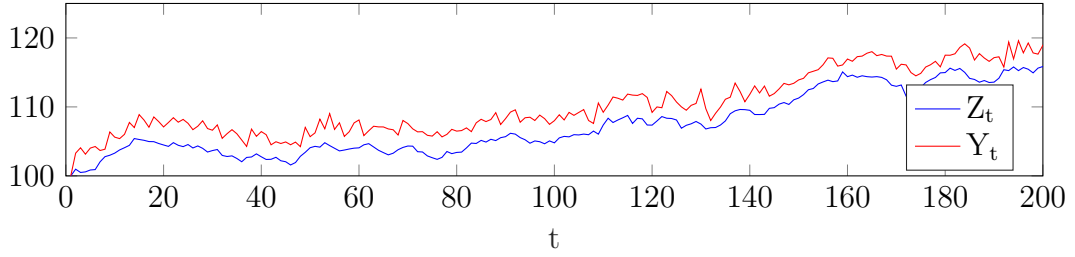
- Case 1: process Z_t follows a geometric Brownian Motion (GBM) and process $Y_t = c + Z_t + \epsilon_t$ where c is a constant and $\epsilon \sim N(0, 0.5)$ is white noise. The parameters of the GBM are $\mu = 0.04, \sigma = 0.2$ and $Z_0 = 100$.
- Case 2: Z_t and Y_t both follow a geometric Brownian Motion with correlation matrix $\Sigma = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$, drift and volatility are $\mu_1 = 0.01, \mu_2 = 0.03, \sigma_1 = 0.05, \sigma_2 = 0.03$ and $Z_0 = Y_0 = 100$.



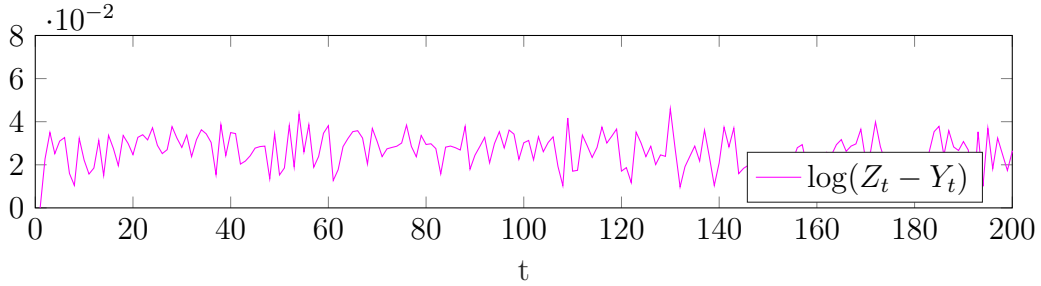
(a) Simulated variables Case 1.



(b) Log spread Case 1.



(c) Simulated variables Case 2.



(d) Log spread Case 2.

Figure 1: Simulated processes.

The simulated processes as well as the log spread are shown in Figure 1. For both cases the assets move together across the corresponding horizon. We observe that for Case 1 the spread of the processes remains stable around a constant mean. We can proceed with a statistical simple statistic analysis. Table 1 shows the correlation results as well as the p-value of augmented Dickey–Fuller test for one lagged period. As we can see, both cases exhibit strong correlation, however only for Case 1 we can reject the null hypothesis of non stationarity for a significance level of 1%. This shows that two variables can be highly correlated, but therefore do not have to be cointegrated.

	Case 1	Case 2
Pearson Correlation	0.9820	0.9911
ADF test (p-value)	0.0062	0.7730

Table 1: Statistics of Case 1 and Case 2

Cointegration is closely related to the pair trading strategy. This strategy groups assets that are closely linked together in pairs. The pairs are traded when their asset prices differ more than a pre-specified amount. The pair trading strategy assumes a certain price relationship between two assets. Any short-term deviation in the market from this relationship allows the investor to take advantage by taking a long position in the undervalued asset and short position in the overvalued asset at the same time. Whenever the price relationship is restored, the investors closes both positions and has realized a profit. In terms of processes S_t and I_t , the strategy yields a profit of

$$\begin{aligned} & (S_t - S_{t-1}) - \beta(I_t - I_{t-1}) \\ & (S_t - \beta I_t) - (S_{t-1} - \beta I_{t-1}) = X_t - X_{t-1} \end{aligned}$$

The profit of the pair trading strategy over period Δt is given by the change process X_t . Since the process X_t is weakly stationary, the investor can use this to construct a strategy of long and short positions.

2.4.3 Stationary Basis

In this section we denote the non-tradable asset by I_t . Define the spread of the log asset prices of I_t and S_t by

$$X_t = \log(S_t) - \log(I_t) \quad (2.11)$$

If I_t and S_t are cointegrated processes then the log spread is stationary. The spread X_t can be thought of as the basis similar to (2.3). This spread is a source of basis risk since the underlying assets are not perfectly correlated. We assume that the log spread follows a mean-reverting Ornstein-Uhlenbeck process

$$dX_t = \kappa(\bar{X} - X_t)dt + \sigma_{X,t}(\rho d\widehat{W}_{1,t} + \sqrt{1 - \rho^2} d\widehat{W}_{2,t}) \quad (2.12)$$

where $\sigma_{X,t} \geq 0$ by assumption. For $\kappa \rightarrow 0$ the log spread X_t converges to a Arithmetic Brownian Motion with drift zero, that is correlated with the tradable asset. Consequently, the variance of the log spread increases for an increasing time to maturity. Note, here ρ quantifies the correlation between the tradable asset and the log spread, instead of the correlation between the tradable and non-tradable asset.

We define $\hat{\rho} = \sqrt{1 - \rho^2}$. From the dynamics of X_t and S_t we derive the dynamics of the non-tradable asset I_t as

$$I_t = S_t e^{-X_t} \quad (2.13)$$

For a positive log spread X_t the asset price of I_t is expected to increase relative to the asset price of S_t . Conversely for a negative log spread the price of I_t decreases

with respect to S_t . Applying Ito's formula to (2.13) shows that the dynamics of I_t satisfy

$$\begin{aligned} dI_t &= \frac{\partial I_t}{\partial t}dt + \frac{\partial I_t}{\partial S_t}dS_t + \frac{\partial I_t}{\partial X_t}dp_t + \frac{1}{2}\frac{\partial^2 I_t}{\partial S_t^2}(dS_t)^2 + \frac{1}{2}\frac{\partial^2 I_t}{\partial X_t^2}(dX_t)^2 + \frac{\partial^2 I_t}{\partial S_t \partial p_t}(dS_t dX_t) \\ &= \mu_t I_t dt + \sigma_t I_t dW_{1t} - \kappa(\bar{X} - X_t)I_t dt - \sigma_{X,t}I_t(\rho d\widehat{W}_{1,t} + \hat{\rho}d\widehat{W}_{2,t}) + \frac{1}{2}I_t\sigma_{X,t}^2 dt - \rho X_t \sigma_{X,t} \sigma_t dt \\ &= \left(\frac{1}{2}\sigma_{X,t}^2 - \kappa(\bar{X} - X_t) + \mu_t - \rho\sigma_{X,t}\sigma_t\right)I_t dt + I_t((\sigma_t - \sigma_{X,t}\rho)d\widehat{W}_{1,t} - \sigma_{X,t}\hat{\rho}d\widehat{W}_{2,t}) \end{aligned}$$

Define the volatility term and the Brownian Motion driving the dynamics of I_t by

$$\begin{aligned} \sigma_{I,t} &= \sqrt{\sigma_t^2 - 2\rho\sigma_t\sigma_{X,t} + \sigma_{X,t}^2} \\ W_{I,t} &= ((\sigma_t - \sigma_{X,t}\rho)\widehat{W}_{1,t} - \sigma_{X,t}\hat{\rho}\widehat{W}_{2,t})/\sigma_{I,t} \end{aligned}$$

Under the asset dynamics above, the correlation between the the assets I_t and S_t is given by

$$\rho_{IX} = \frac{\sigma_t - \sigma_{X,t}\rho}{\sqrt{\sigma_t^2 - 2\rho\sigma_t\sigma_{X,t} + \sigma_{X,t}^2}}$$

The correlation of the two assets is now a function of σ_t, ρ and $\sigma_{X,t}$.

2.5 Trading cost

Most theoretical models assume the market setting to be friction-free and shares of an asset to be infinitely divisible. In reality, trading is often only possible in discrete increments. Moreover, an investors incurs transaction cost when trading on the market. Trading cost consist mainly of: quoted bid-ask spread, commission fees and market impact cost. Other cost could include delay cost, opportunity cost and short borrowing cost. Bid-ask spreads costs are independent of the investor size; commission fees are partly dependent on the volume of trading and investor size; market impact cost are dependent on investor size and liquidity.

Realizing a realistic model for incurred trading cost is not easy and beyond the scope of our research. To simplify the treatment of trading cost, we assume that the trading cost are proportional to the volume of trading. This assumption is mainly valid for large investors; we only take into account bid-ask spread cost; commission fees and market impact cost are assumed to be negligible. Define the trading cost for a discrete time horizon $[0, T]$ by

$$C_T = \sum_{t=0}^T \varrho(\pi_{t+1} - \pi_t)S_t \quad (2.14)$$

where $\pi_0 = 0$ and ϱ are the trading cost of trading a long or short position of one unit of S_t .

3 Hedging strategies

There exist situations where an investor who maximizes a reward function with initial values (t, X_t) finds that the derived strategy is no longer optimal at a later moment in time (s, X_s) for $s > t$. Subsequently, time inconsistency arises. The investor can employ one of three hedging strategies to respond to time inconsistency: naive, pre-commit or equilibrium strategy. We will discuss all three strategies in the next sections.

3.1 Time consistency

The dynamic programming approach applied to our variance-minimize criterion is hard, because in a multi-period setting the conditional variance gives rise to time-inconsistency. To see how this works we look at the law of total variance

$$\text{Var}_t[X_T - M_T] = E_t[\text{Var}_{t+\varepsilon}(X_T - M_T)] + \text{Var}_t(E_{t+\varepsilon}[X_T - M_T]), \quad \varepsilon > 0 \quad (3.1)$$

As we can see, the variance at time t exceeds the expected value of the variance at time $t + \varepsilon$. As a consequence, the hedging strategy θ_t , that is constructed to minimize $\text{Var}_t[X_T - M_T]$ at time t , actually minimizes the future $(t + \varepsilon)$ -expected variance plus the variance of the future $(t + \varepsilon)$ -expectation. The time inconsistency arises as we arrive at time $t + \varepsilon$; the second term in (3.1) vanishes and the hedger is only interested in minimizing the $(t + \varepsilon)$ -variance. However, the hedging strategy chosen at time t minimizes the expected variance *and* the variance of the expectation. Hence from a time $t + \varepsilon$ perspective the hedger may wish to deviate from optimal policy at time t . At time $t + \varepsilon$ the hedger has two options: continue to adopt the original hedging strategy or to devise a new plan that is 'optimal' for him/her at time $t + \varepsilon$. If the latter would be enforced over the complete hedge horizon $[0, T]$ then the hedging strategy is conceived as time consistent.

Basak and Chabakauri (2010) were the first to solve this problem of time inconsistency in the derivation of an optimal dynamic solution. Provided in their research is a novel approach of obtaining a 'consistent' solution to a portfolio selection problem for a mean-variance objective function. They use a game theoretic formulation to define the 'optimal' strategy as an equilibrium solution. Later, Bjork and Murgoci (2014) derived a more thorough solution for the 'consistent' solution for general time-inconsistent problems using a game theory approach. In the next paragraph we illustrate how the problem can be defined in a game theory framework to give a intuitive idea of how the hedging problem of time inconsistency is solved.

The time inconsistency problem can be viewed as a non-cooperative sequential game, with one player for each time t and the sub-game perfect Nash equilibrium as the 'time consistent' solution. A sequential form game is a tuple $\langle \mathcal{I}, (\theta_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$. The hedging problem can be described as a sequential game in discrete time as follows:

1. Discretize the time interval $[t, T]$ into $N \in \mathbb{N}$ equidistant time steps and set $t_i = \frac{i}{N}(T - t)$ for $1 \leq i \leq N$. Consider one player at each time point t_i and call this 'player t_i '. Player t_i can be seen as a future reincarnation of the hedger. Define the set of players as $\mathcal{I} = \{1, \dots, N\}$.
2. Conditional on process X_{t_i} and S_{t_i} player t_i will choose an action. The set of actions for player t_i is denoted by $\vartheta_i = \vartheta(t_i, S_{t_i}, X_{t_i})$. Define $\Theta = \{\theta \mid \theta = (\theta_i)_{i \in \mathcal{I}}, \theta_i \in \vartheta_i, \forall i \in \mathcal{I}\}$ as the set of actions of all players. We can behold Θ as the set of available strategies over $[0, T]$ and ϑ_i as the set of strategies over $[t, t + \frac{i}{N})$.
3. The payoff function of player t_i is $u_i(S_{t_i}, X_{t_i}, \Theta) = \text{Var}_{t_i}[h(X_T) - M_T]$. The payoff functions maps $\Theta \mapsto \mathbb{R}$. Note that the payoff of player t_i does not only depend on the strategy chosen by himself, but also by the strategies chosen by all players coming after him in time.

The sub-game perfect Nash equilibrium of the game above is given by the set of payoff functions $\hat{\Theta}$ that comply with the following reasoning:

1. Pick an arbitrary point t_i in time.
2. Suppose that every player t_k for $k > i$ will play strategy $\hat{\theta}_k^*(t_k, S_{t_k}, X_{t_k})$.
3. The optimal strategy for player t_i is to play strategy $\hat{\theta}_i^*(t_i, S_{t_i}, X_{t_i})$.

Thus, the strategy $\hat{\theta}_t^*(t, S_t, X_t)$ that takes into account that the hedger may wish to deviate from the time- t optimal strategy at time $t + \epsilon$ is a sub-game perfect Nash equilibrium of the game presented above.

3.2 Dynamically time-consistent hedging strategy

Our next step is to apply the line of thought from above to our economic setup as defined in Section 2. The goal is to find a recursive relationship for the value function

$$J_t(S_t, X_t, M_t) = \min_{\theta_t} \text{Var}_t[h(X_T) - M_T] \quad (3.2)$$

By applying the law of total variance as in (3.1) we arrive at

$$J_t(S_t, X_t, M_t) = \min_{\theta_t} E_t[J_{t+\epsilon}] + \text{Var}_t(E_{t+\epsilon}[h(X_T) - M_T]) \quad (3.3)$$

Given the optimal time-consistent hedging strategy θ_s^* with complementary terminal wealth M_T^* for $s \in [t, T]$, the value function of (3.2) becomes

$$J_t(S_t, X_t, M_t) = \text{Var}_t[h(X_T) - M_T^*] \quad (3.4)$$

The hedger at time t plans to follow the optimal strategy θ_s^* for $s \in [t + \epsilon, T]$. Hence the time- t problem of the hedger boils down to finding the optimal hedging strategy θ_s for $s \in [t, t + \epsilon)$ that minimizes the right hand side of (3.3). In the next proposition we state the optimal hedging strategy θ_t^* as derived by Basak and Chabakauri (2011).

Proposition 1. *The optimal hedging policy and the corresponding variance of the hedging error are given by*

$$\theta_t^* = \frac{\rho \nu_t}{\sigma_t} X_t \frac{\partial E_t^*[h(X_T)e^{-r(T-t)}]}{\partial X_t} + S_t \frac{\partial E_t^*[h(X_T)e^{-r(T-t)}]}{\partial S_t} \quad (3.5)$$

$$\text{Var}_t[X_T - M_T^*] = (1 - \rho^2) E_t \left[\int_t^T \nu_s^2 X_s^2 \left(\frac{\partial E_s^*[h(X_T)]}{\partial X_s} \right)^2 ds \right] \quad (3.6)$$

Here, M_T^* represent the wealth level under the optimal hedging strategy. Furthermore, the conditional expectation E_t^* is under the hedge neutral probability measure $\mathcal{P}^*$.

Proof. The wealth process is given by

$$dM_t = [rM_t + \theta_t(\mu_t - r)]dt + \theta_t \sigma_t dW_{1,t} \quad (3.7)$$

Here the wealth process M_T^* is given by

$$M_T^* = M_t e^{r(T-t)} + \int_t^T \theta_s^* (\mu_s - r) e^{r(T-s)} ds + \int_t^T \theta_s^* \sigma_s e^{r(T-s)} dW_{1,s} \quad (3.8)$$

Fix a strategy θ_t over the interval $[t, t + \varepsilon)$ and let ε converge to zero. The law of total variance yields

$$J_t = \min_{\theta_t} E_t[J_{t+\varepsilon}] + \text{Var}_t(E_{t+\varepsilon}[h(X_T) - M_T]) \quad (3.9)$$

Here M_T is the final wealth of using strategy θ_t over the time interval $[t, t + \varepsilon)$ and θ^* over $[t + \varepsilon, T]$. Now define the process G_t as

$$G_t = G_t(X_t, S_t, t) = E_t[h(X_T)e^{-r(T-t)} - \int_t^T \theta_s^* (\mu_s - r) e^{r(T-s)} ds] \quad (3.10)$$

Hence now we can write the following

$$\text{Var}_t(E_{t+\varepsilon}[h(X_T) - M_T]) = \text{Var}_t(G_{t+\varepsilon} - M_{t+\varepsilon}) \quad (3.11)$$

For the process G , Ito's formula gives us

$$dG_t = \dots dt + e^{r(T-t)} \left(\frac{\partial G_t}{\partial X_t} \nu_t X_t dW_{2,t} + \frac{\partial G_t}{\partial S_t} S_t \sigma_t dW_{1,t} \right)$$

Hence we obtain with $\Delta W_{it} = W_{i,(t+\varepsilon)} - W_{i,t}$ for $i = 1, 2$

$$\text{Var}_t(E_{t+\varepsilon}[h(X_T) - M_T]) = \text{Var}_t \left(\frac{\partial G_t}{\partial X_t} \nu_t X_t \Delta M_{2,t} + \frac{\partial G_t}{\partial S_t} S_t \sigma_t \Delta M_{1,t} - \theta_t \sigma_t \Delta W_{1,t} \right) e^{2r(T-t)}$$

Now we can write (3.9) as follows

$$0 = \mathcal{D}J_t + \min_{\theta_t} \left[\left(\frac{\partial G_t}{\partial X_t} \nu_t X_t \right)^2 + \left(\frac{\partial G_t}{\partial S_t} S_t \sigma_t \right)^2 + (\theta_t \sigma_t)^2 + 2\rho \frac{\partial G_t}{\partial S_t} S_t \sigma_t \frac{\partial G_t}{\partial X_t} \nu_t X_t \right. \\ \left. - 2 \frac{\partial G_t}{\partial S_t} S_t \theta_t \sigma_t^2 - 2\rho \frac{\partial G_t}{\partial X_t} \nu_t X_t \theta_t \sigma_t \right] e^{2r(T-t)} \quad (3.12)$$

The minimization gives us

$$\theta_t^* = \frac{\rho \nu_t}{\sigma_t} X_t \frac{\partial G_t}{\partial X_t} + S_t \frac{\partial G_t}{\partial S_t} \quad (3.13)$$

Substituting this back into (3.12) we obtain the following

$$\mathcal{D}J_t + (1 - \rho^2) \left(\nu_t X_t \frac{\partial G_t}{\partial X_t} \right)^2 e^{2r(T-t)} = 0 \quad (3.14)$$

With the terminal expression $J_T = 0$ we obtain an expression for the process J_t

$$J_t = E_t \left[J_T - \int_t^T \mathcal{D}J_s ds \right] = (1 - \rho^2) E_t \left[\int_t^T \left(\nu_s X_s \frac{\partial G_s}{\partial X_s} \right)^2 e^{2r(T-s)} ds \right] \quad (3.15)$$

Next, we show that $G_t = E_t^*[h(X_T)e^{-r(T-t)}]$. We apply the Feynman-Kac theorem to the definition of G_t in (3.10) and substitute the optimal strategy into the partial differential equation. This gives us:

$$\frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} X_t (m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t}) + \frac{\partial G_t}{\partial S_t} r S_t + \frac{1}{2} \left(\frac{\partial^2 G_t}{\partial^2 X_t} + 2\rho \nu_t \sigma_t X_t S_t \frac{\partial^2 G_t}{\partial X_t \partial S_t} + \sigma_t^2 S_t^2 \frac{\partial^2 G_t}{\partial^2 S_t} \right) - r G_t = 0 \quad (3.16)$$

where the terminal condition is given by $G_T = h(X_T)$. By applying the reversed Feynman-Kac solution to equation (3.16) we obtain

$$G_t = E_t^*[h(X_T)e^{-r(T-t)}] \quad (3.17)$$

where the expectation is taken under the probability measure $\mathcal{P}^*$. The probability measure $\mathcal{P}^*$ is defined as the measure on which assets X_t and S_t have drift $(m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t})$ and r respectively. In order to better understand this change of measure look at the drift terms of G_t under $\mathcal{P}$ and $\mathcal{P}^*$.

$$\mathcal{D}G_t = \frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} m_t X_t + \frac{\partial G_t}{\partial S_t} \mu_t S_t + \frac{1}{2} \frac{\partial^2 G_t}{\partial X_t^2} \nu_t^2 X_t^2 + \frac{1}{2} \frac{\partial^2 G_t}{\partial S_t^2} \sigma_t^2 S_t^2 + \rho \frac{\partial^2 G_t}{\partial X_t \partial S_t} \nu_t \sigma_t X_t S_t$$

$$\mathcal{D}G_t^* = \frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} X_t (m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t}) + \frac{\partial G_t}{\partial S_t} r S_t + \frac{1}{2} \frac{\partial^2 G_t}{\partial X_t^2} \nu_t^2 X_t^2 + \frac{1}{2} \frac{\partial^2 G_t}{\partial S_t^2} \sigma_t^2 S_t^2 + \rho \frac{\partial^2 G_t}{\partial X_t \partial S_t} \nu_t \sigma_t X_t S_t$$

Clearly, both drift term are equal if process G_t under the measure $\mathcal{P}^*$ is defined with asset drifts $(m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t})$ and r . Under these alternative drift terms for G_t , the optimal strategy defined in (3.13) solves equation (3.12) and is therefore the solution to the value function in (3.2). Since our goal is to find θ_t such that (3.2) is minimized, the proof is complete. $\square$

An important step in the proof above is the labeling of the hedge-neutral measure $\mathcal{P}^*$. On this measure we define the Brownian motions $W_{1,t}^*$ and $W_{2,t}^*$ with correlation ρ

$$\begin{aligned} dW_{1,t}^* &= dW_{1,t} + \frac{\mu_t - r}{\sigma_t} dt \\ dW_{2,t}^* &= dW_{2,t} + \rho \frac{\mu_t - r}{\sigma_t} dt \end{aligned}$$

The asset process dynamics under $\mathcal{P}^*$ are given by

$$\begin{aligned} dS_t &= rS_t dt + \sigma_t S_t dW_{1,t}^* \\ dX_t &= (m_t - \rho\nu \frac{\mu_t - r}{\sigma_t}) X_t dt + \nu_t S_t dW_{2,t}^* \end{aligned} \tag{3.18}$$

Note that the measure $\mathcal{P}^*$ is a risk-neutral probability measure, since the drift of S_t equals rS_t . Nonetheless, the hedge-neutral measure is not unique; to see this we decompose $dW_{2,t} = \rho d\widehat{W}_{1,t} + \sqrt{1-\rho^2} d\widehat{W}_{2,t}$, where $\widehat{W}_{1,t}$ and $\widehat{W}_{2,t}$ are independent under the real-world measure $\mathcal{P}$. Consequently, any probability measure for which the Brownian motions are defined as $W_{1,t}^* = dW_{1,t} + \frac{\mu_t - r}{\sigma_t} dt$ and $d\widehat{W}_{2,t}^* = d\widehat{W}_{2,t} + q_t dt$ is a risk neutral measure, independently of the process q_t . The hedging strategy in Proposition 1 is compromised of the sensitivity of the time- t risk position under the unique risk-neutral measure to the stochastic processes X_t and S_t . Additionally, the ρ that occurs in the hedging strategy accounts for the incomplete market environment. In Corollary 3.2.1 we derive the optimal dynamic strategy for a function $h(X_T, S_T)$.

Corollary 3.2.1. *The optimal dynamic hedging strategy and corresponding hedging error variance for a risky position $h(X_t, S_t)$ with asset dynamics (5.6) and (5.7) are given by:*

$$\begin{aligned} \theta_t^* &= \frac{\rho\nu_t}{\sigma_t} \frac{\partial E_t^*[h(X_T, S_T)e^{-r(T-t)}]}{\partial X_t} + S_t \frac{\partial E_t^*[h(X_T, S_T)e^{-r(T-t)}]}{\partial S_t} \\ \text{Var}_t[h(X_T, S_T)] - M_T^* &= (1 - \rho^2) E_t \left[\int_t^T \left(\nu_s^2 X_s^2 \frac{\partial E_s^*[h(X_T, S_T)]}{\partial X_s} \right)^2 ds \right] \end{aligned}$$

Proof. The proof is given in Appendix A.1 □

3.2.1 Hedging error variance

The quality of the hedging strategy can be measured by the variance of the hedging error. From Proposition 1 we observe that the hedging error variance decreases quadratically in ρ . The parameter ρ can be interpreted as the measure of market completeness. For the complete market case ($\rho = \pm 1$) we can construct the perfect hedge and the hedging error variance vanishes. For zero correlation ($\rho = 0$) the first term of (3.5) in Proposition 1 disappears; the hedging strategy only depends on the joint distribution of $h(X_t)$ and S_t . Intuitively, hedging for the zero correlation case makes little sense.

3.2.2 Wealth dependence

Note that the hedging strategy θ_t^* , the value function J_t and function G_t are independent of the wealth M_t . θ_t^* , J_t and G_t are only affected by the asset processes X_t and S_t and time t . This observation implicates that the optimal hedging strategy is identical for two hedger with dissimilar wealth levels. This "policy is economically less reasonable because the rich and the poor should not have the same investment strategy" (Bjork and Murgoci, 2014). Bjork and Murgoci (2014) introduced a dynamic mean-variance criterion with wealth dependent mean-variance preference. Later, Dai et al. (2021) extended this framework to conform with the investment wisdom that short-sale is never optimal for a stock investment with a positive stock risk premium.

3.3 Pre-commitment strategy

A hedger that follows the dynamic optimal hedge is concerned with minimizing the time- t variance of the hedging error. The dynamic hedging policy provides a time-consistent solution. The pre-commitment hedging policy disregards the time-consistency of the variance-minimize criterion. The pre-commitment strategy minimizes the hedging error variance at the initial date zero. Conditional of being in time- t , the goal of a pre-committed hedger is to maintain a low time-zero hedging error variance, instead of a time- t hedging error variance.

Duffie and Richardson (1991) derive the optimal pre-commitment mean-variance hedging strategy in continuous time. The asset prices follow a Geometric Brownian Motion with correlated Brownian Motions. The hedger has a future commitment in a risky asset and can trade futures of another correlated asset to minimize the risk. The optimal pre-commitment hedge derived by Duffie and Richardson (1991) holds for the case $r = 0$ and $W_0 = 0$. To obtain the pre-commitment hedge for the more general setting $r > 0$ and $M_0 > 0$ we derive a modified expression of our wealth process M_t .

$$d\tilde{M}_t = \theta_t \tilde{\mu} dt + \theta_t \tilde{\sigma} dW_{1,t}$$

Where $\tilde{M}_t = M_t e^{r(T-t)} - M_0 e^{rT}$, $\tilde{\mu}_t = (\mu - r)e^{r(T-t)}$, $\tilde{\sigma} = \sigma_t e^{r(T-t)}$. For time constant coefficients μ, σ, m, ν the objective function is given by

$$\min_{\theta_t} \text{Var}_0[X_T - \tilde{M}_T]$$

Proposition 2. *The pre-commitment hedging strategy and hedging error variance for two GBM processes is given by*

$$\begin{aligned} \theta_t^{\text{commit}} = & \frac{\rho\nu}{\sigma} X_t e^{(m-r-\frac{\rho\nu(\mu-r)}{\sigma})(T-t)} \\ & - \frac{\mu-r}{\sigma^2} \left((X_0 e^{(m-r-\frac{\rho\nu(\mu-r)}{\sigma})T} - M_0) e^{rt} - (X_t e^{(m-r-\frac{\rho\nu(\mu-r)}{\sigma})T} - M_t^{\text{commit}}) \right) \end{aligned} \quad (3.19)$$

$$\begin{aligned} \text{Var}_t[X_T - M_T^{\text{commit}}] &= (H_t - H_0)^2 c_t (1 - c_t) \\ &\quad + (1 - \rho)^2 \nu^2 X_t^2 e^{2(m - \frac{\rho\nu(\mu-r)}{\sigma})(T-t)} \frac{e^{(\nu^2 - 2\frac{\rho\nu(\mu-r)}{\sigma})(T-t)} - c_t}{\nu^2 + 2\rho\nu\frac{\mu-r}{\sigma} + \frac{\mu-r}{\sigma}} \end{aligned} \quad (3.20)$$

where

$$\begin{aligned} c_t &= e^{-(\frac{\mu-r}{\sigma})^2(T-t)} \\ H_t &= X_t e^{(m-r - \frac{\rho\nu(\mu-r)}{\sigma})(T-t)} - M_t^{\text{commit}} e^{-r(T-t)} \end{aligned}$$

Proof. A proof can be found in Appendix B of Duffie and Richardson (1991). $\square$

The first term of the pre-commitment hedging strategy in (3.19) is equivalent to the dynamic hedging strategy for two GBM derived by Basak and Chabakauri (2011). However, the pre-commitment strategy contains an additional second term that is not present in the dynamic strategy. To understand how this term comes into being we have take a look at the expected hedging error under the dynamic hedging strategy. To derive the expected hedging error we start by looking at the time- t conditional expectation of the wealth process at time of maturity.

$$E_t[M_T] = E_t[M_t e^{r(T-t)} + \int_t^T \theta_s^* (\mu_s - r) e^{r(T-s)} ds + \int_t^T \theta_s^* \sigma_s e^{r(T-s)} dW_{1,s}] \quad (3.21)$$

$$E_t[M_T - M_t e^{r(T-t)}] = E_t[\int_t^T \theta_s^* \sigma_s e^{r(T-s)} dW_{1,s}] \quad (3.22)$$

If we substitute the equation above into the definition of G_t defined in (3.10) and use that $G_t = E_t^*[h(X_T) e^{-r(T-t)}]$ we obtain the following relation:

$$E_t^*[h(X_T) e^{-r(T-t)}] = E_t[h(X_T) e^{-r(T-t)}] - E_t[M_T^* - M_t e^{r(T-t)}] \quad (3.23)$$

Rearranging terms allows us to derive an expression of the hedging error

$$E_t^*[h(X_T) e^{-r(T-t)} - M_t e^{r(T-t)}] = E_t[h(X_T) e^{-r(T-t)} - M_T^*] \quad (3.24)$$

and given that the conditional expectation of the GBM process X_t under the measure P^* is given by:

$$E_t^*[X_T] = X_t e^{(m - \rho\nu(\mu-r)/\sigma)(T-t)} \quad (3.25)$$

the expected hedging error is defined as

$$E_t^*[(h(X_T) - M_T^*) e^{-r(T-t)}] = X_t e^{(m - \rho\nu(\mu-r)/\sigma)(T-t)} - M_t \quad (3.26)$$

As we can see the expected hedging error defined in (3.26) appears twice in the pre-commitment hedging strategy in (3.19): once as time- t value and once as time zero value. A pre-committed hedger accounts in his/her hedging strategy for deviations of the expected time- t hedging error from the time zero expected hedging error. To understand why, we have a look at the law of total variance defined in (3.1)

$$\text{Var}_0[X_T - M_T] = E_0[\text{Var}_t(X_T - M_T)] + \text{Var}_0(E_t[X_T - M_T]), \quad t > 0$$

here it immediately becomes clear that by keeping the variation between $E_t[X_T - M_T]$ and $E_0[X_T - M_T]$ small, results in a smaller $\text{Var}_0(E_t[X_T - M_T])$ and consequently a smaller value of $\text{Var}_0[X_T - M_T]$. A pre-committed hedger will decrease his/her hedging position in S_t whenever the second term in (3.19), given here below, becomes positive.

$$\left((X_0 e^{(m-r-\frac{\rho\nu(\mu-r)}{\sigma})T} - M_0) e^{rt} - (X_t e^{(m-r-\frac{\rho\nu(\mu-r)}{\sigma})T} - M_t^{\text{commit}}) \right)$$

By doing so, the expected wealth at time to maturity M_T decreases; accordingly the expected hedging error $E_t[X_T - M_T]$ will increase and the will be closer to the time zero hedging error. Over the hedging horizon $[0, T]$ the pre-commitment will perform better for small t , nonetheless the dynamically optimal strategy will perform better after a certain time t . Basak and Chabakauri (2011) show for different parameter settings that for tradable and non-tradable GBM asset prices, the pre-commitment strategy is outperformed by the dynamic strategy after a period of time.

3.4 Myopic hedging strategy

The myopic hedging strategy, also called the naive strategy, minimizes at each time point the hedging error over the next moment:

$$\min_{\theta_t} \text{Var}_t[dX_T - dM_t]$$

subject to the budget constraint in (2.5). The myopic strategy can be viewed as a static hedge over an infinitesimally hedging horizon, recurrent over time. The myopic hedge is given by:

$$\theta_t^{\text{myopic}} = \frac{\rho\nu_t}{\sigma_t} X_t$$

The myopic strategy has a similar structure to the first term of the dynamically optimal strategy in (3.5). In contrast to the dynamic strategy the myopic hedge does not account for the impact of future investment returns on the hedging error variance. Generally, the myopic strategy doesn't yield a perfect variance minimum hedge in complete markets. Basak and Chabakauri (2010) provide an short overview of parameter settings of cases where the myopic hedge outperforms the dynamic hedging strategy.

3.5 Delta Tolerance strategy

The dynamic hedging strategies requires frequent adjustment of the asset position. A more frequently adjusted hedging strategy will possibly reduce the hedging error variance, but increase transaction cost. The hedger is faced with a trade off between reducing transaction cost and hedging error variance.

The presence of trading cost results in an incomplete market and therefore renders perfect hedging impossible. A variety of methods have been suggested to deal with this problem. The research of Leland (1985) was the first to provide a solution: the

strategy is based on hedging at fixed regular intervals by following the Black-Scholes hedging strategy using a modified volatility term. Another common method is the delta tolerance strategy first suggested by Whalley and Wilmott (1997). The method prescribes adjusting the hedging strategy to the Black-Scholes delta as soon as the delta of the risk position exceeds a prescribed tolerance distance from the 'perfect' hedge position. We use apply an adaptation of the work of Whalley and Wilmott (1997) to our framework. Recall, the optimal time consistent strategy is defined by

$$\theta_t^* = \frac{\rho\nu_t}{\sigma_t} X_t \frac{\partial E_t^*[h(X_T)e^{-r(T-t)}]}{\partial X_t} + S_t \frac{\partial E_t^*[h(X_T)e^{-r(T-t)}]}{\partial S_t} \quad (3.27)$$

then the series of rebalancing times is recursively given by

$$\tau_1 = t, \quad \tau_{i+1} = \inf \left\{ \tau_i < \tau < T : \left| \Delta_t^i - \Delta_\tau^i \right| > H \right\}, \quad i = 1, 2, \dots,$$

where Δ_t^i for $i \in \{X, S\}$ are the Greeks defined in (3.27) and the parameter H is a proxy for the amount of risk the hedger takes on. The hedger commits to the following strategy

- At time t the hedge position is equal to:

$$\theta_t^* = \frac{\rho\nu_t}{\sigma_t} X_t \Delta_t^X + S_t \Delta_t^S$$

- For time $s > t$ the hedger monitors the absolute difference between Δ_s and Δ_t . If $|\Delta_t - \Delta_s| < H$ then the hedger uses the Δ_t sensitivity

$$\theta_s^* = \frac{\rho\nu_s}{\sigma_s} X_s \Delta_t^X + S_s \Delta_t^S$$

and if $|\Delta_t - \Delta_s| \geq H$ then the hedger uses the Δ_s sensitivity

$$\theta_s^* = \frac{\rho\nu_s}{\sigma_s} X_s \Delta_s^X + S_s \Delta_s^S$$

A risk-averse hedger will choose a large bandwidth H , because this will enable him/her to rebalance the position frequently. Consequently, the hedging error at time of maturity is minimized. A large rebalance frequency could possibly carry large trading costs. A risk tolerant hedger will choose a small H , in order to avoid large trading costs and therefore accepts a larger hedging error.

4 Option Hedging strategies

In this section we will derive the dynamic optimal hedging strategies for options. For the strategies derived in this section we will assume that X_t and S_t follow a Geometric Brownian Motion.

4.1 Plain Vanilla Options

Consider a simple call or put option, this is a derivative with payoff function $h(X_T) = (\omega X_T - \omega K)^+$, where $\omega = 1$ for a call option and $\omega = -1$ for a put option. We assume that the assets S_t and X_t follow a Geometric Brownian Motion. We obtain the following expression for the optimal hedging portfolio of a call and put option

Proposition 3. *The optimal dynamic option hedging strategies for GBM non-tradable and tradable assets (3.20) and (3.21) are given by:*

$$\begin{aligned}\theta_{vanilla,t}^* &= \frac{\rho\nu}{\sigma} X_t \frac{\partial E_t^*[(\omega X_T - \omega K)^+ e^{-r(T-t)}]}{\partial X_t} + S_t \frac{\partial E_t^*[(\omega X_T - \omega K)^+ e^{-r(T-t)}]}{\partial S_t} \\ &= \frac{\rho\nu}{\sigma} e^{\left(m - r - \rho\nu\frac{\mu-r}{\sigma}\right)(T-t)} \Phi(\omega d_1)\end{aligned}\tag{4.1}$$

where

$$d_1 = \frac{\ln\left(\frac{X_t}{K}\right) + \left(m - \rho\nu\frac{\mu-r}{\sigma} + \frac{\nu^2}{2}\right)(T-t)}{\nu\sqrt{T-t}}$$

The derivation of Proposition 3 is provided in Appendix A.2. In their paper, Basak and Chabakauri (2011) derive the optimal dynamic hedging strategy for a call option under the 2GBM model in a general discrete-time setting. In this thesis, we will focus on an adaptation of the plain vanilla option: Asian options.

4.2 Asian Option

Asian options have a large array of applications in commodities, currency, energy, interest rates, equity and insurance markets. The name ‘Asian’ option emerged in 1987 when a Banker’s Trust Tokyo office used it for pricing average options on crude oil contracts. In contrast to vanilla options, the payoff of an Asian option is based on the average of the underlying asset price over an interval of time. Asian options have a lower volatility than plain vanilla options, since the payoff is dependent on an average of the underlying asset instead of a single asset price.

Asian options can be assigned to two main categories: floating strike and fixed strike. The payoff of a floating Asian option is the difference between the average of the

underlying asset and the asset price at time of maturity. The payoff of a fixed Asian option is the difference between the average of the underlying asset and the pre-specified strike price. Moreover, the average of the underlying asset can be taken arithmetically or geometrically. Given an underlying asset price process X_t , the payoff of a Asian arithmetic call option with exercise date T and strike price K is given by

$$A_t = \frac{1}{T} \int_0^t X_t dt$$

$$h(A_T) = (A_T - K)^+$$

and the payoff of a Asian geometric call option is given by

$$G_t = \exp \left(\frac{1}{T} \int_0^t \ln(X_t) dt \right)$$

$$h(G_T) = (G_T - K)^+$$

In this thesis we focus on arithmetic Asian call options with a fixed strike, because geometric Asian options are rarely used in practice.

The optimal hedging strategy for the Asian option is given by

$$\theta_{asian,t}^* = \frac{\rho\nu}{\sigma} X_t \frac{\partial E_t^*[h(A_T)e^{-r(T-t)}]}{\partial X_t} + S_t \frac{\partial E_t^*[h(A_T)e^{-r(T-t)}]}{\partial S_t}$$

There exists no easily tractable closed-form solution for the conditional expected value of an arithmetically averaged Asian option, and thus we have to approximate the expectation numerically. The simulation of a dynamic hedge requires to know the Asian options conditional expectation for many time and state space points. For the simulation of a single hedge portfolio value process, the conditional expectation needs to be calculated at every time point where the hedge portfolio is rebalanced. This approach is extremely time consuming. In practice fast estimation algorithms are desirable. Therefore we resort to the Least-Squares Monte Carlo algorithm presented in Ankirchner, Pigorsch, et al. (2014).

4.3 Least Squares Monte Carlo (LSMC) algorithm

The algorithm that we will present is complementary to the algorithm of Ankirchner, Pigorsch, et al., 2014, however we assume the hedging instrument to be a martingale under our hedge neutral measure $\mathcal{P}^*$ instead of under the real-world measure $\mathcal{P}$. Denote the conditional expectation of the discounted risk position under $\mathcal{P}^*$ by

$$\psi(X_t, t) = E_t^*[e^{-r(T-t)} h(X_T)]$$

and the gradient of the conditional expectation by

$$\nabla \psi(X_t, t) = \frac{\partial \psi(X_t, t)}{\partial X_t}$$

We can express $\nabla\psi(X_t, t)$ as the conditional expectation $\nabla\psi(X_t, t) = E^*[Z_{t,T}|X_t = x]$, where $Z_{t,T}$ are $\mathcal{F}_T$ -measurable random variables. Next, we approximate $E_t^*[Z_{t,T}|X_t = x]$ by $f(x, t)$ at each time step t using a set of chosen basis functions $L_1, \dots, L_n$ for $n \in \mathbb{N}$. The expression for $f(x, t)$ is given by:

$$f(x, t) \approx \sum_{k=1}^n \beta_{t,k} L_k(x)$$

where $\beta_{t,k}$ are regression coefficients minimizing the quadratic distance between $f(X_t, t)$ and $Z_{t,T}$. For sufficiently smooth function h we can use pathwise differentiation to obtain an estimate for $\nabla\psi(X_t, t)$ for the Asian option.

$$\frac{\partial\psi(X_t, t)}{\partial x} = E \left[\frac{\partial h}{\partial x}(X_T) \frac{\partial X_T}{\partial x} | X_t = x \right]$$

4.3.1 Basis functions

Assume that $\nabla\psi(X_t, t)$ is an element of $L^2(\Omega, \mathcal{F}, \mathcal{P})$. Since L^2 is a Hilbert space, it has a countable orthonormal basis and the conditional expectation function can be represented as a linear function of the elements of the basis. Therefore we can represent $\nabla\psi(X_t, t)$ as a linear combination of a countable set of $\mathcal{F}_t$ -measurable basis function. One example suggested by Longstaff and Schwartz (2001) is the set of weighted Laguerre polynomials

$$\begin{aligned} L_0(x) &= e^{-x/2} \\ L_1(x) &= e^{-x/2}(1 - x) \\ L_2(x) &= e^{-x/2}(1 - 2x + \frac{x^2}{2}) \\ &\dots \\ L_n(x) &= e^{-x/2} \frac{e^x}{n!} \frac{d^n}{dx^n} (x^n e^{-x}) \end{aligned}$$

The Laguerre polynomials belong to the set of orthogonal polynomials. Orthogonal polynomials are shown to solve the multicollinearity problem in case of multidimensional problems (Judd., 1999). The accuracy and the robustness of the LSMC using polynomial basis functions cannot be guaranteed for complex products because of the weakness of the polynomials to proxy the value of these products. The selection of basis functions for the regression model is crucial for the accuracy of the LSMC simulation. Therefore, we choose to adopt different sets of basis functions to examine the effect of basis function selection on the reduction of basis risk.

Basis function	Ankirchner et al. (2012)
$L_0(x)$	1
$L_1(x)$	$xe^{-x/c}$
$L_2(x)$	$x^2e^{-x/c}$
$L_3(x)$	$1_{\{u(x)>K\}}$
$L_4(x)$	$x1_{\{u(x)>K\}}e^{-x/c}$
$L_5(x)$	$(u(x) - K)^+e^{-x/c}$

Table 2: Functional form of basis functions. $u(x) = a(x) + (1 - \frac{t}{T})x$ and c is a constant

Basis function	Linear spline	Buckets
$L_0(x)$	1	1
$L_1(x)$	$xe^{-x/c}$	$xe^{-x/c}$
$L_2(x)$	$x^2e^{-x/c}$	$x^2e^{-x/c}$
$L_3(x)$	g_1	$1_{\{u(x) \in B_1\}}k_1$
$L_n(x)$	g_j	$1_{\{u(x) \in B_l\}}k_p$

Table 3: Functional form of basis functions; $g_j(x) = [u(x) - K_i]^+ - [u(x) - K_{i+1}]^+$; $k_p(x) = [x - K_i]^+ - [x - K_{i+1}]^+$

The basis functions of Ankirchner, Pigorsch, et al. (2014) consist of a combination of weighted monomial polynomials, weighted polynomials evaluated for in-the-money paths and the weighted payoffs of in-the-money paths. Longstaff and Schwartz (2001) state that the use of in-the-money paths limits the region over which the conditional expectation must be estimated, hence less basis functions are needed to obtain an accurate approximation of the conditional expectation. The use of all paths would result in estimates with larger standard errors than those obtained by using all of the conditioning (in-the-money) information.

For the linear spline basis functions we have $L_n(x)$ basis functions based on the spread function $g_j(x)$ for $j \in \{1, \dots, J\}$ and K_i for $i \in \{1, \dots, J+1\}$. For the bucket basis functions we have the functions $L_{p,l}(x)$ each representing a combination of a bucket $\{B_1, \dots, B_l\}$ and spread function $k_p(x)$ for $\{1, \dots, L\} \times \{1, \dots, P\}$ where K_i for $i \in \{1, \dots, P+1\}$. The choice of linear spline basis functions is motivated by shape of the Asian option payoff function. The functional form of the bucket basis functions is peculiar; in the basis functions we subdivide the simulation paths into different buckets evaluated for the projected payoff of the simulation path.

4.3.2 The Algorithm

The LSMC method creates scenarios through Monte Carlo simulation, and then performs iteratively, and at each time step a least-squares approximation of the conditional expectation. The LSMC algorithm aims at approximating the conditional expectation, without iterative simulation at each time step.

Preparation Let $\Delta t = \frac{T}{N}$ denote the step size for horizon $[0, T]$ and $N \in \mathbb{N}$. Define $t = \Delta i$ for $i \in \{0, \dots, N\}$.

Choose n basis functions $L_1, \dots, L_n \in \mathbb{R}$.

Step 1 Simulate $m \in \mathbb{N}$ trajectories of X and S under the $\mathcal{P}^*$ measure via Euler approximation. The j th Euler approximation of X and S at time $t + \Delta t$ is defined by:

$$\begin{aligned}\hat{X}_{t+\Delta t}^j &= \hat{X}_t^j + (m - \rho \frac{\mu - r}{\sigma}) \hat{X}_t^j \Delta t + \nu \hat{X}_t^j \Delta W_{2t}^j \\ \hat{S}_{t+\Delta t}^j &= \hat{S}_t^j + r \hat{S}_t^j \Delta t + \sigma \hat{S}_t^j \Delta W_{1t}^j\end{aligned}$$

where $\Delta W_{1,t}^j, \Delta W_{2,t}^j \sim N(0, \Delta t)$.

Step 2 Calculate

$$\hat{D}_t^j(\hat{X}_t^j, t) = \frac{\partial h}{\partial \hat{X}_t^j}(X_T^j) \frac{\partial X_T^j}{\partial \hat{X}_t^j}$$

Step 3 For every time point $i \in \{1, \dots, N-1\}$ regress $\hat{D}_t^j$ onto $L_1(\hat{X}_t^j), \dots, L_n(\hat{X}_t^j)$ using least-squares regression; the output is the set of estimated coefficients $\hat{\beta}_{t,1}, \dots, \hat{\beta}_{t,n}$ that minimize the sum of squared errors:

$$\min_{\beta} \left[\sum_{j=1}^m (y_j - \hat{y}_j)^2 \right]$$

where $y_j = \hat{D}_t^j$ and $\hat{y}_j = \sum_{k=1}^n \beta_t^k L_k(\hat{X}_t^j)$.

Step 4 Use the conventional Monte Carlo estimator for $i = 0$:

$$f(\hat{X}_0^j, 0) = \hat{D}_0^j(\hat{X}_0^j, 0)$$

For $i \in \{1, \dots, N\}$:

$$f(\hat{X}_t^j, t) = \sum_{k=1}^n \hat{\beta}_{t,k} L_k(\hat{X}_t^j)$$

where the vector $f(\hat{X}_t, t)$ is an approximation of $\nabla \psi(t, \hat{X}_t)$.

Step 5 Simulate m new trajectories of X and S via Euler approximation. Using the approximation of $\nabla \psi(\hat{X}_t, t)$ we calculate the hedging strategy θ_t^* .

4.3.3 Measure of error

In order to evaluate how the set of basis functions approximate the conditional expectation we use a measure of error. To use this measure we evaluate the basis functions for the case where S_t and X_t under the are martingales (i.e. driftless) under the hedge neutral measure $\mathcal{P}^*$. A similar approach is used in the paper of Ankirchner, Pigorsch, et al. (2014).

To explain how we can derive the measure of error, we start by taking the conditional expectation of the wealth process at time of maturity:

$$E_t[M_T] = E_t[M_t e^{r(T-t)} + \int_t^T \theta_s^* (\mu - r) e^{r(T-s)} ds + \int_t^T \theta_s^* \sigma e^{r(T-s)} dW_{1,s}] \quad (4.2)$$

$$E_t[M_T - M_t e^{r(T-t)}] = E_t\left[\int_t^T \theta_s^* \sigma e^{r(T-s)} dW_{1,s}\right] \quad (4.3)$$

If we substitute the equation above into the definition of G_t defined in (3.10) and use that $G_t = E_t^*[h(X_T)e^{-r(T-t)}]$ we obtain the following relation:

$$E_t^*[h(X_T)e^{-r(T-t)}] = E_t[h(X_T)e^{-r(T-t)}] - E_t[M_T^* - M_t e^{r(T-t)}] \quad (4.4)$$

The second term in (4.4) represents the expected gains in wealth, discounted back to time- t , that are obtained by hedging $h(X_t)$ with S_t over $[t, T]$. The left side of equation can also be seen as the expected "hedge neutral" value of $h(X_T)$ under the measure $\mathcal{P}^*$. Ultimately, the expected value of our risk position $h(X_T)$ equals the expected value under the real world measure $\mathcal{P}$ plus expected gains obtained by hedging. If we now assume that the non-tradable asset is a martingale, the right hand side of (4.4) is also a martingale. Consequently, the second term in (4.4) vanishes and the hedging strategy at time- t is not influenced by future hedging strategies anymore. Therefore, the hedging strategy becomes time-consistent. The hedging error variance under the martingale is given by:

$$\begin{aligned} \mathcal{V}_t &= \text{Var}_t[X_T - M_T^*] = (1 - \rho^2) E_t \left[\int_t^T \nu_s^2 X_s^2 \left(\frac{\partial E_s[h(X_T)]}{\partial X_s} \right)^2 ds \right] \\ &= (1 - \rho^2) \text{Var}_t \left[\int_t^T \left(\frac{\partial E_s[h(X_T)]}{\partial X_s} \right)^2 \nu X_s dW_{2,s} \right] \\ &= (1 - \rho^2) \text{Var}_t \left[\int_t^T \left(\frac{\partial E_s[h(X_T)]}{\partial X_s} \right)^2 dX_s \right] \\ &= (1 - \rho^2) \text{Var}_t[h(X_T)] \end{aligned}$$

Hence we can approximate $\text{Var}_t[X_T - M_T^*]$ by generating sample paths of $h(X_T)$ and multiply the empirical variance of $h(X_T)$ by $(1 - \rho^2)$; we call this estimate $\mathcal{V}_t$. We express the measure of error $\mathcal{M}_1$ as a fraction of the 'true' value.

$$\mathcal{M}_1 = \frac{1}{m} \sum_{n=1}^m \frac{|\hat{V}_t^n - \mathcal{V}_t^n|}{\mathcal{V}_t^n}$$

where $\hat{V}_t^n$ is the hedging error variance obtained by using the LSMC approximation of the conditional expectation.

5 Models

In this section we derive the hedging strategy for two simple models. In the first model the assets X_t and S_t are Geometric Brownian motions, hereupon referred to as the 2GBM model. In the second model the log asset prices of X_t and S_t are cointegrated, hereupon referred to as the stationary spread model.

5.1 2GBM model

For the 2GBM model the dynamics of X_t and S_t satisfy:

$$dX_t = mX_t dt + \nu X_t dW_{2,t} \quad (5.1)$$

$$dS_t = \mu S_t dt + \sigma S_t dW_{1,t} \quad (5.2)$$

where m, ν, μ and σ are constants. For now, consider the simple case of $h(X_T) = X_T$. The optimal dynamic hedging strategy for X_T can be obtained from Proposition 1 given that the conditional expectation of X_t under the measure P^* is given by:

$$E_t^*[X_T] = X_t e^{(m - \rho\nu(\mu - r)/\sigma)(T-t)} \quad (5.3)$$

Proposition 4. *The optimal dynamic hedging strategy and corresponding hedging error variance for the 2GBM model (3.20)-(3.21) are given by:*

$$\theta_t^{2GBM} = \frac{\rho\nu}{\sigma} X_t e^{(m - r - \rho\nu(\mu - r)/\sigma)(T-t)} \quad (5.4)$$

$$\begin{aligned} \text{Var}_t[X_T - M_T^*] &= (1 - \rho^2) \nu^2 X_t^2 e^{2(m - \rho\nu(\mu - r)/\sigma)(T-t)} \frac{e^{(\nu^2 + 2\rho\nu(\mu - r)/\sigma)(T-t)} - 1}{\nu^2 + 2\rho\nu(\mu - r)/\sigma} \\ &= (1 - \rho^2) \nu^2 X_t^2 e^{2(m - \rho\nu S_a)(T-t)} \frac{e^{(\nu^2 + 2\rho\nu S_a)(T-t)} - 1}{\nu^2 + 2\rho\nu S_a} \end{aligned} \quad (5.5)$$

where $S_a = \frac{\mu - r}{\sigma}$ is the Sharpe ratio of the hedge instrument.

The Sharpe ratio describes how much return is achieved per unit of risk by holding the underlying asset. Or put differently, how much risk an investor is willing to take on to achieve a level of return. The higher the Sharpe ratio of an asset, the better its returns are relative to the amount of risk taken on. A Sharpe ratio of zero implicates that the returns of asset S_t match the returns of the risk-free bond B_t . A negative Sharpe ratio implicates that the asset returns are smaller than the risk-free returns or that the asset returns are negative. If we were to obtain a more realistic indicator of the risk that we take on by holding asset S_t we could use the Sortino ratio; we then divide excess return by downside volatility instead of total volatility.

To assess the relative performance of the hedged position with respect to the unhedged position we define:

$$\Lambda_t = \sqrt{\frac{\text{Var}_t[X_T - M_T^*]}{\text{Var}_t[X_T]}}$$

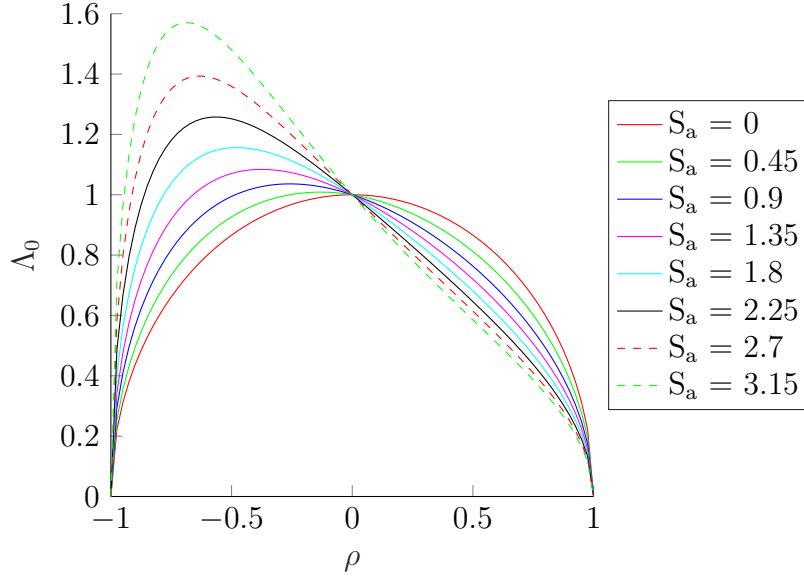


Figure 2: Ratio of the standard deviation of hedged ($X_T - M_T^*$) and unhedged position X_T plotted for $\rho \in [-1, 1]$. Evaluated for $T = 2$, $X_0 = 10$, $S_0 = 10$, $r = 0.01$, $m = 0.2$, $\nu = 0.2$ and Sharpe ratio $S_a \in [0.45, 3.15]$

Figure 2 displays the performance measure Λ_t for $\rho \in [-1, 1]$ from a time zero perspective. Evidently, if absolute correlation is high ($|\rho| > 0.9$), then a small change in the correlation leads to a large change reduction in basis risk relative to total risk. According to (5.4), the optimal hedging strategy for no correlation $\rho = 0$ is to not invest in the tradable asset at all; in Figure 2 we can see that the hedging error variance equals the variance of the unhedged risk position. For $\rho = \pm 1$ we are in the complete market setting and therefore it is possible to construct the perfect hedge. Accordingly, the hedging error variance equals zero.

Λ_0 becomes more skewed as S_a increases. For a positive S_a the measure Λ_0 increases in S_a for $\rho \in [-1, 0)$ and decreases in S_a for $\rho \in (0, 1]$. For $S_a < 0$ the contrary is true: Λ_0 increases in S_a for $\rho \in [-1, 0)$ and decreases in S_a for $\rho \in (0, 1]$. For example, assume we have two correlated assets with $\rho = 0.5$, then for a Sharpe ratio of 3.15 the basis risk is reduced by roughly 40% and for a Sharpe ratio of 0.9 only 20%. A higher Sharpe ratio implicates lower volatility per unit of return of asset S_t . For $\rho > 0$ basis risk decreases in S_a because the asset S_t is less volatile and therefore the hedger is better able to replicate the risk position $h(X_T) = X_T$.

The optimal hedging policy $\theta_{2GBM,t}$ is to take a long position in S_t for $\rho > 0$ and a short position for $\rho < 0$. Next, we evaluate the sensitivity of the hedging strategy

with respect to the asset volatility parameters. First take the derivative with respect to σ

$$\frac{d\theta_t^{2GBM}}{d\sigma} = \frac{e^{3(m-r-\frac{\nu\rho(\mu-r)}{\sigma})}(3(\mu-r)\nu^2\rho^2 - \nu\rho\sigma)}{\sigma^3}$$

For $S_a > \frac{1}{3\rho\nu}$ and $\rho \in (-1, 0]$ the strategy θ_t^{2GBM} increases in σ ; vice versa for $\rho \in [0, 1)$ and $S_a > \frac{1}{3\rho\nu}$ the optimal hedging strategy decreases in σ . Next, we take the derivative with respect to ν

$$\frac{d\theta_t^{2GBM}}{d\nu} = \frac{e^{3(m-r-\frac{\nu\rho(\mu-r)}{\sigma})}(3(\mu-r)\nu\rho^2 - \rho\sigma)}{\sigma^2}$$

Given that $\nu > \frac{1}{3\rho S_a}$ and $\rho \in [0, 1)$ the hedging position $\theta_{2GBM,t}$ decreases in ν ; conversely, for $\nu > \frac{1}{3\rho S_a}$ and $\rho \in (-1, 0]$ we have that θ_t^{2GBM} increases in ν .

ρ	S_a	$\frac{d\theta_t^{2GBM}}{d\sigma}$	$\frac{d\theta_t^{2GBM}}{d\nu}$
$(-1, 0]$	$> \frac{1}{3\rho\nu}$	> 0	> 0
$(-1, 0]$	$< \frac{1}{3\rho\nu}$	< 0	< 0
$[0, 1)$	$< \frac{1}{3\rho\nu}$	> 0	> 0
$[0, 1)$	$> \frac{1}{3\rho\nu}$	< 0	< 0

Table 4: Hedging strategy sensitivity with respect to volatility parameters.

Table 4 summarizes the volatility sensitivities for the different parameter conditions. To illustrate: given a fixed value of correlation $\rho > 0$ and asset volatility $\nu > 0$, as volatility σ increases the optimal strategy for the hedger is to increase his/her hedging position as long as S_a is larger than $\frac{1}{3\rho\nu}$ (note that S_a decreases in σ).

We are also interested in the parameters for which we have $\Lambda_0 > 1$, i.e. the hedged position $X_T - M_T^*$ has a higher variance than the unhedged position X_T . To this end, we will evaluate the derivative

$$\begin{aligned} \Pi_t &= \left. \frac{d\text{Var}_t[X_T - M_T^*]}{d\rho} \right|_{\rho=0} \\ &= - \frac{2e^{2m(T-t)}S_a(e^{\nu^2(T-t)} - 1) + 2e^{2m(T-t)}S_a(T-t)\nu^2}{\nu} \end{aligned}$$

for the parameters S_a, m, ν and T . For $\Pi_t > 0$ we have $\Lambda_0 > 1$ on $\rho \in (0, 1]$ and vice versa. We find that $\Pi_t < 0$ holds for a positive Sharpe ratio and $\Pi_t > 0$ for a negative Sharpe ratio. That is, for a positive Sharpe ratio the measure Λ_0 decreases in $\rho = 0$ and thus for positive correlated assets the hedging strategy will yield a lower variance than the unhedged position. The other way around, for a negative Sharpe ratio only negatively correlated variable will yield a lower variance than the unhedged position.

5.2 Stationary spread model

For the stationary spread model the dynamics of X_t and S_t satisfy:

$$dS_t = \mu S_t dt + \sigma S_t dW_{1,t} \quad (5.6)$$

$$dX_t = \kappa(\bar{X} - X_t)dt + \sigma_X dW_{2,t} \quad (5.7)$$

where μ, σ_X, κ and $\bar{X}$ are constants. In contrast to the 2GBM model, the market here consists of three stochastic processes. Due to the incompleteness imposed by the existing spread, the hedging strategy includes a term that captures the sensitivity of the hedge neutral asset I_t with respect to stochastic spread. This term quantifies the source of incompleteness due to level of co-integration. We consider hedging the risk position $h(X_T, S_T) = S_T e^{-X_T} = I_T$. In Corollary 3.2.1 we define the optimal dynamic hedging strategy in the scenario where the risk position explicitly depends on S_t , i.e. $h(X_T, S_T)$. Next, in Lemma 5.2.1 we derive closed form solutions for the Greeks of $h(X_T, S_T)$. Using Corollary 3.2.1 and Lemma 5.2.1 we present in Proposition 5 the dynamic optimal hedging strategy for assets (5.6) and (5.7)

Lemma 5.2.1. *Let X_t and S_t follow the dynamics of (5.12) and (5.13) then*

$$\frac{\partial E_t^*[e^{-r(T-t)} S_T e^{-X_T}]}{\partial S_t} = e^{-r(T-t)} \psi_s(x, s, t)$$

$$\frac{\partial E_t^*[e^{-r(T-t)} S_T e^{-X_T}]}{\partial X_t} = e^{-r(T-t)} e^{-\kappa(T-t)} S_t \psi_x(x, s, t)$$

where

$$\begin{aligned} \psi_s(x, s, t) &= \exp \left(r(T-t) - X_t e^{-\kappa(T-t)} - (\bar{X} + \rho \sigma_X \sigma \frac{1}{\kappa} + \rho \sigma_X \frac{1}{\kappa} \frac{\mu - r}{\sigma})(1 - e^{-\kappa(T-t)}) \right) \\ &\times \exp \left(\sigma_p^2 \frac{1}{4\kappa} (1 - e^{-2\kappa(T-t)}) \right) \end{aligned}$$

Proof. Under probability measure $\mathcal{P}^*$ the spread X_t is defined as

$$X_t = X_0 e^{-\kappa t} + \bar{X}(1 - e^{-\kappa t}) - \frac{1}{\kappa} \rho \sigma_X \frac{\mu - r}{\sigma} (1 - e^{-\kappa t}) + \int_0^t e^{-\kappa(t-u)} \sigma_X dW_{2,u}$$

Now we mathematically define $h(X_t, S_t) = S_t e^{-X_t}$

$$\begin{aligned} S_t e^{-X_t} &= S_t \exp \left((r - \frac{1}{2} \sigma^2) t - X_0 e^{-\kappa t} - \bar{X}(1 - e^{-\kappa t}) - \frac{1}{\kappa} \rho \sigma_X \frac{\mu - r}{\sigma} (1 - e^{-\kappa t}) \right) \\ &\times \exp \left(\int_0^t (\sigma - \rho e^{-\kappa(t-u)} \sigma_X) d\hat{W}_{1u} \right) \\ &\times \exp \left(- \int_0^t \hat{\rho} e^{-\kappa(t-u)} \sigma_X d\hat{W}_{2u} \right) \end{aligned} \quad (5.8)$$

Note that for a deterministic function $g(u)$ we have

$$\int_0^t g(u) dW_u \sim N(0, \eta^2) \quad \text{where} \quad \eta^2 = \int_0^t g^2(u) du$$

Due to the fact that $d\hat{W}_{1u} \perp d\hat{W}_{2u}$ the last two exponential terms of equation (5.8) are independently and lognormally distributed. The expectation of a lognormal distributed variable $\exp\left(\int_0^t g(u) dW_u\right)$ is given by $\exp\left(\frac{1}{2}\eta^2\right)$. Subsequently we have for the last two terms of the equation above:

$$\begin{aligned} E_t^* \left[\exp \left(- \int_0^t \hat{\rho} e^{-\kappa(t-u)} \sigma_X d\hat{W}_{2,u} \right) \right] &= \exp \left(\frac{1}{2} \int_0^t \left(\hat{\rho} e^{-\kappa(t-u)} \sigma_X \right)^2 du \right) \\ E_t^* \left[\exp \left(\int_0^t (\sigma - \rho e^{-\kappa(t-u)} \sigma_X) d\hat{W}_{1,u} \right) \right] &= \exp \left(\frac{1}{2} \int_0^t \left(\sigma - \rho e^{-\kappa(t-u)} \sigma_X \right)^2 du \right) \end{aligned}$$

By working out the integrals the expectation of $h(X_T, S_T) = S_T e^{-X_T}$ is defined as

$$\begin{aligned} E_t^*[S_T e^{-X_T}] &= S_t \exp \left(r(T-t) - X_t e^{-\kappa(T-t)} - (\bar{X} + \rho \sigma_X \sigma \frac{1}{\kappa} + \rho \sigma_X \frac{1}{\kappa} \frac{\mu - r}{\sigma})(1 - e^{-\kappa(T-t)}) \right) \\ &\quad \times \exp \left(\sigma_X^2 \frac{1}{4\kappa} (1 - e^{-2\kappa(T-t)}) \right) \end{aligned}$$

Now define $\psi(x, s, t) = E_t^*[S_T e^{-X_T}]$.

$$\psi_s(t, x, s) = \frac{1}{S_t} E_t^*[S_T e^{-X_T}]$$

$$\psi_x(t, x, s) = -e^{-\kappa(T-t)} E_t^*[S_T e^{-X_T}]$$

□

Using the result of Lemma 5.3.1 we can derive the optimal hedging strategy.

Proposition 5. *The optimal dynamic hedging strategy and corresponding hedging error variance for a risky position $h(X_t, S_t) = S_T e^{-X_T}$ with asset dynamics (5.6) and (5.7) are given by:*

$$\begin{aligned} \theta_t^{spread} &= \frac{\rho \sigma_X}{\sigma} \frac{\partial E_t^*[e^{-r(T-t)} S_T e^{-X_T}]}{\partial X_t} + S_t \frac{\partial E_t^*[e^{-r(T-t)} S_T e^{-X_T}]}{\partial S_t} \\ &= \left(1 - \frac{\rho \sigma_X}{\sigma} e^{-\kappa(T-t)} \right) e^{-r(T-t)} E_t^*[S_T e^{-X_T}] \\ &= \left(1 - \frac{\rho \sigma_X}{\sigma} e^{-\kappa(T-t)} \right) S_t \psi_s(x, s, t) \end{aligned} \tag{5.9}$$

$$\begin{aligned}
 \text{Var}_t[I_T - M_T^*] = & \sigma_X^2(1 - \rho^2)S_t^2 \\
 & \times \exp\left(2(\bar{X} + \rho\sigma_X \frac{1}{\kappa} \frac{\mu - r}{\sigma} - X_t)e^{-\kappa(T-t)} - 2(\bar{X} + \rho\sigma_X \frac{1}{\kappa} \frac{\mu - r}{\sigma})\right) \\
 & \times \exp\left(-2\rho\sigma \frac{\sigma_X}{\kappa}(1 - 2e^{-\kappa(T-t)}) + 2\frac{\sigma_X^2}{4\kappa}(1 - 2e^{-\kappa(T-t)})\right) \\
 & \times \int_t^T \exp\left(-2\kappa(T-u) + \sigma^2u - 2\rho\sigma\sigma_X \frac{1}{\kappa}e^{-\kappa(T-u)} + \sigma_X^2 \frac{1}{2\kappa}e^{-2\kappa(T-u)}\right) du
 \end{aligned} \tag{5.10}$$

The hedging strategy in Proposition 5 has a similar functional form as the hedging strategy in Ankirchner et al. (2012). The optimal strategy in (5.9) consists of two components. The first component consists of the sensitivity of the expected risk position with respect to the tradable asset. The second component is a function of time to maturity:

$$y(T-t) = 1 - \frac{\rho\sigma_X}{\sigma}e^{-\kappa(T-t)}$$

The function $y(T-t)$ converges to 1 for a large horizon $T-t$ or large mean-reversion term κ . For a large horizon or mean reversion term we expect the log spread X_t to return to its mean reversion level $\bar{X}$ before the time of maturity. Consequently, the basis risk that arises through the log spread is eliminated. The optimal hedging strategy equals

$$\theta_t^{spread} = S_t\psi_s(x, s, t)$$

which translates to taking a hedging position equal to the sensitivity of asset X_t with respect to S_t . Additionally, we observe that for small values of mean reversion κ the function $y(T-t)$ converges to

$$y(T-t) = 1 - \frac{\rho\sigma_X}{\sigma} = \frac{\sigma - \rho\sigma_X}{\sigma} = \frac{\sigma_I}{\sigma}\rho_{IX} \tag{5.11}$$

and the optimal hedging strategy equals

$$\theta_t^{spread} = \frac{\sigma_I}{\sigma}\rho_{IX}S_t\psi_s(x, s, t) \tag{5.12}$$

The same goes for times close to time to maturity; for $\lim(T-t) \rightarrow 0$ the function $y(T-t)$ converges to the value in (5.12) as well.

6 Simulation results

We use the Monte Carlo method to simulate m asset paths. Define the following

$$V_j = (h(X_{j,T}) - M_{j,T}) \quad (6.1)$$

where $X_1, \dots, X_m$ are i.i.d random variables. Denote the sample mean by a_m and sample standard deviation by b_m .

$$a_m = \frac{1}{m} \sum_{j=1}^m V_j(X_{j,T}, S_{j,T}) \quad b_m = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (V_j(X_{j,T}, S_{j,T}) - a_m)^2}$$

The Law of Large numbers implies

$$\lim_{m \rightarrow \infty} a_m \rightarrow E[h(X_T) - M_T] \quad \lim_{m \rightarrow \infty} b_m^2 \rightarrow \text{Var}[h(X_T) - M_T]$$

Under the assumption that $m \rightarrow \infty$ and $\text{Var}[h(X_T) - M_T] < \infty$ the Central Limit Theorem states that

$$\sqrt{m}(a_m - E[h(X_T) - M_T]) \xrightarrow{a} N(0, \text{Var}[h(X_T) - M_T])$$

Using the Central Limit Theorem we can construct a 95% confidence interval for our Monte Carlo estimator a_m

$$[a_m - 1.96 \frac{1}{\sqrt{m}} b_m, a_m + 1.96 \frac{1}{\sqrt{m}} b_m]$$

The rate of convergence of the Monte Carlo estimator equals $\sqrt{m}$. As we try to achieve greater accuracy using Monte Carlo simulation, we must increase the number of simulation paths drastically. To reduce the sampling standard error by half we have to quadruple the number of simulations. This relatively slow convergence is one of the weaknesses of the Monte Carlo simulation technique.

In this numerical experiment the hedging portfolio is rebalanced at the monitoring dates of the (underlying of the) risk position. From a theoretical perspective the hedging strategies derived in this paper require continuous rebalancing of the hedging position. Discrete rebalancing originates an error, however continuous rebalancing would incur an infinite amount of transaction cost for the hedger. To find a balance in the trade-off between accuracy and validity of our hedging strategy, we have adopted discrete rebalancing.

To assess the performance of a given hedging strategy we compare the variance of the hedged position to the variance of the unhedged position. Define the measure HE (hedging effectiveness) as the percentage increase or decrease in the hedging error

variance from using strategy θ_t compared the unhedged strategy. Mathematically, the measure HE is defined as

$$HE(\theta) = 1 - \left(\frac{\text{Var}[h(X_T) - M_T]}{\text{Var}[h(X_T)]} \right) \quad (6.2)$$

A positive value implies that the quality of the evaluated hedge is higher than that of the alternative hedge. The higher the measure HE , the better the quality of the hedge. The measure $HE\%$ can be interpreted as the percentage reduction of basis risk due to hedging; for a complete market this would be one hundred percent.

6.1 Model error

There is a possibility that a statistical test leads to the wrong conclusion, or different test lead to different conclusions. Consider the following situation: assets X_t and I_t are cointegrated, however due to ignorance or model uncertainty the hedger chooses to his/her risk position $h(X_T, S_T)$ using the dynamic optimal 2GBM hedging strategy. Consequently, the question arises: how does the performance of θ_t^{2GBM} compare to θ_t^{spread} under a model with cointegrated assets? The 2GBM hedging strategy under this model is given by

$$\theta_t^{2GBM} = \frac{\rho_{IX}\sigma_I}{\sigma} I_t \exp \left((\mu_I - \rho_{IX}\sigma_I \frac{\mu - r}{\sigma})(T - t) \right) \quad (6.3)$$

where the true parameters of I_t are given by

$$\begin{aligned} \mu_I &= \left(\frac{1}{2}\sigma_X^2 - \kappa(\bar{X} - X_t) + \mu - \rho\sigma_X\sigma_t \right) \\ \sigma_I &= \sqrt{\sigma^2 - 2\rho\sigma\sigma_X + \sigma_X^2} \end{aligned}$$

Note, in the 2GBM strategy above we use the true parameters of the cointegrated 'true' asset model. This choice is motivated by the attempt to isolate the effect of using the wrong model assumptions, without interference of parameter risk. Furthermore, in a practical application where the asset evolve according to our 'true' cointegrated model and the hedger uses a 2GBM hedging strategy, we assume that the hedger is able to estimate the drift and volatility parameters from historical observations and plug them into his/her hedging strategy (6.3). We simulate both strategies under the cointegrated model for different time of maturities ($T - t$). Define the function $\Delta HE = HE_{spread} - HE_{2GBM}$ as the difference in hedging effectiveness of strategies θ_t^{spread} and θ_t^{2GBM} under the cointegrated asset model. The difference in hedging effectiveness measure by ΔHE increases for larger values of correlation and mean reversion as can be seen in Figure 3. Stronger levels of mean reversion translate to a higher relative performance of strategy θ_t^{spread} to θ_t^{2GBM} , even for 'relatively' small values of correlation ρ . Clearly, for short time to maturity hedges the long-term equilibrium of the cointegrated assets is not reached. However, for hedging strategies with a longer horizon, we observe that the contrast in hedging performance between the 2GBM and spread strategy is considerable. Strong correlation and a large mean reversion term expedite the time it takes for the θ_t^{spread} strategy to reach this long

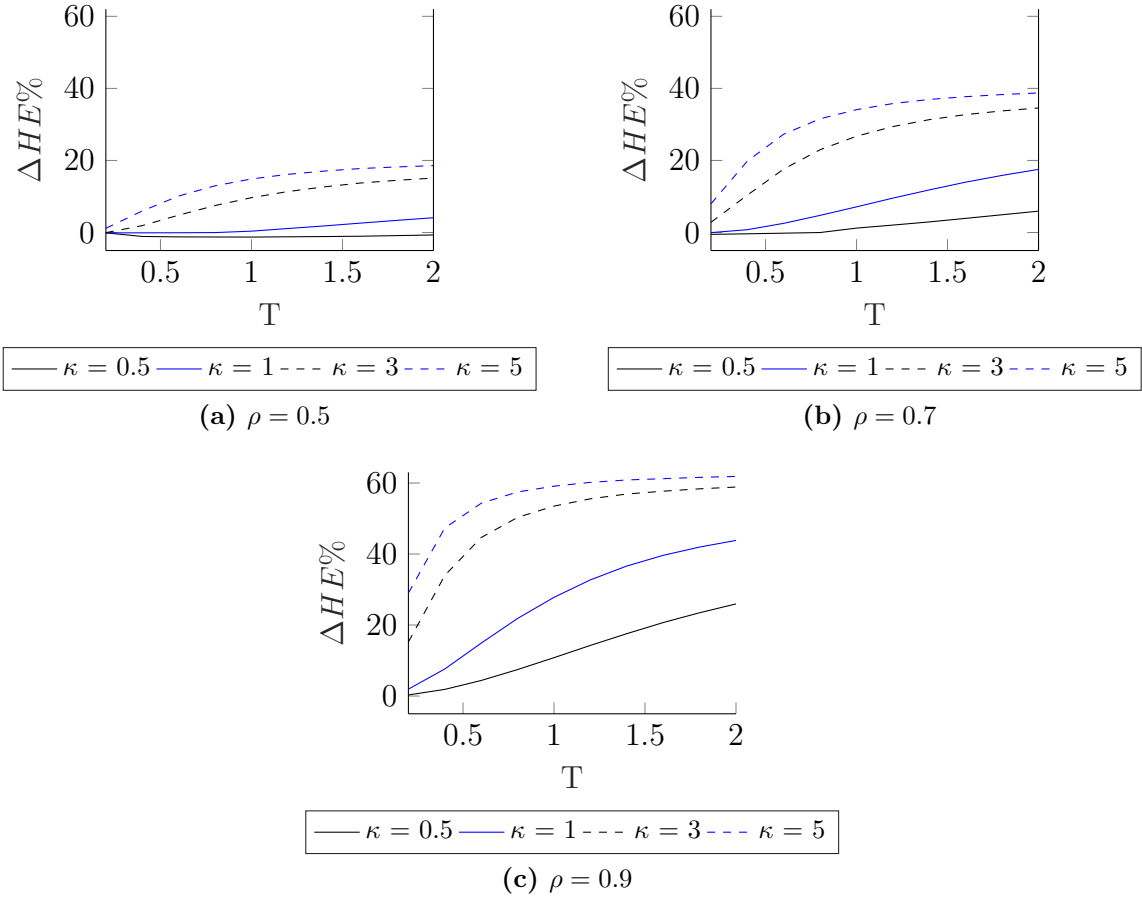


Figure 3: The difference in hedging effectiveness for different correlation values and mean reversion parameters. Asset parameters: $\mu = 0.1$, $\sigma = 0.2$, $\bar{X} = 0.1$, $I_0 = 12$, $S_0 = 10$, $X_0 = 0.08$. Monte Carlo parameters: $T \in [0.2, 2]$, $m = 100000$ and $N \in [50, 500]$.

term state. In Figure 4 we plot the hedging strategy of θ_t^{spread} and θ_t^{2GBM} over time for a time to maturity $T = 1$, mean reversion $\kappa = 3$ and correlation $\rho = 0.9$. The difference in hedging effectiveness measure ΔHE is 70% as we can see in Figure 3c. The 2GBM strategy clearly under hedges the risk position as the hedging position is significantly smaller than the spread strategy. As the θ_t^{spread} strategy approaches the time to maturity, the quantiles converge to the θ_t^{2GBM} strategy quantiles.

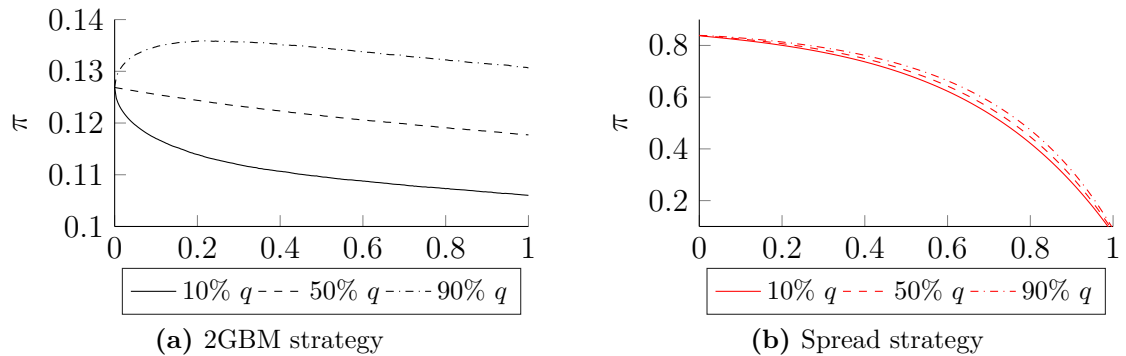


Figure 4: Hedging position quantiles for $\rho = 0.9$, $\kappa = 3$ and $T = 1$.

6.2 Option hedging

In this section we analyze dynamic hedging strategy for $h(A_T)$ where the risk position is an Asian call option. We will evaluate a base case Asian option that is based on the following asset parameters

μ	m	σ	ν	r	X_0	S_0	T	K
0.1	0.1	0.2	0.2	0.02	10	10	1	[8,12]

Table 5: Parameter values for base case Asian option

We define the following input parameters for the basis functions.

- Parameter c in the exponential dampening factor is set to 2.
- The buckets B_l are determined using the q -% quantiles of the simulated asset payoff $h(A_T)$ under the hedge neutral measure. The range of q is equal to $q = \{0, 25, 50, 75, 1\}$.
- The strike prices K_i in the spread functions are set to a range with increments of one.

Figure 5 illustrates the hedging effectiveness of the simulated Asian option using four different sets of basis functions. The hedging effectiveness is evaluated at different levels of moneyness, which is the strike price divided by the asset price at time zero. The four basis sets perform approximately equal across correlation and moneyness in terms of the hedging effectiveness. Next, we evaluate the basis functions in terms of measure of error $\mathcal{M}_1$. In order to evaluate the measure of error, we simulate the Asian option for the scenario where X_t and S_t are both martingales. The results are displayed in Figure 6. Recall, $\mathcal{M}_1$ is defined as the absolute difference between the approximated hedging error variance of the LSMC algorithm and plain Monte Carlo, as a fraction of the approximated plain Monte Carlo value. This measure of error should give us some insight into the performance of our basis functions. The measure of error increases in moneyness for all four basis function sets, as we can observe in Figure 6. A possible explanation for this occurrence could be that for higher strike prices the number of nonzero payoffs will decrease. Consequently, it will become harder for the LSMC to approximate the conditional expectation. The Linear Spline basis function set reveals the most consistent measure of error over the levels of moneyness, therefore we select this basis function set for our next analyses.

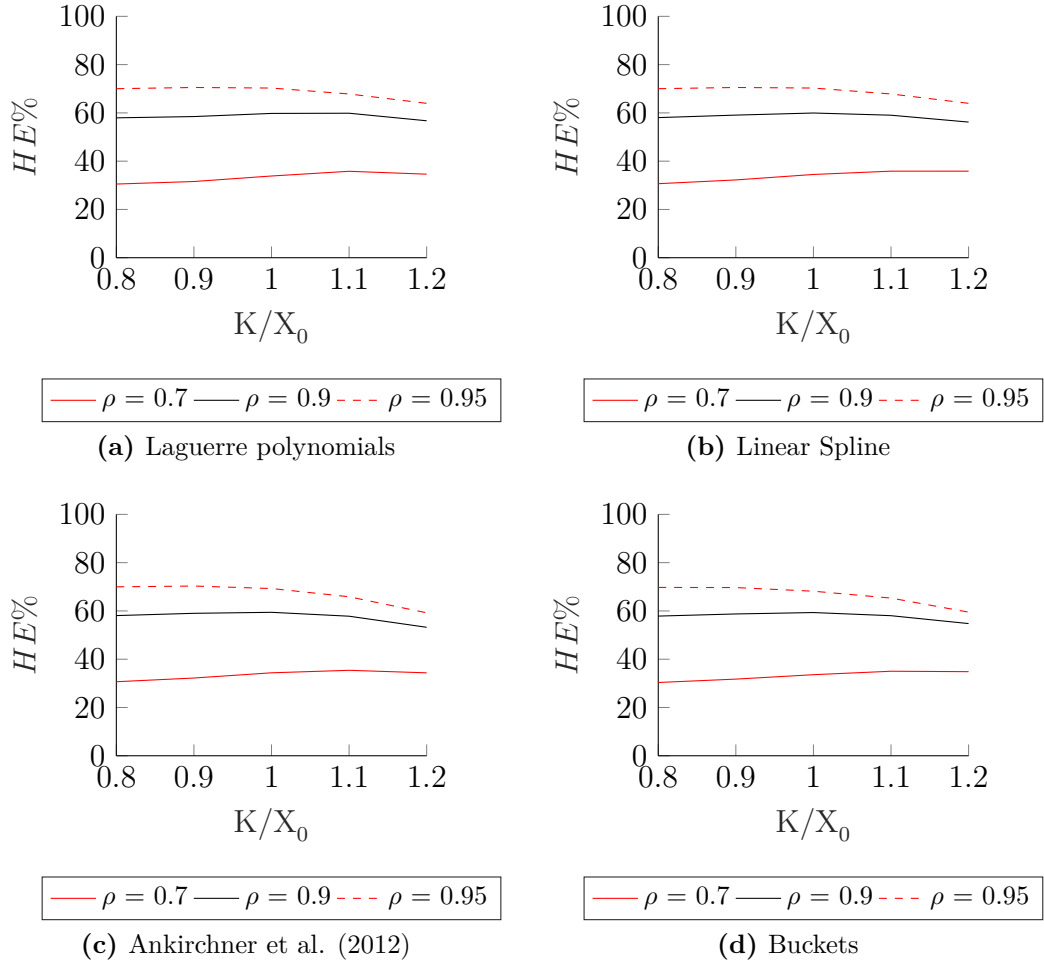


Figure 5: HE as function of moneyness for different basis function sets. Monte Carlo parameters: $N = 250$ and $m = 100000$. Basis functions parameters: $J = 22$, $L = 4$, $P = 32$; $3 \leq K_i \leq 25$; for the linear spline basis functions; $3 \leq K_i \leq 35$. for the bucket functions.

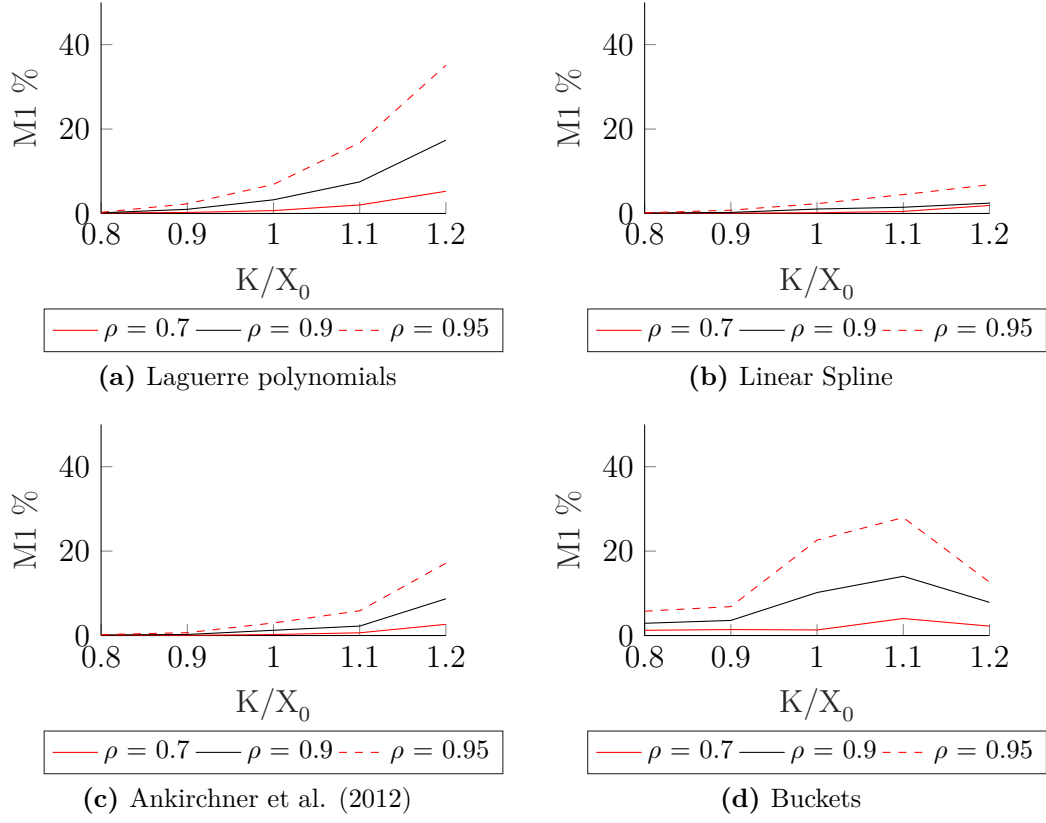


Figure 6: Measure of error $\mathcal{M}_1$ as function of moneyiness for different basis function sets under the driftless $\mu = m = 0$ scenario. Monte Carlo parameters: $N = 250$ and $m = 100000$.

6.2.1 Parameter risk

In this section we perform a small numerical experiment to assess the impact of a misspecified drift parameter. We assume that the hedger computes the hedging strategy using erroneous estimates of the drift parameters, but the tradable and non-tradable assets are simulated according to the 2GBM model with the correct values. Assume that the correlation between the assets is 0.9. In order to systematically evaluate the impact of misspecified drift parameter we define the estimates μ_1 and μ_2 which are respectively 50% lower and 50% higher than the correct parameter value μ^* . Similarly, define m_1 and m_2 for m^* . Table 6 displays the parameter sensitivity for different strike prices. Asian option hedging strategies are fairly robust to misestima-

		<i>HE%</i>		
		<i>K</i> = 8	<i>K</i> = 10	<i>K</i> = 12
μ^*	m^*	57.88	59.94	57.09
μ_1	m_1	57.89	59.74	55.78
μ_2	m_2	57.86	60.06	57.83
μ_1	m_2	57.43	56.48	48.48
μ_2	m_1	57.89	50.79	32.11

Table 6: Parameters: $T = 1$, $S_0 = 10$, $X_0 = 10$, $M_0 = 1$, $r = 0.02$ and $\sigma = \nu = 0.2$, $\mu^* = 0.1$ and $m^* = 0.1$. Simulated using $m = 100000$ paths and $N = 250$ time steps.

tion of parameters as we observe in Table 6, especially if the misestimation of μ and m is in the same direction. For the misspecified parameters μ_2 and m_1 the effect is moderately more apparent for out-of-the-money options; the hedger believes that the option is less risky than in reality and that the asset S_t is a more effective hedging instrument than in reality. For the parameters μ_1 and m_2 the hedger assumes that the hedging instrument is less effective than in reality and that the option is riskier than in reality.

6.2.2 Hedging strategies

Using the linear spline basis function set we analyse Asian options with three different strike prices: ATM ($K = 10$), OTM ($K = 12$) and ITM ($K = 8$). The correlation is fixed at $\rho = 0.9$ and hedging horizon $T = 1$. A graphical representation of the quantiles of portfolio weights π_t over time for $q \in \{0.1, 0.5, 0.9\}$ is presented in Figure 7.

The Asian call option Delta represents the change in expected option payoff with respect to a change in the underlying asset. Delta can be interpreted as the probability of a positive payoff value; for more in-the-money the options this probability is larger and thereby also the value of Delta. The dynamic Asian option hedging strategy is driven by this Delta under the hedge neutral measure, consequently for more in-the-money options the hedging position will be larger. The dynamic hedging strategy

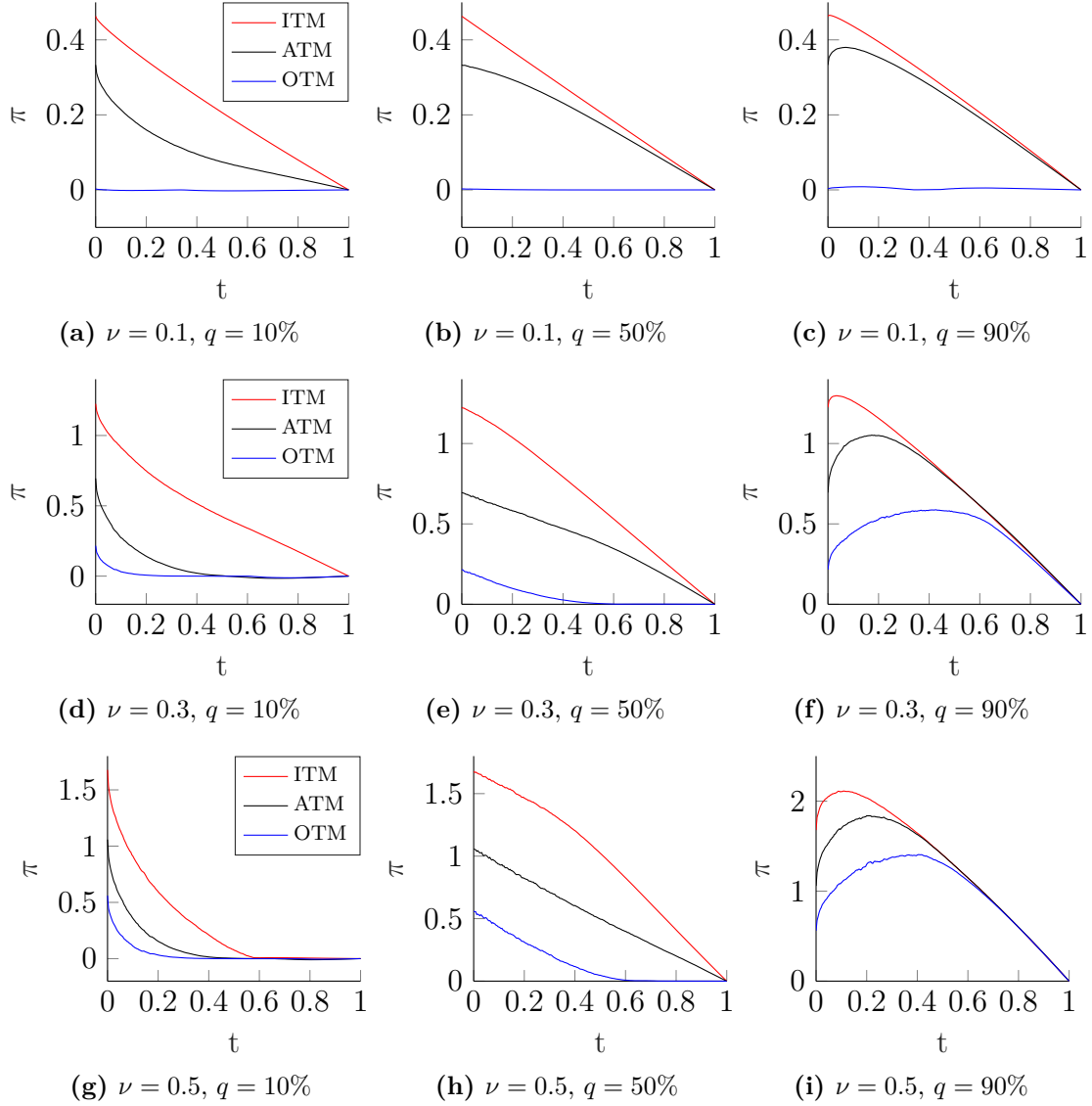


Figure 7: Quantiles of optimal dynamic hedging strategies for different levels of moneyness and volatility. Monte Carlo parameters: $N = 250$ and $m = 100000$.

is thus dependent on the moneyness of the option. By looking at the options over the different quantile scenarios we observe that the ATM and OTM options are more sensitive to the simulation path of the underlying asset. This sensitivity is called Gamma and is denoted by

$$\Gamma_t = \frac{\partial^2 E_t^*[h(A_T)e^{-r(T-t)}]}{\partial^2 X_t}$$

The optimal hedging position for ATM and OTM options will require rebalancing more often to reduce basis risk. We expect that the hedging strategy to perform better if we use hedge using Gamma instead of Delta, since the hedging strategy is more sensitive with respect to the underlying assets for ATM and OTM options compared to ITM options. The sensitivity of the Asian option payoff with respect volatility of the underlying asset is called vanna. An increase in volatility results in increment in

the probability that OTM options will have a positive payoff. Consequently, in more volatile market conditions the optimal hedging strategy for OTM is bound to change more than ATM and ITM options.

Furthermore, the decay rate of the hedging position changes across all three dimensions: moneyness, quantile scenario and volatility. For example, for $\nu > 0.1$ all the 90% quantile scenarios exhibit a certain time point $\tau \in [0, T]$ for which the optimal strategy is to decrease the hedge position. This time point τ increases with the strike price K . This is due the nature of the Asian option. As moneyness increases, it will take longer before there is a reduction of the uncertainty about the terminal payoff.

To illustrate the dynamics of the hedging strategy more clearly we plot asset paths and corresponding hedging strategies for an arbitrarily chosen simulation. For a correlation $\rho = 0.9$ we graphically display the strategies in Figure 8b. In Figure 8a we introduce the averaging process

$$Q_t = \frac{1}{t} \int_0^t X_s ds$$

In Figure 8c we present the time difference operator for the Delta of the Asian option, where the operator is defined as $\phi(\Delta_t) = \Delta_t - \Delta_{t-1}$. Consistent with our intuition, Delta displays more variation for more out-the-money options.

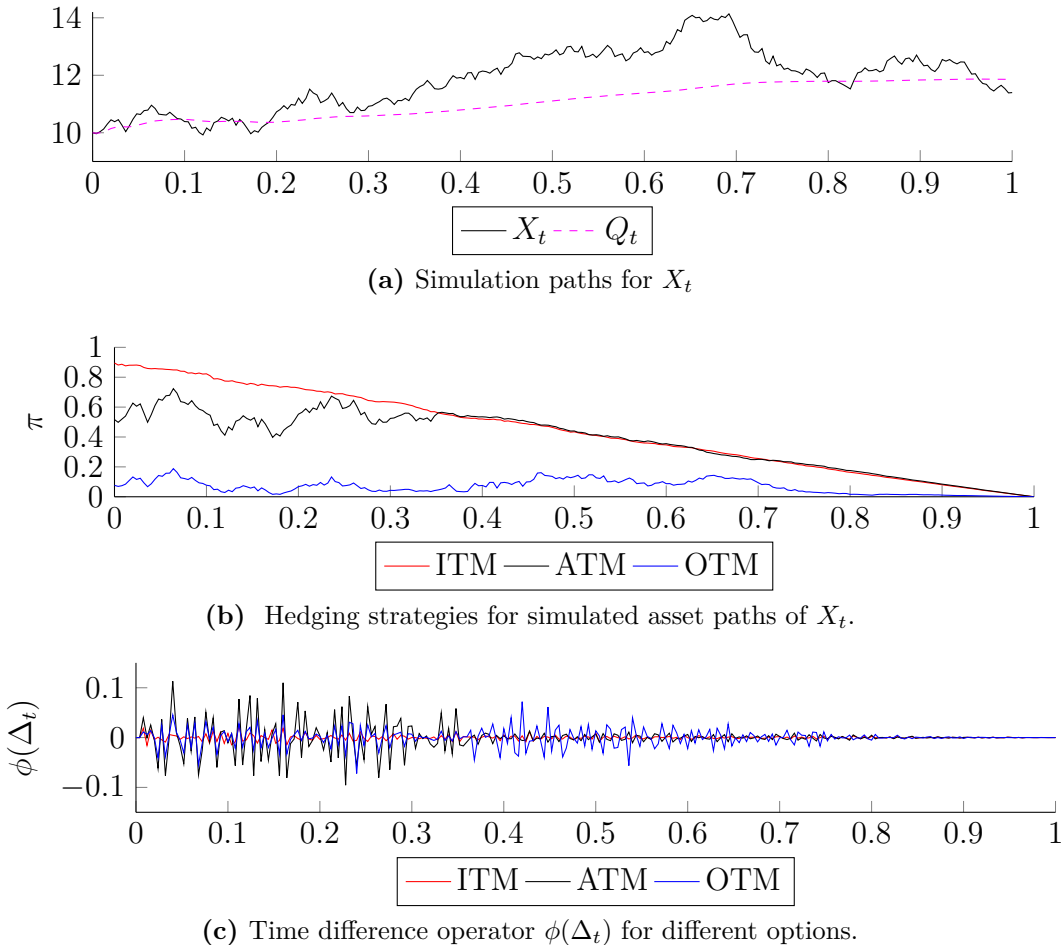


Figure 8: Simulated paths for X_t and ATM and OTM hedging strategies.

6.2.3 Delta Tolerance strategy

In this section we will introduce transaction cost into our model. The transaction cost are defined as a percentage of the volume of the trade. We will let the transaction cost ϱ vary over a hypothetical range of zero to 200 bps. One basis point (bps) is equal to 0.01%. With the introduction of transaction cost we are able to evaluate the trade-off between reducing transaction cost and hedging error variance. Furthermore, we will test the effect of the Delta Tolerance strategy on the performance of our hedge. In Table 7 we present the hedging effectiveness of our hedging strategy for a OTM option. Our choice to evaluate the OTM is motivated by the fact that OTM options are more sensitive to a change in the underlying asset than ITM and ATM options, and therefore require more frequent rebalancing. The correlation is set to $\rho = 0.9$, and all other asset parameter are equal to the values in Table 5. In Table 7 we present the hedging effectiveness of our Delta Tolerance strategy for different levels of transaction cost and different values of the bandwidth H .

bps	0	50	100	200
H	$HE\%$			
0	56.09	44.79	29.92	-4.24
0.05	54.35	47.21	38.11	16.68
0.1	43.17	38.41	32.26	17.04
0.15	39.72	35.60	30.28	17.01
0.2	37.80	34.00	29.10	16.82

Table 7: Hedging effectiveness as function of bandwidth H for different transaction cost scenario's. Calculated for an OTM option. Asset parameter as given in Table 5.

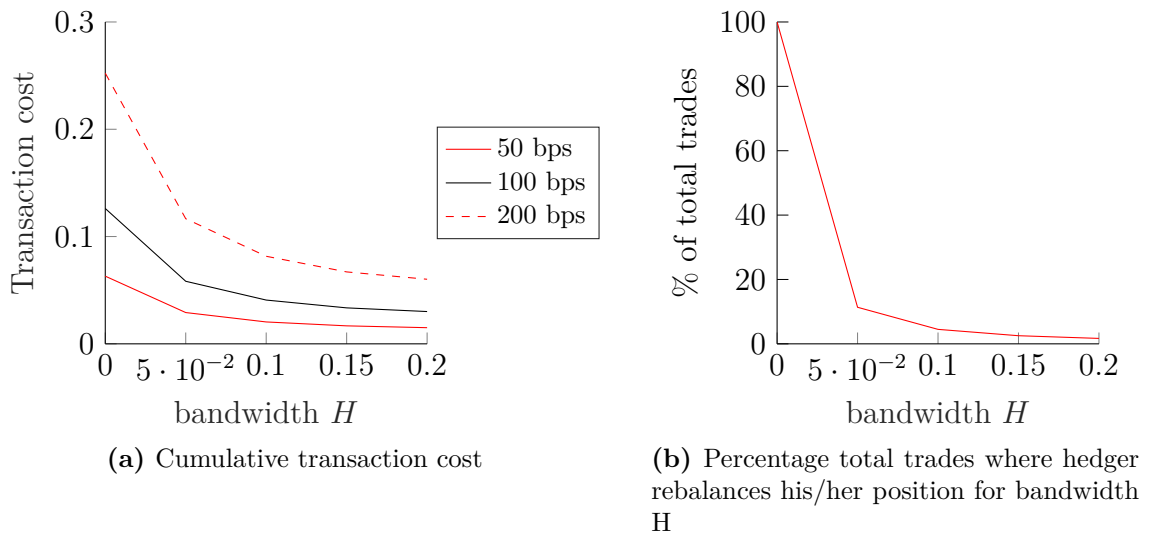


Figure 9: Illustration of the cumulative transaction cost and percentage of trades for different values of bandwidth H .

For zero transaction cost, increasing bandwidth H has a negative impact on the hedging effectiveness. Intuitively, this makes sense since the Delta Tolerance strategy for bandwidth zero equals the dynamic optimal hedging strategy. Under zero transaction cost, increasing the bandwidth results in a sub-optimal strategy. The hedging effectiveness is negatively impacted by an increase of transaction cost. For $H = 0$ and sufficiently high transaction cost ($\text{bps} \geq 200$) our dynamic hedging strategy becomes subordinate to the unhedged strategy. Transaction cost affect the ability of our hedging strategy to replicate the payoff of our OTM Asian option. In the Delta Tolerance strategy the hedging position is only corrected if the change of the asset with respect to the expected value of the Asian option is larger than our preset bandwidth. Therefore, the number of periods that the hedging strategy is rebalanced over the hedging horizon is drastically reduced, as can be seen in Figure 9b. For a bandwidth larger than 0.2, the dynamic strategy will converge to a static one.

7 Conclusion

In this research we studied the dynamic time consistent hedging strategy of Basak and Chabakauri (2011) in an incomplete market setting. We derived closed form solutions for this strategy under an asset model with two Geometric Brownian Motions and an asset model with two cointegrated assets. Furthermore, we constructed a numerical hedging strategy for an Asian option using a LSMC algorithm in a market setting with two Geometric Brownian Motions.

The objective of this research was to evaluate the performance of the two extensions of the Basak and Chabakauri (2011) paper using simulation techniques. The performance of the cointegrated asset model is evaluated in the context of model error. We find that the relative hedging performance increases in correlation, mean reversion and time to maturity, in the case that hedger employs the correct hedging strategy under the cointegrated asset model. In that respect, a hedger can significantly improve his/her hedging effectiveness by accounting for incorporating the possibility of cointegration between assets. Furthermore, we find that Asian option hedging strategies perform approximately equal in terms of reduction of basis risk for the range of strike prices we evaluated in our experiment. However, the hedging positions themselves differ considerably for the different strike prices. At-the-money and out-of-the-money options are more sensitive to the simulation path of the underlying asset and require more frequent rebalancing of the hedging position.

Although we did not find surprising results in general, future research may provide greater insights into basis risk and the effects of hedging. Some possible other research directions could be to extend our numerical hedging strategy to more complex risk positions or exotic options. In line with this research direction, a comparison could be made between the performance of the LSMC algorithm and the implementation of a neural network to approximate the conditional expectation of the risk position. Additionally, it would be interesting to leverage such a numerical approach to investigate the performance of a risk position that is hedged using an imperfectly correlated diversified portfolio of assets.

A Appendix

A.1 Proof of Corollary 3.2.1

Proof. The wealth process M_T^* is given by

$$M_T^* = M_t e^{r(T-t)} + \int_t^T \theta_s^* (\mu_s - r) e^{r(T-s)} ds + \int_t^T \theta_s^* \sigma_s e^{r(T-s)} dW_{1,s}$$

Using the law of total variance we obtain a recursive representation of the value function

$$J_t = \min_{\theta_t} E_t[J_{t+\varepsilon}] + \text{Var}_t(E_{t+\varepsilon}[h(X_T, S_T) - M_T])$$

Now define the process G_t as

$$G_t = G_t(X_t, S_t, t) = E_t[h(X_T, S_T) e^{-r(T-t)} - \int_t^T \theta_s^* (\mu_s - r) e^{r(T-s)} ds]$$

Applying Ito's Lemma to process G :

$$dG_t = \dots dt + e^{r(T-t)} \left(\frac{\partial G_t}{\partial X_t} \nu_t dW_{2,t} + \frac{\partial G_t}{\partial S_t} S_t \sigma_t dW_{1,t} \right)$$

Now we can write the value function as:

$$\begin{aligned} 0 = \mathcal{D}J_t + \min_{\theta_t} & \left[\left(\frac{\partial G_t}{\partial X_t} \nu_t \right)^2 + \left(\frac{\partial G_t}{\partial S_t} S_t \sigma_t \right)^2 + (\theta_t \sigma_t)^2 + 2\rho \frac{\partial G_t}{\partial S_t} S_t \sigma_t \frac{\partial G_t}{\partial X_t} \nu_t \right. \\ & \left. - 2 \frac{\partial G_t}{\partial S_t} S_t \theta_t \sigma_t^2 - 2\rho \frac{\partial G_t}{\partial X_t} \nu_t \theta_t \sigma_t \right] e^{2r(T-t)} \end{aligned}$$

Minimizing the equation above in θ results in:

$$\theta_t^* = \frac{\rho \nu_t}{\sigma_t} \frac{\partial G_t}{\partial X_t} + S_t \frac{\partial G_t}{\partial S_t}$$

Next, we substitute the optimal value of θ back into the PDE:

$$\mathcal{D}J_t + (1 - \rho^2) \left(\nu_t X_t \frac{\partial G_t}{\partial X_t} \right)^2 e^{2r(T-t)} = 0$$

By using the terminal condition $J_T = 0$ we obtain:

$$J_t = (1 - \rho^2) E_t \left[\int_t^T \left(\nu_s X_s \frac{\partial G_s}{\partial X_s} \right)^2 e^{2r(T-s)} ds \right]$$

Next, we show that $G_t = E_t^*[h(X_T)e^{-r(T-t)}]$. We apply the Feynman-Kac theorem to the definition of G_t in (3.10) and substitute the optimal strategy into the partial differential equation. This gives us:

$$\frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} X_t (m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t}) + \frac{\partial G_t}{\partial S_t} r S_t + \frac{1}{2} \left(\frac{\partial^2 G_t}{\partial^2 X_t} + 2\rho \nu_t \sigma_t X_t S_t \frac{\partial^2 G_t}{\partial X_t \partial S_t} + \sigma_t^2 S_t^2 \frac{\partial^2 G_t}{\partial^2 S_t} \right) - r G_t = 0 \quad (\text{A.1})$$

where the terminal condition is given by $G_T = h(X_T, S_T)$. By applying the reversed Feynman-Kac solution to equation (3.16) we obtain

$$G_t = E_t^*[h(X_T, S_T)e^{-r(T-t)}] \quad (\text{A.2})$$

where the expectation is taken under the probability measure $\mathcal{P}^*$. The probability measure $\mathcal{P}^*$ is defined as the measure on which assets X_t and S_t have drift $(m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t})$ and r respectively. In order to better understand this change of measure look at the drift terms of G_t under $\mathcal{P}$ and $\mathcal{P}^*$.

$$\mathcal{D}G_t = \frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} m_t X_t + \frac{\partial G_t}{\partial S_t} \mu_t S_t + \frac{1}{2} \frac{\partial^2 G_t}{\partial X_t^2} \nu_t^2 X_t^2 + \frac{1}{2} \frac{\partial^2 G_t}{\partial S_t^2} \sigma_t^2 S_t^2 + \rho \frac{\partial^2 G_t}{\partial X_t \partial S_t} \nu_t \sigma_t X_t S_t$$

$$\mathcal{D}G_t^* = \frac{\partial G_t}{\partial t} + \frac{\partial G_t}{\partial X_t} X_t (m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t}) + \frac{\partial G_t}{\partial S_t} r S_t + \frac{1}{2} \frac{\partial^2 G_t}{\partial X_t^2} \nu_t^2 X_t^2 + \frac{1}{2} \frac{\partial^2 G_t}{\partial S_t^2} \sigma_t^2 S_t^2 + \rho \frac{\partial^2 G_t}{\partial X_t \partial S_t} \nu_t \sigma_t X_t S_t$$

Clearly, both drift term are equal if process G_t under the measure $\mathcal{P}^*$ is defined with asset drifts $(m_t - \rho \nu_t \frac{\mu_t - r}{\sigma_t})$ and r . Under these alternative drift terms for G_t , the optimal strategy defined in (3.13) solves equation (3.12) and is therefore the solution to the value function in (3.2). Since our goal is to find θ_t such that (3.2) is minimized, the proof is complete. $\square$

A.2 Derivation of Proposition 3

Since the asset X_t follows a GBM, we can write the process X_t under the hedge neutral measure as

$$X_T = X_t \exp \left\{ \left(m - \rho \nu \frac{\mu - r}{\sigma} - \frac{\nu^2}{2} \right) (T - t) + \nu (W_{2,T} - W_{2,t}) \right\}$$

Now let's define the variable z

$$Z = \left\{ \left(m - \rho \nu \frac{\mu - r}{\sigma} - \frac{\nu^2}{2} \right) (T - t) + \nu Y \sqrt{T - t} \right\} \quad Y \sim N(0, 1)$$

We have

$$Z \sim N \left(\left(m - \rho \nu \frac{\mu - r}{\sigma} - \frac{\nu^2}{2} \right) (T - t), \nu^2 (T - t) \right)$$

The time- t discounted payoff of an European call option is given by

$$\begin{aligned} V(t, x) &= e^{-r(T-t)} E_t^* [h(X_T)] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} h(xe^z) f(z) dz \end{aligned}$$

For $h(X_T) = \max(X_T - K, 0)$ we get

$$\begin{aligned} V(t, x) &= e^{-r(T-t)} \int_{-\infty}^{\infty} \max(X_T - K, 0) f(z) dz \\ &= e^{-r(T-t)} \int_{\underline{u}}^{\infty} (xe^z - K) f(z) dz \end{aligned}$$

Where $\underline{u}$ is defined as

$$\underline{u} = \frac{\ln(\frac{K}{x}) - \left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2}\right)(T-t)}{\nu\sqrt{T-t}}$$

Consequently we can write $V(t, x)$ as

$$\begin{aligned} V(t, x) &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \left(x \int_{\underline{u}}^{\infty} e^{\left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2}\right)(T-t) + \nu\sqrt{T-t}y} e^{-\frac{y^2}{2}} dy - K \int_{\underline{u}}^{\infty} e^{-\frac{z^2}{2}} dz \right) \\ &= \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \left(x \int_{\underline{u}}^{\infty} e^{\left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2}\right)(T-t) + \nu\sqrt{T-t}y - \frac{y^2}{2}} dy \right) \\ &\quad - Ke^{-r(T-t)} \Phi\left(-\frac{\ln(\frac{K}{x}) - \left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2}\right)(T-t)}{\nu\sqrt{T-t}}\right) \\ &= \frac{1}{\sqrt{2\pi}} e^{\left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2} - r\right)(T-t)} \left(x \int_{\underline{u}}^{\infty} e^{\nu\sqrt{T-t}y - \frac{y^2}{2}} dy \right) \\ &\quad - Ke^{-r(T-t)} \Phi(d_2) \\ &= \frac{1}{\sqrt{2\pi}} e^{\left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2} - r\right)(T-t)} \left(x \int_{\underline{u}}^{\infty} e^{-\frac{1}{2}(y^2 - 2\nu\sqrt{T-t}y + \nu^2(T-t))} e^{\frac{1}{2}\nu^2(T-t)} dy \right) \\ &\quad - Ke^{-r(T-t)} \Phi(d_2) \\ &= \frac{1}{\sqrt{2\pi}} e^{\left(m - r - \rho\nu\frac{\mu-r}{\sigma}\right)(T-t)} \left(x \int_{\underline{u}}^{\infty} e^{-\frac{1}{2}(y - \nu\sqrt{T-t})^2} dy \right) - Ke^{-r(T-t)} \Phi(d_2) \\ &= xe^{\left(m - r - \rho\nu\frac{\mu-r}{\sigma}\right)(T-t)} \Phi(d_1) - Ke^{-r(T-t)} \Phi(d_2) \end{aligned}$$

Where d_1 and d_2 are defined as:

$$d_2 = \frac{\ln(\frac{x}{K}) + \left(m - \rho\nu\frac{\mu-r}{\sigma} - \frac{\nu^2}{2}\right)(T-t)}{\nu\sqrt{T-t}}$$

$$d_1 = d_2 + \nu\sqrt{T-t}$$

To derive the optimal hedging strategy we take the derivative of function $V(x, t)$ with respect to X_t .

$$\begin{aligned} \theta_t^* &= \frac{\rho\nu}{\sigma} X_t \frac{\partial E_t^*[(X_T - K)^+ e^{-r(T-t)}]}{\partial X_t} \\ &= \frac{\rho\nu}{\sigma} e^{\left(m-r-\rho\nu\frac{\mu-r}{\sigma}\right)(T-t)} \Phi(d_1) + \frac{\rho\nu}{\sigma} x e^{\left(m-r-\rho\nu\frac{\mu-r}{\sigma}\right)(T-t)} \frac{\partial \Phi(d_1)}{\partial X_t} \\ &\quad - \frac{\rho\nu}{\sigma} K e^{-r(T-t)} \frac{\partial \Phi(d_2)}{\partial X_t} \\ &= A_t \Phi(d_1) - A_t \frac{1}{\nu\sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} - \frac{\rho\nu}{\sigma} K e^{-r(T-t)} \frac{1}{\nu X_t \sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_2^2}{2}} \\ &= A_t \Phi(d_1) - A_t \frac{1}{\nu\sqrt{T-t}} \phi(d_1) + \frac{\rho\nu}{\sigma} K e^{-r(T-t)} \frac{1}{\nu X_t \sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(d_1 - \nu\sqrt{T-t})^2}{2}} \\ &= A_t \Phi(d_1) - A_t \frac{1}{\nu\sqrt{T-t}} \phi(d_1) + \frac{\rho\nu}{\sigma} K e^{-r(T-t)} \frac{1}{\nu X_t \sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2 - 2d_1\nu\sqrt{T-t} + \nu^2(T-t)}{2}} \\ &= A_t \Phi(d_1) - A_t \frac{1}{\nu\sqrt{T-t}} \phi(d_1) + \frac{\rho\nu}{\sigma} K e^{-r(T-t)} \frac{1}{\nu X_t \sqrt{T-t}} \phi(d_1) e^{-\frac{\nu^2(T-t)}{2}} e^{\ln(\frac{x}{K}) + (m - \rho\nu\frac{\mu-r}{\sigma} + \frac{\nu^2}{2})(T-t)} \\ &= A_t \Phi(d_1) - A_t \frac{1}{\nu\sqrt{T-t}} \phi(d_1) + \frac{\rho\nu}{\sigma} e^{(m-r-\rho\nu\frac{\mu-r}{\sigma})(T-t)} \frac{1}{\nu\sqrt{T-t}} \phi(d_1) \\ &= A_t \Phi(d_1) \end{aligned}$$

where A_t is defined as

$$A_t = \frac{\rho\nu}{\sigma} e^{\left(m-r-\rho\nu\frac{\mu-r}{\sigma}\right)(T-t)}$$

Bibliography

- Ankirchner, S., Dimitroff, G., Heyne, G., & Pigorsch, C. (2012). Futures cross-hedging with a stationary basis. *Journal of Financial and Quantitative Analysis*, 47(6), 1361–1395. https://EconPapers.repec.org/RePEc:cup:jfnqa:v:47:y:2012:i:06:p:1361-1395_00
- Ankirchner, S., & Imkeller, P. (2011). Hedging with residual risk: A bsde approach. In R. Dalang, M. Dozzi, & F. Russo (Eds.), *Seminar on stochastic analysis, random fields and applications vi* (pp. 311–325). Springer Basel.
- Ankirchner, S., Pigorsch, C., & Schweizer, N. (2014). Estimating residual hedging risk with least-squares monte carlo. *International Journal of Theoretical and Applied Finance*, 17(07), 1450042. <https://doi.org/10.1142/S0219024914500423>
- Ankirchner, S., Schneider, J. C., & Schweizer, N. (2014). Cross-hedging minimum return guarantees: Basis and liquidity risks. *Journal of Economic Dynamics and Control*, 41, 93–109. <https://doi.org/https://doi.org/10.1016/j.jedc.2014.02.010>
- Basak, S., & Chabakauri, G. (2010). Dynamic Mean-Variance Asset Allocation. *The Review of Financial Studies*, 23(8), 2970–3016. <https://doi.org/10.1093/rfs/hhq028>
- Basak, S., & Chabakauri, G. (2011). *Dynamic Hedging in Incomplete Markets: A Simple Solution* (CEPR Discussion Papers No. 8402). C.E.P.R. Discussion Papers. <https://ideas.repec.org/p/cpr/ceprdp/8402.html>
- Bjork, T., & Murgoci, A. (2014). A theory of markovian time-inconsistent stochastic control in discrete time. *Finance and Stochastics*, 18(3), 545–592. <https://EconPapers.repec.org/RePEc:spr:finsto:v:18:y:2014:i:3:p:545-592>
- Dai, M., Jin, H., Kou, S., & Xu, Y. (2021). A dynamic mean-variance analysis for log returns. *Management Science*, 67(2), 1093–1108. <https://EconPapers.repec.org/RePEc:inm:ormnsc:v:67:y:2021:i:2:p:1093-1108>
- Duffie, D., & Richardson, H. R. (1991). Mean-Variance Hedging in Continuous Time. *The Annals of Applied Probability*, 1(1), 1–15. <https://doi.org/10.1214/aoap/1177005978>
- Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276. Retrieved September 22, 2022, from <http://www.jstor.org/stable/1913236>
- Foellmer, H., Sondermann, H., & Sondermann, D. (1985). Hedging of non-redundant contingent claims. <https://doi.org/10.13140/RG.2.1.3298.8322>
- Hull, J. (2009). Options, futures and other derivatives. <https://books.google.nl/books?id=sEmQZoHoJCcC>
- Judd, K. (1999). Judd, kenneth l. (1998). numerical methods in economics. *Kyklos*, 52(1), 118–120. <https://doi.org/https://doi.org/10.1111/j.1467-6435.1999.tb02815.x>
- Leland, H. (1985). Option pricing and replication with transactions costs. *Journal of Finance*, 40(5), 1283–1301. <https://EconPapers.repec.org/RePEc:bla:jfnan:v:40:y:1985:i:5:p:1283-1301>

- Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies*, 14(1), 113–147. <https://doi.org/10.1093/rfs/14.1.113>
- Monoyios, M. (2004). Performance of utility-based strategies for hedging basis risk. *Quantitative Finance*, 4(3), 245–255. <https://doi.org/10.1088/1469-7688/4/3/001>
- Monoyios, M. (2007). Optimal hedging and parameter uncertainty. *IMA Journal of Management Mathematics*, 18(4), 331–351. <https://doi.org/10.1093/imaman/dpm022>
- Schweizer, M. (1994). Approximating random variables by stochastic integrals. *The Annals of Probability*, 22(3), 1536–1575. Retrieved September 23, 2022, from <http://www.jstor.org/stable/2245032>
- Whalley, A. E., & Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3), 307–324. <https://EconPapers.repec.org/RePEc:bla:mathfi:v:7:y:1997:i:3:p:307-324>