Mean–Variance Hedging in an Illiquid Market

Melusi Manqoba Mavuso

A dissertation submitted to the Department of Actuarial Science, Faculty of Commerce, at the University of the Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

October 10, 2014

Master of Philosophy specializing in Mathematical Finance,
University of the Cape Town,
Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy in the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

Signed by candidate

Signature removed

April 8, 2015

Abstract

Consider a market consisting of two correlated assets: one liquidly traded asset and one illiquid asset that can only be traded at time 0. For a European derivative written on the illiquid asset, we find a hedging strategy consisting of a constant (time 0) holding in the illiquid asset and dynamic trading strategies in the liquid asset and a riskless bank account that minimizes the expected square replication error at maturity. This mean–variance optimal strategy is first found when the liquidly traded asset is a local martingale under the real world probability measure through an application of the Kunita–Watanabe projection onto the space of attainable claims. The result is then extended to the case where the liquidly traded asset is a continuous square integrable semimartingale, and we again use the Kunita–Watanabe decomposition, now under the variance optimal martingale measure, to find the mean–variance optimal strategy in feedback form. In an example, we consider the case where the two assets are driven by correlated Brownian motions and the derivative is a call option on the illiquid asset. We use this example to compare the terminal hedging profit and loss of the optimal strategy to a corresponding strategy that does not use the static hedge in the illiquid asset and conclude that the use of the static hedge reduces the expected square replication error significantly (by up to 90% in some cases). We also give closed form expressions for the expected square replication error in terms of integrals of well–known special functions.

Acknowledgements

First, I would like to acknowledge the help and guidance from my supervisor A/Prof. Tom McWalter. Your suggestions were very helpful!

I would also like to thank my co-supervisors from Old Mutual, Mr Dylan Flint and Dr. Marchand van Rooyen, for structuring a problem that has enabled me to learn so much just from trying to solve it.

I would like to thank Mr Ralph Rudd for taking his time to proofread my dissertation and help me correct some errors.

I would also like to thank the UCT department of statistical sciences for the warm treatment throughout the year.

Finally, I would like to thank Mr Armin van Buuren for keeping me in 'a state of trance' during the long nights of typing this dissertation.

I dedicate this work to my friends, colleagues and everyone who was getting tired of hearing the phrase "my dissertation is essentially completed". Thank you for asking, it kept me going!

Thank you mom and grandma for the support you've given me throughout my academic career.

Contents

List of Figures

Chapter 1

Introduction

One of the recent major breakthroughs in the mathematics of finance is the idea that one can (in theory) eliminate the risk carried in writing an option by trading in the underlying security and a riskless asset. This idea of replication was first presented by Black and Scholes (1973) and Merton (1973) and was further studied by Harrison and Kreps (1979), Kreps (1981), Harrison and Pliska (1981) and Harrison and Pliska (1983) in their seminal papers of the late 70s and early 80s.

However, in order for such a perfect replication to be possible, many assumptions are made. Some of these include continuous trading, absence of transaction costs and unlimited short selling. Relaxing these assumptions has been the subject of enormous research in the last three decades and substantial progress has been made. This dissertation focuses on solving a problem where continuous trading is restricted to one of two assets.

1.1 Problem Formulation

Consider a company (e.g. a bank or insurance company) that has an obligation to make a payment at some future time $T > 0$ represented by the random variable H. Such a random payment could, for instance, be an option on a share or an annuity to policyholders. Ideally, the company would want to construct a portfolio of assets to trade that will have the same value as H at time T, under all circumstances. If this is not possible, the requirement might be relaxed to finding a portfolio of assets whose value is as 'close as possible' to H at time T.

To put this into perspective, assume that H is an option written on a stock whose (discounted) price process is denoted by $\bar{U}$. Under certain models for $\bar{U}$ (and integrability conditions on H), the theory of Black and Scholes (1973) shows that it is possible to construct a dynamic trading strategy in $\bar{U}$ and the riskless bank account that will replicate H under all circumstances. This replication process makes use of all the 'frictionless' market assumptions of Harrison and Pliska (1983) and our aim is to relax the assumption of continuous trading in $\bar{U}$.

We consider the case when the asset $\bar{U}$ is not liquid and continuous trading in $\bar{U}$ is thus not possible. We further assume that, together with a riskless bank account, there is another asset X that is correlated to $\bar{U}$, which can be traded continuously without any restrictions (so that all the assumptions of Black and Scholes (1973) hold for trading the asset X). The question arises: how does one utilize the ability to trade partly in the illiquid asset $\bar{U}$ and the liquid asset X together with the riskless asset to reduce the risk of writing (or even holding) the derivative H? Such a hedging strategy will generally not replicate H, except in trivial cases such as when X and $\bar{U}$ are perfectly correlated. So, in general, we can only hope to make the portfolio value as 'close as possible' to H at time T.

At this stage the problem is still not well–defined since, for instance, there are many ways in which the asset $\bar{U}$ may be illiquid. Also, we need to make precise what it means to make the portfolio value as 'close as possible' to H at time T.

One interpretation of $\bar{U}$ being illiquid might be that trading in $\bar{U}$ is not possible at all times. In that case we have to construct the 'best' strategy using only asset X (and the riskless bank account). This problem has been studied extensively in the literature under the topic of *Hedging in an Incomplete Market.* Some references in that direction include Schweizer (1991), Duffie and Richardson (1991), Schweizer (1992), Davis (2006) and Monoyios (2004). A specific example of this *hedging of basis risk* problem is also considered by McWalter (2007) and Hulley and McWalter (2008) in the case when X and $\bar{U}$ are correlated geometric Brownian motions.

The aim of this dissertation is to consider a slight change to the above by assuming that the asset $\bar{U}$ can be traded only once — at the inception of the contract. Thus, any 'admissible' trading strategy will consist of a static hedge in $\bar{U}$ and a dynamic hedge in both X and the bank account.

As mentioned before, the new trading strategy described above will generally not replicate the payoff H. We therefore need a criteria by which to choose the 'best' trading strategy. An obvious way of doing this is to define a utility function and then choose a strategy that maximizes the expected utility. This type of reasoning has been followed by (among others) Davis (2006).

In this dissertation we choose to *minimize* a certain kind of quadratic utility functional. To be more precise, if V_T is the value of a trading strategy at time T, we will choose a strategy that minimizes the *expected square replication error*:

$$\mathbb{E}((V_T - H)^2)$$

over all admissible trading strategies. This is known as *mean–variance hedging* and was first considered by Föllmer and Sondermann (1986) in the case when X is a martingale under the real world probability measure $\mathbb{P}$. It was later generalized by (among others) Duffie and Richardson (1991) in the Brownian motion setting and by Schweizer (1988), Föllmer and Schweizer (1991), Schweizer (1991), Schweizer (1992), Rheinlander and Schweizer (1997) and Gouriéroux *et al.* (1998) to the case when X is a semimartingale under $\mathbb{P}$. This criterion is also used by McWalter (2007) and Hulley and McWalter (2008) in the context of a two dimensional Brownian motion model for X and $\bar{U}$, which provided us with the initial motivation to extend their results.

To formulate the problem mathematically, we want to find a constant holding $\hat{\theta}$ in $\bar{U}$, a dynamic trading strategy $\hat{\varphi} = \{\hat{\varphi}_t : 0 \leq t \leq T\}$ in X and a dynamic strategy $\hat{\eta} = \{\hat{\eta}_t : 0 \leq t \leq T\}$ in the bank account B that will minimize

$$\mathbb{E}((V_T - H)^2) = \mathbb{E}((\theta \bar{U}_T + \varphi_T X_T + \eta_T B_T - H)^2)$$

over all such trading strategies (θ, φ, η). Without loss of generality, we will work with discounted assets and set $B \equiv 1$. The problem is then to minimize

$$\mathbb{E}((V_T - H)^2) = \mathbb{E}((\theta \bar{U}_T + \varphi_T X_T + \eta_T - H)^2)$$

over all such trading strategies (θ, φ, η). Also, confining ourselves to self–financing strategies, the problem simplifies to minimizing

$$\mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t + \theta \bar{U}_T - H\right)^2\right)$$

over all strategies (θ, φ, v_0), where v_0 is the initial capital required for the dynamic strategy in X and the bank account. The dependence on η is dropped due to the self–financing condition.

If $H \in L^2$ and $V_T \in L^2$ for each admissible strategy (θ, φ, v_0), then this problem is equivalent to finding an L^2 projection of H onto the space of attainable claims; therefore if this space is closed in L^2, a unique solution will exist by the *Projection Theorem*. Whether or not this space is closed turns out to be very different when X is a local martingale to when X is only a semimartingale under $\mathbb{P}$. The local martingale case is easy and transparent — in that case the space of attainable claims is closed in L^2 by the Itô isometry. It is for this reason that we first consider the martingale case.

In both cases, we use ideas from the solution of hedging in an incomplete market with no trading constraints. The solution to this problem was found by Föllmer and Sondermann (1986) when X is a square integrable martingale and later generalized by Schweizer (2001) to when X is a local martingale. For the semimartingale case, more assumptions on X are required for the space of attainable claims to be closed. However in both cases, the *Kunita–Watanabe decomposition* proves to be vital in solving the problem. When X is a local martingale, this decomposition is used to find a projection onto the space of stochastic integrals with respect to X, while after a change of measure to the *variance–optimal martingale measure*, the same projection is used when X is only a semimartingale.

In general, even the optimal strategy will not necessarily replicate the payoff exactly. Another aim of this dissertation is to derive closed form expressions for the expected square replication error. This quantity is normally found using simulations in the literature (e.g. see Heath *et al.* (2001b) and Hulley and McWalter (2008)), but we manage to write integral expressions of this quantity in terms of well–known special functions for Brownian motion stocks and when H is a call option on $\bar{U}$.

1.2 The Structure of the Dissertation

The dissertation is divided into five chapters. The second chapter gives a summary of the main definitions, notation and preliminary results that we use throughout the dissertation. This is done in a quick way and proofs are not given. The reader is referred to relevant sources in the literature. This chapter can be skimmed through by the well informed reader.

In Chapter 3 we solve the problem stated above in the special case when the liquidly traded asset X is a local martingale under $\mathbb{P}$. We start by discussing general mean–variance hedging in an incomplete market when X is a local martingale and make a presentation similar to Föllmer and Sondermann (1986) and Schweizer (2001). We then use a similar argument to solve the illiquid market hedging problem. The solution turns out to be a simple application of the projection theorem to an appropriate closed subspace. We also consider an example when X and $\bar{U}$ are driven by correlated Brownian motion processes and H is a call option on $\bar{U}$. In this example we derive closed–form expressions for the expected square replication error in terms of well–known special functions.

The main purpose of Chapter 4 is to extend the results of Chapter 3 to the case when X is a continuous semimartingale. This chapter is very similar in style to Chapter 3. We again start off by discussing mean–variance hedging in an incomplete market when X is a semimartingale. We follow presentations by Gouriéroux *et al.* (1998) and Heath *et al.* (2001a) and observe that different assumptions on X guarantee the closedness of the space of attainable claims. We then apply these results to find a solution to our problem by first fixing the constant holding in $\bar{U}$ to be θ, treating the problem as an incomplete market hedging problem of a modified claim and then finding the optimal value of θ. These results are then applied to a two–dimensional geometric Brownian motion model for X and $\bar{U}$ similar to the one used in Chapter 3. Closed–form expressions for the expected square replication error are found and we also provide some numerical results for specific model parameters.

The last chapter concludes by giving some possible extensions of these results and discussing some comparisons. This dissertation should be accessible to a graduate student with a background in stochastic calculus and elementary mathematical finance.

Chapter 2

Preliminaries

This chapter presents definitions and results that will be used throughout the dissertation. It is assumed that the reader is familiar with these concepts and no proofs will be provided; instead we make reference to relevant papers and textbooks.

2.1 Hilbert Spaces

In this section we quickly recall some important results and definitions in Hilbert spaces that will be used throughout. We simply state the main results without proof and refer the reader to any standard textbook on analysis for any proofs. Some examples include Brezis (2011), Zeidler (1995) and Conway (1990). We assume that the reader is familiar with basic analysis in metric and normed spaces. We also assume that all vector spaces considered are over the field $\mathbb{R}$ of real numbers. Both the zero vector and the number zero are denoted by 0.

2.1.1 Elementary Results and Definitions

Definition 2.1.1. *Let V be a vector space. An* ***inner product*** *on V is a function $(\cdot,\cdot) : V \times V \longrightarrow \mathbb{R}$ such that for every $u, v, w \in V$ and $\alpha \in \mathbb{R}$,*

1. $(u, u) \geq 0$ *and* $(u, u) = 0$ *if and only if* $u = 0$
2. $(u, v) = (v, u)$
3. $(\alpha u, v) = \alpha\,(u, v)$
4. $(u + v, w) = (u, w) + (v, w)$.

The pair $(V, (\cdot,\cdot))$ is called an ***inner product space*** *or* ***pre-Hilbert space****.*

We will often refer to V (on its own) as an inner product space if the inner product is clear from the context.

Theorem 2.1.1. *If $(V, (\cdot,\cdot))$ is an inner product space, then for every $u, v \in V$*

$$|(u, v)\,|^2 \leq (u, u)\,(v, v)\,. \tag{2.1}$$

Inequality (2.1) is known as the *Cauchy–Schwartz inequality* and is used to show the following:

Theorem 2.1.2. *If $(V, (\cdot,\cdot))$ is an inner product space, then the function $||\cdot||: V \longrightarrow \mathbb{R}$ defined by*

$$||v|| := \sqrt{(v, v)} \quad \text{for every } v \in V \tag{2.2}$$

is a norm on V. This makes $(V, ||\cdot||)$ a normed space and a metric space under the metric d defined by

$$d(u, v) := ||u - v|| \quad \text{for every } u, v \in V.$$

A sequence (v_n) in a normed space V is a function $v : \mathbb{N}^+ \longrightarrow V$. We say that (v_n) converges to $v \in V$ if for every $\epsilon > 0$ there exists $N \in \mathbb{N}^+$ such that for every $n \geq N$, we have $||v_n - v|| < \epsilon$. We will write $v_n \longrightarrow v$ or $\lim_{n \to \infty} v_n = v$ to say that (v_n) converges to v.

Definition 2.1.2. *A sequence (v_n) in a normed space V is a* ***Cauchy sequence*** *if for every $\epsilon > 0$ there exist $N \in \mathbb{N}^+$ such that for every $n, m \in \mathbb{N}^+$, if $n, m \geq N$, then $||v_n - v_m|| < \epsilon$.*

It is clear that every convergent sequence is Cauchy. Indeed, assume that (v_n) converges to v and let $\epsilon > 0$. Choose $N \in \mathbb{N}^+$ such that if $n \geq N$, then $||v_n - v|| < \frac{\epsilon}{2}$. Now if $n, m \geq N$, then $||v_n - v_m|| = ||v_n - v + v - v_m|| \leq ||v_n - v|| + ||v_m - v|| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. The converse, however, does not hold in general.

Definition 2.1.3. *An inner product space V is* ***complete*** *if for every Cauchy sequence (v_n) in V, there exists $v \in V$ such that $v_n \longrightarrow v$; i.e. V is complete if every Cauchy sequence in V converges to an element of V. A subset U of V is complete if every Cauchy sequence in U converges to an element of U.*

A complete inner product space is called a ***Hilbert space****.*

Theorem 2.1.3. *If V is a finite dimensional inner product space, then V is a Hilbert space; i.e. every finite dimensional inner product space is complete.*

An analogous result is also true for normed spaces in general; complete normed spaces are called **Banach spaces**.

Definition 2.1.4. *A subset U of a normed space V is* ***closed*** *in V (or simply closed if the normed space is clear from the context) if every convergent sequence in U has a limit in U; i.e. U is closed if it satisfies the condition that if (v_n) is a sequence with $v_n \in U$ for every $n \in \mathbb{N}^+$ and $v_n \longrightarrow v \in V$, then $v \in U$.*

Theorem 2.1.4. *If V is a normed space and $U \subseteq V$ is a finite dimensional subspace of V, then U is closed; i.e. every finite dimensional subspace of a normed space V is closed.*

Theorem 2.1.5. *If V is a Hilbert space, then $U \subseteq V$ is closed if and only if U is complete.*

Let V and W be vector spaces. Recall that a map $T : V \longrightarrow W$ is linear if

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) \text{ for every } u, v \in V \text{ and } \alpha, \beta \in \mathbb{R}.$$

Definition 2.1.5. *Let $(V, ||\cdot||_V)$ and $(W, ||\cdot||_W)$ be normed spaces and $T : V \longrightarrow W$ be a linear map.*

1. *We call T an* ***isomorphism*** *if T is one to one and onto; i.e. bijective.*
2. *We call T an* ***isometry*** *if*

$$||T(v)||_W = ||v||_V \quad \textit{for every } v \in V.$$

We say that T is an ***isometric isomorphism*** *if T is both an isomorphism and an isometry.*

If there exist an isometric isomorphism $T : V \longrightarrow W$, then the normed spaces V and W are said to be **isometrically isomorphic**. Normed spaces that are isometrically isomorphic are essentially the same as normed spaces.

Theorem 2.1.6. *Let V and W be inner product spaces (and therefore normed spaces) that are isometrically isomorphic. If V is a Hilbert space, then so is W and conversely.*

Again, this is also true for Banach spaces.

2.1.2 Direct Sums, Orthogonality and Projections

If V is a vector space and U_1, U_2 are subsets of V, then we define $U_1 + U_2$ as

$$U_1 + U_2 := \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}.$$

Definition 2.1.6. *Let V be a vector space, U_1 and U_2 be subspaces (non-empty, closed under vector addition and scalar multiplication) of V. Then*

1. *V is a **sum** of U_1 and U_2 (written as $V = U_1 + U_2$) if for every $v \in V$ there exist $u_1 \in U_1$ and $u_2 \in U_2$ such that $v = u_1 + u_2$.*
2. *V is a **direct sum** of U_1 and U_2 (written as $V = U_1 \oplus U_2$) if for every $v \in V$ there exist unique $u_1 \in U_1$ and $u_2 \in U_2$ such that $v = u_1 + u_2$.*

Theorem 2.1.7. *Let V be a vector space and U_1 and U_2 be subspaces of V. The following are equivalent:*

1. *$V = U_1 \oplus U_2$*
2. *$V = U_1 + U_2$ and $U_1 \cap U_2 = \{0\}$.*

One question that we need to answer for later chapters is the following: if U_1 and U_2 are closed subspaces of a Hilbert space V, is $U_1 + U_2$ closed? It turns out that this is not true in general unless at least one of them is finite dimensional.

Theorem 2.1.8. *If U_1 and U_2 are closed subspaces of a Hilbert space V and U_1 (or U_2) is finite dimensional, then $U_1 + U_2$ is closed.*

Definition 2.1.7. *Let V be an inner product space. We say that $u, v \in V$ are **orthogonal** (written as $u \perp v$) if $(u, v) = 0$. If $U \subseteq V$ we define the **orthogonal compliment** of U by*

$$U^\perp := \{v \in V : (u, v) = 0 \textit{ for all } u \in U\} = \{v \in V : v \perp u \textit{ for all } u \in U\}.$$

Theorem 2.1.9. *Let U be a subset of an inner product space V. Then $U^\perp$ is a closed subspace of V.*

We are now in a position to solve the following problem: Given a Hilbert space V, a subset U of V and an element $v \in V$, find $u \in U$ that best approximates v in U. That is, we want to find $u \in U$ such that

$$||v - u|| = \inf_{w \in U} ||v - w||.$$

Such an element u, if it exists, is called the **projection** of v onto U. This problem is more interesting of course when $v \in V \backslash U$, since we may choose $u = v$ if $v \in U$. Here u is the best estimator of v in U, where 'best' is measured in terms of minimizing the norm. The question of existence (and uniqueness) of such a best estimator for an arbitrary $v \in V \backslash U$ depends on both the topology and geometry of U.

If V is a vector space, then $U \subseteq V$ is convex if for every $u, v \in U$ and $\lambda \in [0, 1]$, we have $\lambda u + (1 - \lambda)v \in U$. Subspaces provide good examples of convex subsets since they are closed addition and scalar multiplication.

Theorem 2.1.10. *Let V be a Hilbert space and $U \subseteq V$ be a non–empty closed and convex subset of V. Then for every $v \in V$, there exist a unique $u \in U$ such that*

$$||v - u|| = \inf_{w \in U} ||v - w||.$$

Furthermore, u is the unique element of U that satisfies the following (variational) inequality

$$(v - u, w - u) \leq 0 \quad \textit{for all } w \in U.$$

The above theorem gives a sufficient condition for a projection on a subset U of V to exist for every element. An important special case is when U is actually a closed *subspace* of a V. In that case we get the celebrated projection theorem.

Theorem 2.1.11. *Let U be a closed subspace of a Hilbert space V.*

1. *For every $v \in V$, there exists a unique projection u of v onto U. Furthermore, u is such that $v - u \in U^{\perp}$ (i.e. $(v-u, w) = 0$ for every $w \in U$).*
2. *$V = U \oplus U^{\perp}$; i.e. for every $v \in V$, there exist unique elements $u \in U$ and $u^{\perp} \in U^{\perp}$ such that $v = u + u^{\perp}$.*

The projection theorem tells us that if U is a closed subspace of a Hilbert space V, then each $v \in V$ can be associated with a unique $P_U(v) := u \in U$ such that

$$||v - P_U(v)||= \inf_{w \in U} ||v - w||.$$

This defines a map $P_U : V \longrightarrow U \subseteq V$. This map is called the **orthogonal projection onto** U and it has the following properties:

1. P_U is linear; i.e. $P_U(\alpha v_1 + \beta v_2) = \alpha P_U(v_1) + \beta P_U(v_2)$ for every $v_1, v_2 \in V$ and $\alpha, \beta \in \mathbb{R}$
2. P_U is continuous; i.e. If $v_n \longrightarrow v$ in V, then $P_U(v_n) \longrightarrow P(v)$
3. P_U is non-expansive; i.e. $||P_U(v)||\leq ||v||$ for every $v \in V$
4. P_U is idempotent; i.e. $P_U(P_U(v)) = P_U(v)$ for every $v \in V$.

The *orthogonal decomposition* of $v \in V$ as $v = P_U(v) + (v - P_U(v))$, where $P_U(v)$ and $v - P_U(v)$ are orthogonal, is at the heart of this dissertation.

2.2 Stochastic Processes

Much of what we will be doing deals with stochastic processes. In this short section we give brief definitions and important results in stochastic processes and stochastic calculus. The main references used are Protter (2004), Rheinlander (2011), Karatzas and Shreve (1991), Jacod and Shiryaev (2003) and Durrett (1996). We again assume that the reader is familiar with all the theory of stochastic processes, including probability and measure theory, together with modes of convergence of sequences of random variables.

2.2.1 Stochastic Processes

Throughout we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $1 \leq p < \infty$ we denote by $L^p(\mathbb{P})$ the vector space of equivalence classes of random variables X with $\mathbb{E}(|X|^p) < \infty$. With the understanding that we identify random variables that agree almost surely (a.s.), we will simply think of $L^p(\mathbb{P})$ as a space of random variables. We will also write L^p if the measure $\mathbb{P}$ is clear from the context.

Theorem 2.2.1. *For $1 \leq p < \infty$, the space L^p is a Banach space with norm $||X||_{L^p}= (\mathbb{E}(|X|^p))^{1/p}$ and L^2 is a Hilbert space with inner product defined by*

$$(X, Y) := \mathbb{E}(XY) \text{ for every } X, Y \in L^2.$$

A **stochastic process** X on $(\Omega, \mathcal{F}, \mathbb{P})$ is an indexed collection of random variables. We will deal only with stochastic processes that are indexed by time in some interval $I \subseteq \mathbb{R}^+ := [0, \infty)$. For now we set $I = \mathbb{R}^+$. Then

$$X := \{X_t : t \in \mathbb{R}^+\},$$

where each X_t is a random variable. If we fix $\omega \in \Omega$, then $\{X_t(\omega) : t \in \mathbb{R}^+\}$ is called a **sample path** of X. Two stochastic processes X and Y are **indistinguishable** if

$$\mathbb{P}\left(\left\{\omega \in \Omega : X_t(\omega) = Y_t(\omega)\ t \in \mathbb{R}^+\right\}\right) = 1.$$

We will use this to define equality of stochastic processes and write $X = Y$ if and only if X and Y are indistinguishable. A **filtration** $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{R}^+\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing collection of sub–sigma algebras of $\mathcal{F}$. We will assume that $\mathbb{F}$ always satisfies the following **usual conditions** of continuity and completeness:

1. $\mathcal{F}_0$ contains all the $\mathbb{P}$–null sets of $\mathcal{F}$ (this is called completeness)
2. $\mathbb{F}$ is right continuous in the sense that for every $t \in \mathbb{R}^+$,

$$\mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t.$$

For all our applications it will be convenient to further assume that $\mathcal{F}_0$ is *trivial* in the sense that it contains only events of probability zero or one. We then call $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ a **filtered probability space**. X is said to be **adapted** to $\mathbb{F}$ (or simply adapted, if $\mathbb{F}$ is known) if X_t is $\mathcal{F}_t$–measurable for each t. All processes we deal with will be assumed to be adapted unless explicitly stated otherwise. We denote by $\mathbb{F}^X$ the **filtration generated** by X. That is $\mathbb{F}^X = \sigma(X)$.

The process X is said to be **continuous** if almost all the sample paths of X are a continuous function of $t \in \mathbb{R}^+$ and X is **càdlàg** (resp. **càglàd**) if X is right continuous with finite left limits (resp. left continuous with finite right limits). For a càdlàg process X, we will define the process $X_- := \{X_{t^-} : t \in \mathbb{R}^+\}$ by

$$X_{t^-} := \lim_{s \to t^-} X_s$$

and the **jump process** $\Delta X = X - X_-$ with $\Delta X_0 = 0$. We note that if X is continuous then $\Delta X \equiv 0$.

A random variable $\tau : \Omega \longrightarrow [0, \infty]$ is a **stopping time** if $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \in \mathbb{R}^+$. If τ is a stopping time, then the process X^τ defined by setting $X_t^\tau := X_{\tau \wedge t} = X_{\min(\tau,t)}$ is called the **stopped process.** If $\mathcal{E}$ is a set of stochastic processes, X is said to be **locally** in $\mathcal{E}$ if there exist an increasing sequence of stopping times (τ_n) such that $\tau \to \infty$ $\mathbb{P}$–a.s. and $X^{\tau_n} \in \mathcal{E}$ for every positive integer n. The sequence (τ_n) is called a **localizing sequence** for X and the set of all processes that are locally in $\mathcal{E}$ is denoted by $\mathcal{E}_{\text{loc}}$. We say that $\mathcal{E}$ is **stable under stopping** if $X \in \mathcal{E} \Rightarrow X^\tau \in \mathcal{E}$ for any stopping time τ.

Let $\mathbb{L}$ be the collection of all adapted càglàd (left continuous with right limits) processes on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. The sigma field generated by elements of $\mathbb{L}$ is called the **predictable sigma field**. This is a sigma field on $\Omega \times \mathbb{R}^+$ and is denoted by $\mathcal{P}$. That is

$$\mathcal{P} := \sigma(\mathbb{L}).$$

It is shown in Revuz and Yor (1999) that if $\mathbb{F}$ satisfies the usual conditions (which is what we will always assume), then $\mathcal{P}$ is also generated by the continuous adapted processes. A process

X is said to be **predictable** if the map $(\omega, t) \mapsto X_t(\omega)$ from $\Omega \times \mathbb{R}^+$ to $\mathbb{R}$ is $\mathcal{P}$–measurable.

Now let $t \in \mathbb{R}^+$ and consider a partition of $[0, t]$: $\pi_t^k := \{t_0, t_1, \ldots, t_k\}$, where k is a positive integer and $0 = t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_k = t$. We define the **total variation process** $V^{(1)}(X)$ of X as (supremum over all partitions π_t^k of $[0, t]$)

$$V_t^{(1)}(X) := \sup_{\pi_t^k} \left(\sum_{i=1}^{k} |X_{t_i} - X_{t_{i-1}}| \right).$$

If $V_t^{(1)}(X) < \infty$ a.s. for every t, we say X is of **finite variation**. We will also write the total variation of X as

$$V_t^{(1)}(X) := \int_0^t |dX|_s.$$

If X is of finite variation, we can define the integral of a locally bounded process φ with respect to X as a path-wise Lebesgue–Stieltjes integral. We denote this integral by

$$(\varphi \cdot X)_t := \int_0^t \varphi_s \, dX_s,$$

so that

$$(\varphi \cdot X)_t(\omega) := \int_0^t \varphi_s(\omega) \, dX_s(\omega) \quad \text{for every } \omega \in \Omega.$$

Definition 2.2.1. *A process M is a* ***martingale*** *with respect to $\mathbb{F}$ if*

1. *M is adapted to $\mathbb{F}$*
2. *$\mathbb{E}(|M_t|) < \infty$ for every $t \in \mathbb{R}^+$; i.e. M_t is integrable for every $t \in \mathbb{R}^+$*
3. *$\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ for $s < t$.*

We say that M is a ***submartingale*** *(resp.* ***supermartingale****) if it satisfies 1 and 2 and $\mathbb{E}(M_t|\mathcal{F}_s) \geq M_s$ (resp. $\mathbb{E}(M_t|\mathcal{F}_s) \leq M_s$) for $s < t$.*

M is a **local martingale** if there exist an increasing sequence of stopping times (τ_n) with $\tau_n \to \infty$ such that the stopped process $M^{\tau_n} = \{M_{\tau_n \wedge t} : t \in \mathbb{R}^+\}$ is a martingale for each $n \in \mathbb{N}^+$. Local martingales are generally not of finite variation, thus the integral with respect to a local martingale cannot be defined as a path–wise Lebesgue–Stieltjes integral. The construction of this *stochastic integral* is beyond the scope of this presentation and is discussed by many textbooks (see for example Protter (2004) or Jacod and Shiryaev (2003)). We will not go through this construction but simply mention that this 'stochastic integral' with respect to a local martingale M is well defined for predictable processes φ and is also a local martingale. It will also be denoted by

$$(\varphi \cdot M)_s := \int_0^t \varphi_s \, dM_s.$$

We define a **semimartingale** to be a process X that can be written as

$$X = X_0 + M + A,$$

where X_0 is finite, M is a local martingale and A is an adapted process of finite variation, both null at 0 (i.e. their values are zero at $t = 0$). If A is also predictable, then X is called a **special semimartingale**. We will denote the vector space of semimartingales by $\mathcal{S}$ and the space of special semimartingales by $\mathcal{S}_p$. The stochastic integral of a locally bounded predictable process φ with respect to a semimartingale X is defined as

$$(\varphi \cdot X)_t := (\varphi \cdot M)_t + (\varphi \cdot A)_t$$

i.e.

$$\int_0^t \varphi_s \, dX_s := \int_0^t \varphi_s \, dM_s + \int_0^t \varphi_s \, dA_s. \tag{2.3}$$

We will sometimes write (2.3) in differential notation as

$$d(\varphi \cdot X)_t = \varphi_t \, dX_t = \varphi_t \, dM_t + \varphi_t \, dA_t.$$

We say φ is X–**integrable** if $(\varphi \cdot X)$ is finite. The space of X–integrable processes is denoted by $L(X)$. For $\varphi \in L(X)$, the process $(\varphi \cdot X)$ is also a semimartingale. The stochastic integral can also be generalized to d–dimensional processes. For further details, see Protter (2004).

If X and Y are martingales, we define the **cross variation** process of X and Y as

$$\langle X, Y \rangle_t := \sup_{\pi_t^k} \left(\sum_{i=1}^{k} \left(X_{t_i} - X_{t_{i-1}} \right) \left(Y_{t_i} - Y_{t_{i-1}} \right) \right),$$

where the supremum is taken over all partitions π_t^k of $[0, t]$. We call $\langle X \rangle = \langle X, X \rangle$ the **quadratic variation** of X. We will also define the cross variation and quadratic variation this way when X and Y are *continuous* semimartingales.

Theorem 2.2.2. *Let X and Y be continuous semimartingales and $\varphi \in L(X)$, $\eta \in L(Y)$ be continuous. Then*

$$\left\langle \int_0^t \varphi_s \, dX_s, \int_0^t \eta_s \, dY_s \right\rangle = \int_0^t \varphi_s \eta_s \, d \langle X, Y \rangle_s .$$

In particular, this result tells us that

$$\left\langle \int_0^t \varphi_s \, dX_s \right\rangle = \int_0^t \varphi_s^2 \, d \langle X \rangle_s .$$

We also define the **stochastic exponential** of a continuous semimartingale X as

$$\mathcal{E}(X)_t := \exp\left(X_t - \frac{1}{2} \langle X \rangle_t \right).$$

We end this section with a very useful formula for continuous semimartingales. The following is Theorem 3.3 of Revuz and Yor (1999).

Theorem 2.2.3 (Itô's formula)**.** *Let $X = (X^1, \ldots, X^d)$ be a d–dimensional continuous semimartingale and $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Then $f(X)$ is also a continuous semimartingale and*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s) \, dX_s^i + \frac{1}{2} \sum_{j=1}^{d} \sum_{i=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \, d \left\langle X^i, X^j \right\rangle_s .$$

In differential notation, this can be written as

$$df(X_t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(X_s) \, dX_s^i + \frac{1}{2} \sum_{j=1}^{d} \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \, d \left\langle X^i, X^j \right\rangle_s .$$

2.2.2 Square Integrable Martingales

A process X is **uniformly integrable** (UI) if

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}^+} \int_{|X_t|\geq n} X_t \, d\mathbb{P} = 0.$$

We will denote the space of UI martingales by $\mathcal{M}$. A useful result about UI martingales is the following:

Theorem 2.2.4. *If $M \in \mathcal{M}$, then there exist an integrable random variable M_∞ such that*

$$\lim_{t\to\infty} M_t = M_\infty \quad \textit{almost surely.}$$

Furthermore, $M_t = \mathbb{E}(M_\infty|\mathcal{F}_t)$ for every $t > 0$.

If $p \geq 1$, we say that X is bounded in L^p if $\sup_{t\in\mathbb{R}^+} \mathbb{E}(|X_t|^p) < \infty$. We will mainly be concerned with $p = 2$.

Definition 2.2.2. *A martingale M is **square integrable** if M is bounded in L^2. The vector space of square integrable martingales is denoted by $\mathcal{M}^2$.*

Theorem 2.2.5. *If $M \in \mathcal{M}^2$ then $M \in \mathcal{M}$ and there exist $M_\infty \in L^2$ such that*

$$\lim_{t\to\infty} M_t = M_\infty \quad \textit{almost surely and in } L^2.$$

Furthermore, $M_t = \mathbb{E}(M_\infty|\mathcal{F}_t)$ for every $t > 0$.

It is easy to see that if $H \in L^2$, then the process M defined by $M_t := \mathbb{E}(H|\mathcal{F}_t)$ is a square integrable martingale. Thus each $M \in \mathcal{M}^2$ can be associated with a unique $H \in L^2$. One can then define an inner product on $\mathcal{M}^2$ by

$$(M, N) := \mathbb{E}(M_\infty N_\infty) \quad \text{for every } M, N \in \mathcal{M}^2.$$

It then follows that the mapping $M \mapsto M_\infty$ is an isometry from the inner product space $\mathcal{M}^2$ to the Hilbert space L^2. By Theorem 2.1.6, we get that $\mathcal{M}^2$ is also a Hilbert space.

Theorem 2.2.6. *$\mathcal{M}^2$ is a Hilbert space with the inner product*

$$(M, N) := \mathbb{E}(M_\infty N_\infty) \quad \textit{for every } M, N \in \mathcal{M}^2,$$

and norm

$$||M|| := \sqrt{\mathbb{E}(M_\infty^2)} \quad \textit{for every } M \in \mathcal{M}^2.$$

From now onwards, we will work on the index set $I = [0, T]$ for some $T > 0$ and set $M_T := M_\infty$ for any $M \in \mathcal{M}$. We define $\mathcal{M}_{\text{loc}}$ to be the space of local martingales. If $M \in \mathcal{M}_{\text{loc}}$, define $L^2(M)$ to be the set of all predictable processes $\varphi \in L(M)$ such that

$$\mathbb{E}\left(\int_0^T \varphi_t^2 \, d\langle M\rangle_t\right) < \infty.$$

We have the Itô–isometry.

Theorem 2.2.7 (Itô isometry)**.** *For every $M \in \mathcal{M}_{\text{loc}}$ and $\varphi \in L^2(M)$, the stochastic integral $(\varphi \cdot M)$ is square integrable (i.e. $(\varphi \cdot M) \in \mathcal{M}^2$) and*

$$\mathbb{E}\left(\left(\int_0^T \varphi_t \, dM_t\right)^2\right) = \mathbb{E}\left(\int_0^T \varphi_t^2 \, d\langle M\rangle_t\right).$$

So we can define a norm on $L^2(M)$ by

$$||\varphi||_M^2 := \mathbb{E}\left(\int_0^T \varphi_t^2 \, d\langle M\rangle_t\right).$$

The following is Lemma 2.1 of Schweizer (2001).

Theorem 2.2.8. *If $M \in \mathcal{M}_{\text{loc}}$, then the space*

$$\mathcal{S}(M) := \{(\varphi \cdot M) : \varphi \in L^2(M)\}$$

is a closed subspace of $\mathcal{M}^2$. Furthermore, $\mathcal{S}(M)$ is stable under stopping.

A result of this form was proved by Kunita and Watanabe (1967) when M is square integrable. The remarkable feature here is that M is only assumed to be a local martingale. See Schweizer (2001) for an elaboration of this.
The space $\mathcal{S}(M)$ is called the **stable subspace** generated by M. Since we know that $\mathcal{S}(M)$ is a closed subspace of $\mathcal{M}^2$, by the projection theorem, we can write $\mathcal{M}^2$ as

$$\mathcal{M}^2 = \mathcal{S}(M) \bigoplus \mathcal{S}(M)^\perp.$$

This kind of 'decomposition' is used extensively in the dissertation. Before we state it as a theorem, we need some terminology.

Definition 2.2.3. *Two martingales $M, N \in \mathcal{M}^2$ are **strongly orthogonal** if $\langle M, N\rangle = 0$.*

We will now state the well–known Kunita–Watanabe decomposition of a square integrable martingale. We mention that there are many versions and generalizations of this powerful result. We state the most general version for our purposes.

Theorem 2.2.9 (Kunita–Watanabe decomposition)**.** *Let $N \in \mathcal{M}^2$ and $M \in \mathcal{M}_{\text{loc}}$. Then there exist unique $\varphi^N \in L^2(M)$ and $L^N \in \mathcal{S}(M)^\perp \subseteq \mathcal{M}^2$ such that*

$$N_t = N_0 + \int_0^t \varphi_s^N \, dM_s + L_t^N \quad 0 \le t \le T.$$

Furthermore, L^N is null at zero ($L_0^N = 0$) and strongly orthogonal to every element of $\mathcal{S}(M)$.

Now, if we are given $H \in L^2$, then we know that N defined by $N_t = \mathbb{E}(H|\mathcal{F}_t)$ is a martingale in $\mathcal{M}^2$. Thus we get a decomposition of any square integrable random variable.

Corollary 2.2.1. *If $H \in L^2$ and $M \in \mathcal{M}_{\text{loc}}$, then there exist a unique $\varphi^H \in L^2(M)$ and $L^H \in \mathcal{S}(M)^\perp \subseteq \mathcal{M}^2$ such that*

$$H = \mathbb{E}(H) + \int_0^T \varphi_s^H \, dM_s + L_T^H.$$

Furthermore, L^H is null at zero and strongly orthogonal to every element of $\mathcal{S}(M)$.

As noted in Pham (2000), this result can be extended to the case when $H \in L^1$ and M is continuous. We state a one–dimensional version of Theorem 1.2.10 of Pham (2009) (see also Galtchouk (1976) and Kunita and Watanabe (1967)).

Theorem 2.2.10. *If M is a continuous local martingale and N is a càdlàg local martingale, then there exist $\varphi \in L^2_{\text{loc}}(M)$ and L a càdlàg local martingale null at zero and orthogonal to M, such that*

$$N_t = N_0 + \int_0^t \varphi_s^N \, dM_s + L_t^N \quad 0 \le t \le T.$$

This can also be applied to any $H \in L^1(M)$ to find a decomposition of the form

$$H = \mathbb{E}(H) + \int_0^T \varphi_s^H \, dM_s + L_T^H.$$

We now consider the special case when the process L^N in the decomposition is identically equal to zero.

Definition 2.2.4. *A continuous local martingale M has the* ***predictable representation property*** *(PRP) if for any $\mathbb{F}^M$–local martingale N, there exist a predictable process $\varphi^N \in L(M)$ such that*

$$N_t = N_0 + \int_0^t \varphi_s^N \, dM_s \quad \text{for every } 0 \leq t \leq T.$$

An example of a stochastic process that has PRP is Brownian motion.

Theorem 2.2.11 (Martingale Representation)**.** *Let $W = (W^1, \dots, W^d)$ be a (independent) d–dimensional Brownian motion process and M be a local martingale with respect to the Brownian motion filtration (denoted $\mathbb{F}^W$). Then there exists a predictable process $\varphi = (\varphi^1, \dots, \varphi^d) \in L(W)$ such that*

$$M_t = M_0 + \int_0^t \varphi_s \, dW_s = M_0 + \sum_{i=1}^d \int_0^t \varphi_s^i \, dW_s^i.$$

The proof of this result can be found in Revuz and Yor (1999). We end this section with a statement of a theorem by Yor (1978) taken from Rheinlander (2011).

Theorem 2.2.12 (L^1–Martingale Representation)**.** *Let M be a locally bounded local martingale, and $(\varphi^n) \subseteq L(M)$ be a sequence such that $(\varphi^n \cdot M)$ is a martingale for each $n \in \mathbb{N}^+$. If $(\varphi^n \cdot M)_T$ converges to ϑ in L^1, then there exists $\varphi \in L(M)$ such that $(\varphi \cdot M)$ is a martingale and*

$$\vartheta = \int_0^T \varphi_t \, dM_t.$$

2.3 Continuous Trading

We now introduce some results, notation and terminology on continuous trading in a frictionless market. We follow the presentation by Rheinlander (2011) with some minor differences.

2.3.1 Theory of Continuous Trading

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, where the filtration $\mathbb{F}$ satisfies the usual conditions. We consider a financial market consisting of $d+1$ primary assets that are described by a $d+1$–dimensional càdlàg semimartingale (B, S) that is adapted to $\mathbb{F}$. Here B is one–dimensional and represents a riskless asset (bank account), while $S = (S^1, S^2, \dots, S^d)$ is a d–dimensional process representing d risky assets. We will assume that B is strictly positive and take it as a numeraire and consequently work with the assets $(1, X) = (1, X^1, X^2, \dots, X^d)$ where

$$X^i := \frac{S^i}{B} \text{ for } i = 1, 2, \dots, d.$$

Now, assume we have written a derivative H that expires at time $T > 0$. By this we mean that H is an $\mathcal{F} = \mathcal{F}_T$–measurable random variable. Our aim is to construct a trading strategy in the primary assets that will replicate H almost surely.

Definition 2.3.1. *A* ***strategy*** *is a pair* (η, φ) *such that* η *is a real valued adapted process and* $\varphi = (\varphi^1, \ldots, \varphi^d)$ *is a d–dimensional predictable X–integrable process (i.e.* $\varphi \in L(X)$*). For a strategy* (η, φ)*, we define the* ***value process*** $V(\eta, \varphi)$ *as*

$$V_t(\eta, \varphi) := \eta_t + \varphi_t X_t = \eta_t + \sum_{i=1}^{d} \varphi_t^i X_t^i \quad \text{for } 0 \leq t \leq T,$$

the ***gains process*** $G(\varphi)$ *as*

$$G_t(\varphi) := \int_0^t \varphi_s \, dX_s \quad \text{for } 0 \leq t \leq T$$

and the ***cost process*** C *as*

$$C(\eta, \varphi) = V(\eta, \varphi) - G(\varphi).$$

A strategy (η, φ) *is said to be* ***self–financing*** *if*

$$V_t(\eta, \varphi) = V_0(\eta, \varphi) + G_t(\varphi) = V_0(\eta, \varphi) + \int_0^t \varphi_s \, dX_s \quad \text{for } 0 \leq t \leq T.$$

Notice that for a self–financing strategy (η, φ), once we know the starting value $V_0(\eta, \varphi) = V_0$ and φ, then η can be determined from the self–financing condition

$$\eta_t + \varphi_t X_t = V_0 + G_t(\varphi)$$

by

$$\eta_t = V_0 + G_t(\varphi) - \varphi_t X_t.$$

Thus, a self–financing strategy (η, φ) can also be denoted by (V_0, φ), and we will often do so. Also note that a strategy is self–financing if and only if its cost process C is constant. For the remainder of this section we will assume that all the strategies we deal with are self–financing.

Definition 2.3.2. *A probability measure* Q *on* $(\Omega, \mathcal{F})$ *is called an* ***equivalent local martingale measure*** *(ELLM) for* X *if*

1. Q *is equivalent to* $\mathbb{P}$ $(Q \sim \mathbb{P})$*; i.e.* $Q(A) = 0 \Longleftrightarrow \mathbb{P}(A) = 0$ *for every* $A \in \mathcal{F}$
2. X *is a local martingale under* Q*.*

We will denote by $\mathbf{P}$ *the (convex) set of equivalent local martingale measures for* X*; i.e.*

$$\mathbf{P} := \{Q \sim \mathbb{P} : X \text{ is a } Q \text{ local martingale}\}.$$

The strategies that we have defined above turn out to be too general for many results to hold. Thus we will put further restrictions on which processes are to be allowed as strategies.

Definition 2.3.3. *A (self–financing) strategy* (η, φ) *is an* ***admissible strategy*** *if* $G(\varphi) \geq -C$ *for some* $C > 0$*.*

We note that if (η, φ) is an admissible strategy, then by Fatou's lemma (see Ansel and Stricker (1994)) $G(\varphi)$ is a Q–supermartingale for every $Q \in \mathbf{P}$. This is particularly useful when proving the Fundamental Theorem of Asset Pricing discussed below.

Definition 2.3.4. *The market admits* ***no arbitrage opportunities*** *if there is no strategy* (v_0, φ) *such that:*

1. $V_0(v_0, \varphi) = v_0 \leq 0$

2. $V_T(v_0, \varphi) \geq 0$ $\mathbb{P}$–*a.s.*

3. $\mathbb{P}(V_T(v_0, \varphi) > 0) > 0$.

We will make the usual assumption that the market admits no arbitrage opportunities. This requirement turns out to be linked to the existence of an equivalent local martingale measure for X. Indeed, assume $\mathbf{P} \neq \emptyset$ and let φ be an (supposed) arbitrage opportunity. For $Q \in \mathbf{P}$, $V(v_0, \varphi) = v_0 + G(\varphi)$ is a Q–supermartingale. Thus taking expectations under Q, we get

$$\mathbb{E}^Q(V_T) \leq v_0 + 0 = V_0 \leq 0,$$

and since $V_T \geq 0$ Q–a.s. (since $Q \sim \mathbb{P}$), it follows that $Q(V_T > 0) = 0 \iff \mathbb{P}(V_T > 0) = 0$. A contradiction. Thus, we have just proved:

Theorem 2.3.1. *If there exists an equivalent local martingale measure Q for X (i.e. $\mathbf{P} \neq \emptyset$) then there are no arbitrage opportunities.*

This theorem does have a partial converse. Harrison and Kreps (1979) proves the converse of the theorem for a finite probability space. In the general case, Delbaen and Schachermayer (1994) have shown that the converse is true provided that one is willing to strengthen the definition of no arbitrage opportunities to a concept called *No Free Lunch with Vanishing Risk*. The resulting theorem is called the **First Fundamental Theorem of Asset Pricing**.

For the remainder of the section we assume that $\mathbf{P} \neq \emptyset$ and define a strategy (η, φ) to be admissible if $G(\varphi)$ is a (cad-lag) martingale for every $Q \in \mathbf{P}$.

Definition 2.3.5. *A contingent claim (derivative) H is* ***attainable*** *if there exists an admissible strategy (v_0^H, φ^H) such that*

$$H = v_0^H + \int_0^T \varphi_t^H \, dX_t. \tag{2.4}$$

A market is ***complete*** *if every contingent claim is attainable.*

We notice that (2.4) is a representation of a random variable in terms of a constant plus a stochastic integral with respect to X. This suggests that an arbitrary claim H will be attainable if and only if X has (PRP).

Now assume that the market is complete. For any $A \in \mathcal{F}$, the claim $\mathbf{1}_A$ is attainable, thus there exist a unique admissible strategy (v_0^A, φ^A) such that

$$\mathbf{1}_A = v_0^A + \int_0^T \varphi_t^A \, dX_t.$$

Hence if $Q \in \mathbf{P}$, then (taking expectations with respect to Q)

$$Q(A) = \mathbb{E}^Q(\mathbf{1_A}) = v_0^A + 0 = v_0^A,$$

which implies that Q is unique. This suggests a link between market completeness and uniqueness of equivalent local martingale measures as stated in the following **Second Fundamental Theorem of Asset Pricing** (see Kreps (1981)).

Theorem 2.3.2. *Assume that $\mathbf{P} \neq \emptyset$. The following are equivalent:*

1. *The market is complete*

2. $\mathbf{P}$ *contains exactly one element*

3. *There exists $Q \in \mathbf{P}$ such that X has the PRP with respect to Q.*

2.3.2 A Complete Market Model

We now consider an example of a complete market model and illustrate how to find the trading strategies in the famous Black–Scholes model. This model has been studied several times and all the results presented below are well known. We use it mainly for the presentation of some arguments we will need for later chapters and as an opportunity to state some important results.

Let W be a one–dimensional Brownian motion and take $\mathbb{F} = \mathbb{F}^W$ to be the filtration generated by this Brownian motion (augmented to satisfy the usual conditions). Consider a risky asset S and a riskless bank account B that satisfy the following *stochastic differential equations* (SDEs):

$$dS_t = S_t \left(\mu \, dt + \sigma \, dW_t\right) \quad \text{and}$$

$$dB_t = r B_t \, dt.$$

Here r is the constant instantaneous risk–free rate of return, μ is the constant return for asset S and σ is the constant volatility for assets S. We will work with the discounted assets $(1, X)$ defined by

$$1 = \frac{B}{B} \text{ and } X := \frac{S}{B}.$$

By Itô's formula, X satisfies the following SDE

$$dX_t = X_t \left((\mu - r) \, dt + \sigma \, dW_t\right).$$

Integrating this SDE, we see that X is a semimartingale:

$$X_t = X_0 + M_t + A_t,$$

with

$$M_t = \int_0^t \sigma X_s \, dW_s \text{ and } A_t = \int_0^t X_s(\mu - r) \, ds.$$

To find an ELMM for X, we need the following result, which is a one dimensional combination of Theorem 5.1 and Corollary 5.13 of Karatzas and Shreve (1991):

Theorem 2.3.3 (Girsanov's Theorem)**.** *Let W be a Brownian motion under $\mathbb{P}$ and λ_t be predictable. Define a measure Q by setting*

$$\frac{dQ}{d\mathbb{P}} := \mathcal{E}\left(\lambda \cdot W\right)_T \tag{2.5}$$

and assume that Novikov's condition holds:

$$\mathbb{E}\left(e^{\frac{1}{2}\int_0^T \lambda_t^2 \, dt}\right) < \infty.$$

Then Q is a probability measure on $\mathcal{F}$ and $\hat{W}_t := W_t - \int_0^t \lambda_s ds$ is a Brownian motion under Q.

This result can be generalized to d–dimensions. See Karatzas and Shreve (1991) for further details. Now notice that the SDE for X can be written as

$$dX_t = X_t \left((\mu - r - \sigma\lambda_t) \, dt + \sigma \, d\hat{W}_t\right).$$

We see that Q defined by (2.5) will be an ELMM for X if and only if we choose $\lambda_t = \lambda := \frac{\mu - r}{\sigma}$; this choice satisfies Novikov's condition. Hence we have

$$dX_t = X_t \sigma \, d\hat{W}_t,$$

confirming that X is a local martingale under Q. So X admits a unique ELMM, which implies that X has the PRP with respect to Q by the Second Fundamental Theorem of Asset Pricing. Thus, if H is any contingent claim, then there exist an admissible strategy (v_0^H, φ^H) such that

$$H = v_0^H + \int_0^T \varphi_t^H \, dX_s \quad \mathbb{P} \text{ a.s.}$$

Note that we could have also applied the martingale representation directly to get this result. To find v_0^H and φ^H, we note that the value of the self–financing and replicating portfolio V^H can be written as

$$V_t^H = v_0^H + \int_0^t \varphi_s^H \, dX_s = \mathbb{E}^Q(H|\mathcal{F}_t),$$

where $\mathbb{E}^Q(\cdot)$ is the expectation under the measure Q. Now if $H = h(X_T)$ for some Borel measurable function $h : \mathbb{R}^+ \longrightarrow \mathbb{R}$, by the Markov property of X, there exists $F^H : [0,T] \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ such that

$$V_t^H = \mathbb{E}^Q(h(X_T)|\mathcal{F}_t) = F^H(t, X_t).$$

Applying Itô's Formula to the Q–martingale $F^H(t, X_t)$ we get

$$\begin{aligned} H = F^H(T, X_T) = F^H(0, X_0) + \int_0^T \frac{\partial F^H}{\partial t}(t, X_t)\, dt + \int_0^T \frac{\partial F^H}{\partial x}(t, X_t)\, dX_t + \frac{1}{2}\int_0^T \frac{\partial^2 F^H}{\partial x^2}(t, X_t)\, d\langle X \rangle_t \\ = F^H(0, X_0) + \int_0^T \left(\frac{\partial F^H}{\partial t}(t, X_t) + \frac{1}{2}\frac{\partial^2 F^H}{\partial x^2}(t, X_t)\sigma^2 X_t^2 \right) dt + \int_0^T \frac{\partial F^H}{\partial x}(t, X_t)\, dX_t \\ = F^H(0, X_0) + \int_0^T \frac{\partial F^H}{\partial x}(t, X_t)\, dX_t. \end{aligned}$$

Thus we see that

$$v_0^H = F^H(0, X_0) = \mathbb{E}^Q(H) \text{ and } \varphi_t^H = \frac{\partial F^H}{\partial x}(t, X_t)$$

by the uniqueness of v_0^H and φ^H.

Chapter 3

Hedging in an Illiquid Market: The Martingale Case

Now that we have covered all the necessary preliminaries, we are ready to tackle the hedging problem stated in Chapter 1. This chapter will solve an easier version of the problem where the dynamically traded asset is a local martingale. We start with this simple case because the key concepts underlying the solution are clearly highlighted without having to deal with the technicalities of working with a semimartingale. First, we take a detour and discuss general market incompleteness.

3.1 Incomplete Markets: a Result by Föllmer and Sondermann

We will present a slight modification to work done by Föllmer and Sondermann (1986) in mean–variance hedging. A further generalization was made by Schweizer (2001).

Consider a market with two (primary) assets: a risky asset S and a riskless asset B (bank account), defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and both adapted to $\mathbb{F}$. We make the assumption that $\mathbb{F} = \{\mathcal{F}_t : t \in \mathbb{R}^+\}$ is augmented to satisfy the usual conditions of continuity and completeness and $\mathcal{F}_0$ is trivial. As before, we will work with the discounted assets $(1, X)$, where

$$X := \frac{S}{B}.$$

Now let H be a European derivative that expires at time $T > 0$. In this chapter we will assume that $H \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and $\mathcal{F}_T = \mathcal{F}$. If the market is complete (e.g. when X is the Black–Scholes model considered in section 2.3.2), then there exists a unique predictable process φ^H and $v_0^H \in \mathbb{R}$ such that

$$H = v_0^H + \int_0^T \varphi_t^H \, dX_t = V_T(v_0^H, \varphi^H) \quad \mathbb{P} - \text{a.s.} \tag{3.1}$$

However, there are many situations in which it is not possible to find v_0^H and φ^H that satisfy (3.1). For instance, in the same geometric Brownian motion model of Section 2.3.2, if the volatility σ is allowed to be stochastic and driven by a second independent factor, the Martingale Representation Theorem does not apply and it is in general not possible to find a self–financing strategy that replicates H almost surely. In this section we deal precisely with the case when the market is not complete.

So if (v_0, φ) is any self–financing strategy and H is non–attainable, then there will generally be a shortfall or surplus of

$$H - v_0 - \int_0^T \varphi_t \, dX_t \quad \left(\text{or } v_0 + \int_0^T \varphi_t \, dX_t - H \text{ if we sold the derivative} \right)$$

at maturity. This section is dedicated to dealing with this random shortfall.

One way to get around the fact that the market is incomplete is to relax the self financing constraint. That is, to include non self financing strategies as admissible strategies. With this relaxation, there are many choices for φ^H and v_0^H (equivalently η) that will replicate H. One trivial choice would be to let $\varphi^H \equiv 0$, $\eta_t^H = 0$ for $t < T$ and $\eta_T^H = H$. This is a non self financing strategy that replicates H. However, this strategy has a high variance at time T, and this is undesirable from a risk management perspective. We would then aim to find a generally non–self financing strategy that has minimal 'risk'.

We first make the following assumption:

Assumption: X is a local martingale under $\mathbb{P}$ (i.e. $X \in \mathcal{M}_{\text{loc}}$).

This assumption is equivalent to saying that $\mathbb{P} \in \mathbf{P}$. This is a strong and unrealistic assumption to make. However, the theory is very simple and transparent under this assumption. We do relax it in the next chapter and consider a more general model for X.

Relaxing the self–financing condition, we will call (η, φ) a *strategy* if φ is predictable with $\varphi \in L^2(X)$, η is adapted and $V_t = \eta_t + \varphi_t X_t$ is right continuous and satisfies $V_t \in L^2(\mathbb{P})$ for every $0 \leq t \leq T$. The distinguishing feature of a non–self–financing strategy is that its cost process is no longer constant. Thus, if the strategy is replicating, then

$$H = V_T = C_T(\eta, \varphi) + G_T(\varphi),$$

where C_T is generally random. So the effectiveness (or riskiness) of a strategy can be defined in terms of the variability of its cost process.

Definition 3.1.1. *Let (φ, η) be a strategy. We define the **risk** of the strategy by*

$$R_t(\eta, \varphi) := \mathbb{E}\left((C_T(\eta, \varphi) - C_t(\eta, \varphi))^2 \,|\mathcal{F}_t \right), \quad 0 \leq t \leq T.$$

*A strategy $(\tilde{\eta}, \tilde{\varphi})$ is said to be **risk minimizing** if*

$$R_t(\tilde{\eta}, \tilde{\varphi}) \leq R_t(\eta, \varphi) \quad \textit{for every} \quad 0 \leq t \leq T \quad \textit{and every strategy} \quad (\eta, \varphi).$$

It is shown in Föllmer and Sondermann (1986) that if a strategy is risk minimizing, then its cost process is a martingale. This reduces the search for a risk minimizing strategy to strategies whose cost processes are martingales under $\mathbb{P}$.

Definition 3.1.2. *A strategy (η, φ) is called **mean self financing** if the cost process C is a martingale.*

Föllmer and Sondermann (1986) also shows that if X is a square integrable martingale, then such a risk minimizing strategy exists and is linked to the Kunita–Watanabe decomposition of H.

We will not pursue the concept of risk minimization any further but take a different route to solving the hedging problem.

3.1.1 Mean–Variance Hedging

We now look at a different kind of hedging problem where we insist on the hedging strategy (η, φ) being self–financing. Recall that such a strategy can be re-parametrized as (v_0, φ), where v_0 is the initial capital; we will use this notation for the rest of the chapter. Since the market is incomplete, such a constraint will generally introduce a hedging shortfall at maturity T given by

$$e_T := V_T(v_0, \varphi) - H = v_0 + \int_0^T \varphi_t \, dX_t - H.$$

Following the work of Föllmer and Sondermann (1986), we aim to find a strategy that minimizes the expected square of this hedging error. First note that if (and only if) a claim H' is attainable, then there exists an initial endowment $v_0' \in \mathbb{R}$ and a strategy φ' in X such that

$$H' = v_0' + \int_0^T \varphi_t' \, dX_t \quad \mathbb{P} - \text{a.s.} \tag{3.2}$$

Since we only consider claims that belong to L^2, we will choose the space of admissible trading strategies φ in X to be $L^2(X)$. Given a non–attainable claim H, the goal of *mean–variance hedging* is to find a claim H' of the form (3.2) (i.e. an attainable claim) that is 'closest' to H with respect to the L^2 norm. That is, we want to solve the following problem:

$$\min_{(v_0,\varphi)\in\mathbb{R}\times L^2(X)} \mathbb{E}\left((V_T(v_0,\varphi) - H)^2\right) = \min_{(v_0,\varphi)\in\mathbb{R}\times L^2(X)} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - H\right)^2\right) \tag{3.3}$$

$$= \min_{(v_0,\varphi)\in\mathbb{R}\times L^2(X)} \left\| v_0 + \int_0^T \varphi_t \, dX_t - H \right\|_{L^2}^2 .$$

Using (3.2), the space $\mathcal{A}$ (subspace of L^2) of all attainable claims can be written as

$$\mathcal{A} := \left\{ x + \int_0^T \varphi_t \, dX_t : x \in \mathbb{R} \text{ and } \varphi \in L^2(X) \right\} = \mathbb{R} + \mathcal{S}(X).$$

Thus we want to find a projection of H onto $\mathcal{A}$. We know from chapter 1 that such a projection exists and is unique if $\mathcal{A}$ is closed. Also, since $\mathbb{R}$ is finite–dimensional, this will be the case if $\mathcal{S}(X)$ is closed in L^2 by Theorem 2.1.8.

Since the stochastic integral with respect to a local martingale is an isometry, by Theorem 2.2.8, $\mathcal{S}(X)$ is a closed and stable subspace of L^2. Thus, the projection problem has a unique solution. We construct the solution by directly minimizing the expression in (3.3). Define $J : \mathbb{R} \times L^2(X) \longrightarrow \mathbb{R}$ by

$$J(v_0, \varphi) := \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - H\right)^2\right) \quad \text{for every } v_0 \in \mathbb{R}, \varphi \in L^2(X)$$

to be the objective function to be minimized. Since $X \in \mathcal{M}_{\text{loc}}$ and $H \in L^2(\mathbb{P})$, we know that H admits a Kunita–Watanabe decomposition:

$$H = \mathbb{E}(H) + \int_0^T \varphi_t^H \, dX_t + L_T^H,$$

where L^H is a martingale null at 0 and strongly orthogonal to $\mathcal{S}(X)$ (see Chapter 2). Substituting this representation of H into J, we obtain

$$J(v_0, \varphi) = \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - \left(\mathbb{E}(H) + \int_0^T \varphi_t^H \, dX_t + L_T^H\right)\right)^2\right)$$

$$= \mathbb{E}\left(\left((v_0 - \mathbb{E}(H)) + \int_0^T (\varphi_t - \varphi_t^H)\, dX_t - L_T^H\right)^2\right)$$

and due to the pairwise orthogonality of all the three terms, we get

$$J(v_0, \varphi) = (v_0 - \mathbb{E}(H))^2 + \mathbb{E}\left(\int_0^T (\varphi_t - \varphi_t^H)^2\, d\langle X\rangle_t\right) + \mathbb{E}((L_T^H)^2)$$

$$\geq \mathbb{E}((L_T^H)^2).$$

So $J(v_0, \varphi) \geq \mathbb{E}((L_T^H)^2)$ for every strategy $\varphi \in L^2(X)$ and initial endowment $v_0 \in \mathbb{R}$. On the other hand $J(v_0, \varphi) = \mathbb{E}((L_T)^2)$ if and only if $v_0 = \mathbb{E}(H)$ and $\varphi = \varphi^H$. Thus, $(\mathbb{E}(H), \varphi^H)$ is the point where the minimum of J occurs in the sense that the minimum expected square replication error is

$$J(\mathbb{E}(H), \varphi^H) = \min_{(v_0,\varphi)\in\mathbb{R}\times L^2(X)} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t\, dX_t - H\right)^2\right) = \mathbb{E}((L_T^H)^2).$$

So we have just proved the following theorem:

Theorem 3.1.1. *The solution to problem (3.3) is given by the trading strategy*

$$v_0 = \mathbb{E}(H) \quad \textit{and} \quad \varphi = \varphi^H.$$

Furthermore, the minimal expected square replication error is given by

$$\mathbb{E}((L_T^H)^2).$$

Proof. Above. ■

We now give a summary of the construction of this strategy:

1. Find the Kunita–Watanabe decomposition of H:

$$H = \mathbb{E}(H) + \int_0^T \varphi_t^H\, dX_t + L_T^H.$$

2. Starting with an initial endowment of $\mathbb{E}(H)$ (i.e. setting $v_0 = v_0^H = \mathbb{E}(H)$), hold φ_t^H units of X at time t and invest $\eta_t^H = V_t - \varphi_t X_t$ in the risk-less asset. Here V_t is the value of the portfolio at time t, given by

$$V_t = \mathbb{E}(H|\mathcal{F}_t) = \mathbb{E}(H) + \int_0^t \varphi_t^H\, dX_t \quad \text{for every} \quad 0 \leq t \leq T.$$

3. This strategy gives a random shortfall of L_T^H at maturity, with variance $\mathbb{E}((L_T^H)^2)$.

We remark that (see Schweizer (2001)) the Kunita–Watanabe decomposition of H

$$H = \mathbb{E}(H) + \int_0^T \varphi_t^H dX_t + L_T^H$$

can be interpreted as follows: $\mathbb{E}(H) + \int_0^T \varphi_t^H\, dX_t$ $(= V_T)$ is the attainable part of H, while L_T^H is the orthogonal unattainable part of H.

3.2 Application to Hedging in an Illiquid Market

We will now use a similar argument to the previous section to solve the problem we posed in the first chapter: Given two correlated assets U and S, where U is illiquid and S is liquid, and a derivative H written on U, find the best self–financing trading strategy consisting of: a time–zero static hedge in U, a dynamic hedging strategy in S and a dynamic hedging strategy in the bank account B that will minimize the expected square replication error.

We work with discounted prices $(1, \bar{U}, X)$, where

$$X := \frac{S}{B} \quad \text{and} \quad \bar{U} := \frac{U}{B}.$$

We will also assume that X is a local martingale under $\mathbb{P}$ and $\bar{U}_T \in L^2(\mathbb{P})$. Suppose we have purchased (or sold) a European derivative $H \in L^2(\mathbb{P})$ on $\bar{U}$ that expires at time $T > 0$. If continuous trading in $\bar{U}$ was possible (and assuming $\bar{U}$ has the PRP with respect to some ELMM for $\bar{U}$), the Martingale Representation Theorem would guarantee the existence of a dynamic hedging portfolio, consisting of holdings in $\bar{U}$ and in the bank account, that replicates H almost surely.

We will consider the case where continuous trading in $\bar{U}$ is not possible, but will be substituted by a time 0 static hedge in $\bar{U}$ and dynamic hedging strategy in the correlated asset X. This new trading strategy will generally not replicate H, unless $\bar{U}$ and X are perfectly correlated; so in general, there will be a profit or loss (P&L) incurred at time T. Our aim is to minimize the expected square of this P&L (equivalently, the variance).

The problem can be stated as follows:

Find a constant holding $\hat{\theta}$ in $\bar{U}$, an initial capital $\hat{v}_0 + \hat{\theta}\bar{U}_0$ and a dynamic (self–financing) strategy $(\hat{\varphi}, \hat{\eta}) = \{(\hat{\varphi}_t, \hat{\eta}_t) : 0 \le t \le T\}$ in X and the bank account respectively that will minimize the expected square replication error (variance).

Note that the total initial capital is expressed as $\hat{v}_0 + \hat{\theta}\bar{U}_0$, where $\hat{v}_0$ is the initial capital for the dynamic hedge in X and the bank account, while $\hat{\theta}\bar{U}_0$ is the initial capital for the static hedge in $\bar{U}$. If we let $V = V(\theta, \varphi, v_0)$ be the value of any such self–financing strategy (θ, φ, v_0), then

$$V_t(\theta, \varphi, v_0) = v_0 + \int_0^t \varphi_s \, dX_s + \theta \bar{U}_t = \varphi_t X_t + \eta_t + \theta \bar{U}_t \quad \text{for all} \quad t \in [0, T].$$

The self–financing condition gives an equivalent mathematical formulation as:

$$\min_{\theta,\varphi,v_0} \mathbb{E}\left[\left(H - v_0 - \int_0^T \varphi_t \, dX_t - \theta \bar{U}_T\right)^2\right] \tag{3.4}$$

$$= \min_{\theta,\varphi,v_0} \mathbb{E}\left[(H - V_T)^2\right] = \min_{\theta,\varphi,v_0} ||H - V_T||^2_{L^2}.$$

We again choose the space of all dynamic trading strategies in X to be $L^2(X)$ and define $G : \mathbb{R} \times L^2(X) \times \mathbb{R} \longrightarrow \mathbb{R}$ by $G(\theta, \varphi, v_0) := \mathbb{E}\left[(H - V_T(\theta, \varphi, v_0))^2\right]$ to be the objective function to be minimized. Then (3.4) can be rewritten as

$$\min_{\theta,\varphi,v_0} G(\theta, \varphi, v_0) = \min_{\theta,\varphi,v_0} \mathbb{E}\left[\left((H - \theta \bar{U}_T) - \left(v_0 + \int_0^T \varphi_t \, dX_t\right)\right)^2\right], \tag{3.5}$$

where the minimum is taken over $\mathbb{R} \times L^2(X) \times \mathbb{R}$. Let $\mathcal{A}^{\bar{U}}$ be the space of attainable claims:

$$\mathcal{A}^{\bar{U}} := \left\{ x + \int_0^T \varphi_t \, dX_t + \theta \bar{U}_T : x, \theta \in \mathbb{R} \text{ and } \varphi \in L^2(X) \right\} = \mathbb{R} + \mathcal{S}(X) + \mathrm{Span}\{\bar{U}_T\}.$$

Minimizing G is equivalent to finding the projection of H onto $\mathcal{A}^{\bar{U}}$. Again we see that $\mathcal{A}^{\bar{U}}$ is a closed subspace of L^2 since $\mathcal{S}(X)$ is closed and $\mathcal{A}^{\bar{U}}$ is a sum of $\mathcal{S}(X)$ and two one–dimensional subspaces of L^2. So there exists a unique projection onto $\mathcal{A}^{\bar{U}}$.

Now to find this projection, first observe from (3.5) that for each fixed $\theta \in \mathbb{R}$, this problem is equivalent to finding the projection of the modified derivative $H - \theta \bar{U}_T$ onto the space $\mathbb{R}+\mathcal{S}(X)$, which is simply the space of attainable claims from the previous section. Thus, following the same kind of reasoning used in the previous section we first find the Kunita–Watanabe decomposition of $H - \theta \bar{U}_T$ (clearly $H - \theta \bar{U}_T \in L^2$ since $\bar{U}_T \in L^2$):

$$H - \theta \bar{U}_T = \mathbb{E}(H - \theta \bar{U}_T) + \int_0^T \varphi_t^{H,\theta} \, dX_t + L_T^{H,\theta},$$

where $\varphi^{H,\theta} \in L^2(X)$ and $L^{H,\theta}$ is strongly orthogonal to $\mathcal{S}(X)$. Substituting this expression into G we get (for every φ, θ and v_0)

$$\begin{aligned} G(\theta, \varphi, v_0) &= \mathbb{E}\left(\left(\mathbb{E}(H) - \theta \mathbb{E}(\bar{U}_T) + \int_0^T \varphi_t^{H,\theta} \, dX_t + L_T^{H,\theta} - v_0 - \int_0^T \varphi_t \, dX_t \right)^2 \right) \\ &= \mathbb{E}\left(\left(\mathbb{E}(H) - v_0 - \theta \mathbb{E}(\bar{U}_T) + \int_0^T (\varphi_t^{H,\theta} - \varphi_t) \, dX_t + L_T^{H,\theta} \right)^2 \right). \end{aligned}$$

Again using orthogonality, we get

$$\begin{aligned} G(\theta, \varphi, v_0) &= \left(\mathbb{E}(H) - v_0 - \theta \mathbb{E}(\bar{U}_T) \right)^2 + \mathbb{E}\left(\int_0^T (\varphi_t^{H,\theta} - \varphi_t)^2 \, d\langle X \rangle_t \right) + \mathbb{E}\left(\left(L_T^{H,\theta} \right)^2 \right) \\ &\geq \mathbb{E}\left(\left(L_T^{H,\theta} \right)^2 \right) =: g(\theta). \end{aligned}$$

So $G(\theta, \varphi, v_0) \geq g(\theta)$ for every admissible strategy $\varphi \in L^2(X)$ and every $v_0, \theta \in \mathbb{R}$. Now assume that there exists a minimizer $\hat{\theta}$ for g. Setting $\varphi = \varphi^{H,\hat{\theta}}$ and $v_0 = \mathbb{E}(H) - \hat{\theta} \mathbb{E}(\bar{U}_T)$ gives

$$G(\hat{\theta}, \varphi^{H,\hat{\theta}}, \mathbb{E}(H) - \hat{\theta} \mathbb{E}(\bar{U}_T)) = \mathbb{E}\left(\left(L_T^{H,\hat{\theta}} \right)^2 \right) = g(\hat{\theta}).$$

Which implies that the minimum of G occurs at $\theta = \hat{\theta}, \varphi = \hat{\varphi} = \varphi^{H,\hat{\theta}}$ and $v_0 = \hat{v}_0 = \mathbb{E}(H) - \hat{\theta} \mathbb{E}(\bar{U}_T)$.

All we need to show is that g does indeed have a point of absolute minimum $\hat{\theta}$. First write the Kunita–Watanabe decompositions of H and $\bar{U}_T$ with respect to X:

$$H = \mathbb{E}(H) + \int_0^T \varphi_t^H \, dX_t + L_T^H \quad \text{and} \quad \bar{U}_T = \mathbb{E}(\bar{U}_T) + \int_0^T \varphi_t^{\bar{U}} \, dX_t + L_T^{\bar{U}}.$$

Due to linearity of the projection, we observe that for every $\theta \in \mathbb{R}$, the terms in decomposition for $H - \theta \bar{U}_T$ are

$$\varphi^{H,\theta} = \varphi^H - \theta \varphi^{\bar{U}} \text{ and } L^{H,\theta} = L^H - \theta L^{\bar{U}}.$$

So we get

$$g(\theta) = \mathbb{E}((L^{H,\theta})^2) = \mathbb{E}((L_T^H - \theta L_T^{\bar{U}})^2) = \mathbb{E}((L_T^H)^2) - 2\theta\mathbb{E}((L_T^H L_T^{\bar{U}})) + \theta^2\mathbb{E}((L_T^{\bar{U}})^2).$$

Thus g is a quadratic function of θ with a positive coefficient of θ^2, hence g has a minimum at

$$\hat{\theta} = \frac{\mathbb{E}((L_T^H L_T^{\bar{U}}))}{\mathbb{E}((L^{\bar{U}})^2)} = \frac{\mathrm{Cov}(L_T^H, L_T^{\bar{U}})}{\mathrm{Var}(L_T^{\bar{U}})}$$

and the minimum variance is

$$g(\hat{\theta}) = \mathbb{E}((L_T^H)^2) - \frac{\left(\mathbb{E}(L_T^H L_T^{\bar{U}})\right)^2}{\mathbb{E}((L_T^{\bar{U}})^2)} = \mathrm{Var}((L_T^H))\left(1 - \left(\mathrm{Corr}(L_T^H, L_T^{\bar{U}})\right)^2\right).$$

Thus we have found the point of minimum for G.

Theorem 3.2.1. *The solution to problem (3.3) is given by the trading strategy*

$$\hat{\theta} = \frac{\mathbb{E}((L_T^H L_T^{\bar{U}}))}{\mathbb{E}((L^{\bar{U}})^2)}, \quad \hat{v}_0 = \mathbb{E}(H) - \hat{\theta}\mathbb{E}(\bar{U}_T) \quad \text{and} \quad \hat{\varphi} = \varphi^{H,\hat{\theta}} = \varphi^H - \hat{\theta}\varphi^{\bar{U}}.$$

Furthermore, the minimal expected square replication error is given by

$$\mathbb{E}((L_T^{H,\hat{\theta}})^2) = \mathbb{E}((L_T^H - \hat{\theta} L_T^{\bar{U}})^2) = \mathbb{E}((L_T^H)^2) - \frac{\left(\mathbb{E}(L_T^H L_T^{\bar{U}})\right)^2}{\mathbb{E}((L_T^{\bar{U}})^2)}.$$

Proof. Above. ∎

The optimization procedure is summarized below:

1. First find the Kunita–Watanabe decompositions of H and $\bar{U}_T$ with respect to the local martingale X:

$$H = \mathbb{E}(H) + \int_0^T \varphi_t^H \, dX_t + L_T^H \quad \text{and} \quad \bar{U}_T = \mathbb{E}(\bar{U}_T) + \int_0^T \varphi_t^{\bar{U}} \, dX_t + L_T^{\bar{U}}.$$

2. Hold a constant $\hat{\theta}$ units of $\bar{U}$, where

$$\hat{\theta} = \frac{\mathbb{E}((L_T^H L_T^{\bar{U}}))}{\mathbb{E}((L^{\bar{U}})^2)}.$$

3. Construct a dynamic trading strategy in X and the bank account with value $V_t^D = \mathbb{E}(H - \hat{\theta}\bar{U}_T | \mathcal{F}_t)$ by holding $\hat{\varphi} = \varphi^{H,\hat{\theta}} = \varphi^H - \hat{\theta}\varphi^{\bar{U}}$ units of X and $\hat{\eta} = V^D - \varphi^{H,\hat{\theta}} X$ units of the riskless bank account. This dynamic strategy uses a starting capital of $\hat{v}_0 = \mathbb{E}(H) - \hat{\theta}\mathbb{E}(\bar{U}_T)$. Note that the value of the whole portfolio is given by $V = V^D + \hat{\theta}\bar{U} = \hat{\eta} + \hat{\varphi}X + \hat{\theta}\bar{U}$, with a total initial capital of $\hat{v}_0 + \hat{\theta}\tilde{U}_0$.

4. This self financing strategy results in a minimum expected square replication error of

$$\mathbb{E}((L_T^H)^2) - \frac{\left(\mathbb{E}(L_T^H L_T^{\bar{U}})\right)^2}{\mathbb{E}((L_T^{\bar{U}})^2)}.$$

3.3 Finding The Best Hedge: General Payoff

We will apply the results of the previous section to the case of two risky assets U and S and a riskless bank account B satisfying the following stochastic differential equations

$$dU_t = U_t\left(\mu_U\, dt + \sigma_U\left(\rho dW_t^1 + \sqrt{1-\rho^2}\, dW_t^2\right)\right) \tag{3.6}$$

$$dS_t = S_t\left(r\, dt + \sigma\, dW_t^1\right) \tag{3.7}$$

$$dB_t = rB_t\, dt, \tag{3.8}$$

where r is the continuously compounded risk–free rate of interest, $\rho \in [-1, 1]$ is the instantaneous correlation between U and S, W^1 and W^2 are two independent Brownian motion processes and $\sigma_U > 0$ and $\sigma > 0$ are the volatilities of U and S respectively. Notice that S is assumed to grow at the risk–free rate, which is in line with the local martingale assumption of the discounted price X from Section 3.2. We also take $\mathcal{F}_t = \sigma(\{W_s^1, W_s^2 : 0 \le s \le t\})$ and $\mathcal{F}_T = \mathcal{F}$.

The discounted assets are defined as:

$$X = \frac{S}{B}, \quad \bar{U} = \frac{U}{B},$$

and satisfy

$$d\bar{U}_t = \bar{U}_t\left((\mu_U - r)\, dt + \sigma_U\left(\rho dW_t^1 + \sqrt{1-\rho^2}\, dW_t^2\right)\right) \tag{3.9}$$

$$dX_t = X_t\sigma\, dW_t^1. \tag{3.10}$$

We assume that the claim H is of the form $H = h(\bar{U}_T)$ for some Borel measurable function $h : \mathbb{R} \longrightarrow \mathbb{R}$ and $\mathbb{E}(H^2) < \infty$. In this section we find simplified expressions for the best hedge parameters of a general claim written on $\bar{U}$. We will follow the four steps given above:

Step 1: We first find the Kunita–Watanabe decompositions of H and $\bar{U}$ with respect to X. That is, we find predictable processes φ^H and $\varphi^{\bar{U}}$ in $L^2(X)$ and square integrable martingales L^H and $L^{\bar{U}}$ that are strongly orthogonal to $\mathcal{S}(X)$, such that

$$H = h(\bar{U}_T) = \mathbb{E}(H) + \int_0^T \varphi_t^H\, dX_t + L_T^H \quad \text{and} \quad \bar{U}_T = \mathbb{E}(\bar{U}_T) + \int_0^T \varphi_t^{\bar{U}}\, dX_t + L_T^{\bar{U}}.$$

To do so, we follow a presentation similar to Hulley and McWalter (2008) in finding the so–called *Föllmer–Schweizer decomposition* (see next chapter).

We will construct both decompositions simultaneously by first finding a decomposition for an arbitrary function f of $\bar{U}_T$ and then substituting $f(\bar{U}_T) = h(\bar{U}_T) = H$ and $f(\bar{U}_T) = \bar{U}_T$. It will be convenient to work with a martingale "tracking process" for $\bar{U}$, thus, set $\tilde{U}_t = e^{(\mu_U - r)(T-t)}\bar{U}_t$ to be the drift adjusted process. Then $\tilde{U}$ satisfies the following SDE

$$d\tilde{U}_t = \tilde{U}_t\sigma_U\left(\rho\, dW_t^1 + \sqrt{1-\rho^2}\, dW_t^2\right).$$

Define $F^f : [0,T] \times [0,\infty) \longrightarrow \mathbb{R}$ by

$$F^f(t, x) = \mathbb{E}(f(\tilde{U}_T)|\tilde{U}_t = x).$$

F^f is well-defined by the Markov property of $\tilde{U}$. We will later consider the following special cases:

$$F^f(t,x) = F^H(t,x) = \mathbb{E}(h(\tilde{U}_T)|\tilde{U}_t = x) \text{ when } f(\bar{U}_T) = h(\bar{U}_T) = H$$

and

$$F^f(t,x) = F^{\bar{U}}(t,x) = \mathbb{E}(\bar{U}_T|\tilde{U}_t = x) = \mathbb{E}(\tilde{U}_T|\tilde{U}_t = x) = x \text{ when } f(\bar{U}_T) = \bar{U}_T$$

since $\tilde{U}$ is a martingale. Note that $F^f(T,x) = f(x)$, thus applying Itô's formula to the process $F^f(t,\tilde{U}_t)$ gives

$$\begin{aligned} f(\bar{U}_T) &= f(\tilde{U}_T) = F^f(T,\tilde{U}_T) \\ &= F^f(0,\tilde{U}_0) + \int_0^T \frac{\partial F^f}{\partial t}(t,\tilde{U}_t)\,dt + \int_0^T \frac{\partial F^f}{\partial x}(t,\tilde{U}_t)\,d\tilde{U}_t + \frac{1}{2}\int_0^T \frac{\partial^2 F^f}{\partial x^2}(t,\tilde{U}_t)\,d[\tilde{U}]_t \\ &= F^f(0,\tilde{U}_0) + \int_0^T \left(\frac{\partial F^f}{\partial t} + \frac{1}{2}\tilde{U}_t^2\sigma_U^2\frac{\partial^2 F^f}{\partial x^2}\right) dt + \int_0^T \frac{\partial F^f}{\partial x}\rho\tilde{U}_t\sigma_U\,dW_t^1 \\ &\quad + \int_0^T \frac{\partial F^f}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2. \end{aligned}$$

By the martingale property of the process $F^f(t,\tilde{U}_t)$, F^f satisfies the following PDE

$$\frac{\partial F^f}{\partial t} + \frac{1}{2}x^2\sigma_U^2\frac{\partial^2 F^f}{\partial x^2} = 0, \quad F^f(T,x) = f(x), \tag{3.11}$$

hence

$$f(\bar{U}_T) = F^f(0,\tilde{U}_0) + \int_0^T \frac{\partial F^f}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U\,dX_t + \int_0^T \frac{\partial F^f}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2,$$

which gives a Kunita–Watanabe decomposition of $f(\bar{U}_T)$ with

$$L_T^f = \int_0^T \frac{\partial F^f}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2 \quad \text{and} \quad \varphi_t^f = \frac{\partial F^f}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U.$$

Here both F^f and $\frac{\partial F^f}{\partial x}$ are evaluated at $(t,\tilde{U}_t)$ for $0 \le t \le T$.

Applying this to the derivative $f(\bar{U}_T) = h(\bar{U}_T) = H$ gives

$$\begin{aligned} H = h(\bar{U}_T) &= F^H(0,\tilde{U}_0) + \int_0^T \frac{\partial F^H}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U\,dX_t + \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2 \\ &= \mathbb{E}(H) + \int_0^T \frac{\partial F^H}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U\,dX_t + \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2 \end{aligned}$$

with

$$L_T^H = \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2 \quad \text{and} \quad \varphi_t^H = \frac{\partial F^H}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U.$$

Similarly if we apply this decomposition to $f(\bar{U}_T) = \bar{U}_T$, we get

$$\bar{U}_T = F^{\bar{U}}(0,\tilde{U}_0) + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U\,dX_t + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U\,dW_t^2$$

$$= \mathbb{E}(\bar{U}_T) + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x} \rho \frac{\tilde{U}_t}{\sigma X_t} \sigma_U \, dX_t + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x} \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2$$

with

$$L_T^{\bar{U}} = \int_0^T \frac{\partial F^{\bar{U}}}{\partial x} \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2 \quad \text{and} \quad \varphi_t^{\bar{U}} = \frac{\partial F^{\bar{U}}}{\partial x} \rho \frac{\tilde{U}_t}{\sigma X_t} \sigma_U.$$

Since $F^{\bar{U}}(t,x) = x$, it follows that $\frac{\partial F^{\bar{U}}}{\partial x} = 1$, hence

$$L_T^{\bar{U}} = \int_0^T \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2 \quad \text{and} \quad \varphi_t^{\bar{U}} = \rho \frac{\tilde{U}_t}{\sigma X_t} \sigma_U.$$

This completes step 1.

Step 2: From the previous section, the expression for $\hat{\theta}$ is given by

$$\hat{\theta} = \frac{\mathbb{E}\left((L_T^H L_T^{\bar{U}})\right)}{\mathbb{E}\left((L^{\bar{U}})^2\right)}$$

$$= \frac{\mathbb{E}\left(\left(\int_0^T \frac{\partial F^H}{\partial x} \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2\right)\left(\int_0^T \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2\right)\right)}{\mathbb{E}\left(\left(\int_0^T \sqrt{1-\rho^2} \tilde{U}_t \sigma_U \, dW_t^2\right)^2\right)}.$$

$$= \frac{\int_0^T \mathbb{E}\left(\frac{\partial F^H}{\partial x} \tilde{U}_t^2\right) dt}{\int_0^T \mathbb{E}\left(\tilde{U}_t^2\right) dt}$$

by Fubini's Theorem. The expression for $\hat{\theta}$ can be interpreted as a weighted 'Black–Scholes delta', since if continuous trading in $\bar{U}$ was possible, we saw from Chapter 2 that the holding in $\bar{U}$ would be $\frac{\partial F}{\partial x}$. In this case the holding is some kind of weighted average, weighted by $\mathbb{E}(\tilde{U}_\cdot^2)$. It is also interesting to note that this value for $\hat{\theta}$ does not depend on the correlation ρ.

Step 3: Expressions for the holding in X and the initial capital for the dynamic strategy are given by

$$\hat{\varphi}_t = \varphi_t^{H,\hat{\theta}} = \varphi_t^H - \hat{\theta}\varphi_t^{\bar{U}} = \left(\frac{\partial F^H}{\partial x} - \hat{\theta}\right) \rho \frac{\tilde{U}_t}{\sigma X_t} \sigma_U$$

and

$$\hat{v_0} = \mathbb{E}(H) - \hat{\theta}\mathbb{E}(\bar{U}_T) = F^H(0, \tilde{U}_0) - \hat{\theta} F^{\bar{U}}(0, \tilde{U}_0) = F^H(0, \tilde{U}_0) - \hat{\theta}\tilde{U}_0.$$

The holding in the bank account is given by

$$\eta_t = F^H(t, \tilde{U}_t) - \hat{\theta}\bar{U}_t - \hat{\varphi}_t X_t.$$

Step 4: Finally, for any $\theta \in \mathbb{R}$, the minimal (with respect to φ and v_0) expected square replication error $g(\theta)$ is given by

$$g(\theta) = a_1\theta^2 - 2a_2\theta + a_3,$$

where

$$a_1 = \mathbb{E}\left(\left(L_T^{\bar{U}}\right)^2\right) = \sigma_U^2(1-\rho^2) \int_0^T \mathbb{E}\left(\tilde{U}_t^2\right) dt,$$

$$a_2 = \mathbb{E}\left(L_T^H L_T^{\bar{U}}\right) = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\frac{\partial F^H}{\partial x}\tilde{U}_t^2\right)\,dt$$

and

$$a_3 = \mathbb{E}\left(\left(L_T^H\right)^2\right) = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\left(\frac{\partial F^H}{\partial x}\right)^2\tilde{U}_t^2\right)\,dt.$$

This has a minimum value of

$$g(\hat{\theta}) = \mathbb{E}\left(\left(L_T^H\right)^2\right) - \frac{\left(\mathbb{E}\left(L_T^H L_T^{\bar{U}}\right)\right)^2}{\mathbb{E}\left(\left(L_T^{\bar{U}}\right)^2\right)} = a_3 - \frac{a_2^2}{a_1}.$$

This completes step 4 and the whole 4–step procedure.

3.4 Finding The Best Hedge: Call Option

Now consider the case of a call option with strike price K, written on $\bar{U}$. In this case, we have

$$H = h(\bar{U}_T) = \max(\bar{U}_T - K, 0).$$

It can be shown that

$$F^H(t,x) := \mathbb{E}(H|\tilde{U}_t = x) = \mathbb{E}((\bar{U}_T - K)^+|\tilde{U}_t = x) = x\Phi(d_1(t,x)) - K\Phi(d_2(t,x)),$$

where

$$d_1(t,x) = \frac{\log\left(\frac{x}{K}\right) + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}} \quad \text{and} \quad d_2(t,x) = d_1(t,x) - \sigma_U\sqrt{T-t}.$$

Here Φ is the cumulative distribution function (cdf) of the standard normal distribution defined by

$$\Phi(x) := \int_{-\infty}^x \phi(t)\,dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\,dt.$$

This implies that

$$\frac{\partial F^H}{\partial x} = \Phi(d_1(t,x)).$$

So the optimal portfolio consists of holding $\hat{\varphi}$ units of X and a constant $\hat{\theta}$ units in $\bar{U}$, where

$$\hat{\theta} = \frac{\int_0^T \mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right)\,dt}{\int_0^T \mathbb{E}\left(\tilde{U}_t^2\right)\,dt} \quad \text{and} \quad \hat{\varphi}_t = \left(\Phi(d_1(t,\tilde{U}_t)) - \hat{\theta}\right)\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U, \quad 0 \le t \le T.$$

We now analyse the behaviour of the expected square replication error g for different values of θ and also find closed form expressions for the terms in g. Recall that for a given $\theta \in \mathbb{R}$,

$$g(\theta) = \mathbb{E}\left(\left(L_T^H - \theta L_T^{\bar{U}}\right)^2\right) = a_1\theta^2 - 2a_2\theta + a_3$$

where

$$a_1 = \mathbb{E}\left(\left(L_T^{\bar{U}}\right)^2\right) = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\tilde{U}_t^2\right)\,dt,$$

$$a_2 = \mathbb{E}\left(L_T^H L_T^{\bar{U}}\right) = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\frac{\partial F^H}{\partial x}\tilde{U}_t^2\right)\,dt = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right)\,dt$$

and

$$a_3 = \mathbb{E}\left(\left(L_T^H\right)^2\right) = \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\left(\frac{\partial F^H}{\partial x}\right)^2 \tilde{U}_t^2\right)\,dt$$
$$= \sigma_U^2(1-\rho^2)\int_0^T \mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2 \tilde{U}_t^2\right)\,dt.$$

Thus, for the call option we get

$$g(\theta) = \sigma_U^2(1-\rho^2)\left[\int_0^T \mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2 \tilde{U}_t^2\right)\,dt - 2\theta\int_0^T \mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right)\,dt\right.$$
$$\left.+\theta^2\int_0^T \mathbb{E}\left(\tilde{U}_t^2\right)\,dt\right]$$

To simplify these expressions we need the following results. First define

$$\Phi_2(x,y;\rho_0) = \int_{-\infty}^{y}\int_{-\infty}^{x} \frac{1}{2\pi\sqrt{(1-\rho_0^2)}} e^{-\frac{1}{2(1-\rho_0^2)}\left(t^2+s^2-2\rho_0 st\right)}\,dtds$$

to be the joint cdf (evaluated at x and y) of random variables X and Y having a Bivariate normal distribution with $X \sim N(0,1)$, $Y \sim N(0,1)$ and correlation ρ_0.

Lemma 3.4.1. *For any $a, b \in \mathbb{R}$,*

$$\int_{-\infty}^{\infty} \Phi(ax+b)\phi(x)\,dx = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right).$$

Proof. Let X and Y be independent standard normal random variables. Note that

$$\Pr(Y \le aX+b) = \int_{-\infty}^{\infty} \Pr(Y \le aX+b | X = x)\phi(x)\,dx = \int_{-\infty}^{\infty} \Pr(Y \le ax+b)\phi(x)\,dx$$
$$= \int_{-\infty}^{\infty} \Phi(ax+b)\phi(x)\,dx.$$

Now

$$\Pr(Y \le aX+b) = \Pr\left(\frac{Y-aX}{\sqrt{1+a^2}} \le \frac{b}{\sqrt{1+a^2}}\right) = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right)$$

since $\frac{Y-aX}{\sqrt{1+a^2}} \sim N(0,1)$. ■

Lemma 3.4.2. *For any $a, b \in \mathbb{R}$,*

$$\int_{-\infty}^{\infty} \left(\Phi(ax+b)\right)^2 \phi(x)\,dx = \Phi_2\left(\frac{b}{\sqrt{1+a^2}}, \frac{b}{\sqrt{1+a^2}}; \frac{a^2}{1+a^2}\right)$$

Proof. Let X, Y and Z be independent standard normal random variables. Then

$$\Pr\left(\{Y \le aX+b\} \cap \{Z \le aX+b\}\right) = \int_{-\infty}^{\infty} \Pr\left(\{Y \le aX+b\} \cap \{Z \le aX+b\} | X = x\right)\phi(x)\,dx$$
$$= \int_{-\infty}^{\infty} \Pr\left(\{Y \le ax+b\} \cap \{Z \le ax+b\}\right)\phi(x)\,dx = \int_{-\infty}^{\infty} \Pr\left(Y \le ax+b\right)\left(Z \le ax+b\right)\phi(x)\,dx$$
$$= \int_{-\infty}^{\infty} \left(\Phi(ax+b)\right)^2 \phi(x)\,dx.$$

Now

$$\Pr\left(\{Y \le aX+b\} \cap \{Z \le aX+b\}\right)$$

$$= \Pr\left(\left\{\frac{Y-aX}{\sqrt{1+a^2}} \le \frac{b}{\sqrt{1+a^2}}\right\} \bigcap \left\{\frac{Z-aX}{\sqrt{1+a^2}} \le \frac{b}{\sqrt{1+a^2}}\right\}\right)$$

where $\frac{Y-aX}{\sqrt{1+a^2}}$ and $\frac{Z-aX}{\sqrt{1+a^2}}$ are standard normal variables with

$$\mathtt{Corr}\left(\frac{Y-aX}{\sqrt{1+a^2}}, \frac{Z-aX}{\sqrt{1+a^2}}\right) = \frac{a^2}{1+a^2}.$$

Hence

$$\int_{-\infty}^{\infty} (\Phi(ax+b))^2\,\phi(x)\,dx = \Phi_2\left(\frac{b}{\sqrt{1+a^2}}, \frac{b}{\sqrt{1+a^2}}; \frac{a^2}{1+a^2}\right).$$

■

We use these results to prove the following propositions. First let

$$a(t) = \sqrt{\frac{t}{T-t}}, \quad \beta(t) = \frac{\log\left(\frac{\tilde{U}_0}{K}\right) - \sigma_U^2 t + \frac{1}{2}\sigma_U^2 T}{\sigma_U\sqrt{T-t}} \quad \text{and} \quad b(t) = \beta(t) + 2\sigma_U\sqrt{t}a(t).$$

Proposition 3.4.1.

$$\mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\sigma_U^2 t}\Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right)$$

Proof. Since $\tilde{U}_t = \tilde{U}_0 e^{-\frac{1}{2}\sigma_U^2 t + \sigma_U\left(\rho W_t^1 + \sqrt{1-\rho^2}W_t^2\right)}$, then for fixed t we can write $\tilde{U}_t$ as $\tilde{U}_t = \tilde{U}_0 e^{-\frac{1}{2}\sigma_U^2 t + \sigma_U\sqrt{t}Z}$ where $Z \sim N(0,1)$. It then follows that

$$d_1(t,\tilde{U}_t) = \frac{\log\left(\frac{\tilde{U}_t}{K}\right) + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}} = \frac{\log\left(\frac{\tilde{U}_0}{K}\right) - \sigma_U^2 t + \frac{1}{2}\sigma_U^2 T + \sigma_U\sqrt{t}Z}{\sigma_U\sqrt{T-t}}$$
$$= a(t)Z + \beta(t).$$

Now

$$\mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right) = \int_{-\infty}^{\infty} \Phi(a(t)z + \beta(t))\tilde{U}_0^2 e^{-\sigma_U^2 t + 2\sigma_U\sqrt{t}z}\phi(z)\,dz$$
$$= \tilde{U}_0^2 e^{-\sigma_U^2 t}\int_{-\infty}^{\infty} \Phi(a(t)z + \beta(t))e^{2\sigma_U\sqrt{t}z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,dz$$
$$= \tilde{U}_0^2 e^{\sigma_U^2 t}\int_{-\infty}^{\infty} \Phi(a(t)z + \beta(t))\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(z-2\sigma_U\sqrt{t})^2}\,dz$$

after completing the square. Substituting $u = z - 2\sigma_U\sqrt{t}$ yields

$$\tilde{U}_0^2 e^{\sigma_U^2 t}\int_{-\infty}^{\infty} \Phi(a(t)u + b(t))\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(u)^2}\,du = \tilde{U}_0^2 e^{\sigma_U^2 t}\int_{-\infty}^{\infty} \Phi(a(t)u + b(t))\phi(u)\,du.$$

Hence by Lemma 4.4.1 we get

$$\tilde{U}_0^2 e^{\sigma_U^2 t}\int_{-\infty}^{\infty} \Phi(a(t)u + b(t))\phi(u)\,du = \tilde{U}_0^2 e^{\sigma_U^2 t}\Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right)$$

as required. ■

Proposition 3.4.2.

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2\tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\sigma_U^2 t}\Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right)$$

Proof. Following similar steps to the proof of Proposition 4.4.1, we get

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2 \tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\sigma_U^2 t} \int_{-\infty}^{\infty} (\Phi(a(t)u + b(t)))^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du.$$

Hence by Lemma 4.4.2

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2 \tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\sigma_U^2 t} \Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right)$$

■

It is also clear that $\mathbb{E}(\tilde{U}_t^2) = \tilde{U}_0^2 e^{\sigma_U^2 t}$, hence we get an expression for g given by

$$g(\theta) = a_1\theta^2 - 2a_2\theta + a_3$$

where

$$a_1 = \tilde{U}_0^2 \sigma_U^2 (1-\rho^2) \int_0^T e^{\sigma_U^2 t} dt, \quad a_2 = \tilde{U}_0^2 \sigma_U^2 (1-\rho^2) \int_0^T e^{\sigma_U^2 t} \Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right) dt$$

and

$$a_3 = \tilde{U}_0^2 \sigma_U^2 (1-\rho^2) \int_0^T e^{\sigma_U^2 t} \Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right) dt.$$

The value of θ which minimizes g is

$$\hat{\theta} = \frac{a_2}{a_1} = \frac{\displaystyle\int_0^T e^{\sigma_U^2 t} \Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right) dt}{\displaystyle\int_0^T e^{\sigma_U^2 t}\, dt}.$$

These expressions can then be evaluated using common numerical integration techniques. Also, it is possible to switch the order of integration and leave the expressions in terms of special functions only (no integrals), but those resulting expressions are messy, so we decided not to include them.

Chapter 4

Hedging in an Illiquid Market: The Semimartingale Case

The aim of this chapter is to solve the hedging problem introduced in Chapter 3 in a more general setting of an illiquid market where the dynamically traded stock is a continuous semimartingale. In the previous chapter we made a simplifying and limiting assumption that the continuously traded asset X was a local martingale. This, of course, has many limitations for practical use. We now relax this assumption and assume that X is a semimartingale under $\mathbb{P}$. Following a similar style to the previous chapter, we first present general mean–variance hedging in an incomplete market with no trading constraints.

4.1 Mean–Variance Hedging

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and X be a special semimartingale with canonical decomposition

$$X = X_0 + M + A,$$

where X_0 is constant, M is a local martingale null at 0 and A is a predictable process of finite variation, also null at 0. We assume that $\mathcal{F}_0$ is trivial and again work in one dimension ($d = 1$). X is taken to represent the discounted price process of a risky asset. We consider a market with continuous trading in two primary assets: the risky asset X and a risk-less bank account whose value is 1 at all times (this is after discounting). The market is not complete in the sense that not every derivative can be replicated by a self–financing trading strategy. So given a contingent claim $H \in L^2(\mathbb{P})$, we want to find an initial endowment $\hat{v}_0 \in \mathbb{R}$ and a trading strategy $\hat{\varphi}$ that will minimize the expected square replication hedging error (ESRE). If we let Θ denote the set of all admissible trading strategies in X, then we want to find a pair $(\hat{v}_0, \hat{\varphi}) \in \mathbb{R} \times \Theta$ such that

$$\mathbb{E}\left(\left(\hat{v}_0 + \int_0^T \hat{\varphi}_t \, dX_t - H\right)^2\right) = \min_{(v_0,\varphi)\in\mathbb{R}\times\Theta} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - H\right)^2\right). \tag{4.1}$$

Such a pair $(\hat{v}_0, \hat{\varphi})$ (if it exists) is called the *mean–variance optimal* strategy for H and $\hat{v}_0$ is called the *approximation price* of the claim.

Problem (4.1) has been solved by many authors in the literature under different assumptions on X and Θ. It was first solved by Duffie and Richardson (1991) when X is a diffusion and $H = kX$ for some constant k. This was later generalized by Schweizer (1992) to the case of a general payoff. Schweizer (2001) gives an overview of results in this direction.

Recall that $L(X)$ denotes the space of all predictable X–integrable processes. For each $\varphi \in L(X)$, the process $(\varphi \cdot X)$ is also a semimartingale. We want to choose a suitable subspace Θ of $L(X)$ in which all admissible strategies should belong. Since we consider only square integrable claims, one necessary condition on Θ is that for each $\varphi \in \Theta$, the random variable $(\varphi \cdot X)_T$ must be in $L^2(\mathbb{P})$. Now, for a chosen subspace Θ, define

$$G_T(\Theta) := \left\{ \int_0^T \varphi_t \, dX_t : \varphi \in \Theta \right\}$$

and

$$\mathcal{A} := \mathbb{R} + G_T(\Theta) = \left\{ x + \int_0^T \varphi_t \, dX_t : \varphi \in \Theta \text{ and } x \in \mathbb{R} \right\}.$$

Here $G_T(\Theta)$ represents the space (subspace of $L^2(\mathbb{P})$) of contingent claims that are attainable at 0 initial cost, while $\mathcal{A}$ is the space of all attainable claims (by self–financing strategies). Problem (4.1) is therefore equivalent to finding the L^2 projection of H onto the space $\mathcal{A}$ of attainable claims. From Chapter 2 we know that if $\mathcal{A}$ is a closed subspace of $L^2(\mathbb{P})$, then such a projection exists and is unique. Since $\mathbb{R}$ is finite dimensional, it is enough for $G_T(\Theta)$ to be closed in $L^2(\mathbb{P})$ (see Theorem 2.1.8). As expected, the closedness of $G_T(\Theta)$ depends a lot on X and Θ. In the literature, there are two main choices for Θ:

1. Gouriéroux *et al.* (1998) chooses $\Theta = \Theta'$ to be the the space of processes $\varphi \in L(X)$ such that $(\varphi \cdot X)_T = \int_0^T \varphi_t dX_t \in L^2(\mathbb{P})$ and $(\varphi \cdot X)$ is a martingale under every equivalent local martingale measure Q for X with $dQ/d\mathbb{P} \in L^2(\mathbb{P})$.

2. Rheinlander and Schweizer (1997) chooses $\Theta = \tilde{\Theta}$ to be the space of processes $\varphi \in L(X)$ such that the stochastic integral $(\varphi \cdot X)$ is in the space of square integrable semimartingales

$$\mathcal{S}^2 := \left\{ X = X_0 + M + A : M \in \mathcal{M}^2_{\text{loc}} \text{ and } \mathbb{E}\left(\left(\int_0^T |dA|_s \right)^2 \right) < \infty \right\}.$$

This is a stronger requirement and yields results under stronger requirements on X, but as remarked by Heath *et al.* (2001a), it is a more natural extension of the martingale case.

Rheinlander and Schweizer (1997) prove that in general, $\tilde{\Theta} \subseteq \Theta'$ and if $G_T(\tilde{\Theta})$ is closed in L^2, then $\tilde{\Theta} = \Theta'$. We will work with Θ' due to its flexibility and follow a presentation similar to Heath *et al.* (2001a).

Recall that $\mathbf{P}$ denotes the set of equivalent local martingale measures (ELMMs) for X. We will assume that X has at least one ELMM (i.e. $\mathbf{P} \neq \emptyset$, which implies absence of arbitrage by the First Fundamental Theorem of Asset Pricing). Due to market incompleteness (and convexity of $\mathbf{P}$), $\mathbf{P}$ therefore contains infinitely many elements. Define

$$\mathbf{P}^2_e := \left\{ Q \in \mathbf{P} : \frac{dQ}{d\mathbb{P}} \in L^2(\mathbb{P}) \right\}$$

to be the set of ELMMs for X with square integrable density. The following assumption is key to the main results of this section:

Assumption: $\mathbf{P}^2_e \neq \emptyset$.

We define the admissible trading strategies as follows (using the same notation as Heath *et al.* (2001a)):

$$\Theta' = \left\{ \varphi \in L(X) : \int_0^T \varphi_t \, dX_t \in L^2(\mathbb{P}) \text{ and } (\varphi \cdot X) \text{ is a } Q\text{–martingale for every } Q \in \mathbf{P}^2_e \right\}.$$

The mean variance hedging problem can then be formulated as

$$\min_{(v_0,\varphi)\in\mathbb{R}\times\Theta'} \mathbb{E}\left(\left(v_0+\int_0^T \varphi_t\,dX_t - H\right)^2\right) = \min_{v_0\in\mathbb{R}}\left(\min_{\varphi\in\Theta'} \mathbb{E}\left(\left(v_0+\int_0^T \varphi_t\,dX_t - H\right)^2\right)\right)$$

$$= \min_{v_0\in\mathbb{R}}\left(\min_{\varphi\in\Theta'}\left\|v_0+\int_0^T \varphi_t\,dX_t - H\right\|_{L^2}^2\right).$$

To establish that $G_T(\Theta')$ is a closed subspace of $L^2(\mathbb{P})$, we first make the following assumptions:

1. X is continuous and $X \in \mathcal{S}^2$.
2. We assume that A is absolutely continuous with respect to $\langle M\rangle$ in the sense that

$$A_t = \int_0^t \alpha_s\, d\langle M\rangle_s$$

for some predictable process α such that the *Mean–Variance Tradeoff process* (MVT) K satisfies

$$K_t := \int_0^t \alpha_s^2\, d\langle M\rangle_s < \infty \ \mathbb{P} \text{ a.s. for each } t\in[0,T].$$

This is called the *structure condition (SC).*

Remark: Since $\mathbf{P} \neq \emptyset$, the assumption that X is continuous already implies that SC is satisfied. Thus SC is a natural consequence of the no arbitrage assumption (Delbaen and Schachermayer (1996)).

The following is Proposition 5.2 of Pham (2000).

Theorem 4.1.1. *$G_T(\Theta')$ is closed in $L^2(\mathbb{P})$.*

Proof. Let $\vartheta_n = \int_0^T \varphi_t^n dX_t$ be a sequence in $G_T(\Theta')$ that converges to some $\vartheta \in L^2(\mathbb{P})$. We need to show that $\vartheta \in G_T(\Theta')$ by finding $\varphi\in\Theta'$ such that $\vartheta = \int_0^T \varphi_t dX_t$. Now for each $Q\in\mathbf{P}_e^2$, by the Cauchy–Schwartz inequality, we have

$$\mathbb{E}^Q(|\vartheta_n-\vartheta|) = \mathbb{E}\left(|\vartheta_n-\vartheta|\frac{dQ}{d\mathbb{P}}\right) \le \left(\mathbb{E}\left(|\vartheta_n-\vartheta|^2\right)\right)^{\frac{1}{2}}\left(\mathbb{E}\left(\left(\frac{dQ}{d\mathbb{P}}\right)^2\right)\right)^{\frac{1}{2}} \longrightarrow 0$$

since $\mathbb{E}\left(\left(\frac{dQ}{d\mathbb{P}}\right)^2\right) < \infty$. Hence $\vartheta_n = \int_0^T \varphi_t^n dX_t$ converges to ϑ in $L^1(Q)$. Now, we apply the L^1–Martingale Representation Theorem of Yor (1978) (Theorem 2.2.11) to deduce that there exists a predictable process $\varphi\in L(X)$ such that $\vartheta = \int_0^T \varphi_t dX_t$ and $(\varphi\cdot X)$ is a Q–martingale. Since $Q\in\mathbf{P}_e^2$ was arbitrary, we conclude that $\varphi\in\Theta'$ as required. ■

The closedness of $G_T(\Theta')$ implies that $\mathcal{A}$ is also closed in $L^2(\mathbb{P})$ and hence the projection problem has a unique solution by the Projection Theorem. However, at this stage we know nothing much about this unique solution $(\hat{v}_0,\hat{\varphi})$ other than the fact that $H-\hat{v}_0-\int_0^T \hat{\varphi}_t dX_t \in \mathcal{A}^\perp$.

Duffie and Richardson (1991) were the first to find an expression for $\hat{\varphi}$ when $\hat{v}_0 = 0$ and X is a diffusion by conjecturing it from discrete time arguments. Using the fact that $\hat{\varphi}$ is the unique element of $\tilde{\Theta}$ that satisfies

$$\mathbb{E}\left(\left(H-\int_0^T \hat{\varphi}_t\,dX_t\right)\int_0^T \varphi_t\,dX_t\right) = 0 \quad \text{for every } \varphi\in\tilde{\Theta},$$

they derive an ODE for

$$h(t) := \mathbb{E}\left(\left(H - \int_0^t \hat{\varphi}_s \, dX_s\right) \int_0^t \varphi_s \, dX_s\right) \quad \text{for every } 0 \leq t \leq T$$

and show that the only solution to this ODE is $h \equiv 0$, implying that $h(T) = 0$ as required. This was later generalized by Schweizer (1992), Schweizer (1994) and Rheinlander and Schweizer (1997).

We now make an attempt to characterize the optimal strategy $(\hat{v}_0, \hat{\varphi})$.

Definition 4.1.1. *The variance optimal ELMM $\tilde{P}$ is the unique element of $\mathbf{P}_e^2$ that minimizes*

$$\left\|\frac{dQ}{d\mathbb{P}}\right\|_{L^2} = \sqrt{1 + Var\left(\frac{dQ}{d\mathbb{P}}\right)}$$

over all $Q \in \mathbf{P}_e^2$.

A more general definition of $\tilde{P}$ involves signed measures, implying that $\tilde{P}$ is generally a signed measure. However, Delbaen and Schachermayer (1996) show that if $\mathbf{P}_e^2 \neq \emptyset$ and X is continuous (as assumed), then $\tilde{P}$ exists and is in $\mathbf{P}_e^2$.

The definition of this measure may seem arbitrary, but the variance optimal martingale measure turns out to be very useful in solving the hedging problem.

First, let

$$\tilde{Z}_t := \tilde{\mathbb{E}}\left(\frac{d\tilde{P}}{d\mathbb{P}}\middle| \mathcal{F}_t\right),$$

where $\tilde{\mathbb{E}}$ is the expectation under the variance optimal martingale measure. It is shown in Gouriéroux *et al.* (1998) that there exists $\xi \in \Theta'$ such that

$$\tilde{Z}_t = \tilde{Z}_0 + \int_0^t \xi_s \, dX_s.$$

Since $\frac{d\tilde{P}}{d\mathbb{P}} \in L^2(\mathbb{P})$, it follows from the Cauchy–Schwartz inequality that $H \in L^1(\tilde{P})$. Since X is a continuous local martingale under $\tilde{P}$, from Chapter 2 we saw that we can write the generalized version of the Kunita–Watanabe decomposition of H with respect to X:

$$H = \tilde{\mathbb{E}}(H) + \int_0^T \varphi_t^{H,\tilde{P}} \, dX_t + L_T^{H,\tilde{P}} = V_T^{H,\tilde{P}},$$

with

$$V_t^{H,\tilde{P}} := \tilde{\mathbb{E}}(H|\mathcal{F}_t) = \tilde{\mathbb{E}}(H) + \int_0^t \varphi_s^{H,\tilde{P}} \, dX_s + L_t^{H,\tilde{P}}, \quad 0 \leq t \leq T.$$

The following is Theorem 1 of Heath *et al.* (2001a).

Theorem 4.1.2. *The mean–variance optimal strategy $(\hat{v}_0, \hat{\varphi})$ is given by*

$$\hat{v}_0 = \tilde{\mathbb{E}}(H)$$

and

$$\hat{\varphi}_t = \varphi_t^{H,\tilde{P}} - \frac{\xi_t}{\tilde{Z}_t}\left(V_{t^-}^{H,\tilde{P}} - \tilde{\mathbb{E}}(H) - \int_0^t \hat{\varphi}_s \, dX_s\right), \quad 0 \leq t \leq T.$$

Proof. See Heath *et al.* (2001a). ■

Now let

$$Z_t^{\tilde{P}} := \mathbb{E}\left(\frac{d\tilde{P}}{d\mathbb{P}}\middle|\mathcal{F}_t\right)$$

be the density process of $\tilde{P}$ with respect to $\mathbb{P}$. Proposition 2 of Heath *et al.* (2001a) gives the minimal ESRE:

Proposition 4.1.1. *The minimal expected square replication error of H is given by*

$$\mathbb{E}\left(\left(\hat{v}_0 + \int_0^T \hat{\varphi}_t \, dX_t - H\right)^2\right) = \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\tilde{P}}\rangle_t\right).$$

Proof. See Heath *et al.* (2001a). ■

The variance optimal martingale measure is in general difficult to find. Pham *et al.* (1998) shows that in the special case when K_T is deterministic, $\tilde{P}$ is equal to the so called *minimal martingale measure* $\hat{P}$, defined by setting

$$\frac{d\hat{P}}{d\mathbb{P}} := \mathcal{E}\left(-(\alpha \cdot M)\right)_T = \exp\left(-\int_0^T \alpha_s \, dM_s - \frac{1}{2}K_T\right).$$

Theorem 4.1.3. *If K_T is deterministic, then $\tilde{P} = \hat{P}$,*

$$Z_t^{\tilde{P}} = \mathcal{E}\left(-(\alpha \cdot M)\right)_t \quad 0 \leq t \leq T,$$

$$\tilde{Z}_t = \tilde{\mathbb{E}}\left(\frac{d\hat{P}}{d\mathbb{P}}\middle|\mathcal{F}_t\right) = e^{K_T}\mathcal{E}\left(-(\alpha \cdot M)_t\right) \quad 0 \leq t \leq T,$$

$$\xi_t = -\tilde{Z}_t \alpha_t \quad 0 \leq t \leq T,$$

and

$$\frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} = e^{-(K_T - K_t)} \quad 0 \leq t \leq T.$$

Proof. See Heath *et al.* (2001a) ■

Notice that when K_T is deterministic, then the mean–variance optimal strategy can be written as

$$\hat{\varphi}_t = \varphi_t^{H,\tilde{P}} + \alpha_t\left(V_{t-}^{H,\tilde{P}} - \tilde{\mathbb{E}}(H) - \int_0^t \hat{\varphi}_s \, dX_s\right), \quad 0 \leq t \leq T.$$

4.2 Application to Hedging in an Illiquid Market

We apply the results of the previous section to the problem of static–dynamic hedging we solved in Chapter 3 to the case where the liquidly tradeable asset X is a continuous semimartingale rather than a local martingale. We will make the same assumptions as in Section 4.1 about the properties of X. We also assume that $\bar{U}_T \in L^2(\mathbb{P})$.

Just to recall, we have two correlated assets whose discounted prices are X and $\bar{U}$. X is liquidly traded while $\bar{U}$ is illiquid and can only be traded at time 0. We have a derivative $H \in L^2(\mathbb{P})$ that is written on $\bar{U}$ and we want to hedge it using a static hedge in $\bar{U}$, a dynamic hedging strategy in X and a dynamic hedging strategy in the bank account (whose value is 1 at all times). If we again measure the 'success' of a trading strategy by its ability to minimize the expectation of the square of the terminal hedging error e_T, our aim is then to find $\hat{\theta}, \hat{v_0} \in \mathbb{R}$ and $\hat{\varphi} \in \Theta'$ such that

$$\mathbb{E}\left(\left(\hat{v_0} + \int_0^T \hat{\varphi}_t \, dX_t + \hat{\theta}\bar{U}_T - H\right)^2\right) = \min_{(\theta,\varphi,v_0)\in\mathbb{R}\times\Theta'\times\mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t + \theta\bar{U}_T - H\right)^2\right). \tag{4.2}$$

Notice that (4.2) can be written as

$$\min_{\theta\in\mathbb{R}} \left(\min_{(\varphi,v_0)\in\Theta'\times\mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t + \theta\bar{U}_T - H\right)^2\right)\right) = \min_{\theta\in\mathbb{R}} g(\theta),$$

where $g : \mathbb{R} \longrightarrow \mathbb{R}$ is the minimum ESRE for a fixed $\theta \in \mathbb{R}$ defined by

$$\begin{aligned} g(\theta) &:= \min_{(\varphi,v_0)\in\Theta'\times\mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t + \theta\bar{U}_T - H\right)^2\right) \\ &= \min_{(\varphi,v_0)\in\Theta'\times\mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - \left(H - \theta\bar{U}_T\right)\right)^2\right). \end{aligned}$$

The advantage of writing the problem in this form is that we can now think of first finding the mean–variance optimal hedging strategy for the modified claim $H - \theta\bar{U}_T$ when θ is fixed and then minimize g over all $\theta \in \mathbb{R}$. The function g is well–defined since the minimum

$$\min_{(\varphi,v_0)\in\Theta'\times\mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - \left(H - \theta\bar{U}_T\right)\right)^2\right) \tag{4.3}$$

exists if $H - \theta\bar{U}_T$ is in $L^2(\mathbb{P})$ (Theorem 4.1.1). So, this expression suggests that we follow the following steps:

Step 1: First find the minimum ESRE of the modified claim $H - \theta\bar{U}_T$ (with trading only in X and the bank account, thus making the market incomplete) for each fixed $\theta \in \mathbb{R}$. This gives a value for $g(\theta)$. This is essentially an application of Theorem 4.1.1 with this modified claim.

Step 2: Then minimize g over all values of $\theta \in \mathbb{R}$ to find the complete mean–variance optimal strategy.

We now execute the two steps.

Step 1: To find this minimum ESRE for a fixed $\theta \in \mathbb{R}$, we write the Kunita–Watanabe decomposition of $H - \theta\bar{U}_T$ with respect to the local martingale X under the variance optimal martingale measure $\tilde{P}$:

$$H - \theta\bar{U}_T = \tilde{\mathbb{E}}(H - \theta\bar{U}_T) + \int_0^T \varphi_t^{H,\theta} \, dX_t + L_T^{H,\theta},$$

where $L^{H,\theta}$ is a $\tilde{P}$ martingale strongly $\tilde{P}$ orthogonal to $\mathcal{S}(X)$. Theorem 4.1.2 tells us that this minimum (for a fixed $\theta \in \mathbb{R}$) is achieved at the point $(\varphi^\theta, v_0^\theta)$, where

$$v_0^\theta = \tilde{\mathbb{E}}(H - \theta\bar{U}_T) = \tilde{\mathbb{E}}(H) - \theta\tilde{\mathbb{E}}(\bar{U}_T)$$

and

$$\varphi_t^\theta = \varphi_t^{H,\theta} - \frac{\xi_t}{\tilde{Z}_t}\left(V_{t-}^{H,\theta} - v_0^\theta - \int_0^t \varphi_s^\theta \, dX_s\right), \quad 0 \leq t \leq T.$$

Here $V_t^{H,\theta} = \tilde{\mathbb{E}}(H - \theta \bar{U}_T | \mathcal{F}_t)$. Now by Proposition 4.1.1, the minimum ESRE $g(\theta)$ is given by

$$g(\theta) = \min_{(\varphi, v_0) \in \Theta' \times \mathbb{R}} \mathbb{E}\left(\left(v_0 + \int_0^T \varphi_t \, dX_t - \left(H - \theta \bar{U}_T\right)\right)^2\right)$$

$$= \mathbb{E}\left(\left(v_0^\theta + \int_0^T \varphi_t^\theta \, dX_t - \left(H - \theta \bar{U}_T\right)\right)^2\right)$$

$$= \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\theta} \rangle_t\right).$$

So we have found an expression for $g(\theta)$ for a fixed $\theta \in \mathbb{R}$.

Step 2: Now we need to find the value of θ which minimizes g. Following similar steps to the previous chapter, we first find the (generalized) Kunita–Watanabe decomposition of both H and $\bar{U}_T$ with respect to the continuous $\tilde{P}$ local martingale X:

$$H = \tilde{\mathbb{E}}(H) + \int_0^T \varphi_t^{H,\tilde{P}} \, dX_t + L_T^{H,\tilde{P}} \quad \text{and} \quad \bar{U}_T = \tilde{\mathbb{E}}(\bar{U}_T) + \int_0^T \varphi_t^{\bar{U},\tilde{P}} d\, X_t + L_T^{\bar{U},\tilde{P}}.$$

Due to linearity of the projection, it follows that

$$L^{H,\theta} = L^{H,\tilde{P}} - \theta L^{\bar{U},\tilde{P}} \quad \text{and} \quad \varphi^{H,\theta} = \varphi^{H,\tilde{P}} - \theta \varphi^{\bar{U},\tilde{P}}.$$

From the properties of quadratic variation,

$$\langle L^{H,\theta} \rangle = \langle L^{H,\tilde{P}} \rangle - 2\theta \langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}} \rangle + \theta^2 \langle L^{\bar{U},\tilde{P}} \rangle.$$

Hence we get

$$g(\theta) = \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\theta} \rangle_t\right) = \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\left(\langle L^{H,\tilde{P}} \rangle - 2\theta \langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}} \rangle + \theta^2 \langle L^{\bar{U},\tilde{P}} \rangle\right)_t\right)$$

$$= \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\tilde{P}} \rangle_t\right) - 2\theta \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}} \rangle_t\right) + \theta^2 \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{\bar{U},\tilde{P}} \rangle_t\right).$$

So, g is again a quadratic function of θ. Thus the minimum occurs at

$$\hat{\theta} = \frac{\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}} \rangle_t\right)}{\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} \, d\langle L^{\bar{U},\tilde{P}} \rangle_t\right)}. \tag{4.4}$$

Hence the mean–variance optimal strategy is given implicitly by

$$\hat{\varphi}_t = \varphi^{\hat{\theta}} = \varphi_t^{H,\hat{\theta}} - \frac{\xi_t}{\tilde{Z}_t}\left(V_{t-}^{H,\hat{\theta}} - \tilde{\mathbb{E}}(H - \hat{\theta} \bar{U}_T) - \int_0^t \hat{\varphi}_s \, dX_s\right), \quad 0 \le t \le T$$

$$= (\varphi_t^{H,\tilde{P}} - \hat{\theta} \varphi_t^{\bar{U},\tilde{P}}) - \frac{\xi_t}{\tilde{Z}_t}\left(V_{t-}^{H,\hat{\theta}} - \tilde{\mathbb{E}}(H - \hat{\theta} \bar{U}_T) - \int_0^t \hat{\varphi}_s \, dX_s\right), \quad 0 \le t \le T. \tag{4.5}$$

and

$$\hat{v_0} = v_0^{\hat{\theta}} = \tilde{\mathbb{E}}(H) - \hat{\theta} \tilde{\mathbb{E}}(\bar{U}_T). \tag{4.6}$$

Theorem 4.2.1. *The mean–variance optimal strategy is given by equations (4.4), (4.5) and (4.6).*

Proof. Above. ■

We put everything together and summarize the optimization procedure:

1. First find the variance optimal martingale measure $\tilde{P}$ and the dynamics of all processes under $\tilde{P}$.

2. Find the Kunita–Watanabe decompositions of H and $\bar{U}_T$ under $\tilde{P}$ with respect to X:

$$H = \tilde{\mathbb{E}}(H) + \int_0^T \varphi_t^{H,\tilde{P}}\, dX_t + L_T^{H,\tilde{P}} \quad \text{and} \quad \bar{U}_T = \tilde{\mathbb{E}}(\bar{U}_T) + \int_0^T \varphi_t^{\bar{U},\tilde{P}}\, dX_t + L_T^{\bar{U},\tilde{P}}.$$

3. Choose $(\hat{\theta}, \hat{\varphi}, \hat{v}_0)$ as in equations (4.4), (4.5) and (4.6).

4. The (minimum) expected square replication error of this strategy is

$$g(\hat{\theta}) = \mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t}\, d\langle L^{H,\tilde{P}}\rangle_t\right) - \frac{\left(\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t}\, d\langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}}\rangle_t\right)\right)^2}{\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t}\, d\langle L^{\bar{U},\tilde{P}}\rangle_t\right)}.$$

We will now apply this to a specific model.

4.3 The Brownian Motion Model

We apply the results of the previous section to the case when X and $\bar{U}$ are Itô processes that satisfy the following SDEs

$$d\bar{U}_t = \bar{U}_t\left((\mu_U - r)\, dt + \sigma_U\left(\rho\, dW_t^1 + \sqrt{1-\rho^2}\, dW_t^2\right)\right) \tag{4.7}$$

$$dX_t = X_t((\mu - r)\, dt + \sigma\, dW_t^1). \tag{4.8}$$

Here W^1 and W^2 are independent standard Brownian motion processes and $\mathbb{F} = \mathbb{F}^W = \sigma(W^1, W^2)$ (augmented to satisfy the usual conditions). This is the same model from Chapter 4 but with a non–zero drift for X. Integrating (4.8), we observe that X can be written as

$$X = X_0 + M + A$$

where

$$M_t = \int_0^t X_s \sigma\, dW_s^1 \quad \text{and} \quad A_t = \int_0^t X_s(\mu - r)\, ds.$$

It can be shown that with this decomposition, X satisfies all the assumptions made in the previous section. We also have that A is absolutely continuous with respect to $\langle M\rangle$ with

$$A_t = \int_0^t \alpha_s d\langle M\rangle_s,$$

where

$$\alpha_s = \frac{(\mu - r)}{\sigma^2 X_s^2}.$$

The mean–variance trade–off process can be expressed as

$$K_t = \int_0^t \alpha_s^2 \, d\langle M\rangle_s = \int_0^t \left(\frac{\mu - r}{\sigma}\right)^2 ds = \left(\frac{\mu - r}{\sigma}\right)^2 t = \lambda^2 t,$$

where

$$\lambda := \left(\frac{\mu - r}{\sigma}\right)$$

is the market price of risk associated with asset X. Since X is continuous and K_T is deterministic, X satisfies (SC) and $\tilde{P} = \hat{P}$. Thus

$$\frac{d\tilde{P}}{d\mathbb{P}} = \frac{d\hat{P}}{d\mathbb{P}} = \mathcal{E}\left((\alpha \cdot M)\right)_T = \exp\left(-\int_0^T \alpha_s \, dM_s - \frac{1}{2}K_T\right) = \exp\left(-\lambda W_T^1 - \frac{1}{2}\lambda^2 T\right).$$

Now set

$$\hat{W}_t^1 := W_t^1 + \lambda t, \quad \hat{W}_t^2 := W_t^2, \quad 0 \le t \le T.$$

It can be shown that Novikov's condition is satisfied, thus applying a two–dimensional version of Girsanov's theorem (see Karatzas and Shreve (1991)), we get that $\hat{W}^1$ and $\hat{W}^2$ are independent Brownian motion processes under $\tilde{P}$. The SDE for $\bar{U}$ can be written as

$$\begin{aligned} d\bar{U}_t &= \bar{U}_t \left((\mu_U - r)\, dt + \sigma_U \left(\rho(\, d\hat{W}_t^1 - \lambda\, dt) + \sqrt{1-\rho^2}\, d\hat{W}_t^2\right)\right) \\ &= \bar{U}_t \left((\mu_U - r - \rho\sigma_U\lambda)\, dt + \sigma_U \left(\rho\, d\hat{W}_t^1 + \sqrt{1-\rho^2}\, d\hat{W}_t^2\right)\right) \\ &= \bar{U}_t \left(\gamma\, dt + \sigma_U \left(\rho\, d\hat{W}_t^1 + \sqrt{1-\rho^2}\, d\hat{W}_t^2\right)\right) \end{aligned}$$

where

$$\gamma = \mu_U - r - \rho\sigma_U\lambda.$$

On the other hand

$$dX_t = X_t \left((\mu - r)\, dt + \sigma(\, d\hat{W}_t^1 - \lambda\, dt)\right) = X_t \sigma\, d\hat{W}_t^1,$$

which is expected since $\tilde{P}$ is an ELMM for X. This completes step 1. We will now find the best strategy and minimum ESRE for a general pay-off and for a call option.

4.3.1 General Payoff

Let $H \in L^2(\mathbb{P})$ be a claim written on $\bar{U}$. We assume that $H = h(\bar{U}_T)$ for some real valued Borel measurable function h. We want to find the best hedge parameters $(\hat{\theta}, \hat{\varphi}, \hat{v}_0)$ and the minimum ESRE $g(\hat{\theta})$. This will follow the same steps as in Section 3.3, so to avoid repetition, we will skip some steps and refer the reader to the appropriate sections. The ideas used below are taken from McWalter (2007) and Hulley and McWalter (2008), where the decomposition also corresponds to the Föllmer–Schweizer decomposition (see Monat and Stricker (1995)).

The first step is to find the Kunita–Watanabe decompositions of H and $\bar{U}_T$ under $\tilde{P}$ with respect to X. First define the process $\tilde{U}$ by

$$\tilde{U}_t = e^{\gamma(T-t)}\bar{U}_t, \quad 0 \le t \le T.$$

Notice that $\tilde{U}_T = \bar{U}_T$ and

$$d\tilde{U}_t = -\gamma\tilde{U}_t\, dt + \tilde{U}_t \left(\gamma\, dt + \sigma_U \left(\rho\, d\hat{W}_t^1 + \sqrt{1-\rho^2}\, d\hat{W}_t^2\right)\right)$$

$$= \tilde{U}_t \sigma_U \left(\rho \, d\hat{W}_t^1 + \sqrt{1-\rho^2} \, d\hat{W}_t^2\right).$$

So $\tilde{U}$ is a local martingale under $\tilde{P}$. Solving the SDE for $\tilde{U}$ gives (for $t > s$):

$$\tilde{U}_t = \tilde{U}_s \exp\left(-\frac{1}{2}\sigma_U^2(t-s) + \sigma_U\left(\rho(\hat{W}_t^1 - \hat{W}_s^1) + \sqrt{1-\rho^2}(\hat{W}_t^2 - \hat{W}_s^2)\right)\right).$$

Now let $f = f(\bar{U}_T) = f(\tilde{U}_T)$ be a function of $\bar{U}_T$ (or of $\tilde{U}_T$ since $\tilde{U}_T = \bar{U}_T$). We will later consider the following two special cases: $f(\bar{U}_T) = h(\bar{U}_T) = H$ and $f(\bar{U}_T) = \bar{U}_T = \tilde{U}_T$. For a given f we define $F^f : [0,T] \times [0,\infty) \longrightarrow \mathbb{R}$ by

$$F^f(t,x) := \tilde{\mathbb{E}}(f(\bar{U}_T)|\tilde{U}_t = x) = \tilde{\mathbb{E}}(f(\tilde{U}_T)|\tilde{U}_t = x).$$

Since $F^f(T,x) = f(x)$, from Itô's formula we get (refer to Section 3.3 for details)

$$f(\bar{U}_T) = F^f(0,\tilde{U}_0) + \int_0^T \frac{\partial F^f}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U \, dX_t + \int_0^T \frac{\partial F^f}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, d\hat{W}_t^2,$$

which gives a Kunita–Watanabe decomposition of $f(\bar{U}_T)$ with

$$L_T^{f,\tilde{P}} = \int_0^T \frac{\partial F^f}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, dW_t^2 \quad \text{and} \quad \varphi_t^{f,\tilde{P}} = \frac{\partial F^f}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U.$$

Here both F^f and $\frac{\partial F^f}{\partial x}$ are evaluated at $(t,\tilde{U}_t)$ for $0 \le t \le T$.

When $f(\bar{U}_T) = h(\bar{U}_T) = H$, we let

$$F^H(t,x) := \tilde{\mathbb{E}}(h(\bar{U}_T)|\tilde{U}_t = x) = \tilde{\mathbb{E}}(H|\tilde{U}_t = x).$$

The decomposition for H is then given by

$$H = \tilde{\mathbb{E}}(H) + \int_0^T \frac{\partial F^H}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U \, dX_t + \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, d\hat{W}_t^2,$$

giving

$$L_T^{H,\tilde{P}} = \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, dW_t^2 \quad \text{and} \quad \varphi_t^{H,\tilde{P}} = \frac{\partial F^H}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U$$

since $\hat{W}^2 = W^2$. While $f(\bar{U}_T) = \bar{U}_T$ gives

$$F^{\bar{U}}(t,x) = \tilde{\mathbb{E}}(\tilde{U}_T|\tilde{U}_t = x) = x \quad \text{and} \quad \frac{\partial F^{\bar{U}}}{\partial x} = 1,$$

thus

$$\begin{aligned}\bar{U}_T &= \tilde{\mathbb{E}}(\bar{U}_T) + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x}\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U \, dX_t + \int_0^T \frac{\partial F^{\bar{U}}}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, d\hat{W}_t^2 \\ &= \tilde{\mathbb{E}}(\bar{U}_T) + \int_0^T \rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U dX_t + \int_0^T \sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, d\hat{W}_t^2\end{aligned}$$

with

$$L_T^{\bar{U},\tilde{P}} = \int_0^T \sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, dW_t^2 \quad \text{and} \quad \varphi_t^{\bar{U},\tilde{P}} = \rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U \; 0 \le t \le T.$$

This gives (for each $\theta \in \mathbb{R}$)

$$L_T^{H,\theta} = L_T^{H,\tilde{P}} - \theta L_T^{\bar{U},\tilde{P}} = \int_0^T \frac{\partial F^H}{\partial x}\sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, dW_t^2 - \theta\int_0^T \sqrt{1-\rho^2}\tilde{U}_t\sigma_U \, dW_t^2$$

and

$$\varphi_t^{H,\theta} = \varphi^{H,\tilde{P}} - \theta\varphi^{\bar{U},\tilde{P}} = \left(\frac{\partial F^H}{\partial x} - \theta\right)\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U.$$

Since K_T is deterministic, by Proposition 4.1.3,

$$\frac{Z_t^{\tilde{P}}}{\tilde{Z}_t} = e^{-(K_T - K_t)} = e^{-\lambda^2(T-t)},$$

hence for every $\theta \in \mathbb{R}$

$$g(\theta) = \mathbb{E}\left(\int_0^T e^{-\lambda^2(T-t)}\, d\langle L^{H,\theta}\rangle_t\right)$$

$$= \mathbb{E}\left(\int_0^T e^{-\lambda^2(T-t)}\left(\frac{\partial F^H}{\partial x}\right)^2 \tilde{U}_t^2\sigma_U^2(1-\rho^2)\, dt\right) - 2\theta\mathbb{E}\left(\int_0^T e^{-\lambda^2(T-t)}\frac{\partial F^H}{\partial x}\tilde{U}_t^2\sigma_U^2(1-\rho^2)\, dt\right)$$

$$+\theta^2\mathbb{E}\left(\int_0^T e^{-\lambda^2(T-t)}\tilde{U}_t^2\sigma_U^2(1-\rho^2)\, dt\right).$$

Using Fubini's Theorem we get

$$g(\theta) = \sigma_U^2(1-\rho^2)\left[\int_0^T e^{-\lambda^2(T-t)}\mathbb{E}\left(\left(\frac{\partial F^H}{\partial x}\right)^2\tilde{U}_t^2\right)dt - 2\theta\int_0^T e^{-\lambda^2(T-t)}\mathbb{E}\left(\frac{\partial F^H}{\partial x}\tilde{U}_t^2\right)dt\right.$$
$$\left.+\theta^2\int_0^T e^{-\lambda^2(T-t)}\mathbb{E}\left(\tilde{U}_t^2\right)dt\right].$$

Hence the minimum ESRE occurs at

$$\hat{\theta} = \frac{\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t}\, d\langle L^{H,\tilde{P}}, L^{\bar{U},\tilde{P}}\rangle_t\right)}{\mathbb{E}\left(\int_0^T \frac{Z_t^{\tilde{P}}}{\tilde{Z}_t}\, d\langle L^{\bar{U},\tilde{P}}\rangle_t\right)} = \frac{\int_0^T e^{-\lambda^2(T-t)}\mathbb{E}\left(\frac{\partial F^H}{\partial x}\tilde{U}_t^2\right)dt}{\int_0^T e^{-\lambda^2(T-t)}\mathbb{E}\left(\tilde{U}_t^2\right)dt},$$

which again represents a 'weighted delta'. Since $V_t^{H,\hat{\theta}} = \tilde{\mathbb{E}}\left(H - \hat{\theta}\bar{U}_T|\mathcal{F}_t\right) = F^H(t,\tilde{U}_t) - \hat{\theta}\tilde{U}_t$, the expressions for $\hat{\varphi}$ and $\hat{v_0}$ are (for every $0 \le t \le T$)

$$\hat{v_0} = \tilde{\mathbb{E}}(H) - \hat{\theta}\tilde{\mathbb{E}}(\bar{U}_T) = F^H(0,\tilde{U}_0) - \hat{\theta}\tilde{U}_0$$

and

$$\hat{\varphi}_t = (\varphi_t^{H,\tilde{P}} - \hat{\theta}\varphi^{\bar{U},\tilde{P}}) - \frac{\xi_t}{\tilde{Z}_t}\left(V_{t^-}^{H,\hat{\theta}} - \tilde{\mathbb{E}}(H - \hat{\theta}\bar{U}_T) - \int_0^t \hat{\varphi}_s\, dX_s\right)$$

$$= \left(\frac{\partial F^H}{\partial x} - \hat{\theta}\right)\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U + \alpha_t\left(F^H(t,\tilde{U}_t) - \hat{\theta}\tilde{U}_t - \hat{v_0} - \int_0^t \hat{\varphi}_s\, dX_s\right).$$

We take an example when H is a call option written on $\bar{U}$.

4.3.2 Call Option

For a call option we have

$$H = h(\bar{U}_T) = \max(\bar{U}_T - K, 0) = (\bar{U}_T - K)^+.$$

It follows that

$$F^H(t,x) := \tilde{\mathbb{E}}(H|\tilde{U}_t = x) = \tilde{\mathbb{E}}((\tilde{U}_T - K)^+|\tilde{U}_t = x) = x\Phi(d_1(t,x)) - K\Phi(d_2(t,x))$$

where

$$d_1(t,x) = \frac{\ln\left(\frac{x}{K}\right) + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}} \text{ and } d_2(t,x) = d_1(t,x) - \sigma_U\sqrt{T-t}.$$

Also,

$$\frac{\partial F^H}{\partial x}(t,x) = \Phi(d_1(t,x)).$$

We now calculate both the ESRE and minimum ESRE and simplify these expressions using the same integration tricks we used in Section 3.4. First note that the expression for the minimum ESRE (for each $\theta \in \mathbb{R}$) can be written as

$$g(\theta) = \sigma_U^2(1-\rho^2)e^{-\lambda^2 T}\left[\int_0^T e^{\lambda^2 t}\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2\tilde{U}_t^2\right)dt - 2\theta\int_0^T e^{\lambda^2 t}\mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right)dt\right.$$
$$\left. + \theta^2\int_0^T e^{\lambda^2 t}\mathbb{E}\left(\tilde{U}_t^2\right)dt\right],$$

with the minimum ESRE occurring at

$$\hat{\theta} = \frac{\int_0^T e^{\lambda^2 t}\mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right)dt}{\int_0^T e^{\lambda^2 t}\mathbb{E}\left(\tilde{U}_t^2\right)dt}.$$

The optimal initial capital used for the dynamic hedge in X and the bank account is

$$\hat{v_0} = F^H(0,\tilde{U}_0) - \hat{\theta}\tilde{U}_0 = \tilde{U}_0\Phi(d_1(0,\tilde{U}_0)) - K\Phi(d_2(0,\tilde{U}_0)) - \hat{\theta}\tilde{U}_0$$

and the optimal holding in X is

$$\hat{\varphi}_t = \left(\frac{\partial F^H}{\partial x}(t,\tilde{U}_t) - \hat{\theta}\right)\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U + \alpha_t\left(F^H(t,\tilde{U}_t) - \hat{\theta}\tilde{U}_t - \hat{v_0} - \int_0^t \hat{\varphi}_s\,dX_s\right)$$

$$= \left(\Phi(d_1(t,\tilde{U}_t)) - \hat{\theta}\right)\rho\frac{\tilde{U}_t}{\sigma X_t}\sigma_U + \alpha_t\left(\tilde{U}_t\Phi(d_1(t,\tilde{U}_t)) - K\Phi(d_2(t,\tilde{U}_t)) - \hat{\theta}\tilde{U}_t - \hat{v_0} - \int_0^t \hat{\varphi}_s\,dX_s\right).$$

We will now simplify the expression for g by evaluating the expectations inside the integrals. First notice that under $\mathbb{P}$ (and for a fixed $t \in [0,T]$), $\tilde{U}_t$ can be written as

$$\tilde{U}_t = \tilde{U}_0\exp\left(-\frac{1}{2}\sigma_U^2 t + \sigma_U\left(\rho(W_t^1 - \lambda t) + \sqrt{1-\rho^2}W_t^2\right)\right)$$

$$= \tilde{U}_0\exp\left(-\left(\frac{1}{2}\sigma_U^2 + \sigma_U\lambda\rho\right)t + \sigma_U\sqrt{t}Z\right)$$

where $Z \sim N(0,1)$. This implies that

$$d_1(t,\tilde{U}_t) = \frac{\ln\left(\frac{\tilde{U}_t}{K}\right) + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}} = \frac{\ln\left(\frac{\tilde{U}_0}{K}\right) - \left(\frac{1}{2}\sigma_U^2 + \sigma\lambda\rho\right)t + \sigma_U\sqrt{t}Z + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}}$$

$$= a(t)Z + \beta(t),$$

where

$$a(t) = \sqrt{\frac{t}{T-t}} \text{ and } \beta(t) = \frac{\ln\left(\frac{\tilde{U}_0}{K}\right) - \left(\frac{1}{2}\sigma_U^2 + \sigma\lambda\rho\right)t + \frac{1}{2}\sigma_U^2(T-t)}{\sigma_U\sqrt{T-t}}.$$

Proposition 4.3.1. *For every non–negative integer k,*

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^k \tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U\lambda\rho\right)t}\int_{-\infty}^{\infty}\left(\Phi(a(t)z + b(t)\right)^k \phi(z)\,dz.$$

where $b(t) = \beta(t) + 2a(t)\sigma_U\sqrt{t}$.

Proof. Let k be a non–negative integer,

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^k \tilde{U}_t^2\right) = \mathbb{E}\left(\left(\Phi(a(t)Z + \beta(t))\right)^k \tilde{U}_0^2 \exp\left(-\left(\sigma_U^2 + 2\sigma_U\lambda\rho\right)t + 2\sigma_U\sqrt{t}Z\right)\right)$$

$$= \tilde{U}_0^2 e^{-\left(\sigma_U^2 + 2\sigma_U\lambda\rho\right)t}\mathbb{E}\left(\left(\Phi(a(t)Z + \beta(t))\right)^k e^{2\sigma_U\sqrt{t}Z}\right).$$

Now

$$\mathbb{E}\left(\left(\Phi(a(t)Z + \beta(t))\right)^k e^{2\sigma_U\sqrt{t}Z}\right) = \int_{-\infty}^{\infty}\left(\Phi(a(t)z + \beta(t))\right)^k e^{2\sigma_U\sqrt{t}z}\phi(z)\,dz$$

$$= \int_{-\infty}^{\infty}\left(\Phi(a(t)z + \beta(t))\right)^k e^{2\sigma_U\sqrt{t}z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,dz$$

$$= \int_{-\infty}^{\infty}\left(\Phi(a(t)z + \beta(t))\right)^k \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(z^2 - 4\sigma_U\sqrt{t}z\right)}\,dz$$

$$= \int_{-\infty}^{\infty}\left(\Phi(a(t)z + \beta(t))\right)^k \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left((z-2\sigma_U\sqrt{t})^2 - 4\sigma_U\sqrt{t}\right)}\,dz$$

$$= e^{2\sigma_U^2 t}\int_{-\infty}^{\infty}\left(\Phi(a(t)u + b(t))\right)^k \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\,du$$

after substituting $u = z - 2\sigma_U\sqrt{t}$. Hence

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^k \tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U\lambda\rho\right)t}\int_{-\infty}^{\infty}\left(\Phi(a(t)z + b(t)\right)^k \phi(z)\,dz.$$

■

Substituting $k = 0$ we get

$$\mathbb{E}(\tilde{U}_t^2) = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U\lambda\rho\right)t}\int_{-\infty}^{\infty}\phi(z)\,dz = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U\lambda\rho\right)t}.$$

Substituting $k = 1$ by Lemma 3.4.1 we get

$$\mathbb{E}\left(\Phi(d_1(t,\tilde{U}_t))\tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U\lambda\rho\right)t}\int_{-\infty}^{\infty}\Phi\left(a(t)z + b(t)\right)\phi(z)\,dz$$

$$= \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U \lambda\rho\right)t} \Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right).$$

Finally letting $k = 2$ and using Lemma 3.4.2 we get

$$\mathbb{E}\left(\left(\Phi(d_1(t,\tilde{U}_t))\right)^2 \tilde{U}_t^2\right) = \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U \lambda\rho\right)t} \int_{-\infty}^{\infty} \left(\Phi\left(a(t)z + b(t)\right)\right)^2 \phi(z)\, dz.$$

$$= \tilde{U}_0^2 e^{\left(\sigma_U^2 - 2\sigma_U \lambda\rho\right)t} \Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right).$$

Thus the expression $g(\theta)$ becomes

$$g(\theta) = \sigma_U^2(1-\rho^2)\tilde{U}_0^2 e^{-\lambda^2 T}\left[\int_0^T e^{\kappa t}\Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right) dt \right.$$
$$\left. -2\theta \int_0^T e^{\kappa t}\Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right) dt + \theta^2 \int_0^T e^{\kappa t}\, dt\right],$$

with minimum ESRE occurring at

$$\hat{\theta} = \frac{\displaystyle\int_0^T e^{\kappa t}\Phi\left(\frac{b(t)}{\sqrt{1+a(t)^2}}\right) dt}{\displaystyle\int_0^T e^{\kappa t}\, dt},$$

where $\kappa = \sigma_U^2 - 2\sigma_U\rho\lambda + \lambda^2$.

Note that in particular, when $\theta = 0$ we get an expression for the ESRE of the basis risk problem considered by McWalter (2007) and Hulley and McWalter (2008) as

$$\sigma_U^2(1-\rho^2)\tilde{U}_0^2 e^{-\lambda^2 T}\left[\int_0^T e^{\kappa t}\Phi_2\left(\frac{b(t)}{\sqrt{1+a(t)^2}}, \frac{b(t)}{\sqrt{1+a(t)^2}}; \frac{a(t)^2}{1+a(t)^2}\right) dt\right].$$

4.3.3 Some Numerical Results

Now consider specific values of the parameters. We let

$$\bar{U}_0 = 100,\ X_0 = 100,\ r = 10\%,\ \mu = 0.13,\ \sigma = 0.2,\ \mu_U = 0.12,\ \sigma_U = 0.25,\ T = 1,\ K = 98$$

and keep $\rho \in [-1, 1]$ as a variable parameter. We performed a simulation using Matlab (version R2013a) on 100 000 paths of X and $\bar{U}$, with trading 200 times per year. The following are 3–dimensional plots of the ESRE $g(\theta;\rho)$ as a function of $\theta \in [-1, 2]$ and $\rho \in [-1, 1]$.

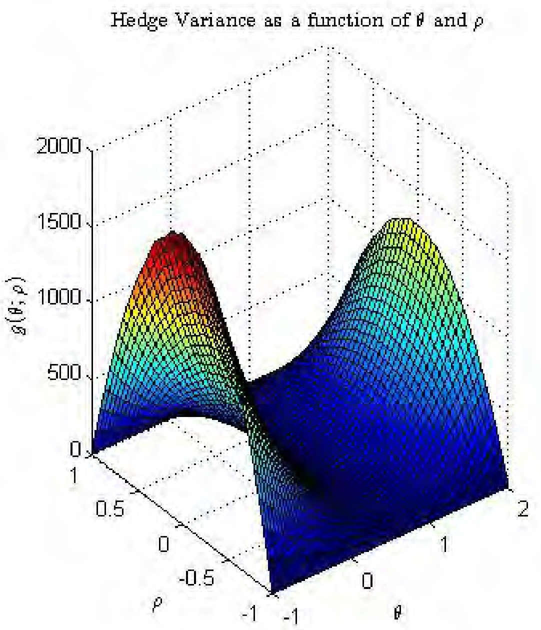

Fig. 4.1: Expected square replication error (variance) as a function of both θ and ρ from 100 000 simulations.

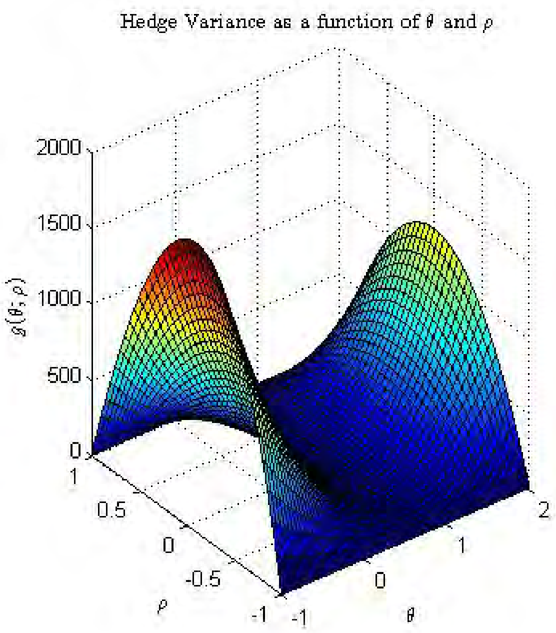

Fig. 4.2: Expected square replication error as a function of both θ and ρ from exact integration.

Figure 4.1 calculates g using Monte Carlo simulation, while Figure 4.2 uses the integral expression derived above. These graphs show the expected properties of the function g: as $|\rho|$ tends to 1, the ESRE tends to 0, which is expected since there is perfect replication in that case. Also, the two graphs agree very closely, justifying the accuracy of the simulation method used in the literature.

We also produce a graph for $g(\theta; \rho)$ as a function of ρ for two particular values of θ. We choose $\theta = \hat{\theta}$ which is the value of θ that minimizes g and $\theta = 0$, which corresponds to when the static hedge in g is not used.

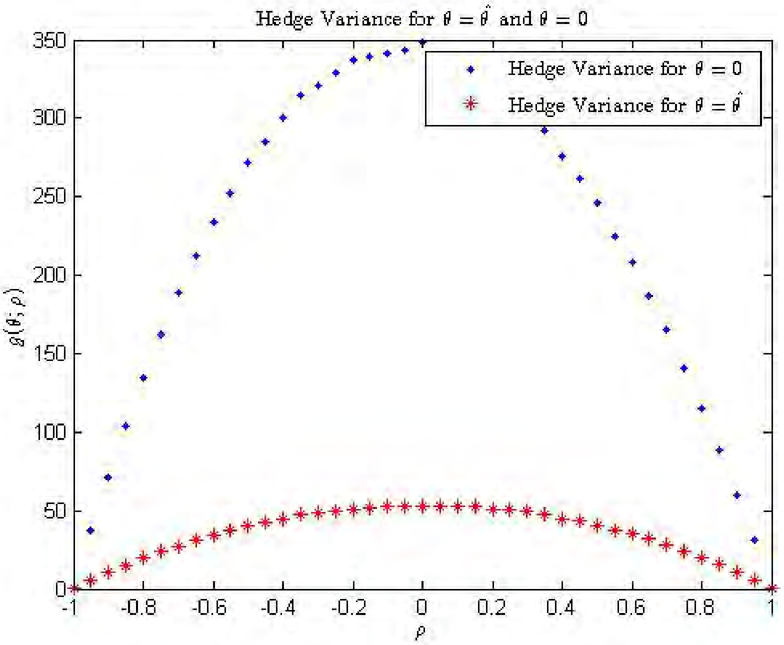

Fig. 4.3: Expected square replication error as a function of ρ for $\theta = \hat{\theta}$ and $\theta = 0$.

The first obvious thing we observe is that the graph of g when $\theta = \hat{\theta}$ is below that of $\theta = 0$. This is expected from the definition of $\hat{\theta}$. Furthermore, the graph clearly highlights the effect of the static hedge as compared to not using it at all. The static hedge in $\bar{U}$ clearly makes a huge difference in reducing the expected square hedging error. Here is a table showing the percentage reduction in the ESRE g (this is $((g(0;\rho) - g(\hat{\theta};\rho))/g(0;\rho)) \times 100\%$):

Correlation ρ	Percentage Reduction
-0.95	89.02%
0	84.5278%
0.95	79.14%

Fixing $\rho = 0.8$ we compare the distributions of the hedging error at maturity e_T for when the static hedge is used optimally ($\theta = \hat{\theta}$) and when the static hedge is not used at all ($\theta = 0$).

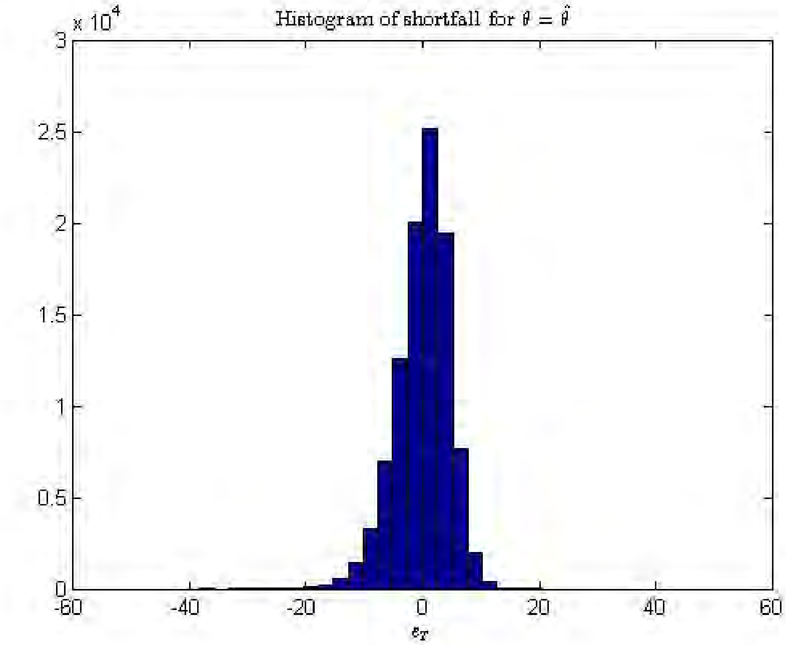

Fig. 4.4: Histogram of the terminal hedging error e_T for $\theta = \hat{\theta}$ when $\rho = 0.8$.

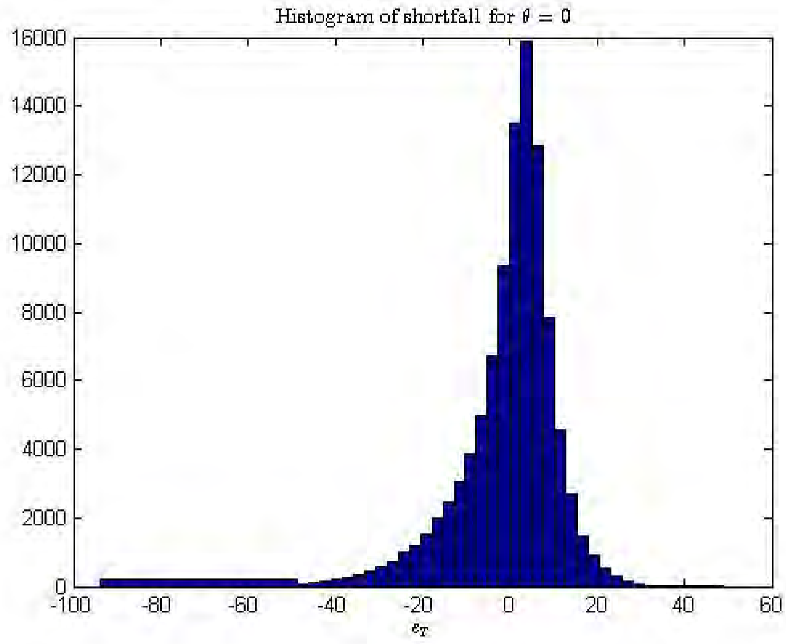

Fig. 4.5: Histogram of the terminal hedging error e_T for $\theta = 0$ when $\rho = 0.8$.

Again the histogram shows a more dispersed distribution for $\theta = 0$ than for $\theta = \hat{\theta}$. Finally, we compute two sample paths of $\hat{\varphi}$, the holding in X, when using the optimal static hedge and the holding in X when no static hedge is used. The strategy is approximated recursively using the formula from Heath *et al.* (2001b):

$$\hat{\varphi}_{t_i} = \varphi_{t_i}^{H,\hat{\theta}} + \alpha_{t_i} \left(F^H(t_i, \tilde{U}_{t_i}) - \hat{\theta}\bar{U}_{t_i} - \hat{v}_0 - \sum_{j=1}^{i-1} \hat{\varphi}_{t_j}(X_{t_j} - X_{t_{j-1}}) \right),$$

where $0 = t_0 < t_1 < \ldots < t_n = T$ are the discrete hedging times.

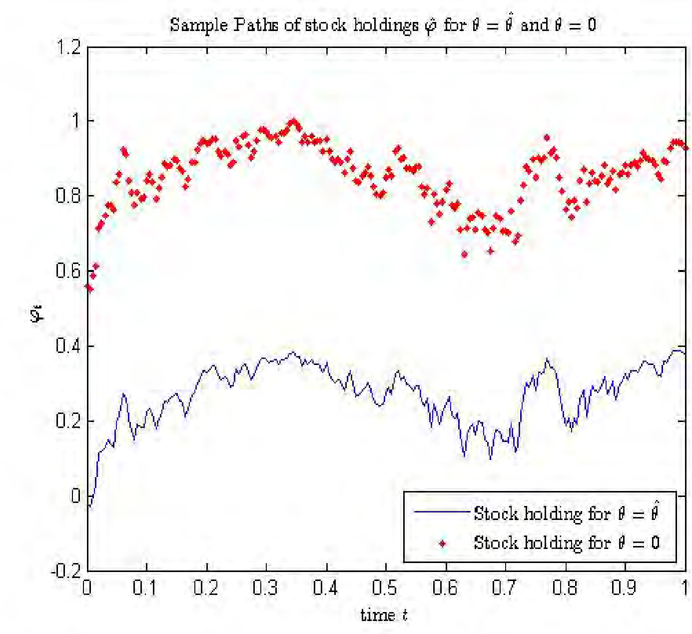

Fig. 4.6: Two sample paths for the holding $\hat{\varphi}$ in X for $\theta = \hat{\theta}$ and $\theta = 0$ when $\rho = 0.8$.

The two strategies look similar except for a constant shift.

Chapter 5

Conclusion

An interesting observation made from the results of the previous two chapters is that the static hedge in the illiquid asset proved to be very effective — reducing the expected square replication error by up to 90% in some cases. The value of the optimal holding $\hat{\theta}$ in $\bar{U}$ is not far from intuitive. If continuous trading in $\bar{U}$ was possible, we would invest $\frac{\partial F^H}{\partial x}$ (delta) units of $\bar{U}$ at all times. Now when $\bar{U}$ can only be traded at the beginning of the contract, the holding in $\bar{U}$ is a 'weighted average' of these 'deltas', weighted by $\bar{U}^2$.

An interesting extension of these results that might be worthwhile to consider is when the illiquid stock $\bar{U}$ may be traded only at fixed points in time $0 < t_1 < t_2 < \ldots < t_n$. In that case, we would then be looking for processes $\theta_0, \theta_1, \ldots, \theta_n$ such that each θ_i is at least adapted to $\mathcal{F}_{t_i}$ (we might need them to be predictable). One could further extend this to the case where the t_i's are random stopping times. The new problem would then be to choose both the times and the sizes of the trades.

One criticism of the method we employed is the use of a quadratic loss function as the objective to be minimized. This has been criticized extensively in the literature, mainly for the fact that it penalizes both profits and losses equally. The argument is that as a risk manager, you are more concerned with making huge losses than you are with making profits and thus losses should be penalized more heavily. Suggestions to use an asymmetric functional and other more general utility functions have also been made. However, the main motivation for the use of a quadratic functional given by Schweizer (2001) is the fact that one does not know before performing the optimization procedure whether one is dealing with a buyer or a seller. Also, a quadratic loss functional produces simple projection results.

We also mention that apart from solving the main hedging problem, we also managed to calculate (in expression 4.2) closed–form expressions for the expected square replication error in the case of a call option. In the literature, this quantity is usually found through simulations, but our results show that the value calculated from the simulations is very close to the exact expected square replication error, even with just 100 000 simulations.

Finally, we mention that the results on mean–variance hedging (in an incomplete market) have also been obtained by other methods. Jeanblanc *et al.* (2012) and Bobrovnytska and Schweizer (2004) use dynamic programming to arrive at similar conclusions. Also, Czichowsky *et al.* (2012) have solved a much more general version of our problem (hedging under convex trading constraints — *predictable correspondences*) using convex duality techniques.

Bibliography

Ansel, J.-P. and Stricker, C. (1994). Couverture des actifs contingents et prix maximum, *Annales de l'Institut Henri Poincaré* **30**(2): 303–315.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): 637–659.

Bobrovnytska, O. and Schweizer, M. (2004). Mean-variance hedging and stochastic control: beyond the brownian setting, *Automatic Control, IEEE Transactions on* **49**(3): 396–408.

Brezis, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer.

Conway, J. B. (1990). *A Course in Functional Analysis*, Vol. 96, Springer.

Czichowsky, C., Schweizer, M. *et al. (2012). Convex duality in mean-variance hedging under convex trading constraints,* Advances in Applied Probability **44***(4): 1084–1112.*

Davis, M. H. (2006). Optimal hedging with basis risk, From stochastic calculus to mathematical finance*, Springer, pp. 169–187.*

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing, Mathematische Annalen **300***: 463–520.*

Delbaen, F. and Schachermayer, W. (1996). The variance-optimal martingale measure for continuous processes, Bernoulli **2***(1): 81–105.*

Duffie, D. and Richardson, H. R. (1991). Mean-variance hedging in continuous time, Annals of Applied Probability **1***: 1–15.*

Durrett, R. (1996). Stochastic Calculus: A Practical Introduction*, CRC Press.*

Föllmer, H. and Schweizer, M. (1991). Hedging of contingent claims under incomplete information, in *M. H. A. Davis and R. J. Elliott (eds),* Applied Stochastic Analysis*, Gorden and Breach Science Publishers.*

Föllmer, H. and Sondermann, D. (1986). Hedging of non-redundant contingent claims, in *W. Hildenbrand and A. Mas-Colell (eds),* Contributions to Mathematical Economics*, North-Holland, chapter 12, pp. 205–223.*

Galtchouk, L. (1976). Représentation des martingales engendrées par un processus à accroissements indépendants (cas des martingales de carré intégrable, Annales de l'Institut Henri Poincaré (B) **12***(3): 199–211.*

Gouriéroux, C., Laurent, J. P. and Pham, H. (1998). Mean-variance hedging and numéraire, Mathematical Finance **8***(3): 179–200.*

Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets, Journal of Economic Theory **20***(3): 381–408.*

Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading, Stochastic Processes and their Applications **11***: 215–260.*

Harrison, J. M. and Pliska, S. R. (1983). A stochastic calculus model of continuous trading: Complete markets, Stochastic Processes and their Applications **15***: 313–316.*

Heath, D., Platen, E. and Schweizer, M. (2001a). A comparison of two quadratic approaches to hedging in incomplete markets, Mathematical Finance **11***(4): 385–413.*

Heath, D., Platen, E. and Schweizer, M. (2001b). Numerical comparison of local risk-minimisation and mean-variance hedging, in *E. Jouini, J. Cvitanić and M. Musiela (eds),* Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, chapter 14, pp. 509–537.*

Hulley, H. and McWalter, T. (2008). Quadratic hedging of basis risk.

Jacod, J. and Shiryaev, A. N. (2003). Limit Theorems for Stochastic Processes*, 2nd edn, Springer-Verlag.*

Jeanblanc, M., Mania, M., Santacroce, M., Schweizer, M. et al. (2012). Mean-variance hedging via stochastic control and bsdes for general semimartingales, The Annals of Applied Probability **22***(6): 2388–2428.*

Karatzas, I. and Shreve, S. E. (1991). Brownian Motion and Stochastic Calculus*, 2nd edn, Springer-Verlag.*

Kreps, D. M. (1981). Arbitrage and equilibrium in economics with infinitely many commodities, Journal of Mathematical Economics **8***: 15–35.*

Kunita, H. and Watanabe, S. (1967). On square integrable martingales, Nagoya Mathematical Journal **30***: 209–245.*

McWalter, T. A. (2007). Quadratic criteria for optimal martingale measures in incomplete markets.

Merton, R. C. (1973). Theory of rational option pricing, Bell Journal of Economics and Managment Science **4***(1): 141–183.*

Monat, P. and Stricker, C. (1995). Föllmer-Schweizer decomposition and mean-variance hedging for general claims, Annals of Probability **23***(2): 605–628.*

Monoyios, M. (2004). Performance of utility-based strategies for hedging basis risk, Quantitative Finance **4***: 245–255.*

Pham, H. (2000). On quadratic hedging in continuous time, Mathematical Methods of Operations Research **51***: 315–339.*

Pham, H. (2009). Continuous-time Stochastic Control and Optimization with Financial Applications*, Vol. 1, Springer.*

Pham, H., Rheinlander, T. and Schweizer, M. (1998). Mean-variance hedging for continuous processes: New proofs and examples, Finance and Stochastics **2***(2): 173–198.*

Protter, P. (2004). Stochastic Integration and Differential Equations*, 2nd edn, Springer-Verlag.*

Revuz, D. and Yor, M. (1999). Continuous Martingales and Brownian Motion*, 3rd edn, Springer-Verlag.*

Rheinlander, T. (2011). Hedging Derivatives*, Vol. 15, World Scientific.*

Rheinlander, T. and Schweizer, M. (1997). On $\mathcal{L}^2$-projections on a space of stochastic integrals, Annals of Probability **25***(4): 1810–1831.*

Schweizer, M. (1988). Hedging of Options in a General Semimartingale Model*, Diss. 8615, ETH Zürich.*

Schweizer, M. (1991). Option hedging for semimartingales, Stochastic Processes and their Applications **37***: 339–363.*

Schweizer, M. (1992). Mean-variance hedging for general claims, Annals of Applied Probability **2***(1): 171–179.*

Schweizer, M. (1994). Approximating random variables by stochastic integrals, Annals of Applied Probability **22***: 1536–1575.*

Schweizer, M. (2001). A guided tour through quadratic hedging approaches, in *E. Jouini, J. Cvitanić and M. Musiela (eds),* Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, chapter 15, pp. 538–574.*

Yor, Marc, S. L. (1978). Sous-espaces denses dans l^1 ou h^1 et reprsentation des martingales, Sminaire de probabilits de Strasbourg **12***: 265–309.*
URL: *http://eudml.org/doc/113154*

Zeidler, E. (1995). Applied Functional Analysis*, Vol. 108, Springer.*