Carl von Ossietzky Universität Oldenburg

Optimal Hedging Strategies in Robust Market Models under Capital Constraint

Von der Fakultät V für Mathematik und Naturwissenschaften

der Carl von Ossietzky Universität Oldenburg

zur Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation

von

Herrn Miguel Hinrichs

geboren am 28.08.1997 in Nordenham.

Betreuende Gutachterin: Prof. Dr. Angelika May

Zweiter Gutachter: Prof. Dr. Marcus Christiansen

Tag der Disputation: 07.05.2025

Zusammenfassung

(German)

Diese Dissertation behandelt optimale Hedging-Strategien unter Kapitalbeschränkungen im Rahmen einer robusten Marktmodellierung. Die Modellunsicherheit bei der Beschreibung von Finanzmärkten kann reduziert werden, indem nicht nur ein einzelnes Modellmaß, sondern eine Menge von Maßen simultan betrachtet wird – ein Ansatz, der unter anderem als robuste Marktmodellierung bezeichnet wird.

In dieser Arbeit wird gezeigt, dass die Konstruktion optimaler Hedging-Strategien unter Kapitalbeschränkungen in robusten Marktmodellen zur Entstehung einer Indifferenzkurve optimaler Strategien führt. Diese Indifferenzkurve veranschaulicht die Kompromisse, die zwischen verschiedenen Marktannahmen und Modellvarianten eingegangen werden müssen. Es wird gezeigt, dass alle Hedging-Strategien entlang dieser Indifferenzkurve mithilfe von worst-case Martingalmaßen hergeleitet werden können. Zudem werden Stetigkeitseigenschaften dieser worst-case Maße bewiesen, wodurch sich der numerische Aufwand zur Bestimmung der optimalen Hedging-Strategien reduziert.

Darüber hinaus werden einige Probleme dargestellt, die bei der Interpretation der mathematischen Lösungen aus praktischer Sicht auftreten, insbesondere bei der Verwendung nichtäquivalenter Maße. Es wird gezeigt, dass sich diese Probleme durch einen Modellierungsansatz mit Bildmaßen lösen lassen. Darüber hinaus wird untersucht, inwiefern sich die Ergebnisse unter diesem Ansatz von den ursprünglichen Resultaten unterscheiden beziehungsweise unter welchen Bedingungen die Ergebnisse übereinstimmen.

Abstract

(English)

This thesis investigates optimal hedging strategies under capital constraints within a robust market modelling framework. Uncertainty in financial market modelling can be reduced by considering not just a single model measure but a set of measures simultaneously – an approach known as robust market modelling.

This thesis shows that constructing optimal hedging strategies under capital constraints in robust market models leads to an indifference curve of optimal strategies. This indifference curve illustrates the trade-offs that must be made between different market assumptions and model variants. It is shown that all hedging strategies along this indifference curve can be derived using worst-case martingale measures. Furthermore, continuity properties of these worst-case measures are proven, which helps to reduce the computational effort required to determine optimal hedging strategies.

Additionally, challenges in interpreting the mathematical solutions from a practical perspective are discussed, especially in view of non-equivalent measures. It is shown that these issues can be addressed through a modelling approach based on pushforward measures. Moreover, the thesis examines how the results under this approach differ from the original results and under what conditions they coincide.

Contents

List of Figures

Chapter 1.

Introduction

1.1. Motivation

The choice of a market model plays a central role in the pricing of stock options and other derivatives. Apart from the stochastic model risk, which consists of the stochastic behaviour of the underlying, the choice of a market model introduces an additional risk, also known as uncertainty. There is always a risk that the model chosen is not a good approximation of reality or does not reflect certain risks well. It is difficult to develop models that reflect all risks, especially as financial markets become more complex, for example due to globalisation, digitalisation and climate change. This distinction between quantifiable and unquantifiable risks was made by Frank Knight as early as 1921, see Knight [40], which is why unquantifiable risks, including model risk, are also known as Knightian risks. It can therefore be useful to adopt a robust approach to pricing and hedging portfolios. In this context, a robust approach means considering not just a single market model, but a set of market models simultaneously. In this set, an investor can combine all the market models that seem realistic to him. The resulting set is called $\mathcal{P}$ in this thesis. Approaches to reduce model uncertainty seem necessary, especially after the increasing frequency of financial crises, such as the dot-com bubble in 2000, the housing bubble in 2007, the eurozone crisis in 2010, the coronavirus crisis in 2020 or, most recently, the war in Ukraine.

It is not only modelling that contributes to the security of financial markets, but also the hedging of financial products. This thesis mainly deals with the hedging of options in the above mentioned robust market model. Since robust modelling of financial markets usually leads to an incomplete market, hedging financial products in particular is a challenging mathematical problem, as there are non-replicable options and thus a difference between the option and the corresponding hedging strategy, the so called hedging error. One of the problems in financial mathematics consist in developing suitable hedging strategies in order to minimise the risk arising from this error. The hedging error and the resulting risk are mainly related to the chosen market model. Therefore, in view of the non-quantifiable model uncertainty, it seems reasonable to consider risk-minimising hedging strategies under robust market modelling. One of the most important statements in this context is the pricing-hedging duality, which states that the supremum over all prices of a financial product in an incomplete market is sufficient to set up a hedging strategy that hedges the financial product without risk, this strategy is called superhedging strategy. However, since on the one hand superhedging prices are unrealistic high and on the other hand participants in the financial market generally have to take on risk in order to earn money, the more relevant practical question is how risk can be minimised under capital constraints. In addition, the marginal utility of the capital invested in a hedging strategy decreases, which further devalues the benefit of a 100% hedging strategy and emphasises the benefits of hedging strategies under capital constraints.

From a mathematical perspective an option can be modelled as a random variable C representing a payoff in some future time T. For an adapted cadlag process S, representing the underlying stock, a hedging strategy H can be described as an S integrable predictable process. Consequently, the corresponding hedging error in T is the random variable $x + \int_0^T HdS - C$, whereas x is the initial investment and the integral is the Itô integral, representing the gains of the hedging strategy. Defining a risk measure ρ, we are facing the optimisation problem of minimising the risk $\rho(x + \int_0^T HdS - C)$ in H under the constraint $x \leq \tilde{V}_0$ simultaneously under all market models that were considered realistic, whereas $\tilde{V}_0$ represents a capital constraint.

In this thesis we aim to examine how optimal hedging strategies can be constructed in robust market models under capital constraints and especially with focus on how they behave under different market models. We will show that there exists an indifference curve of optimal hedging strategies and how we can determine all optimal solutions to the above motivated optimisation problem.

Since optimality of hedging strategies strongly depends on the chosen risk measure we will focus on regulatory important risk measures such as the value at risk and the expected shortfall, but we also show that the presented theory can be generalised up to general coherent risk measures.

1.2. Theoretical background

Due to the importance of this topic for the financial world, there is a large amount of literature examining the problems motivated above. Therefore, we will first give a brief overview of the most important publications on this topic and results that are particularly relevant for this thesis.

One of the first publications to introduce the modelling of financial market models using a geometric Brownian motion is by Samuelson [60], which significantly influenced the seminal work of Black and Scholes [8] on the Black-Scholes model and laid the foundations of classical financial market modelling using a single market measure $\mathbb{P}$. The classical Black-Scholes model was further developed due to its popularity, leading to a comprehensive theory of complete and later incomplete markets, including the fundamental theorem of asset pricing in Harrison and Kreps [28] or Delbaen and Schachermayer [17], the introduction of jump models in Merton [48], closed form solutions in stochastic volatility models in Heston [30], or contributions to interest rate modelling in Vasicek [66] or Ingersoll and Ross [13] just to present some of the most notable developments. Despite its great success and widespread use, financial crises led to increasing criticism of classical modelling for its failure to address model uncertainty. A natural response to this criticisms has been the development of robust mathematical frameworks that account for uncertainty in model parameters and probability measures.

One possible framework that was developed completely omits the physical probability measures and defines only the set of possible paths for the underlying. This approach is called a pathwise approach or model free set up and often relies on the fact that many options, especially European type, are so liquidly traded that prices can be used as an input for modelling instead of an output. One of the first works using a model free approach is Hobson [32], where price bounds of lookback options are derived without a specific probabilistic model. The approach was further generalised in numerous works, e.g. Hobson [31] with the idea of reconstructing models using option prices observed at the market. Most importantly, this approach allows for a robust version of the fundamental theorem of asset pricing, see Accacio et al [1], and a robust version of the pricing-hedging duality, see Hou and Obłój [33].

For many risk measures, however, a physical market measure is important. Since we eventually want to describe the optimality of hedging strategies with respect to a risk measure, this work uses the second approach to address model uncertainty, the quasi sure approach, which uses a set of measures $\mathcal{P}$, that each defines a possible or realistic market measure. The mathematical framework for a quasi sure stochastic analysis regarding a set of measures, in contrast to almost sure stochastics regarding a single probability measure, was developed by Laurent and Martini [18], among others. As also shown in this thesis, this approach allows for a robust fundamental theorem of asset pricing, see e.g. Biagini [5] using a pathwise continuity assumption, and a robust pricing-hedging duality, see Nutz [51] for the time continuous case or Bouchard and Nutz [9] for the time discrete case. While in this set up the extreme case of a singular set $\mathcal{P} = \{\mathbb{P}\}$ refers to the classical modelling, the other extreme case of all possible probability measures refers to the model free approach.

The pricing-hedging duality shows that the largest of all arbitrage free prices is sufficient to find an admissible superhedging strategy, i.e. a strategy that always exceeds the financial product. As already mentioned in the motivation, this approach is not suitable for real-world applications, because these superhedging prices are excessively large especially in stochastic volatility models. As a consequence Föllmer and Leukert [23, 22] presented a solution to the problem of optimal hedging under capital constraints in classical market modelling regarding shortfall probability and expected shortfall, also called quantile hedging. This idea was generalised to the class of coherent risk measures in Rudloff [58] and Nakano [49]. Note here that there are many other suitable risk measures, e.g. the quadratic hedging error that can be optimally reduced using so-called mean variance hedging strategies, that are mostly based on the closedness of the space of all attainable claims and orthogonality arguments, see Schweizer [62] for an overview. This thesis will concentrate on quantile hedging and thus on risk measures that are of regulatory importance.

Surprisingly, there has been limited research on the combination of optimal hedging under capital constraints and robust market modelling, which is the main topic of this thesis. We will examine how optimality with respect to a risk measure can be understood in robust markets and will show that there is an indifferent curve of optimal hedging strategies. Furthermore, we will present some problems that arise when comparing the mathematical results with a practical point of view and how these problems can be solved.

1.3. Outline

This thesis is divided into 5 chapters. The second chapter establishes the mathematical framework by introducing the notation used in this thesis, some basic results from stochastic analysis and some additional mathematical results that will be referred to later.

The third chapter introduces the basics for the derivation of optimal hedging strategies. First, robust modelling, as opposed to classical market modelling, is introduced and examined for typical properties such as arbitrage freedom or the fundamental theorem of asset pricing. In addition, the pricing-hedging duality is formulated and proved in a robust form. Finally, the Neyman-Pearson theory is introduced and extended in a way that is suitable for financial mathematics and robust markets.

In the fourth chapter, these results are used to derive optimal hedging strategies not only for single market models, but also in the robust market models introduced. Optimality is understood here from a regulatory perspective, i.e. with a focus on the shortfall probability and the expected shortfall, but can also be extended to general coherent risk measures. In particular, it is shown how optimal strategies can be determined for robust market models, that an indifference curve of optimal hedging strategies emerges and how these hedging strategies can be constructed.

In the fifth chapter, some examples of optimal hedging strategies are derived both analytically for the Black-Scholes model and numerically in more complex models where an analytical solution does not exist or is too complex to calculate.

The sixth and final chapter describes problems that can arise when interpreting the mathematical solution from a real-world perspective and presents a way in which mathematical modelling can be adapted to solve these problems. In addition, a summary of the work and an outlook on further open research questions are presented.

Chapter 2.

Preliminaries

We first establish the basic mathematical concepts and notations that will be used throughout, before moving on to the core topics of this thesis, namely the robust modelling of financial markets and optimal hedging strategies in this framework. In particular, we introduce key definitions from stochastic analysis that form the basis for modelling financial markets and formulating hedging strategies.

A precise specification of the underlying probability space and stochastic processes is crucial, as different conventions exist in the literature. Since our results depend on these definitions, we explicitly outline the relevant notation, with a focus on stochastic processes, the measurable spaces they are defined on and the corresponding probability measures. These concepts are essential for modelling stock price dynamics and play a central role in our theoretical framework.

This section provides only a brief overview of the mathematical framework used in this thesis and introduces some results that will be referred to in the thesis. Since most of the following definitions and constructions are based on the works of Jacod and Shiryaev [36] and Protter [56], further information can be found in these books, that can be considered as standard literature in the stochastic analysis.

2.1. Stochastic processes

Definition 2.1.1. *Filtration*
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be a right continuous family of sub-σ-algebras of $\mathcal{F}$, i.e $\mathcal{F}_s \subseteq \mathcal{F}_t$ for every $s \leq t$ and $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$. Then

$$\mathbb{F} := (\mathcal{F}_t)_{t\in\mathbb{R}_+}$$

is called a filtration and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a filtered probability space.

For the following definitions we assume the filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ to be fixed.

Definition 2.1.2. *Adaptation*
Let $X : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ be a stochastic process. If $X(\cdot, t)$ is a $\mathcal{F}_t$-measurable random variable for every $t \in \mathbb{R}_+$, X is called an $\mathbb{F}$-adapted process.

Definition 2.1.3. *Cadlag*
A process $(X_t)_{t\geq 0}$ is called a cadlag process if the function of sample paths $t \mapsto X_t$ is right continuous with existing left limits almost surely. Equivalently, a process is called caglad process if $t \mapsto X_t$ is left continuous with existing right limits almost surely.

Definition 2.1.4. *Predictability*
A stochastic process $(X_t)_{t\geq 0}$ is called predictable if it is adapted to the σ-algebra generated by all caglad processes, the predictable σ-field.

Definition 2.1.5. *Process of finite variation*
Let X be a cadlag process. The process $V_a^b(X) = \sup_{\pi \in P} \sum_{t_i \in \pi} |X_{t_i} - X_{t_{i+1}}|$, where P is the set of all partitions on $[a, b]$, is called the total variation process of X. The set $\mathcal{V}$ is the set of all adapted cadlag processes with finite total variation, i.e. a total variation process that is finite on each compact interval of $[0, \infty)$.

Proposition 2.1.6.
For every continuous process $A \in \mathcal{V}$ there exists a predictable process θ such that

$$\int_0^t \theta dA$$

is increasing and $|\theta| = 1$.

Proof. The proposition is a direct consequence of the Hahn decomposition theorem, where the predictability of θ follows from the continuity of A. For a detailed proof, see Lemma 4 in Lowther [44]. $\square$

2.2. Semimartingales

Definition 2.2.1. *Martingales, supermartingales and submartingales*
Let M be an adapted process with $\mathbb{E}[|M_t|] < \infty$ and let $s \leq t$. M is called martingale with respect to the filtration $\mathbb{F}$ if

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s.$$

We call M a supermartingale if $\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$ or a submartingale if $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s$. Note that these properties depend on the choice of the measure $\mathbb{P}$ and the filtration $\mathbb{F}$.

Definition 2.2.2. *Localized classes*
If $\mathcal{C}$ is a class of processes, we define the larger class $\mathcal{C}_{loc}$ as the set of all processes such that there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \to \infty$ and such that for every $X \in \mathcal{C}_{loc}$ the stopped process X^{τ_n} is in $\mathcal{C}$.

Proposition 2.2.3.
A local martingale bounded from below is a supermartingale.

Proof. Without loss of generality, let the local martingale be non-negative. This can be assumed since a constant can be added to any lower bounded process until zero is a lower bound. Then, using Fatou's lemma and the localizing sequence $(\tau_n)_{n\in\mathbb{N}}$, we get:

$$\mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[\lim_{n\to\infty} X_t^{\tau_n}|\mathcal{F}_s] \leq \liminf_{n\to\infty} \mathbb{E}[X_t^{\tau_n}|\mathcal{F}_s] = X_s.$$

□

Definition 2.2.4. *Semimartingale*
A process X is called a semimartingale if it can be decomposed into $X = X_0 + M + A$ with X_0 a finite valued $\mathcal{F}_0$-measurable random variable, M a local martingale with $M_0 = 0$ and $A \in \mathcal{V}$.

Remark 2.2.5.
Note that the space of all semimartingales is stable under many transformations, e.g. stopping, localization and, most importantly for this thesis, under absolutely continuous changes of measures. This space is one of the most important sets in the stochastic analysis mainly because it is the largest class of processes that can be used to construct stochastic integrals, see Jacod and Shiryaev [35].

Remark 2.2.6.
Note that in general one often uses cadlag versions of semimartingales, whose existence can be guaranteed in complete filtered probability spaces using a right continuous filtration under mild conditions, see Jacod and Shiryaev [36]. Since we will work on the Skorokhod space throughout this work, as introduced in Section 2.5, we will assume every semimartingale to be cadlag if not stated otherwise.

Definition 2.2.7. *Special Semimartingale*
Let X be a semimartingale with $X = X_0 + M + A$. If there exists a decomposition such that A is predictable, then X is called a special semimartingale. This decomposition is unique.

Definition 2.2.8. *Quadratic covariation*
Let X and Y be cadlag semimartingales. The quadratic covariation $[X, Y]$ is defined by

$$[X, Y] = XY - \int X_- dY - \int Y_- dX.$$

Definition 2.2.9. *Predictable quadratic variation*
The predictable quadratic variation, or sometimes called conditional quadratic variation, $\langle X, Y\rangle$ can be defined as the compensator to the quadratic variation process $[X, Y]$.

Remark 2.2.10.
A good example to see the difference between quadratic variation and predictable quadratic variation is a Poisson process N of intensity $\lambda > 0$. For this process the quadratic variation can be described as the process itself:

$$[N, N]_t = \sum_{s \leq t} (\Delta N_s)^2 = \sum_{s \leq t} \Delta N_s = N_t.$$

Since the Poisson process is not predictable, it should be clear that the predictable quadratic variation will differ from $[N, N]$. Indeed, the compensator of a Poisson process is known as

$$\langle N, N\rangle_t = \lambda t,$$

which is predictable because it is deterministic and the process $N_t - \lambda t$ is a martingale.

Definition 2.2.11. *Compensator*
Let X be an adapted cadlag process. The compensator of X, denoted X^p, is the predictable process of finite variation such that $X - X^p$ is a local martingale.

Remark 2.2.12.
The existence of a compensator is not generally guaranteed, but by definition the compensator exists for all special semimartingales.

Proposition 2.2.13.
A semimartingale X with bounded jumps, i.e. $\Delta X \leq c$ for a $c \in \mathbb{R}$, is a special semimartingale.

Proof. This result is proved as Lemma I.4.24 in Jacod and Shiryaev [36]. □

Definition 2.2.14. *Characteristics of semimartingales*
Let S be a $\mathbb{R}^d$-valued semimartingale and let $h : \mathbb{R}^d \to \mathbb{R}^d$ be an arbitrary bounded function with $h(x) = x$ for x close to zero. We call a triplet $(B(h), C, \nu)$ the characteristics of S relative to the truncation function h if

(a) $B(h)$ is a predictable and adapted cadlag process in $\mathbb{R}^d$ of finite variation, such that

$$S(h) = S_0 + M + B(h)$$

for a martingale M, where $S(h)$ is defined as

$$S(h)_t = S_t - \sum_{u \leq t} (\Delta S_u - h(\Delta S_u)),$$

which can be interpreted as the semimartingale S without large jumps. Since $S(h)$ is a special semimartingale, by Proposition 2.2.13, $B(h)$ is uniquely defined.

(b) C is a continuous, adapted process of locally finite variation in $\mathbb{R}^{d \times d}$, such that

$$C = \langle S^c, S^c \rangle,$$

where S^c is the continuous martingale part of S (as defined in Proposition 4.27 in Jacod and Shiryaev [36]).

(c) ν is the compensator of the measure

$$\mu^S(\omega; dt, dx) = \sum_{u \geq 0} \mathbb{1}_{\{\Delta S_u \neq 0\}} \delta_{(u, \Delta S_u(\omega))}(dt, dx),$$

where δ is the Dirac measure.

Proposition 2.2.15.
Let M be a local martingale. M is constant on an interval if and only if $[M, M]$ is constant on the same interval.

Proof. The quadratic variation of a process M can be represented via

$$[M, M]_t = M_t^2 - 2\int_0^t M dM.$$

Now let $[M, M]$ be constant on an interval $[s, t]$. First, $M_{u\wedge t} - M_{u\wedge s}$ is also a local martingale with constant quadratic variation on $[s, t]$. This is true because the properties of a local martingale transfer to a sum of two local martingales and because the following equation holds for the quadratic variation:

$$[M_{u\wedge t} - M_{u\wedge s}, M_{u\wedge t} - M_{u\wedge s}] = [M_{u\wedge t}, M_{u\wedge t}] + [M_{u\wedge s}, M_{u\wedge s}] - 2[M_{u\wedge s}, M_{u\wedge t}],$$

where each of the right-hand summands is constant on $[s, t]$.

Let $X_u := M_{u\wedge t} - M_{u\wedge s}$, then using the above representation of the quadratic variation, we get

$$\begin{aligned}
\mathbb{E}[X_t^2|\mathcal{F}_s] &= \mathbb{E}\left[[X, X]_t + 2\int_0^t X dX \middle| \mathcal{F}_s\right] \\
&= \mathbb{E}\left[[X, X]_s + 2\int_0^s X dX + 2\int_s^t X dX \middle| \mathcal{F}_s\right] \\
&= [X, X]_s + 2\int_0^s X dX + 2\underbrace{\mathbb{E}\left[\int_s^t X dX \middle| \mathcal{F}_s\right]}_{=0} \\
&= X_s^2
\end{aligned}$$

and thus

$$0 = \mathbb{E}[X_u^2 - X_s^2|\mathcal{F}_s] = \mathbb{E}[(M_{u\wedge t} - M_{u\wedge s})^2|\mathcal{F}_s],$$

where $M_{u\wedge t} = M_{u\wedge s}$ almost surely. □

Definition 2.2.16.

Let X be a real-valued cadlag semimartingale. The stochastic exponential of X, denoted by $\mathcal{E}(X)$, is defined as the solution of the stochastic differential equation

$$d\mathcal{E}(X) = \mathcal{E}(X)_- dX \text{ with } \mathcal{E}(X)_0 = 1.$$

Proposition 2.2.17.

Let X be a cadlag semimartingale. For the stochastic exponential it holds that

$$\mathcal{E}(X)_t = \exp(X_t - X_0 - \frac{1}{2}\langle X^c, X^c\rangle_t) \prod_{u\leq t}(1 + \Delta X_u)\exp(-\Delta X_u).$$

In addition, $\mathcal{E}(X)$ is a local martingale if X is a local martingale.

Proof. This statement can be found in Theorem I.4.61 in Jacod and Shiryaev [36]. □

Proposition 2.2.18.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let S be an integrable $\mathbb{R}^d$-valued random variable. Let $Z \geq 0$ be a uniformly bounded random variable with $\mathbb{P}(Z \geq c) > 0$ for a $c > 0$. Then for any sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ there exists a random variable $X \geq 0$ and a uniformly bounded $\mathcal{G}$-measurable random variable $Y \geq c$, such that

$$X \leq Y, \ \mathbb{P}(X = Y) > 0 \text{ and } \mathbb{E}[XS|\mathcal{G}] = \mathbb{E}[ZS|\mathcal{G}].$$

Proof. A proof can be found in [46]. □

2.3. Stopping times and decompositions

Definition 2.3.1. *Totally inaccessible stopping time*
A stopping time τ is called totally inaccessible if for any predictable stopping time σ we have

$$\mathbb{P}(\tau = \sigma < \infty) = 0.$$

Proposition 2.3.2.
For every semimartingale X there exist countable sequences $(\tau_n)_{n\in\mathbb{N}}$ of predictable stopping times and $(\sigma_n)_{n\in\mathbb{N}}$ of totally inaccessible, such that the the union of both sequences contains all jumps of X

$$\{t \in \mathbb{R}_+ : \Delta X_t \neq 0\} \subseteq \bigcup_{n=1}^{\infty} \{\tau_n\} \cup \bigcup_{n=1}^{\infty} \{\sigma_n\}.$$

In particular one can decompose X into

$$X_t = X_0 + X_t^c + \sum_{n=1}^{\infty} \Delta X_{\tau_n} \mathbb{1}_{[\![\tau_n, \infty[\![} + \sum_{n=1}^{\infty} \Delta X_{\sigma_n} \mathbb{1}_{[\![\sigma_n, \infty[\![}.$$

Proof. Proposition I.2.26 in Jacod and Shiryaev [36] states that the jump times of quasi-left-continuous semimartingales can be represented by a sequence of totally inaccessible stopping times. Furthermore, Proposition 3.15 in Cerny and Ruf [11] states that every semimartingale can be decomposed into a quasi-left-continuous semimartingale and a process with jumps only at predictable stopping times. □

Proposition 2.3.3.
Let X be a process with compensator X^p and let τ be a stopping time.

(a) If τ is predictable, then $\Delta(X^p)_\tau = \mathbb{E}[\Delta X_\tau | \mathcal{F}_{\tau-}]$.

(b) If τ is totally inaccessible, then $\Delta(X^p)_\tau = 0$ almost surely.

Proof. The first statement is proved as Theorem I.2.28 in Jacod and Shiryaev [36]. The second statement follows from Proposition I.2.24 in Jacod and Shiryaev [36], since the compensator is predictable. □

Proposition 2.3.4.
Let M be a cadlag and bounded martingale and let τ be a predictable stopping time. Then

$$\mathbb{E}[\mathbb{1}_{\tau<\infty} \Delta M_\tau] = 0,$$

which implies

$$\mathbb{E}[\Delta M_\tau | \mathcal{F}_{\tau-}] = 0.$$

Proof. This is a direct consequence of Proposition 2.3.3. □

2.4. Changes of measures

Proposition 2.4.1. *Radon-Nikodym theorem*
Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathbb{P}$ and Q be two σ-finite measures. If $\mathbb{P} \ll Q$, i.e. $\mathbb{P}$ is absolutely continuous with respect to Q, then there exists a function $f : \Omega \to \mathbb{R}$ such that

$$\mathbb{P}(A) = \int_A f dQ.$$

We use the notation $f = \frac{d\mathbb{P}}{dQ}$. The function f is unique up to $\mathbb{P}$ null sets.

Remark 2.4.2.
For easier notation in the rest of this thesis we want to use an extension of the usual Radon-Nikodym derivative. For measures $\mu \ll \nu$ defined on Ω the corresponding Radon-Nikodym derivative $\frac{d\mu}{d\nu}$ is usually defined ν-a.s. For the rest of this thesis we define the Radon-Nikodym derivative to be zero on the complement $supp(\nu)^c$ when $\Omega \neq supp(\nu)$, where *supp* denotes the support of measures. In particular this implies that we set $\frac{0}{0} = 0$ in this context.

Remark 2.4.3.
In view of stochastic processes, i.e. on a complete, filtered probability space, we can define the uniformly integrable martingale $Z \in L^1(\mathbb{P})$ as a right continuous version of

$$Z_t = \mathbb{E}^{\mathbb{P}}\left[\frac{dQ}{d\mathbb{P}}\middle|\mathcal{F}_t\right].$$

If $\mathcal{Q} \sim \mathbb{P}$, i.e. Q and $\mathbb{P}$ are equivalent, then $\frac{dQ}{d\mathbb{P}} = \left(\frac{d\mathbb{P}}{dQ}\right)^{-1}$. We also adopt the notation

$$\left.\frac{dQ}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathbb{E}^{\mathbb{P}}\left[\frac{dQ}{d\mathbb{P}}\middle|\mathcal{F}_t\right].$$

Proposition 2.4.4.
Let Q and $\mathbb{P}$ be equivalent probability measures with $Z_t = \mathbb{E}^{\mathbb{P}}[\frac{dQ}{d\mathbb{P}}|\mathcal{F}_t]$ and let X be an adapted cadlag process. Then X is a Q-martingale if and only if ZX is a $\mathbb{P}$-martingale.

Proof. See Chapter 6 in Protter [56]. □

Proposition 2.4.5. *Girsanov theorem*
Let $\mathbb{P}$ and Q be equivalent probability measures with $Z_t = \mathbb{E}^{\mathbb{P}}[\frac{dQ}{d\mathbb{P}}|\mathcal{F}_t]$ and let X be a semimartingale with $X = X_0 + M + A$ under $\mathbb{P}$, i.e. M is a local martingale under $\mathbb{P}$ and A has finite variation under $\mathbb{P}$. Then X has the decomposition $X = X_0 + \tilde{M} + \tilde{A}$ with $\tilde{A}$ a finite variation process under Q and

$$\tilde{M}_t = M - \int_0^t \frac{1}{Z_s} d[Z, M]_s,$$

which is a Q-local martingale.

Proof. This proposition is proven in Chapter 3 as Theorem 20 in Protter [56]. □

Proposition 2.4.6.
Let X_t be a $\mathcal{F}_t$-measurabel random variable. For two probability measures $Q \sim \mathbb{P}$ it holds that

$$\mathbb{E}^Q[X_t|\mathcal{F}_s] = \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{P}}{dQ}\middle|\mathcal{F}_s\right]\mathbb{E}^{\mathbb{P}}\left[\left.\frac{dQ}{d\mathbb{P}}\right|_{\mathcal{F}_t} X_t\middle|\mathcal{F}_s\right].$$

Proof. This statement is also known as Bayes rule for conditional expectations, see Chapter 7 in Liptser and Shiryaev [43]. □

Proposition 2.4.7.
Let $A \in \mathcal{V}$ be a continuous predictable process and let $B \in \mathcal{V}$ be a continuous, predictable and increasing process. If

$$\int_0^t \theta dA = 0 \text{ a.s.}$$

for every $t \in [0,T]$ and every $\theta \geq 0$, which satisfies $\int_0^t \theta dB = 0$, then there exists a predictable process α such that

$$A_t = A_0 + \int_0^t \alpha dB \quad \text{and} \quad \int_0^t |\alpha| dB < \infty.$$

Proof. The following proof uses the idea of the proof of Lemma 4 in [45], but generalizes the statement. For clearer notation, let $A_0 = B_0 = 0$ and let the processes A and B be of integrable variation. Let two measures μ and ν be the signed measures on the predictable σ-algebra on the measure space $(\mathbb{R}_+ \times \Omega, \mathbb{P})$ with

$$\mu(D) = \mathbb{E}\left[\int_0^T \mathbb{1}_D dA\right] \quad \text{and} \quad \nu(D) = \mathbb{E}\left[\int_0^T \mathbb{1}_D dB\right].$$

The integral $\int_0^\infty \mathbb{1}_D dB$ is always positive since B is increasing. Let D be a set in the predictable σ-algebra on $\Omega \times [0,T]$ with $\nu(D) = 0$. It follows

$$\mathbb{E}\left[\int_0^T \mathbb{1}_D dB\right] = \nu(D) = 0 \text{ and thus } \int_0^T \mathbb{1}_D dB = 0 \text{ almost sure.}$$

So, by the assumption of the proposition, it holds that $\int_0^T \mathbb{1}_D dA = 0$ and hence $\mu(D) = 0$. So μ is a measure dominated by ν and the Radon-Nikodym density $\alpha := \frac{d\mu}{d\nu}$ can be defined. For this density and a predictable process θ it holds

$$\mathbb{E}\left[\int \alpha\theta dB\right] = \int \alpha\theta d\nu = \int \theta d\mu = \mathbb{E}\left[\int \theta dA\right],$$

because as predictable processes A and B are locally bounded and allow the application of the dominated convergence theorem.
Now we can define a process $M := \int \alpha dB - A$, for which it holds that

$$\mathbb{E}\left[\int \theta dM\right] = \mathbb{E}\left[\int \alpha\theta dB - \int \theta dA\right] = 0.$$

Since the process θ is an arbitrary predictable and positive process, M is a martingale. To see this, choose $\theta = \mathbb{1}_F \mathbb{1}_{(s,t]}$ for a $F \in \mathcal{F}_s$ and $s \leq t$. Thus, $\mathbb{E}[\mathbb{1}_F(M_t - M_s)] = 0$. Moreover, since M is also predictable by construction, M is constant as a predictable martingale of finite variation, i.e. $M = 0$ and $A = \int \alpha dB$. □

2.5. Notation and market modelling

Definition 2.5.1. *Skorokhod space*
The Skorokhod space from $\mathbb{R}_+$ to an metric space E is the set of all cadlag paths $\omega : \mathbb{R}_+ \to E$.

Proposition 2.5.2. *Skorokhod topology*
If E is a Polish space, there exists a metrizable topology on the Skorokhod space, such that the Skorokhod space from $\mathbb{R}_+$ to E is Polish.

Proof. For the case $E = \mathbb{R}_+$, which is sufficient for this work, this statement is proven in Chapter IV, Theorem 1.14 in Jacod and Shiryaev [35]. In Kurtz and Ethier [41] the result is generalised on arbitrary Polish spaces E. □

Throughout this thesis, if not stated otherwise, e.g. in Chapter 6, we assume that the sample space Ω is the Skorokhod space with $E = \mathbb{R}_+$. This choice is consistent with most of the literature in financial mathematics. Furthermore, we use the Skorokhod topology and the induced metric such that Ω is a Polish space.

As motivated before, we will work with different probability measures. Therefore, we need a metric on spaces of measures and some properties that remain preserved from the measurable space.

Definition 2.5.3. *Prokhorov metric*
Let (X, d) be a metric space with Borel σ-algebra $\mathcal{B}(X)$. The Prokhorov metric on the space of all probability measures defined on $(X, \mathcal{B}(X))$ is defined as

$$\pi(\mu, \nu) = \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ and } \nu(A) \leq \mu(A^\epsilon) + \epsilon \text{ for all } A \in \mathcal{B}(X)\},$$

where A^ϵ is the ϵ-neighbourhood of A,

$$A^\epsilon = \{x \in X : \exists \tilde{x} \in A \text{ with } d(x, \tilde{x}) < \epsilon)\}.$$

Lemma 2.5.4. *Prokhorov theorem*
Let (X, d) be a metric space. If (X, d) is separable and complete, then the set of all probability measures on the measurable space $(X, \mathcal{B}(X))$ provided with Prokhorov metric is again a complete space.

Proof. This lemma is proven as Theorem 9.2 in Van Gaans [65]. □

Lemma 2.5.5.
A subset of a complete metric space is complete if and only if it is closed.

Proof. This is a fundamental statement one can find for example in Meise [47]. □

Definition 2.5.6. *Value process*
Let H be a left continuous and thus predictable process representing a trading strategy and let S be a semimartingale. we define the value process of the trading strategy H with initial investment x, denoted (H, x), as

$$V_t^{H,x} = x + \int_0^t H dS.$$

Definition 2.5.7. *Quasi-sure*
Let $\mathcal{P}$ be a set of probability measures. The expression $\mathcal{P}$-almost surely is supposed to be understood as $\mathbb{P}$-almost surely for every $\mathbb{P} \in \mathcal{P}$.

Definition 2.5.8. *Stochastic integral*
Throughout this work, for a locally square-integrable and adapted process H and a semimartingale S, we sometimes shorten the notation of stochastic integrals as follows:

$$(H \cdot S)_t = \int_0^t H dS.$$

Note that throughout this work we define $(H \cdot S)_0 = 0$.

Definition 2.5.9. *Superreplicating trading strategy*
For a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, let $C : \Omega \to \mathbb{R}_+$ be a $\mathcal{F}_T$-measurable claim. We call a trading strategy H with initial investment x a superreplication of C or a superhedging strategy of C under $\mathbb{P}$ if

$$V_T^{H,x} \geq C \quad \mathbb{P}\text{-a.s.}$$

Definition 2.5.10. *Equivalent martingale measures*
Throughout this thesis, for a measure $\mathbb{P}$, representing a market model, and a semimartingale we define $\mathcal{Q}^{\mathbb{P}}$ to be the set of all equivalent local martingale measures. For a robust market, i.e. a set $\mathcal{P}$, we define $\mathcal{Q}$ to be the set of all local martingale measures equivalent to a $\mathbb{P} \in \mathcal{P}$, i.e. $\mathcal{Q} = \bigcup_{\mathbb{P} \in \mathcal{P}} \mathcal{Q}^{\mathbb{P}}$.

Chapter 3.

Mathematical foundations

3.1. Robust market modelling

In this section we give a general definition of the term "robust market model" used in this work, which leads to the definition of the set $\mathcal{P}$ of market measures on a measurable space $(\Omega, \mathcal{F})$, i.e. the Skorokhod space in this work. The results and definitions in this section are based on Biagini [5]. We give a definition of arbitrage in the setting of robust market models $\mathcal{P}$ and prove that these robust market models are still arbitrage-free if they contain only arbitrage-free market models. Furthermore, under some additional assumptions we show that if the robust model $\mathcal{P}$ is free of arbitrage, then we can even conclude that every single market model $\mathbb{P} \in \mathcal{P}$ is free of arbitrage. In addition, we show that even the fundamental theorem of asset pricing in a more general form can be applied to robust modelling.

Definition 3.1.1. *Robust market model*
Let I be an index set and let $\mathbb{P}_i$ be a market measure for every $i \in I$. The set containing all these markets $\mathcal{P} = \bigcup_{i \in I}\{\mathbb{P}_i\}$ is called a robust market model.

Remark 3.1.2.
In the following, due to easier notation, we will often consider countable index sets and write $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2, ...\}$, but note that the results and proofs in this chapter and especially in Section 3.2 also hold for general robust market models.

This definition can be interpreted as uncertainty of market choices. All of the markets contained in $\mathcal{P}$ are valid models that have to be taken into account, but it is unknown which one best reflects reality. A good and currently interesting example of this idea is uncertainty about future interest rates. As long as it is unknown whether future interest rates remain high or fall again, market participants want to hedge against both cases. In this case one can use the robust model $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ which contains a model $\mathbb{P}_1$ that reflects a low interest rate and a model $\mathbb{P}_2$ that represents high interest rates.

Definition 3.1.3. *Admissible strategies*
An investment strategy H is said to be admissible under a market model $\mathbb{P}$ if there exists a lower bound $c \in \mathbb{R}$ on the stochastic integral $(H \cdot S)$, such that

$$(H \cdot S)_t \geq c \ \mathbb{P}\text{-a.s. } \forall t \in \mathbb{R}_+.$$

An investment strategy H is said to be admissible under a robust market $\mathcal{P}$ if there exists a lower bound $c \in \mathbb{R}$ for the stochastic integral $(H \cdot S)$, such that

$$(H \cdot S)_t \geq c \ \mathcal{P}\text{-a.s. } \forall t \in \mathbb{R}_+.$$

Demanding admissibility of investment strategies prevents doubling strategies and other arbitrage strategies that require infinitely high capital.

Definition 3.1.4. *Arbitrage in robust markets*
A robust market model $\mathcal{P}$ admits no arbitrage of the first kind if there does not exist a positive, $\mathcal{F}_T$-measurable claim $C \geq 0$ with $\mathbb{P}(C > 0) > 0$ for at least one $\mathbb{P} \in \mathcal{P}$ such that there exists an admissible superhedging-strategy H without initial investment.
Equivalently, this means for every claim $C \geq 0$:

$$V_T^{H,0} \geq C \ \mathcal{P}\text{-a.s. for an admissible strategy } H \ \Rightarrow C = 0 \ \mathcal{P}\text{-a.s.}$$

We also say: $NA_1(\mathcal{P})$ holds or $\mathcal{P}$ satisfies the NA_1-condition.

Proposition 3.1.5.
In the case of single market models $\mathcal{P} = \{\mathbb{P}\}$ Definition 3.1.4 coincides with the usual arbitrage definition of single market models, i.e. the existence of an admissible arbitrage strategy $(H, 0)$ with $\mathbb{P}(V_T^{H,0} \geq 0) = 1$ and $\mathbb{P}(V_T^{H,0} > 0) > 0$.

Proof. If a market model allows for arbitrage according to Definition 3.1.4, then there exists a claim $C \geq 0$ with $\mathbb{P}(C > 0) > 0$ and an admissible investment strategy $(H, 0)$ with $V_T^{H,0} \geq C$. In this case $(H, 0)$ is an arbitrage strategy with $\mathbb{P}(V_T^{H,0} \geq 0) = 1$ and $\mathbb{P}(V_T^{H,0} > 0) > 0$.

On the other hand, if there is an admissible investment strategy $(H, 0)$ with $\mathbb{P}(V_T^{H,0} \geq 0) = 1$ and $\mathbb{P}(V_T^{H,0} > 0) > 0$, then we can define $C := V_T^{H,0}$ which is a claim not almost surely zero and superhedgeable with $(H, 0)$. $\square$

It remains to show that as long as we only use arbitrage-free market models the robust model remains arbitrage-free, which is crucial for working with robust models.

Theorem 3.1.6.
Let $\mathcal{P}$ be a set of market models. If every market model $\mathbb{P} \in \mathcal{P}$ admits no arbitrage of the first kind, then $NA_1(\mathcal{P})$ holds. If the underlying S is continuous $\mathcal{P}$-a.s. and $NA_1(\mathcal{P})$ holds, then there is no arbitrage of the first kind in every single market model $\mathbb{P}$. In short:

- $NA_1(\{\mathbb{P}\})\ \forall \mathbb{P} \in \mathcal{P} \Rightarrow NA_1(\mathcal{P})$.
- $NA_1(\mathcal{P}) \Rightarrow NA_1(\{\mathbb{P}\})\ \forall \mathbb{P} \in \mathcal{P}$, if S is continuous $\mathcal{P}$-a.s.

Proof. This proof follows in large parts ideas of Biagini [5]. It should first be noted that any $\mathcal{P}$-admissible strategy is also a $\mathbb{P}$-admissible strategy for every $\mathbb{P} \in \mathcal{P}$, but admissibility for one $P \in \mathcal{P}$ does not imply $\mathcal{P}$-admissibility.

Then, the assertion $NA_1(\{\mathbb{P}\})\ \forall \mathbb{P} \in \mathcal{P} \Rightarrow NA_1(\mathcal{P})$ follows directly, since every $\mathcal{P}$-admissible strategy must also be admissible for every $\mathbb{P} \in \mathcal{P}$, but by assumption every $\mathbb{P}$-admissible strategy should be arbitrage-free. So let H be a $\mathcal{P}$-admissible strategy, then it follows directly that H is also a $\mathbb{P}$-admissible strategy for every $\mathbb{P} \in \mathcal{P}$ and finally that H is not an arbitrage strategy since $\mathbb{P}$ is arbitrage-free.

For the second statement, we show that under the assumption of pathwise continuity of S, a $\mathbb{P}$-admissible arbitrage strategy can be used to construct a $\mathcal{P}$-admissible arbitrage strategy.

Let $NA_1(\mathcal{P})$ be satisfied and S be continuous $\mathcal{P}$-almost surely. Suppose $NA_1(\{\mathbb{P}\})$ does not hold for a $\mathbb{P} \in \mathcal{P}$. Then there exists a $\mathbb{P}$-admissible arbitrage strategy H with

$$(H \cdot S)_t \geq c \ \mathbb{P}\text{-a.s.},\ (H \cdot S)_T \geq 0 \ \mathbb{P}\text{-a.s. and } \ \mathbb{P}((H \cdot S)_T > 0) > 0.$$

Moreover, it is even possible to find a sequence of $\mathbb{P}$ (but not yet $\mathcal{P}$) arbitrage strategies $(H^n)_{n \in \mathbb{N}}$, for which it holds that

$$\begin{aligned} &(H^n \cdot S)_t \geq -\frac{1}{n}\ \forall t \in [0,T],\ \mathbb{P}-\text{a.s.}, \\ &\mathbb{P}((H^n \cdot S)_T > 0) > 0 \ \ \forall n \in \mathbb{N}. \end{aligned}$$

For this statement, see the Definition of no free lunch with vanishing risk (NFLVR) in Debaen and Schachermayer [17] and note that a market that does not satisfy the NA_1-condition also cannot satisfy the NFLVR property. We can define $C := \lim_{n \to \infty} (H^n \cdot S)_T$, which defines a random variable with $C \geq 0$ and $\mathbb{P}(C > 0) > 0$.

Let $\xi^n := \inf\{t \in \mathbb{R}_+ | V_t^{H^n,1/n} < 0\} \wedge T$ be a stopping time which is predictable due to the continuity of S. Now we can define trading strategies $G^n := H^n \mathbb{1}_{]\!]0,\xi^n]\!]}$. Due to continuity of S it follows that $V_{\xi^n}^{G^n,1/n} = 0$ and hence $V_t^{G^n,1/n} \geq 0$ for all $t \in [0,T]$ $\mathcal{P}$-almost surely.

For the $\mathcal{F}_T$-measurable claim

$$\tilde{C} := \inf_{n \in \mathbb{N}} (G^n \cdot S)_T$$

it holds that $\tilde{C} \geq 0$ $\mathcal{P}$-almost surely and $\mathbb{P}(\tilde{C} > 0) > 0$, since $\tilde{C} \geq C$ $\mathbb{P}$-almost surely with $\xi^n = T$ $\mathbb{P}$-almost surely. With the trading strategy G^n we can show

$$V_T^{G^n,1/n} \geq \tilde{C} \quad \text{and} \quad V_t^{G^n,1/n} \geq 0 \quad \forall n \in \mathbb{N} \quad \mathcal{P}\text{-almost surely,}$$

which means that the claim $\tilde{C}$ can be super-replicated $\mathcal{P}$-almost surely without initial investment, which is a contradiction to $NA_1(\mathcal{P})$. □

Remark 3.1.7.
In general the property $NA_1(\{\mathbb{P}\})$ for all $\mathbb{P} \in \mathcal{P}$ does not follow from $NA_1(\mathcal{P})$ without the assumption of continuity on S, i.e.

$$NA_1(\mathcal{P}) \not\Rightarrow NA_1(\mathbb{P}) \ \forall \mathbb{P} \in \mathcal{P},$$

as the following example shows:

Let the interest rate be zero and let $\mathbb{P}_1$ and $\mathbb{P}_2$ be measures under which S is constant almost surely with the exception of a jump at a deterministic stopping time t_0. At t_0 there is $S_{t_0-} - S_{t_0} = 1$ $\mathbb{P}_1$-a.s. and $S_{t_0-} - S_{t_0} = -1$ $\mathbb{P}_2$-a.s. Both of these market models allow for an arbitrage strategy as it is possible to buy or sell the underlying before t_0 and undo the trade afterwards, which is sufficient to show that $\{\mathbb{P}_1\}$ and $\{\mathbb{P}_2\}$ are not arbitrage free with Proposition 3.1.5. But for the robust market $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ the property $NA_1(\mathcal{P})$ holds as there is no prior knowledge about the direction of the jump at t_0.

In addition to absence of arbitrage, the existence of equivalent local martingale measures follows from the NA_1-condition similar to results for single market models, leading to a robust form of the fundamental theorem of asset pricing. Recall that the set of all local martingale measures equivalent to a measure $\mathbb{P} \in \mathcal{P}$ is denoted by $\mathcal{Q}^{\mathbb{P}}$, while $\mathcal{Q}$ denotes the set of all local martingale measures equivalent to a $\mathbb{P} \in \mathcal{P}$.

Now it remains to show under which additional conditions $NA_1(\mathcal{P})$ is strong enough to conclude the existence of equivalent local martingale measures.

To do so, we need to sufficiently weaken the definition of equivalent local martingale measures and introduce a stopping time τ at which the market is absorbed into a final artificial state, that we call Δ, i.e. we enlarge the image of S to $\mathbb{R}_+ \cup \{\Delta\}$. This allows us to define the equivalent martingale measures up to the stopping time τ, which is a sufficiently weaker property than that of the equivalent local martingale measures. This artificial market state Δ is supposed to be a state that cannot be left again but does not occur under the physical measure, so there is no interpretation from a real-world perspective. Throughout this section, we will adapt the measurable space $(\Omega, \mathcal{F})$ to this setup, but note that we show at the end of this section that we can drop this somewhat unintuitive notation for the rest of this thesis.

Definition 3.1.8. *Market state Δ and stopping time τ*
For this section let $(\Omega, \mathcal{F})$ be the Skorokhod space from $\mathbb{R}_+$ to $\mathbb{R}_+ \cup \{\Delta\}$, with the assumption that $\omega_s = \Delta$ implies $\omega_t = \Delta$ for every $t \geq s$. The stopping time τ is defined as the timepoint at which the market reaches the state Δ:

$$\tau := \inf\{t \in \mathbb{R}_+ : S_t = \Delta\}.$$

We further assume that

$$\mathbb{P}(\tau < \infty) = 0 \ \forall \mathbb{P} \in \mathcal{P}.$$

Remark 3.1.9.
With Lemma 2.5.2, this adaptation of the Skorokhod space remains a Polish space, since the space $\mathbb{R}_+ \cup \{\Delta\}$ is Polish, because every countable union of disjoint Polish spaces remains Polish. One can take the adapted metric

$$\tilde{d}(x,y) = \begin{cases} d(x,y), & x, y \neq \Delta \\ 1, & x = \Delta, y \neq \Delta \text{ or } x \neq \Delta, y = \Delta \\ 0, & x = y = \Delta. \end{cases}$$

Definition 3.1.10. *σ-algebras prior τ*
For a σ-Algebra $\mathcal{F}_t$ on Ω we define the σ-Algebra $\tilde{\mathcal{F}}_t$ on $\{t \leq \tau\} \subseteq \Omega$ to be

$$\tilde{\mathcal{F}}_t := \{A \cap \{t \leq \tau\} | A \in \mathcal{F}_t\}.$$

Definition 3.1.11. *Equivalence prior τ*
Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be measures on $(\Omega, \mathcal{F})$. These two measures are called equivalent prior τ, $\mathbb{P}_1 \sim_\tau \mathbb{P}_2$, if they are equivalent in the usual sense on the measurable space $(\{t \leq \tau\}, \tilde{\mathcal{F}})$.

Definition 3.1.12. *Equivalent local martingale measures prior τ*
Let Q and $\mathbb{P}$ be measures on the filtered measurable space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0})$. Q is called a prior τ equivalent local martingale measure to $\mathbb{P}$, if

$$Q \sim_\tau \mathbb{P}$$

and there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ with

$$\tau_n < \tau, \quad \lim_{n\to\infty} \tau_n = \tau \ Q\text{-a.s.},$$
$$(S_{t\wedge\tau_n})_{t\in\mathbb{R}_+} \text{ is a } (\mathbb{F}_+, Q)\text{-martingale } \forall n \in \mathbb{N}.$$

$\mathcal{Q}_\tau^{\mathbb{P}}$ is defined as the set of all prior τ equivalent local martingale measures Q to $\mathbb{P}$. $\mathcal{Q}_\tau$ is defined as the set of all measures Q, that are prior τ equivalent local martingale measures to a $\mathbb{P} \in \mathcal{P}$.

Since this definition has weaker assumptions than usual local martingale measures, we have a larger set of measures $\mathcal{Q}^{\mathbb{P}} \subseteq \mathcal{Q}_\tau^{\mathbb{P}}$ for each $\mathbb{P} \in \mathcal{P}$ and consequently $\mathcal{Q} \subseteq \mathcal{Q}_\tau$. Using these weaker definitions of equivalence allows for a robust fundamental theorem of asset pricing, very similar to the usual, well-known fundamental theorem of asset pricing in single market modelling:

Theorem 3.1.13. *Robust fundamental theorem of asset pricing*
Let $\mathcal{P}$ be a set of market measures. If $\mathcal{Q}_\tau^{\mathbb{P}} \neq \emptyset$ for every $\mathbb{P} \in \mathcal{P}$, then the market $\mathcal{P}$ is arbitrage-free. If S is almost surely pathwise continuous for every $\mathbb{P} \in \mathcal{P}$, then the converse also holds. That is:

- $\mathcal{Q}_\tau^{\mathbb{P}} \neq \emptyset \ \forall \mathbb{P} \in \mathcal{P} \Rightarrow NA_1(\mathcal{P})$.
- $NA_1(\mathcal{P}) \Rightarrow \mathcal{Q}_\tau^{\mathbb{P}} \neq \emptyset \ \forall \mathbb{P} \in \mathcal{P}$, if S is continuous $\mathcal{P}$-almost sure.

Before proving this theorem, we need two additional auxiliary results. Lemma 3.1.14 shows that the presented weaker form of the martingale property still implies a supermartingale property. Lemma 3.1.15 shows that the assumption NA_1 implies a relation between the continuous martingale part M and the finite variation part A of the Doob-Meyer decomposition of S. In the literature this assumption is sometimes called structure condition on S and becomes important for many other topics, for example in the context of mean-variance hedging. While Lemma 3.1.14 can be found as a statement in [5], but is supplemented here by a proof, Lemma 3.1.15 is already proved in [37], but differs significantly in the proof.

Lemma 3.1.14.

Let Q be a prior τ equivalent local martingale measure for $\mathbb{P}$ with an increasing sequence of stopping times (τ_n) according to Definition 3.1.12, then for any self-financing portfolio H, the value process $V^{H,x}_{t\wedge\tau_n}$ is a local Q-martingale and $V^{H,x}_t 1_{[\![0,\tau[\![}$ is a Q-supermartingale.

Proof. The local martingale property of $V^{H,x}_{t\wedge\tau_n}$ follows directly from the definition of the value process, because $S_{t\wedge\tau_n}$ is a martingale and thus $(H\cdot S^{\tau_n})$ is a local martingale for H locally bounded, see Chapter 3, Theorem 17 in Protter [56]. It holds that

$$V^{H,x}_{t\wedge\tau_n} = x + \int_0^{t\wedge\tau_n} H_t \; dS_t = x + \int_0^t H_t \; dS_{t\wedge\tau_n},$$

where the right integral is a local martingale due to the martingale property of $S_{t\wedge\tau_n}$. The second assertion follows directly from the following calculation. Let $s\le t$, then it holds that

$$\begin{aligned}
E^Q[V^{H,x}_t 1_{[\![0,\tau[\![}(t)|\mathcal{F}_s] &= E^Q[V^{H,x}_{t\wedge\tau} 1_{[\![0,\tau[\![}(t)|\mathcal{F}_s]\\
&= \lim_{n\to\infty} E^Q[V^{H,x}_{t\wedge\tau_n} 1_{[\![0,\tau[\![}(t)|\mathcal{F}_s]\\
&\le \lim_{n\to\infty} E^Q[V^{H,x}_{t\wedge\tau_n} 1_{[\![0,\tau[\![}(s)|\mathcal{F}_s]\\
&= \lim_{n\to\infty} E^Q[V^{H,x}_{t\wedge\tau_n}|\mathcal{F}_s] 1_{[\![0,\tau[\![}(s)\\
&\le \lim_{n\to\infty} V^{H,x}_{s\wedge\tau_n} 1_{[\![0,\tau[\![}(s)\\
&= V^{H,x}_s 1_{[\![0,\tau[\![}(s),
\end{aligned}$$

where the first inequality follows from the fact that $\{t<\tau_n\}\subseteq\{s<\tau_n\}$ for all $s\le t$. The second inequality follows from the fact that every local martingale that is bounded from below is a supermartingale, see Proposition 2.2.3. □

Lemma 3.1.15.

Let S be a semimartingale under $\mathbb{P}$. That is, there exists a decomposition $S=M+A$ into a process of local finite variation A and a local martingale M. If the market model $\mathbb{P}$ is free of arbitrage, then there exists a predictable process θ with

$$A_t = \int_0^t \theta d[M,M] \quad \mathbb{P}\text{-a.s.} \tag{3.1.1}$$

Proof. If we assume that such a θ does not exist, then the conclusion of Proposition 2.4.7 is violated. Thus, by contraposition, there must exist at least one predictable, bounded, and positive process θ and a $t \in [0,T]$ such that

$$\int_0^t \theta d[M,M] = 0 \text{ a.s. and } \mathbb{P}\left(\int_0^t \theta dA \neq 0\right) > 0.$$

So there must be an interval $U \subseteq [0,T]$ with $d[M,M]_t = 0$ almost surely for $t \in U$ while dA is not almost surely zero, since the left expression is increasing. Furthermore, by Proposition 2.1.6, for any process of finite variation, among others, for $X := A\mathbb{1}_U$, there exists a predictable process η with $|\eta_t| = 1$ for every t such that $\int_U \eta dA$ is increasing. Thus, we can define a process $\alpha := \eta\mathbb{1}_U$ which suffices:

$$\int_0^t \alpha d[M,M] = 0 \quad \forall t \in [0,T],$$

$$\int_0^{\cdot} \alpha dA \text{ is increasing,}$$

$$\mathbb{P}\left(\int_0^T \alpha dA > 0\right) > 0, \text{ since } \mathbb{P}\left(\int_0^t \alpha dA \neq 0\right) > 0.$$

Now it remains to show that $\int_U \alpha dM = 0$ holds so that α is an arbitrage strategy. However, by Proposition 2.2.15, $dM_t = 0$ holds for every $t \in U$, since $d[M,M]_t = 0$ holds here. Thus, $\int_0^t \alpha dA = \int_0^t \alpha dS$ and α is an arbitrage strategy, which is a contradiction to the assumed freedom of arbitrage.

So there must be a predictable process θ under which it holds that

$$A_t = \int_0^t \theta d[M,M] \quad \mathbb{P}\text{-a.s.}$$

□

With these definitions and auxiliary results, it is now possible to prove a fundamental theorem of asset pricing in a robust setting, closely related to the results in Biagini [5], but generalised to the context of possible jump models. The proof therefore also follows the idea in Biagini [5].

Proof of Theorem 3.1.13. Let $\mathcal{Q}_\tau^{\mathbb{P}} \neq \emptyset \; \forall \mathbb{P} \in \mathcal{P}$. We show that the robust financial market $\mathcal{P}$ is then free of arbitrage. To do this, let C be a positive $\mathcal{F}_T$-measurable claim with a super-replicating portfolio H whose initial value is $x = 0$. For any arbitrary but fixed market measure $\mathbb{P} \in \mathcal{P}$, it now holds that there exists a measure Q and an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \nearrow \tau$ such that $S_{t\wedge\tau_n}$ is a Q-martingale, see Definition 3.1.12. As a direct consequence, it is also true that

$$V^{H,0}_{t\wedge\tau_n} = 0 + \int_0^{t\wedge\tau_n} H\ dS$$

is a local Q-martingale for all $n \in \mathbb{N}$. By Lemma 3.1.14 it follows that

$$\mathbb{E}^Q[C\mathbb{1}_{\{T<\tau_n\}}] \leq \mathbb{E}^Q[V^{H,0}_T \mathbb{1}_{\{T<\tau_n\}}] \leq \mathbb{E}^Q[V^{H,0}_0 \underbrace{\mathbb{1}_{\{0<\tau_n\}}}_{=1}] = 0 \quad \forall n \in \mathbb{N},$$

so $\mathbb{E}^Q[C\mathbb{1}_{\{T<\tau\}}] = 0$. With the fact $C \geq 0$ it also follows $Q(C > 0,\ T < \tau) = 0$.

It remains to show that the claim C is almost surely zero even under the physical measure $\mathbb{P}$. But from the equivalence $Q \sim_\tau \mathbb{P}$ and $\mathbb{P}(T < \tau) = 1$ we can conclude that

$$\mathbb{P}(C > 0) = \mathbb{P}(C > 0,\ T < \tau) = Q(C > 0,\ T < \tau) = 0,$$

which proves the first assertion: $\mathcal{Q}^{\mathbb{P}}_\tau \neq \emptyset\ \forall \mathbb{P} \in \mathcal{P} \Rightarrow NA_1(\mathcal{P})$.

For the second part let $NA_1(\mathcal{P})$ hold and let S be almost surely pathwise continuous under every $\mathbb{P} \in \mathcal{P}$. By the previously proved Theorem 3.1.6, due to the continuity of S, it directly holds that

$$NA_1(\mathbb{P}) \quad \forall \mathbb{P} \in \mathcal{P}.$$

So it remains to show that for every $\mathbb{P} \in \mathcal{P}$ there is

$$NA_1(\mathbb{P}) \Rightarrow \mathcal{Q}^{\mathbb{P}}_\tau \neq \emptyset.$$

For this, we first prove the existence of a local martingale $Y \geq 0$ with respect to the predictable filtration $(\mathbb{F}_+, \mathbb{P})$ with $Y_0 = 1$ and the property that YS is a local $(\mathbb{F}, \mathbb{P})$-martingale.

By Lemma 3.1.15, for the decomposition $S = M + A$ there is a predictable process θ with

$$A_t = \int_0^t \theta d[M, M] \quad \mathbb{P}\text{-a.s.}$$

In view of the Girsanov theorem, the process

$$Y_t := \exp\left(-\int_0^t \theta dS + \frac{1}{2}\int_0^t \theta^2 d[M, M]\right)$$

is a good candidate to satisfy the required property that Y and YS are local martingales. This can be proven directly using Itô's lemma and the partial integration rule for stochastic integrals as follows:

$$\begin{aligned}
Z &:= ln(Y) = -\int_0^t \theta d(M+A) + \frac{1}{2}\int_0^t \theta^2 d[M,M],\\
dZ &= -\theta dM - \theta dA + \frac{1}{2}\theta^2 d[M,M],\\
dY &= d(e^Z) = e^Z dZ + \frac{1}{2}e^Z (dZ)^2\\
&= Y\big(-\theta dM \quad \underbrace{-\theta dA}_{\overset{(3.1.1)}{=}\theta^2 d[M,M]} \quad + \frac{1}{2}\theta^2 d[M,M]\big) + \frac{1}{2}Y\theta^2 d[M,M]\\
&= -Y\theta dM,
\end{aligned}$$

where M is the local martingale from the decomposition $S = M + A$. Thus, Y is indeed a local martingale.

For the process (YS), the partial integration rule for stochastic integrals together with equation (3.1.1) shows:

$$\begin{aligned}
d(YS) &= YdS + SdY + d[Y,S]\\
&= YdS + SdY + dYdS\\
&= YdM + YdA - SY\theta dM - Y\theta dMdM - Y\theta dAdM\\
&\overset{(3.1.1)}{=} (Y - SY\theta - Y\theta dA)dM + Y\theta d[M,M] - Y\theta d[M,M]\\
&= (Y - SY\theta - Y\theta dA)dM.
\end{aligned}$$

So, by the same argument as before, YS is a local martingale.

This proves that there exists a local martingale Y such that YS is a local martingale under $\mathbb{P}$. The process Y can now be used to construct a measure $Q \sim_\tau \mathbb{P}$, where Y is the density of Q prior τ.

As shown in Theorem 3.1 in Perkowski et al. [53], due to the existence of Y there exists a predictable stopping time ξ and a measure Q^0 on $(\Omega, \mathcal{F}_{\xi-})$, with properties

$$\begin{aligned}
&\mathbb{P}(\xi = \infty) = 1,\\
&Q^0(A_\rho \cap \{\rho < \xi\}) = \mathbb{E}^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho}] = \mathbb{E}^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho}\mathbb{1}_{\{\rho<\xi\}}],
\end{aligned} \tag{3.1.2}$$

for every predictable stopping time ρ and every set $A_\rho \in \mathcal{F}_{\rho+}$. The last equation follows from the fact that $\mathbb{1}_{\{\rho<\xi\}} = 1$ $\mathbb{P}$-almost sure, since $\mathbb{P}(\xi < \infty) = 0$.

Next, we show that the stopping time ξ coincides with the given stopping time τ, i.e. is indistinguishable from it, which means that we already have a local martingale measure equivalent prior τ.

It holds that $A_\rho \cap \{\rho < \xi \wedge \tau\} \in \mathcal{F}_{\rho+}$ for every $A_\rho \in \mathcal{F}_{\rho+}$, since τ and ξ and hence also $\xi \wedge \tau$ are predictable stopping times. Thus, the equation (3.1.2) also holds for $A_\rho \cap \{\rho < \xi \wedge \tau\}$, yielding:

$$\begin{aligned}
Q^0(A_\rho \cap \{\rho < \xi \wedge \tau\}) =& Q^0(A_\rho \cap \{\rho < \xi \wedge \tau\} \cap \{\rho < \xi\}) \\
=& E^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho \cap \{\rho < \xi \wedge \tau\}} \mathbb{1}_{\{\rho < \xi\}}] \\
=& E^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho} \mathbb{1}_{\{\rho < \xi \wedge \tau\}}] \\
=& E^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho} \mathbb{1}_{\{\rho < \xi\}}] \\
=& Q^0(A_\rho \cap \{\rho < \xi\}),
\end{aligned}$$

for every $A_\rho \in \mathcal{F}_{\rho+}$, where the second last equation holds since $\xi = \tau = \infty$ $\mathbb{P}$-almost surely. With $A_\rho = \Omega$ it finally follows:

$$Q^0(\rho < \xi \wedge \tau) = Q^0(\rho < \xi) \text{ and therefore } \xi \leq \tau \; Q^0 \text{ a.s.}$$

Furthermore, for $\rho = 0$ and $A_\rho = \Omega$:

$$Q^0(\xi > 0) = \mathbb{E}^{\mathbb{P}}[Y_0 1_\Omega 1_{0<\xi}] = \mathbb{E}^{\mathbb{P}}[Y_0] = 1 \text{ and thus } 0 < \xi \leq \tau \; Q^0 \text{ a.s.} \tag{3.1.3}$$

Last, it remains to show that another measure Q can be constructed on $(\Omega, \mathcal{F}_{\tau-})$ via the measure Q^0, which still satisfies the above properties and additionally satisfies $Q(\xi = \tau) = 1$. Under this measure, the stopping time ξ coincides with the time τ at which the financial market jumps to the fixed final state Δ.

To do this, take the mapping $\psi : \Omega \to \Omega$ with $\psi_t(\omega) = \omega_t \mathbb{1}_{t<\xi(\omega)} + \Delta \mathbb{1}_{t \geq \xi(\omega)}$ for every $\omega \in \Omega$. The mapping thus generates a path under which the financial market already jumps to the Δ state at time ξ. On the one hand, this mapping is $\mathcal{F}_{\xi-}$-measurable as a composition of measurable mappings, on the other hand, under this mapping

$$\tau \circ \psi = \xi \quad Q^0\text{-almost sure,}$$

because $\xi \leq \tau$ is already shown and $\tau \circ \psi = \inf\{t \in [0, T] : \psi_t(\omega) = \Delta\} \leq \xi$ holds by definition of ψ. Moreover, ψ satisfies the property $\psi \circ \psi = \psi$ because

$$\begin{aligned}
(\psi \circ \psi)(\omega) =& \psi(\omega_t \mathbb{1}_{\{t<\xi(\omega)\}} + \Delta \mathbb{1}_{\{t \geq \xi(\omega)\}}) \\
=& (\omega_t \mathbb{1}_{\{t<\xi(\omega)\}} + \Delta \mathbb{1}_{\{t \geq \xi(\omega)\}}) \mathbb{1}_{\{t<\xi(\omega)\}} + \Delta \mathbb{1}_{\{t \geq \xi(\omega)\}} \\
=& \omega_t \mathbb{1}_{\{t<\xi(\omega)\}} + \Delta \underbrace{\mathbb{1}_{\{t \geq \xi(\omega)\} \cap \{t<\xi(\omega)\}}}_{=0} + \Delta \mathbb{1}_{\{t \geq \xi(\omega)\}} = \psi(\omega).
\end{aligned}$$

With this property it follows

$$\tau \circ \psi = \tau \circ \psi \circ \psi = \xi \circ \psi.$$

Thus Q can be defined as

$$Q := Q^0 \circ \psi^{-1}.$$

For this Q it holds that

$$Q(\xi < \tau) = Q^0(\psi^{-1}(\xi < \tau)) = Q^0(\xi \circ \psi < \tau \circ \psi) = Q^0(\emptyset) = 0,$$

so together with $\xi \leq \tau$ there is $Q(\xi = \tau) = 1$. Q thus satisfies equation (3.1.2) with τ instead of ξ as a stopping time:

$$Q(A_\rho \cap \{\rho < \tau\}) = \mathbb{E}^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho}] = \mathbb{E}^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho} \mathbb{1}_{\rho < \tau}]. \tag{3.1.4}$$

With this equation, the equivalence prior τ follows directly, because it holds that

$$Q(A_\rho \cap \{\rho < \tau\}) = 0 \Rightarrow \mathbb{E}^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho \cap \{\rho < \tau\}}] = 0 \overset{Y \geq 0}{\Rightarrow} \mathbb{E}^{\mathbb{P}}[\mathbb{1}_{A_\rho \cap \{\rho < \tau\}}] = \mathbb{P}(A_\rho \cap \{\rho < \tau\}) = 0,$$
$$\mathbb{P}(A_\rho \cap \{\rho < \tau\}) = 0 \Rightarrow E^{\mathbb{P}}[\mathbb{1}_{A_\rho \cap \{\rho < \tau\}}] = 0 \Rightarrow E^{\mathbb{P}}[Y_\rho \mathbb{1}_{A_\rho \cap \{\rho < \tau\}}] = Q(A_\rho \cap \{\rho < \tau\}) = 0.$$

Thus Q is the desired measure with $Q \sim_\tau \mathbb{P}$ and it only remains to show that there exists a sequence of stopping times $(\tau_n)_{n \in \mathbb{N}}$ as required in Definition 3.1.12 with $\tau_n \nearrow \tau$ Q-almost surely and under which $S_{t \wedge \tau_n}$ is a $(\mathbb{F}, Q)$-martingale.

Let $(\tau_n)_{n \in \mathbb{N}}$ be the localizing sequence of Y, which exists since Y is a local $\mathbb{P}$-martingale. For the limit of this sequence $\tilde{\tau} := \lim_{n \to \infty} \tau_n$ it holds that $\tilde{\tau} = \tau = \infty$ $\mathbb{P}$-almost surely. Moreover, it follows directly via equivalence $Q \sim_\tau \mathbb{P}$ that $Q(\tilde{\tau} < \tau) = 0$, note here that $\{\tilde{\tau} < \tau\} \in \tilde{\mathcal{F}}_\tau$. Thus $\tilde{\tau} \geq \tau$ Q-almost sure. Furthermore, due to the martingale property, either $\tilde{\tau} \leq \tau$ or $(YS)_t = \Delta$ for every $t \in [0, T]$. However, since it has already been shown in (3.1.3) that $\tau > 0$ Q-almost surely, $S_0 \neq \Delta$ and thus $\tilde{\tau} = \tau$ Q-almost sure.

So Q satisfies all the required properties:

- $Q \sim_\tau \mathbb{P}$, according to equation (3.1.4).
- There exists a sequence $\tau_n < \tau$ with $\lim_{n \to \infty} \tau_n = \tau$ Q-almost surely.
- $S_{t \wedge \tau_n}$ is a $(\mathbb{F}_+, Q)$-martingale since (YS) is a local martingale whose localizing sequence converges to τ and Y is the density of Q with respect to $\mathbb{P}$ prior τ.

So there is $Q \in \mathcal{Q}_\tau^{\mathbb{P}}$ and hence $\mathcal{Q}_\tau^{\mathbb{P}} \neq \emptyset$. □

Corollary 3.1.16.
Since the continuity for the second statement in Theorem 3.1.13 is used only for the application of Theorem 3.1.6, the requirement of continuity of S is omitted for single market models $\mathcal{P} = \{\mathbb{P}\}$, that is:

$$NA_1(\{\mathbb{P}\}) \Rightarrow \mathcal{Q}^{\mathbb{P}}_{\tau} \neq \emptyset.$$

Corollary 3.1.17.
Since every equivalent local martingale measure is also a prior τ equivalent local martingale measure, the assumption of the existence of equivalent local martingale measures is sufficient for freedom of arbitrage.

In conclusion, we have shown that even without the need for continuity assumptions, i.e. even in the context of jump models, typical results such as the absence of arbitrage and the fundamental theorem of asset pricing can to some extend be extended to robust market modelling. On the other hand, there are some statements that cannot be generalized into the robust context without continuity assumptions. For the rest of this work for the sake of simplicity in notation, we assume the existence of equivalent martingale measures, i.e. $\mathcal{Q}^{\mathbb{P}} \neq \emptyset$ for each $\mathbb{P} \in \mathcal{P}$, which is sufficient to work in an arbitrage-free context, as discussed in this chapter. Note, however, that the assumption of the existence of equivalent martingale measures can be weakened.

3.2. Pricing-hedging duality for robust market models

In this section we prove a robust form of the well-known superhedging theorem, which states that for an initial investment $x = \sup_{Q \in \mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[C]$ there exists an admissible strategy H with $V_T^{H,x} \geq C$ $\mathbb{P}$-almost surely, see e.g. Theorem 2.4.2 in Delbaen and Schachermayer [16]. In other words, one can say that the largest arbitrage-free price for a claim is sufficient to find a hedging strategy such that there is no risk left. In this section we generalise this statement to a robust market model $\mathcal{P}$ instead of a single market model $\mathbb{P}$. In order to formulate and prove the robust superhedging theorem, we first need to introduce a robust form of the optional decomposition theorem, two auxiliary lemmas and some requirements on the robust market model.

3.2.1. Robust optional decomposition theorem

The following results are mainly based on Nutz [51] and the presented proofs follow the ideas presented in that paper. The following statements and proves are adapted to our notation, supplemented with explanations and presented for the sake of completeness.

Definition 3.2.1. *Dominating diffusion*
Let $\mathbb{P}$ be a probability measure and S a d-dimensional semimartingale with characteristics (B, C, ν) under $\mathbb{P}$. S has dominating diffusion under $\mathbb{P}$ if it holds that

$$\int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu_t(dx, dt) \ll dC_t^{i,i}, \quad \mathbb{P}\text{-a.s. for every } i \leq d,$$

whereas $C^{i,i}$ defines the i-th entry on the main diagonal of C. This implies that for a diffusion of almost surely zero, it is almost surely impossible for there to be a jump.

Theorem 3.2.2. *Optional decomposition theorem*
Let $\mathcal{R}$ be a set of equivalent local martingale measures to a physical measure $\mathbb{P}$ and an adapted cadlag process S. Every process X that is a local supermartingale with respect to every $Q \in \mathcal{R}$ has a decomposition of the form

$$X_t = X_0 + \int_0^t H_u dS_u - K_t,$$

with an adapted, increasing process K and an S integrable predictable process H.

Proof. A proof to this well known statement is published in Föllmer [21]. □

In the context of financial mathematics, the integral $\int_0^t H_u dS_u$ can be interpreted as the profit or loss of a trading strategy H and the process K as a consumption process modelling how much money an investor withdraws from the portfolio. This consumption process K is increasing, so money can only be withdrawn but not invested. Thus X can be interpreted as the value process of a self-financing trading strategy with initial investment X_0.

Note that this theorem is limited to a set of equivalent martingale measures $\mathcal{R}$. This fits well with single market models where $\mathcal{Q}^{\mathbb{P}}$ contains only measures equivalent to $\mathbb{P}$, but cannot be applied to robust market models. For this purpose, the statement is extended to robust market models in the following theorem.

Theorem 3.2.3. *Robust optional decomposition theorem*
Let $\mathcal{P}$ be a set of physical measures and $\mathcal{Q}$ be the set of all local martingale measures regarding a semimartingale S equivalent to a $\mathbb{P} \in \mathcal{P}$. If S has dominating diffusion among all $Q \in \mathcal{Q}$ and X is a local supermartingale for every $Q \in \mathcal{Q}$, then for each $\mathbb{P} \in \mathcal{P}$ there exists a decomposition

$$X_t = X_0 + \int_0^t H_u dS_u - K_t^{\mathbb{P}} \quad \mathcal{Q}^{\mathbb{P}}\text{-a.s. for every } \mathbb{P} \in \mathcal{P}$$

where H is a $\mathbb{P}$-independent, S-integrable and predictable process and $K^{\mathbb{P}}$ is a $\mathbb{P}$-dependent, increasing process for every $\mathbb{P} \in \mathcal{P}$.

In the context of financial mathematics, we can interpret the theorem to mean that there is always a trading strategy H such that the difference $X - (H \cdot S)$ is decreasing, which means that after an initial investment of X_0 there can only be a potential profit that is described as $K^{\mathbb{P}}$. The most interesting part is that only the profit $K^{\mathbb{P}}$ depends on the actual market model, while the trading strategy remains unchanged for all $\mathbb{P} \in \mathcal{P}$.

Proof. The semimartingale (S, X) is a special semimartingale. Thus, for any $Q \in \mathcal{Q}$, one can find characteristics (B^Q, C^Q, ν^Q). According to Neufeld and Nutz [50], there exist characteristics (B^Q, C, ν^Q) with a $\mathbb{R}^{d+1,d+1}$-valued process C independent of Q and a Borel-measurable mapping defining B^Q and ν^Q depending on the measure Q. Furthermore, let $C^S := \langle S^c \rangle$, which means that C^S is the $d \times d$ submatrix of C, which represents the quadratic covariation of S. Also, let $C^{SX} := (C^{1,d+1}, ..., C^{d,d+1})^T$, i.e. $C^{SX} = \langle S^c, X^c \rangle$.

Now let $\mathbb{P} \in \mathcal{P}$ be arbitrary but fixed. Since X is a supermartingale with respect to every $Q \in \mathcal{Q}^{\mathbb{P}}$, according to the Optional Decomposition Theorem 3.2.2 there exists a decomposition

$$X = X_0 + (H^{\mathbb{P}} \cdot S) - K^{\mathbb{P}} \quad \mathbb{P}\text{-almost surely,} \tag{3.2.1}$$

where $H^{\mathbb{P}}$ and $K^{\mathbb{P}}$ depend on $\mathbb{P}$. It remains to show that there exists a process H such that $H = H^{\mathbb{P}}$ holds for every $\mathbb{P} \in \mathcal{P}$.

Let the process A be defined as the trace of C^S, i.e. $A := tr(C^S)$. Since the increments of the main diagonals of C are always non-negative, $dA = 0$ implies $dC = 0$. This means that $dC^S \ll dA$ and $dC^{SX} \ll dA$. According to the Radon-Nikodym Theorem, processes c^S and c^{SX} can be found, such that

$$dC^S = c^S dA \text{ and } dC^{SX} = c^{SX} dA. \tag{3.2.2}$$

It is known from the proof of the Optional Decomposition Theorem 3.2.2 (see proof of Theorem 1 in Föllmer [21]), that

$$H := c^{SX} (c^S)^+ \tag{3.2.3}$$

is a suitable candidate for the choice of the investment strategy. $(c^S)^+$ denotes the pseudoinverse of the matrix c^S, sometimes called Moore-Penrose inverse. We will first show that on the one hand $(H \cdot S^c) = (H^{\mathbb{P}} \cdot S^c)$ and on the other hand also $(H \cdot (S - S^c)) = (H^{\mathbb{P}} \cdot (S - S^c))$ $\mathcal{Q}^{\mathbb{P}}$-almost surely for every $\mathbb{P} \in \mathcal{P}$.
To show that $(H \cdot S^c) = (H^{\mathbb{P}} \cdot S^c)$, the continuous martingale part can be taken on both sides of the equation (3.2.1), leading to

$$X^c = (H^{\mathbb{P}} \cdot S^c).$$

For the quadratic variation, we obtain

$$dC^{SX} = d\langle S^c, X^c \rangle = d\langle S^c, (H^{\mathbb{P}} \cdot S^c) \rangle = H^{\mathbb{P}} d\langle S^c, S^c \rangle = H^{\mathbb{P}} dC^S \quad \mathcal{Q}^{\mathbb{P}}\text{-almost surely,}$$

which, with (3.2.2), leads to the equation

$$c^{SX} = H^{\mathbb{P}} c^S \quad \mathcal{Q}^{\mathbb{P}} \times dA\text{-almost everywhere.}$$

Using $dS^c \ll dA$ it already follows

$$(H \cdot S^c) = (H^{\mathbb{P}} \cdot S^c) \quad \mathcal{Q}^{\mathbb{P}}\text{-almost surely} \tag{3.2.4}$$

and by Itô isometry it follows that H is integrable with respect to S^c. Since c^S does not need to be invertible in general, this is not yet sufficient to show $H = H^{\mathbb{P}}$. Therefore, the equality $(H \cdot (S - S^c)) = (H^{\mathbb{P}} \cdot (S - S^c))$ must still be shown.

Let A^* be defined by $dA^* := \min_{i=1,\dots,d} d(C^S)^{i,i}$. Then $dA^* \ll dA$ holds and it follows

$$c^{SX} = H^{\mathbb{P}} c^S \quad \mathcal{Q}^{\mathbb{P}} \times dA^*\text{-almost everywhere.}$$

In addition, it holds that

$$dA_t^* \neq 0 \Rightarrow d(C_t^S)^{i,i} = (c_t^S)^{i,i} dA_t \neq 0 \;\; dA^*\text{-almost sure} \;\; \forall i \leq d.$$

Thus (c^S) is a dA^*-almost surely positive definite matrix and hence dA^*-almost surely invertible. Since under dA^* the pseudo inverse of c^S becomes the real inverse,

$$H = H^{\mathbb{P}} \quad \mathcal{Q}^{\mathbb{P}} \times dA^* \text{ almost everywhere, for all } \mathbb{P} \in \mathcal{P}. \tag{3.2.5}$$

The process $S - S^c$ has the characteristics $(B^Q, 0, \nu^Q)$ under a measure $Q \in \mathcal{Q}$. Due to the condition that S has dominating diffusion among all $Q \in \mathcal{Q}$, ν^Q is dominated by dA^*. Moreover, with the property $B_t^Q = \int (h(x) - x)\nu^Q(\{t\}, dx)$, which follows by Theorem II.2.34 in Jacod and Shiryaev [36], together with the local martingale property of S, $dB^Q \ll dA^*$ holds. Since dA^* dominates the characteristics of $S - S^c$, (3.2.5) is already sufficient to claim

$$(H \cdot (S - S^c)) = (H^{\mathbb{P}} \cdot (S - S^c)) \quad \mathcal{Q}^{\mathbb{P}}\text{-almost-sure.} \tag{3.2.6}$$

So with the two equations (3.2.4) and (3.2.6), it is proven

$$(H \cdot S) = (H^{\mathbb{P}} \cdot S) \quad \mathcal{Q}^{\mathbb{P}}\text{-almost sure for every } \mathbb{P} \in \mathcal{P},$$

where H is the trading strategy we are looking for. □

3.2.2. Robust pricing-hedging duality

Having introduced the robust form of the optional decomposition theorem, we can proceed to prove the pricing-hedging duality in its robust form. One of the main arguments in the proof of the pricing-hedging duality is the construction of the superhedging strategy that satisfies the desired equality. The superhedging strategy is constructed using the essential supremum of conditional expectations, which must first be defined. Note that the main ideas of the proofs here are again introduced in Nutz [51] and [52].

First, we will introduce some notation, definitions and assumptions needed to formulate the proof of the pricing-hedging duality. Let $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ be the filtration generated by $(t, \omega) \mapsto \omega_t$. The following notation corresponds to the notation used in Nutz [52].

Note, that the introduced notation and definitions are necessary to introduce a consistent meaning to conditional expectations on a space of measures ($\mathcal{P}$ in this case), which strictly speaking introduces a sublinear space in view of the mapping

$$X \mapsto \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[X].$$

In the literature of stochastic control theory and quasi-sure modelling it is common to work with concatenations of paths to define the set of associated martingale measures. Since we are working in a quasi-sure model approach, we will first introduce the commonly used set-up, see [5, 51, 52].

Definition 3.2.4.
Let $\omega, \tilde{\omega} \in \Omega$. The concatenation at time $t \in \mathbb{R}_+$ of two paths $\omega \otimes_t \tilde{\omega}$ is defined as

$$(\omega \otimes_t \tilde{\omega})_s := \omega_s \mathbb{1}_{[0,t)}(s) + (\omega_t + \tilde{\omega}_{s-t})\mathbb{1}_{[t,\infty)}(s).$$

Definition 3.2.5.
Let $Q \in \mathcal{Q}$ and $\omega \in \Omega$, then, as shown in Theorem 1.1.8 in Stroock and Varadhan [63], there exists a regular conditional probability distribution $\{Q_t^\omega\}_{\omega\in\Omega}$ conditional on $\mathcal{F}_t$ under which it holds that

$$\begin{aligned} &Q_t^\omega \left(\{\tilde{\omega} \in \Omega : \tilde{\omega}|_{[0,t]} = \omega|_{[0,t]}\}\right) = 1, \\ &\mathbb{E}^{Q_t^\omega}[f] = \mathbb{E}^Q[f|\mathcal{F}_t](\omega), \end{aligned}$$

for $f : \Omega \to \mathbb{R}$ bounded and $\mathcal{F}$-measurable. Then, for an event $A \in \mathcal{F}$, let the probability measure $Q^{t,\omega}$ be defined as

$$Q^{t,\omega}(A) := Q_t^\omega(\{\omega \otimes_t \tilde{\omega} : \tilde{\omega} \in A\}).$$

Definition 3.2.6.
Let f be a function on Ω, let $\omega \in \Omega$. The function $f^{t,\omega}$ is defined as

$$f^{t,\omega}(\tilde{\omega}) := f(\omega \otimes_t \tilde{\omega}).$$

Definition 3.2.7.
Let $\omega \in \Omega$ and $t \in \mathbb{R}_+$ and define a set $\{\mathcal{Q}(t,\omega)\}_{(t,\omega)\in\mathbb{R}_+\times\Omega}$, such that

$$\mathcal{Q}(t,\omega) = \mathcal{Q}(t,\tilde{\omega}) \text{ if } \omega|_{[0,t]} = \tilde{\omega}|_{[0,t]}.$$

In this notation, $\mathcal{Q} = \mathcal{Q}(0,\omega)$ for all $\omega \in \Omega$, since $\omega_0 = 0$. So instead of choosing $\mathcal{Q}$ directly, a family $\{\mathcal{Q}(t,\omega)\}$ can be chosen which induces $\mathcal{Q}$.

The main idea is that these sets are supposed to represent all conditional probabilites depending on ω up to time t. Therefore the sets $\mathcal{Q}(t,\omega)$ are supposed to fulfil the following assumptions.

Assumption 3.2.8.
Let $0 \le s \le t$, $\tilde{\omega} \in \Omega$ and $Q \in \mathcal{Q}(s,\tilde{\omega})$ and let $prob(\Omega)$ denote the set of all probability measures on Ω. Then let the following conditions on $\{\mathcal{Q}(t,\omega)\}_{(t,\omega)\in\mathbb{R}_+\times\Omega}$ be satisfied:

- Measurability: the set $\{(\tilde{Q},\omega) : \omega \in \Omega,\ \tilde{Q} \in \mathcal{Q}(t,\omega)\}$ is analytic.
  Note that every Borel set is analytic in a Polish space.
- Invariance: $Q^{t-s,\omega} \in \mathcal{Q}(t, \tilde{\omega} \otimes_s \omega)$ for Q almost all $\omega \in \Omega$.
- Stability: For a $\mathcal{F}_{t-s}$ measurable mapping $\mu : \Omega \to prob(\Omega)$ with $\mu(\omega) \in \mathcal{Q}(t, \tilde{\omega} \otimes_s \omega)$ it holds for Q-almost all $\omega \in \Omega$ that
$$\tilde{Q}(A) := \int\int \mathbb{1}_A(\omega \otimes_{t-s} \omega')\mu_\omega(d\omega')Q(d\omega),\ A \in \mathcal{F}$$
  is a measure in $\mathcal{Q}(s,\tilde{\omega})$.

Definition 3.2.9.
Let $f : \Omega \to \bar{\mathbb{R}}$ be an upper semianalytic function, $\omega \in \Omega$ and $t \in \mathbb{R}_+$, then let $\mathcal{E} : [0,T] \times \Omega \to \mathbb{R}$ be defined as the following sublinear expected value:
$$\mathcal{E}_t(f)(\omega) := \sup_{Q\in\mathcal{Q}(t,\omega)} \mathbb{E}^Q[f^{t,\omega}].$$

Remark 3.2.10.
Definition 3.2.9 does indeed define a conditional expectation, because on the one hand
$$\mathbb{E}^{Q^{t,\omega}}[f^{t,\omega}] = \mathbb{E}^{Q^\omega_t}[f] = \mathbb{E}^Q[f|\mathcal{F}_t](\omega),$$
while on the other hand the Assumption 3.2.8 implies that the set $\mathcal{Q}(t,\omega)$ includes all conditional probabilities $\{Q^{t,\omega} : Q \in \mathcal{Q}\}$.

Assumption 3.2.11.
The underlying S has dominating diffusion among all measures $Q \in \mathcal{Q}$, as introduced in Definition 3.2.1.

Assumption 3.2.12.
The set $\mathcal{Q}$ contains all equivalent local martingale measures, i.e. for every local martingale measure Q with $Q \sim \mathbb{P}$ for at least one $\mathbb{P} \in \mathcal{P}$ it holds that $Q \in \mathcal{Q}$.

Assumption 3.2.12 guarantees that the remaining assumptions are indeed assumptions on the physical measures $\mathcal{P}$, rather than on an arbitrary selection of equivalent local martingale measures.

After introducing two additional lemmas and using the previous assumptions and notation we can finally formulate and prove the robust pricing-hedging duality in Theorem 3.2.15.

Lemma 3.2.13.
Let $\omega \in \Omega$, $0 \leq s \leq t$, $Q \in \mathcal{Q}$ and let $C : \Omega \to \bar{\mathbb{R}}$. Under the Assumptions 3.2.8, the sublinear expected value $\mathcal{E}_t(\cdot)$ from Definition 3.2.9 fulfils

$$\begin{aligned}\mathcal{E}_t(C) &= \operatorname*{ess\,sup}_{\tilde{Q} \in \mathcal{Q}(t,Q)} \mathbb{E}^{\tilde{Q}}[C|\mathcal{F}_t],\\ \mathcal{E}_s(C) &= \operatorname*{ess\,sup}_{\tilde{Q} \in \mathcal{Q}(s,Q)} \mathbb{E}^{\tilde{Q}}[\mathcal{E}_t(C)|\mathcal{F}_s],\end{aligned}$$

where $\mathcal{Q}(s,Q) := \{\tilde{Q} \in \mathcal{Q} : \tilde{Q}(A) = Q(A)\ \forall A \in \mathcal{F}_s\}$.

Proof. A proof of this statement can be found in Theorem 2.3 in Nutz and Van Handel [52]. The property that the space Ω with the chosen topology is a Polish space is necessary for this result. □

Lemma 3.2.14.
The random variable $\mathcal{E}_t(C)$ is a supermartingale for every $Q \in \mathcal{Q}$.

Proof. As shown in the previous lemma, $\mathcal{E}_s(C) = \operatorname{ess\,sup}_{\tilde{Q} \in \mathcal{Q}(s,Q)} \mathbb{E}^{\tilde{Q}}[\mathcal{E}_t(C)|\mathcal{F}_s]$, so the following inequality holds:

$$\mathbb{E}^{Q}[\mathcal{E}_t(C)|\mathcal{F}_s] \leq \operatorname*{ess\,sup}_{\tilde{Q} \in \mathcal{Q}(s,Q)} \mathbb{E}^{\tilde{Q}}[\mathcal{E}_t(C)|\mathcal{F}_s] = \mathcal{E}_s(C) \quad \forall Q \in \mathcal{Q}.$$

□

With this statement, the robust Optional Decomposition Theorem can also be applied to $\mathcal{E}_s(C)$, which is crucial for the proof of a robust pricing-hedging duality. This expression becomes the basis of the superhedging portfolio, that is supposed to be constructed.

Theorem 3.2.15. *Robust pricing-hedging duality*
Let $\mathcal{P}$ be a set of measures with equivalent measures $\mathcal{Q}$ such that the sets $\{\mathcal{Q}(t,\omega)\}_{(t,\omega)\in\mathbb{R}_+\times\Omega}$ satisfy the assumptions 3.2.8 and 3.2.12 and such that $\mathcal{P}$ satisfies the Assumption 3.2.11. Moreover, let $C:\Omega\to\bar{\mathbb{R}}$ be an upper analytic, $\mathcal{F}_T$-measurable function with finite expected value with respect to all $Q\in\mathcal{Q}$.

Then the robust pricing-hedging duality holds:

$$\sup_{Q\in\mathcal{Q}}\mathbb{E}^Q[C]=\min\{x\in\mathbb{R}:\text{There exists an admissible hedging strategy } H$$
$$\text{with } x+\int_0^T H\ dS\geq C,\ \mathcal{P}-a.s.\}.$$

Proof. The direction "$\leq$" follows from the local martingale property of S under all $Q\in\mathcal{Q}$, because $(H\cdot S)$ as a local martingale is also a supermartingale. Assuming that there exists an admissible investment strategy H and $x\in\mathbb{R}$ with $x+(H\cdot S)_T\geq C$ $\mathcal{Q}$-almost surely, then for all $Q\in\mathcal{Q}$ it follows:

$$x=x+(H\cdot S)_0\geq\mathbb{E}^Q[x+(H\cdot S)_T]\geq\mathbb{E}^Q[C]. \tag{3.2.7}$$

So it remains to prove the direction "$\geq$". This direction is proven by constructing a superhedging strategy that satisfies the inequality using $\mathcal{E}_t(C)$. Given Assumption 3.2.8, with Lemma 3.2.14 it holds that the expression $\mathcal{E}_t(C)=\sup_{Q\in\mathcal{Q}(t,\omega)}\mathbb{E}[C^{t,\omega}]$ is a supermartingal with respect to all $Q\in\mathcal{Q}$. Moreover, with Assumption 3.2.11 and the robust Optional Decomposition Theorem 3.2.3, any supermartingale X for any $Q\in\mathcal{Q}$ can be decomposed into an almost sure representation $X=X_0+(H\cdot S)-K^Q$.

This means that it suffices to show that $\mathcal{E}_t(C)$ can be used to construct a supermartingale X satisfying $X_0\leq\sup_{Q\in\mathcal{Q}}E^Q[C]$ and $X_T\geq C$. To do this, we define the following process:

$$\tilde{X}_t:=\limsup_{t\in\mathbb{Q}_+,\ r\searrow t}\mathcal{E}_r(C)\quad\forall t\in[0,T)\text{ and }\tilde{X}_T:=\mathcal{E}_T(C).$$

For a countable, dense subset of $\mathbb{R}_+$ ($\mathbb{Q}_+$ in this case), Föllmer's Lemma (Theorem 2.44 in [29]) guarantees for the supermartingale $\tilde{X}$ the existence of a $\mathbb{F}_+$-adapted supermartingale X with the following properties:

a) $X_t(\omega)=\lim_{s\in\mathbb{Q}_+,\ s\searrow t}\tilde{X}_s(\omega)$.

b) $X_{t-}(\omega):=\lim_{s\in\mathbb{Q}_+,\ s\nearrow t}X_s(\omega)$ exists and is finite for almost all $\omega\in\Omega$ and for all $t\in\mathbb{R}_+\setminus\{0\}$. Moreover, $X_{t-}(\omega)=\lim_{s\in\mathbb{Q}_+,\ s\nearrow t}\tilde{X}_s(\omega)$.

c) $\tilde{X}_t\geq\mathbb{E}[X_t|\mathcal{F}_t]$ for all $t\in\mathbb{R}_+$.

This means that X is a cadlag version of $\tilde{X}$. Moreover, under this result, it is not necessary that $\mathcal{F}_0$ must contain all $\mathbb{P}$-zero sets, as it is required in [51].

It remains to show that $X_T \geq C$ $\mathcal{P}$-almost surely and $X_0 \leq \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[C]$.

The first inequality is derived directly from Lemma 3.2.13, because with the $\mathcal{F}_T$-measurability of the claim C, it holds Q-almost surely for all $Q \in \mathcal{Q}$ that

$$X_T = \mathcal{E}_T(C) = \operatorname*{ess\,sup}_{\tilde{Q} \in \mathcal{Q}(s,Q)} \underbrace{\mathbb{E}^{\tilde{Q}}[C|\mathcal{F}_T]}_{=C} = C.$$

The second inequality is more challenging to show. It should be noted that X_0 is not $\mathcal{F}_0$, but only $\mathcal{F}_{0+}$ measurable. So X_0 is not a deterministic quantity, but stochastic. By definition of the cadlag version X, the following inequality holds via the third property $c)$ for each $Q \in \mathcal{Q}$:

$$\mathbb{E}^Q[X_0] = \mathbb{E}^Q[X_0|\mathcal{F}_0] \leq \tilde{X}_0 = \mathcal{E}_0(C). \tag{3.2.8}$$

Since the superhedging portfolio is itself an upper bound, X_0 is a bounded random variable. Let $x_0^Q \in \mathbb{R}$ be the smallest upper bound for X_0 under Q, such that $X_0 \leq x_0^Q$ Q-almost surely. It remains to show that $x_0^Q = \sup_{Q \in \mathcal{Q}} \mathbb{E}[X_0]$ holds, so that we can finally conclude from (3.2.8) that

$$X_0 \leq x_0^Q = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[X_0] \leq \mathcal{E}_0(C) = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[C], \tag{3.2.9}$$

so X actually satisfies the required properties.

The equality $x_0^Q = \sup_{Q \in \mathcal{Q}} \mathbb{E}[X_0]$ follows by the following argumentation:

Let $Q \in \mathcal{Q}$ be arbitrary but fixed. Let $\mathcal{R}^Q$ be the set of all measures on $\mathcal{F}_{0+}$ that are equivalent to Q. Now a sequence of measures in $\mathcal{R}^Q$ can be constructed which converges to the Dirac measure $\delta_{\{X_0 = x_0^Q\}}$. Take for example the following probability measures:

$$\begin{aligned}
\mu_n &:= \frac{n-1}{n} \delta_{\{X_0 = x_0^Q\}} + \frac{1}{n} Q, \\
\nu_{n,\epsilon} &:= \frac{n-1}{n} \mathrm{Unif}([x_0^Q - \epsilon, x_0^Q)) + \frac{1}{n} Q.
\end{aligned}$$

For the case $Q(X_0 = x_0^Q) = 0$, $\nu_{n,\epsilon}$ is a measure equivalent to Q for all $n \in \mathbb{N}$ and $\epsilon > 0$ small enough, so $\nu_{n,\epsilon}$ is contained in $\mathcal{R}^Q$ and $\lim\limits_{n \to \infty, \epsilon \to 0} \mathbb{E}^{\nu_{n,\epsilon}}[X_0] = x_0^Q$. For the case $Q(X_0 = x_0^Q) > 0$ μ_n is in $\mathcal{R}^Q$ for all $n \in \mathbb{N}$ and it holds that $\lim\limits_{n \to \infty} \mathbb{E}^{\mu_n}[X_0] = x_0^Q$. We can conclude

$$x_0^Q = \sup_{\mathbb{P} \in \mathcal{R}^Q} \mathbb{E}^{\mathbb{P}}[X_0].$$

Moreover, there exists a measure $Q' \in \mathcal{Q}$ such that $\mathbb{P} \in \mathcal{R}^Q$ is the restriction of Q' to $\mathcal{F}_{0+}$, i.e. $\mathbb{P} = Q'|_{\mathcal{F}_{0+}}$. This is shown as follows: For every $\mathbb{P} \in \mathcal{R}^Q$ there exists a Radon-Nikodym derivative $Z = \frac{dQ}{dP}$, which is also $\mathcal{F}_{0+}$-measurable. Then $\mathbb{P} = Q'|_{\mathcal{F}_{0+}}$ for the measure Q' with $dQ' = ZdQ$ on $\mathcal{F}_T$. For the random variable Z we have

$$Z = \mathbb{E}^Q[Z|\mathcal{F}_{0+}] = \mathbb{E}^{Q'}[1|\mathcal{F}_{0+}] = 1 \quad Q\text{-a.s.}$$

and thus, due to the right-continuity of S, it follows

$$\mathbb{E}^{Q'}[S_t|\mathcal{F}_s] = \mathbb{E}^{Q'}[S_t|\mathcal{F}_{s+}] = \mathbb{E}^Q[ZS_t|\mathcal{F}_{s+}] = \underbrace{Z}_{=1}\mathbb{E}^Q[S_t|\mathcal{F}_{s+}].$$

Since Q is a local martingale measure, Q' is also a local martingale measure to S and since it was required in Assumption 3.2.12 that $\mathcal{Q}$ contains all local martingale measures equivalent to a $Q \in \mathcal{Q}$, there must also be $Q' \in \mathcal{Q}$. Thus the inequality in (3.2.9) is satisfied and the supermartingale X satisfies the desired properties $X_T \geq C$ and $X_0 \leq \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[C]$. According to Theorem 3.2.3, this supermartingale can be represented as a self-financing trading strategy

$$X = X_0 + (H \cdot S) + K^{\mathbb{P}}$$

for each $\mathbb{P} \in \mathcal{P}$. It should be noted that $X_0 = \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[C]$, since it has already been shown in (3.2.7) that the superhedging portfolio requires at least the initial capital $\sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[C]$.

It remains to examine the admissibility of the trading strategy. The trading strategy is representable as $X_0 + (H \cdot S) = X - K$, where K is a decreasing process and X_t is bounded from below Q-almost surely by $\mathbb{E}^Q[C|\mathcal{F}_t]$ according to the construction and Lemma 3.2.13. The expected value of C is assumed to be finite, which proves admissibility, and the constructed H justifies the equation

$$\sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[C] = \min\{x \in \mathbb{R} : \text{There exists an admissible hedging strategy } H \\ \text{with } x + \int_0^T H\ dS \geq C,\ \mathcal{P} - f.s.\}.$$

The explicit construction of the superhedging portfolio thereby shows that the minimum is indeed reached. □

It is worth noting, which of the commonly used market models satisfy the assumptions 3.2.8, 3.2.11 and 3.2.12 and therefore satisfy the pricing-hedging duality.

Lévy processes are defined to have independent increments, which means that for any Lévy process the set of conditional equivalent local martingale measures is independent of (t, ω), or in other words $\mathcal{Q} = \mathcal{Q}(t, \omega)$ for every $(t, \omega) \in [0, T] \times \Omega$. This means that Lévy processes by definition fulfil the invariance and stability assumption in 3.2.8. Measurability can be seen as a more technical assumption, a proof of the measurability assumption for Lévy processes can be found in Nutz [51]. For the dominating diffusion (Assumption 3.2.11) we need to have a look at the characteristics (dB^Q, dC^Q, ν^Q) of a semimartingale S under a measure $Q \in \mathcal{Q}$. If we assume that we have increments that include a Brownian motion, which includes many commonly used models, we get $dC_t = d\langle S^c, S^c\rangle_t \geq \sigma^2 dS_t > 0$, which means the quadratic variation of S^c is strictly increasing and therefore $dC^Q \neq 0$ almost surely. In these cases, dominating diffusion is obviously satisfied. The same argument can be applied to many other models that are not pure jump models.

In fact, the pricing-hedging duality is satisfied not only by any continuous Lévy model, but also by the Merton-Jump model and a lot of other jump diffusion models.

3.3. Robust Neyman-Pearson lemma

One subfield of statistics is test theory. Among other things, this field deals with the existence and structure of optimal hypothesis tests, i.e. how to optimally decide between two hypotheses H_0 and H_1 for the distribution of a random variable X. A test $\phi : \Omega \to [0,1]$ specifies the probability of accepting the null hypothesis H_0. In general, the probability of a correct decision for H_0, i.e. $\mathbb{E}[\phi|H_0]$, is to be maximised, while the probability of a wrong decision for H_0 is limited to $\mathbb{E}[\phi|H_1] \leq \alpha$ by a confidence level $\alpha \in (0,1)$.

It will be shown that the search for optimal hedging strategies can also be put into the form of a hypothesis test. Therefore, this chapter examines hypothesis testing and the Neyman-Pearson test as an optimal test, especially with respect to robust forms of hypothesis testing.

It is generally known that many hypothesis tests are optimally solved by a Neyman-Pearson test, or 0-1 test, which has many advantages in application.

Definition 3.3.1.
Let $\mathbb{P}_0$ and $\mathbb{P}_1$ be two probability measures with a dominating measure μ and density functions $f_0 = \frac{d\mathbb{P}_0}{d\mu}$, $f_1 = \frac{d\mathbb{P}_1}{d\mu}$. A test $\phi : \Omega \to [0,1]$ is called a Neyman-Pearson test if there exists a constant $c \in \mathbb{R}$ and $r \in (0,1)$ with

$$\phi = \begin{cases} 1, & \frac{f_1}{f_0} > c \\ r, & \frac{f_1}{f_0} = c \\ 0, & \frac{f_1}{f_0} < c. \end{cases}$$

In other words, a Neyman-Pearson test will, except on the set $\{\frac{f_1}{f_0} = c\}$, almost surely accept either the null hypothesis H_0, i.e. $\phi = 1$, or the counter-hypothesis H_1, i.e. $\phi = 0$. Only on the set $\{\frac{f_1}{f_0} = c\}$, a null set most of the time, the test makes a random decision, which is very convenient for many applications. It should be noted that in most applications $\mathbb{P}_0(\frac{f_1}{f_0} = c) = 0 = \mathbb{P}_1(\frac{f_1}{f_0} = c)$, so there is no need to randomise.

The Neyman-Pearson lemma describes the optimality of those tests:

Theorem 3.3.2. *Neyman-Pearson lemma*
Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be two probability measures. Let μ be a dominating measure with respect

to $\mathbb{P}_1$ and $\mathbb{P}_2$. The optimal test function $\phi : \Omega \to [0,1]$ regarding confidence level $\alpha \in (0,1)$ between the two measures, i.e. the solution to the optimisation problem

$$\max_{\phi} \mathbb{E}^{\mathbb{P}_1}[\phi] \tag{3.3.1}$$
$$\text{under the constraint } \mathbb{E}^{\mathbb{P}_2}[\phi] \leq \alpha,$$

is a Neyman-Pearson test $\tilde{\phi}$ with suitable constants $r \in (0,1)$ and $c \in \mathbb{R}_+$:

$$\tilde{\phi} = \mathbb{1}_{\{\frac{d\mathbb{P}_1}{d\mu} > c\frac{d\mathbb{P}_2}{d\mu}\}} + r\mathbb{1}_{\{\frac{d\mathbb{P}_1}{d\mu} = c\frac{d\mathbb{P}_2}{d\mu}\}}.$$

3.3.1. Minimax optimisation

The Neyman-Pearson lemma initially describes only the decision between a simple null hypothesis and a simple alternative. In many applications, such a restriction is not sufficient, especially with respect to the robust formulation of financial markets, which consist of multiple physical measures and in the case of incomplete markets, already have an infinite number of equivalent martingale measures. This motivates the transition to sets of hypotheses. For this purpose, we define $\mathfrak{P}$ and $\mathfrak{Q}$ as two sets of measures to be tested against each other. We can then rewrite the optimisation problem (3.3.1) as a maximin problem under constraints that must be satisfied simultaneously by all measures in $\mathfrak{Q}$:

$$\sup_{\phi} \left(\inf_{\mathbb{P} \in \mathfrak{P}} \mathbb{E}^{\mathbb{P}}[\phi] \right) \tag{3.3.2}$$
$$\text{under the constraint } \sup_{Q \in \mathfrak{Q}} \mathbb{E}^{Q}[\phi] \leq \alpha.$$

Thus, the smallest probability correctly choosing $\mathfrak{P}$ over all measures in $\mathfrak{P}$ is to be maximised.

The following assumptions are needed to formulate a robust form of the Neyman-Pearson theorem:

Assumption 3.3.3.

Let $\mathfrak{P}$, $\mathfrak{Q}$ and α fulfil the following assumptions:

I. The confidence level is positive:

$$\alpha > 0.$$

II. There exists a measure μ dominating $\mathfrak{P}$ and $\mathfrak{Q}$:

$$\mathbb{P} \ll \mu \ \forall \mathbb{P} \in \mathfrak{P} \text{ and } Q \ll \mu \ \forall Q \in \mathfrak{Q}.$$

III. The intersection of both sets $\mathfrak{P}$ and $\mathfrak{Q}$ is the empty set:

$$\mathfrak{P} \cap \mathfrak{Q} = \emptyset.$$

IV. The set of every Radon-Nikodym derivative of $\mathfrak{P}$ regarding μ

$$\left\{ \frac{d\mathbb{P}}{d\mu} : \mathbb{P} \in \mathfrak{P} \right\}$$

is convex and weakly*-closed.

With these assumptions one can show a more general formulation of the Neyman-Pearson lemma.

Theorem 3.3.4. *Robust Neyman-Pearson lemma*
With Assumption 3.3.3 the maximin problem

$$\sup_{\phi} \left(\inf_{\mathbb{P} \in \mathfrak{P}} \mathbb{E}^{\mathbb{P}}[\phi] \right) \tag{3.3.3}$$
$$\text{under the constraint } \sup_{Q \in \mathfrak{Q}} \mathbb{E}^{Q}[\phi] \leq \alpha$$

is solved by a test function

$$\tilde{\phi} = \mathbb{1}_{\{ \frac{d\tilde{\mathbb{P}}}{d\mu} > c \int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\tilde{\lambda}(Q) \}},$$

with a worst case pair $(\tilde{\mathbb{P}}, \tilde{\lambda}) \in \mathfrak{P} \times prob(\mathfrak{Q})$, where $prob(\mathfrak{Q})$ is the set of all probability measures on the measurable space $\mathfrak{Q}$ provided with σ-field generated by integrals $\int_{\Omega} f dQ$, where f is a bounded and measurable function on $(\Omega, \mathcal{F}_T)$ and with a constant $c \in \mathbb{R}$.

Proof. This statement was proven in Rudloff [58]. Roughly summarised the idea for the proof is to solve the corresponding dual problem and prove strong duality. This can be done by taking the space Λ of all finite, signed measures on $(\mathfrak{Q}, S)$ (with a σ-algebra S), as the dual space to the space $\mathcal{L}$ of all bounded and measurable functions on $(\mathfrak{Q}, S)$, see Example 1.63 in [69]. With the bilinear form $\langle f, \lambda \rangle = \int_{\mathfrak{Q}} f d\lambda$ for $f \in \mathcal{L}$, $\lambda \in \Lambda$ and the dual problem that is supposed to be optimised in λ one can see why the solution is in the form of a mixture of measures. □

Remark 3.3.5.
This statement is similar to a theorem in Cvitanic and Karatzas [14], which both have in common that the set $\mathfrak{Q}$ does not have to be closed. The solution in [14] is of the form $\tilde{\phi} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > c\frac{d\tilde{R}}{d\mu}\}}$ with a random variable $\tilde{R} \in \mathfrak{R} = \{R \in L^1(\mu) : \mathbb{E}^\mu[R\phi] \le \alpha\}$, which is a compact set containing the convex hull of $\mathfrak{Q}$. In both cases one avoids the issue of lacking closedness of the set $\mathfrak{Q}$ either by introducing a new larger set $\mathfrak{R}$ or by giving a solution in terms of an integral mixture. Using this larger set $\mathfrak{R}$ is inconvenient in view of future applications in financial mathematics, since we want to use Radon-Nikodym derivatives of physical market measures with respect to corresponding equivalent martingale measures.

3.3.2. Measure-convexity

Since we want to apply Neyman-Pearson theory to financial applications, we also face the problem that one of the most important sets, the set of all equivalent martingale measures, is in general not closed (see Proposition 3.3.15). However, we will use a different approach to overcome this problem by introducing the concept of measure-convexity.

Note that the formulation in terms of a mixture or an integral over $\mathfrak{Q}$, as in Theorem 3.3.4, is inconvenient since we eventually want to apply this theory to financial markets where we usually know the Radon-Nikodym derivative $\frac{d\mathbb{P}}{dQ}$ of a physical measure $\mathbb{P}$ with respect to corresponding equivalent martingale measures Q. This raises the question of whether we can simplify the solution to make use of this knowledge. This question leads to the introduction of the so-called measure-convexity.

Remark 3.3.6.
It should be noted that this is only a problem because in the financial context we consider probability measures as a subset of the infinite dimensional Hilbert space of signed measures. In finite dimensional spaces the concept of measure-convexity introduced in this section coincides with the usual definition of convexity.

Remark 3.3.7. *σ-algebras on sets of probability measures*
Since we are mixing over a set of probability measures it may be useful to note how to define the natural σ-algebra on such spaces: Let $(\Omega, \mathcal{F})$ be a measurable space with $prob(\Omega)$ the set of all probability measures on $(\Omega, \mathcal{F})$. Now for each $A \in \mathcal{F}$ we can define a function $\chi_A : prob(\Omega) \to \mathbb{R}$ by $\chi_A(\mathbb{P}) = \mathbb{P}(A)$. The σ-algebra generated by $\{\chi_A^{-1}(I) : A \in \mathcal{F}, I \subseteq \mathbb{R} \text{ Borel set}\}$ defines a σ-algebra on $prob(\Omega)$. See Gaudard and

Hadwin [25] for other constructions of suitable σ-algebras and relations between them, depending on whether the set Ω is separable or complete.

Definition 3.3.8. *Barycenter*
Let X be a locally convex space with topological dual space X^*. An element $x \in X$ is called a barycenter of the probability measure μ on X if

(a) each $l \in X^*$ is measurable regarding μ.

(b) $l(x) = \int\limits_X l(x) d\mu$ for every $l \in X^*$.

Definition 3.3.9. *Measure-convexity*
A subset $Y \subseteq X$ is called measure-convex if for every probability measure μ on Y the corresponding barycenter is still an element of Y.

Remark 3.3.10.
Measure-convexity for a set Y implies in particular for every probability measure μ on Y that

$$\int\limits_Y y d\mu(y) \in Y, \tag{3.3.4}$$

which is exactly what is needed in Theorem 3.3.4 to simplify the optimal solution. In general for infinite dimensional spaces one can only say that the mixture $\int_Y y d\mu(y)$ lies in the closed convex hull $\overline{co}(Y)$.

One can think of measure-convexity as a weaker form of completeness and boundedness for convex spaces. These three properties are sufficient, but not necessary, for measure-convexity.

Lemma 3.3.11.
Every complete, bounded and convex subspace Y of a locally convex space X is measure-convex.

Proof. Winkler [67] proves this statement as Corollary 1.2.4. □

On the other hand, not every measure-convex space has to satisfy such strong properties. One can find weaker necessary and sufficient conditions for measure-convexity, which is known as Fremlin-Pryce Theorem.

Theorem 3.3.12. *Fremlin-Pryce theorem*
A subset Y of a locally convex space X is measure-convex if and only if

(a) Y is bounded and

(b) The closed convex hull $\overline{co}(U)$ for every compact subset $U \subseteq Y$ is a compact subset of Y.

Proof. In Winkler [67] this statement can be found as Proposition 1.2.5. □

Before applying measure-convexity to the already introduced robust Neyman-Pearson theory and proving that the assumption of measure-convexity is not too strict for financial applications, it can be shown that the assumption is non-trivial, i.e. that there are indeed convex sets that are not measure-convex. In Alfsen [2] one can find the following example which is convex but not measure-convex.

Example 3.3.13.
Let X be the compact set of all probability measures on $[0,1]$ and let Y be the subset of all atomic measures in X, i.e. $Y = \{\sum_{i=1}^{\infty} c_i \delta_{x_i}, x_i \in [0,1], \sum_{i=1}^{\infty} c_i = 1\}$. The set Y contains all Dirac measures $E = \{\delta_x : x \in [0,1]\}$, which is also known to be the set of extreme points of X, see Theorem 2.1 in Winkler [68]. Now Y is convex but not measure-convex, because from the Choquet Theorem, see e.g. Phelps [54], we know that for any measure $p \in X$ there exists a measure μ supported on the extreme points of X, i.e. on $E \subseteq Y$, such that $x = \int_E y dp(y) \notin Y$.

Now we can formulate the main theorem of this chapter, a solution to the optimisation problem (3.3.2), that is suitable for financial application.

Theorem 3.3.14.
Let $\mathfrak{Q}$ be a measure-convex set. Let Assumption 3.3.3 hold. The optimisation problem (3.3.2) is solved by a function

$$\tilde{\phi} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > c \frac{d\tilde{Q}}{d\mu}\}} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\tilde{Q}} > c\}}$$

with a worst case pair $(\tilde{\mathbb{P}}, \tilde{Q}) \in \mathfrak{P} \times \mathfrak{Q}$.

Proof. From Theorem 3.3.4 we know that there exists an optimal solution of the form

$$\tilde{\phi} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > \int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\tilde{\lambda}(Q)\}}$$

with a worst case measure $\tilde{\mathbb{P}} \in \mathfrak{P}$ and a finite measure $\tilde{\lambda}$ on the set $\mathfrak{Q}$.

With normalization of $\tilde{\lambda}$ we can define $\tilde{\lambda}$ to be a probability measure and get a constant $\tilde{c} = ||\tilde{\lambda}||$ such that there is a solution

$$\tilde{\phi} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > \tilde{c} \int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\tilde{\lambda}(Q)\}}$$

with a probability measure $\tilde{\lambda}$. We define the set

$$\mathfrak{Q}' = \left\{ \frac{dQ}{d\mu} : Q \in \mathfrak{Q} \right\}.$$

One can show that $\mathfrak{Q}'$ is measure-convex if and only if $\mathfrak{Q}$ is measure-convex, since for any probability measure λ on $\mathfrak{Q}$ one can show that

$$\int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\lambda(Q) = \frac{d \int_{\mathfrak{Q}} Q d\lambda(Q)}{d\mu} \quad \mu\text{-a.s.},$$

since for $\tilde{Q} := \int_{\mathfrak{Q}} Q d\lambda(Q)$ and every $A \in \mathcal{F}$ we have

$$\begin{aligned} \int\limits_A \left(\frac{d \int_{\mathfrak{Q}} Q d\lambda(Q)}{d\mu} \right) d\mu &= \int\limits_A \frac{d\tilde{Q}}{d\mu} d\mu = \tilde{Q}(A) = \Big(\int\limits_{\mathfrak{Q}} Q d\lambda(Q) \Big)(A) = \int\limits_{\mathfrak{Q}} Q(A) d\lambda(Q) \\ &= \int\limits_{\mathfrak{Q}} \int\limits_A \underbrace{\frac{dQ}{d\mu}}_{\geq 0} d\mu d\lambda(Q) = \int\limits_A \Big(\int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\lambda(Q) \Big) d\mu. \end{aligned}$$

Finally with measure-convexity of $\mathfrak{Q}$ one can find $\tilde{Q} \in \mathfrak{Q}$ such that

$$\int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\lambda(Q) = \frac{d\tilde{Q}}{d\mu}.$$

This leaves us with

$$\tilde{\phi} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > \tilde{c} \int\limits_{\mathfrak{Q}} \frac{dQ}{d\mu} d\tilde{\lambda}(Q)\}} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\mu} > \tilde{c} \frac{d\tilde{Q}}{d\mu}\}} = \mathbb{1}_{\{\frac{d\tilde{\mathbb{P}}}{d\tilde{Q}} > \tilde{c}\}}.$$

□

3.3.3. Financial application

First, the following remark shows, that closedness of $\mathfrak{Q}$ for the Neyman-Pearson theorem is a too strict requirement for financial applications, since in Proposition 3.3.15 it is shown that the set of equivalent martingale measures is never closed in incomplete markets. However, it can be shown that the requirement for measure-convexity does indeed fit to the financial settings defined in this work.

Proposition 3.3.15.
The set of all equivalent martingale measures $\mathcal{Q}^{\mathbb{P}}$ regarding a market measure $\mathbb{P}$ is either a singleton (in a complete market model) or not closed (for any incomplete market model).

Proof. First, note that if there exists an absolutely continuous but not equivalent martingale measure R (i.e. $R \ll \mathbb{P}$ and $R \nsim \mathbb{P}$), then for any $Q \in \mathcal{Q}^{\mathbb{P}}$ the sequence $\frac{1}{n}Q + \frac{n-1}{n}R \subseteq \mathcal{Q}^{\mathbb{P}}$ converges to $R \notin Q^{\mathbb{P}}$ in L^1, i.e. $\mathcal{Q}^{\mathbb{P}}$ is not closed. This means that it only remains to show existence of such a measure R, which is quite easy in explicit models, but can be challenging in a general form, as shown in [46], whose proof is reformulated and adjusted to our notation in the following proof.

Let $\mathbb{P}$ be such that $\mathcal{Q}^{\mathbb{P}}$ is not empty and not a singleton. Let Q_1 and Q_2 be different measures in $\mathcal{Q}^{\mathbb{P}}$. In this case we can notice that the processes $Z_t = \frac{dQ_2}{dQ_1}|_{\mathcal{F}_t}$ and $S_t Z_t$ are both Q_1 martingales, where S is the modelled underlying, that is again a Q_1 martingale. Using integration by parts, one can see that the quadratic covariation $[Z, S]$ is a local martingale as well:

$$d[Z,S] = d(SZ) - SdZ - ZdS$$

These processes can be used to construct the desired measure R. In particular, as we will show in the following proof, one can construct a non-negative process $\tilde{Z}_t$ with $S\tilde{Z} \geq 0$ being a Q_1-martingale and $\mathbb{P}(\tilde{Z}_T = 0) > 0$. Then $\frac{dR}{dQ_1}|_{\mathcal{F}_t} = \frac{\tilde{Z}_t}{\mathbb{E}^{Q_1}[\tilde{Z}_t]}$ defines a measure R, that is the desired measure.

Since we know from Proposition 2.3.2, that we can characterise all jumptimes with a sequence of predictable and totally inaccessable stopping times, we can divide the proof into 3 cases: Either Z has no negative jumps, or there is a positive probability of negative jumps at a stopping time that is either predictable or totally inaccessible.

1. If $\Delta Z_t \geq 0$ we can use that Z is a non-constant martingale, which means that there must be a constant $0 < c < 1$ such that $\mathbb{P}(\inf_{t \in [0,T]} Z_t \leq c) > 0$; note that we

assume $Z_0 = 1$ almost surely. Using the predictable (since $\Delta Z_t \geq 0$) stopping time $\tau = \inf\{t \in [0, \infty) : Z_t \leq c\}$ with $\tau = \infty$ if $Z_t > c$ for every $t \geq 0$, we can define:

$$\tilde{Z}_t = Z_{t \wedge \tau} - c,$$

which is a martingale as a stopped martingale. $\tilde{Z}S$ remains a martingale as S and ZS are martingales and with $\mathbb{P}(\tilde{Z}_T = 0) = \mathbb{P}(\tau < T) > 0$, $\tilde{Z}$ is the desired process, finishing the first of the three cases.

On the other hand, if $\mathbb{P}(\Delta Z_t < 0) > 0$ for at least one $t \geq 0$, then we cannot use constructions like in the previous case, since $\tilde{Z}$ must be non-negative. In this case, we will use Proposition 2.2.18 to work around this issue.

Before formulating these two cases, note that, without loss of generality, we can assume $0 \leq \frac{dQ_1}{dQ_2} \leq 2$ and thus $\Delta \frac{dQ_1}{dQ_2} \geq -2$, because if necessary we can replace Q_2 by the equivalent martingale measure $\frac{Q_1+Q_2}{2}$ which leads to

$$\frac{dQ_1}{\frac{1}{2}(dQ_1 + dQ_2)} = \frac{1}{(\frac{dQ_1+dQ_2}{2dQ_1})} = \frac{1}{\frac{1}{2} + \frac{dQ_2}{2dQ_1}} \leq 2.$$

2. If there is a predictable stopping time τ with $\mathbb{P}(\Delta Z_\tau < 0) > 0$, we know with Proposition 2.3.4, that $\mathbb{E}^{Q_1}[\Delta Z_\tau | \mathcal{F}_{\tau-}] = 0$, which means that $\mathbb{P}(\Delta Z_\tau \geq \epsilon) > 0$ for $\epsilon > 0$ small enough. To apply Proposition 2.2.18 we use

$$1 + \Delta \frac{Z_\tau}{2} \geq 0 \text{ and } \mathbb{P}\left(1 + \Delta \frac{Z_\tau}{2} \geq 1 + \epsilon\right) > 0$$

with sub-σ-algebra $\mathcal{F}_{\tau-}$ and the random variable $(1, \Delta S_\tau)$. So there exist random variables $X \geq 0$, $\mathcal{F}_\tau$-measurable, and $Y \geq 1 + \epsilon$, $\mathcal{F}_{\tau-}$-measurable, with

$$X \leq Y,\ \mathbb{P}(X = Y) > 0,\ \mathbb{E}^{Q_1}[X \Delta S_\tau | \mathcal{F}_{\tau-}] = \mathbb{E}^{Q_1}\left[\left(1 + \Delta \frac{Z_\tau}{2}\right)\Delta S_\tau \middle| \mathcal{F}_{\tau-}\right]$$
$$\text{and } \mathbb{E}^{Q_1}[X | \mathcal{F}_{\tau-}] = \mathbb{E}^{Q_1}\left[\left(1 + \Delta \frac{Z_\tau}{2}\right)\middle| \mathcal{F}_{\tau-}\right].$$

The process $\tilde{Z}$ we are trying to construct can be defined by

$$\tilde{Z}_t := 1 - \mathbb{1}_{\{t \geq \tau\}} \frac{X - 1}{Y - 1},$$

which is a martingale with

$$\mathbb{E}^{Q_1}[1 - \tilde{Z}_t | \mathcal{F}_{\tau-}] = \mathbb{E}^{Q_1}\left[\mathbb{1}_{\{t \geq \tau\}} \frac{X - 1}{Y - 1} \middle| \mathcal{F}_{\tau-}\right] = \frac{\mathbb{E}^{Q_1}[\mathbb{1}_{\{t \geq \tau\}}(X - 1) | \mathcal{F}_{\tau-}]}{Y - 1}$$
$$= \frac{\mathbb{E}^{Q_1}[\mathbb{1}_{\{t \geq \tau\}} \Delta \frac{Z_\tau}{2} | \mathcal{F}_{\tau-}]}{Y - 1} = 0,$$

and

$$\mathbb{E}^{Q_1}\left[[\tilde{Z}, S]_t \middle| \mathcal{F}_{\tau-}\right] = \frac{\mathbb{E}^{Q_1}[\mathbb{1}_{\{t\geq\tau\}}(X-1)\Delta S_\tau|\mathcal{F}_{\tau-}]}{Y-1} = \frac{\mathbb{E}^{Q_1}[\mathbb{1}_{\{t\geq\tau\}}\Delta\frac{Z_\tau}{2}|\mathcal{F}_{\tau-}]}{Y-1} = 0,$$

which results in $\tilde{Z}$ and $\tilde{Z}S$ being martingales. With $\mathbb{P}(\tilde{Z}_T = 0) = \mathbb{P}(X = Y) > 0$ the process $\tilde{Z}$ is the desired process.

Note that $\mathbb{E}^{Q_1}[\Delta Z_\tau|\mathcal{F}_{\tau-}] = 0$ does not hold in general for a totally inaccessible τ, which rules out the above construction if there are no possible negative jumps at predictable times.

3. Finally, if there are no jumps at predictable stopping times, i.e. Z is quasi-left continuous (see Proposition 2.26 in Jacod and Shiryaev [35]). Then there is a totally inaccessible stopping time σ with $\mathbb{P}(\Delta Z_\sigma < 0) > 0$, we define $\tau = \inf\{\sigma \in [0,T] : \Delta Z_\sigma < 0\}$. In this case we can again apply Proposition 2.2.18 to ΔS_τ and $-\Delta Z_\tau \geq 0$ with $\mathbb{P}(-\Delta Z_\tau > \epsilon) > 0$ for an $\epsilon > 0$ to get $X \geq 0$ and $Y \geq \epsilon$ with $\mathbb{P}(X = Y) > 0$ and $\mathbb{E}[X\Delta S_\tau|\mathcal{F}_{\tau-}] = \mathbb{E}[-\Delta Z_\tau \Delta S_\tau|\mathcal{F}_{\tau-}]$.

With Proposition 2.3.3 we can replace the jump of Z at τ by $-X$ and add its continuous compensator,

$$M_t = Z_{t\wedge\tau} - (X + \Delta Z_\tau)\mathbb{1}_{\{t\geq\tau\}} + \left((X + \Delta Z_\tau)\mathbb{1}_{\{t\geq\tau\}}\right)^p_t,$$

with $\Delta M_\tau = -X$ and $\Delta M_t \geq -\epsilon$ by definition of τ. Since Y is $\mathcal{F}_{\tau-}$-measurable, we can construct A predictable and thus integrable with $A \geq \epsilon$ and $A_\tau = Y \geq \epsilon$. The process $\tilde{M}_t = \int_0^T \frac{1}{A_u} dM_u$ is a martingale because M is a martingale and the exponential $\mathcal{E}(\tilde{M})$ is a martingale:

$$\mathcal{E}(\tilde{M})_t = \exp(\tilde{M}_t - \frac{1}{2}\langle \tilde{M}^c, \tilde{M}^c\rangle_t) \prod_{u\leq t}(1 + \Delta\tilde{M}_u)\exp(-\Delta\tilde{M}_u).$$

Now $\mathcal{E}(\tilde{M})$ is indeed the desired process with:

(a) $\mathcal{E}(\tilde{M}) \geq 0$ as $\Delta\tilde{M}_t \geq -1$ per definition of A and τ,

(b) $\mathbb{P}(\mathcal{E}(\tilde{M})_T = 0) = \mathbb{P}(\Delta\tilde{M}_\tau = -1) = \mathbb{P}(\Delta M_\tau = A_\tau) = \mathbb{P}(X = Y) > 0$,

(c) $\mathcal{E}(\tilde{M})S$ is a martingale where S, $\mathcal{E}(\tilde{M})$ and $[\mathcal{E}(\tilde{M}), S]$ are martingales. The last expression is a martingale as

$$d[\mathcal{E}(\tilde{M}), S] = \mathcal{E}(\tilde{M})d[\tilde{M}, S] = \mathcal{E}(\tilde{M})d([Z^\tau, S] + [\mathbb{1}_{\{t\geq\tau\}}(X + \Delta Z_\tau), S]),$$

where the last bracket is a martingale using $\mathbb{E}[X\Delta S_\tau|\mathcal{F}_{\tau-}] = \mathbb{E}[-\Delta Z_\tau\Delta S_\tau|\mathcal{F}_{\tau-}]$.

Finally, in every case there is a process $\tilde{Z}$ which can be used to construct an absolutely continuous but not equivalent martingale measure, so that the set $\mathcal{Q}^{\mathbb{P}}$ cannot be closed. □

Remark 3.3.16.

Even the larger set of all absolutely continuous martingale measures, which does not allow for the above argumentation, is either a singleton or not compact as shown in Corollary 7.2 in Delbaen [15].

In order to show that the property of measure-convexity introduced here is suitable for the context of financial mathematics, it remains to be shown that most of the sets commonly used in financial mathematics are measure-convex.

Theorem 3.3.17.

The set of all probability measures on the Skorokhod space is measure-convex.

Proof. First, note that the Skorokhod space endowed with the Skorokhod topology is a separable and complete space. Now the Prokhorov Theorem, Lemma 2.5.4, states that the set of all probability measures on the Skorokhod space endowed with the Prokhorov metric is also complete. Since the set of all probability measures is also bounded and convex, Lemma 3.3.11 shows that the set is measure-convex. □

Corollary 3.3.18.

Every convex and closed set of probability measures on the Skorokhod space is measure-convex.

Proof. The set of all probability measures on the Skorokhod space endowed with Skorokhod topology is a complete space using the Prokhorov metric. Every closed subset is again complete due to Lemma 2.5.5. Theorem 3.3.11 ensures measure-convexity. □

Lemma 3.3.19.

Intersections of measure-convex sets are measure-convex.

Proof. Let X_1 and X_2 be measure-convex sets. From Theorem 3.3.12 we know that for any subsets $Y_i \subseteq X_i,\ i = 1, 2$ that are compact we have $\overline{co}(Y_i) \subseteq X_i$ is compact. Now for any subset $Y \subseteq X_1 \cap X_2$ we obviously have $Y \subseteq X_i$, where $\overline{co}(Y) \subseteq X_i$ is compact for $i = 1, 2$, which implies $\overline{co}(Y) \subseteq X_1 \cap X_2$ compact. The argument still holds for arbitrary intersections, not just for finite ones. □

Theorem 3.3.20.
The convex hull of finite unions of measure-convex sets is measure-convex.

Proof. Let X_1 and X_2 be measure-convex sets in a locally convex space X. We can show directly that any barycenter is an element of $co(X_1 \cup X_2)$.

Let y be a barycenter of $co(X_1 \cup X_2)$ with corresponding measure λ, where $\lambda(co(X_1 \cup X_2)) = 1$ and $y = \int_{co(X_1 \cup X_2)} x d\lambda(x)$. Since we can write every element $x \in co(X_1 \cup X_2)$ as a convex combination, we can define continuous functions

$$\begin{aligned} x_1 &: co(X_1 \cup X_2) \to X_1, \\ x_2 &: co(X_1 \cup X_2) \to X_2, \\ a &: co(X_1 \cup X_2) \to [0,1], \end{aligned}$$
$$\text{with } y = a(y)x_1(y) + (1-a(y))x_2(y) \text{ for every } y \in co(X_1 \cup X_2).$$

Now, for every linear functional $f \in X^*$, it holds that

$$\begin{aligned} \int\limits_{co(X_1 \cup X_2)} f(y) d\lambda(y) &= \int\limits_{co(X_1 \cup X_2)} a(y) f(x_1(y)) + (1-a(y)) f(x_2(y)) d\lambda(y) \\ &= \int\limits_{co(X_1 \cup X_2)} f(x_1(y)) a(y) d\lambda(y) + \int\limits_{co(X_1 \cup X_2)} f(x_2(y))(1-a(y)) d\lambda(y) \\ &= \int\limits_{X_1} f(x) \int\limits_{x_1^{-1}(x)} a(y) dy d\lambda(x) + \int\limits_{X_2} f(x) \int\limits_{x_2^{-1}(x)} (1-a(y)) dy d\lambda(x). \end{aligned}$$

We can define new measures $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ on X_1 and X_2 respectively by

$$\begin{aligned} d\tilde{\lambda}_1(x) &= \int\limits_{x_1^{-1}(x)} a(y) dy d\lambda(x), \\ d\tilde{\lambda}_2(x) &= \int\limits_{x_2^{-1}(x)} (1-a(y)) dy d\lambda(x). \end{aligned}$$

Note that $\frac{\tilde{\lambda}_i(x)}{\tilde{\lambda}_i(X_i)}$ are probability measures on X_i with $\tilde{\lambda}_1(X_1) + \tilde{\lambda}_2(X_2) = 1$, while X_1 and X_2 are measure-convex by assumption, which allows us to conclude

$$\begin{aligned} f(y) &= \int\limits_{co(X_1 \cup X_2)} f(y) d\lambda(y) \\ &= \int\limits_{X_1} f(x) d\tilde{\lambda}_1(x) + \int\limits_{X_2} f(x) d\tilde{\lambda}_2(x) \end{aligned}$$

$$= \tilde{\lambda}_1(X_1) \underbrace{\int_{X_1} f(x) d\frac{\tilde{\lambda}_1(x)}{\tilde{\lambda}_1(X_1)}}_{\in X_1} + \tilde{\lambda}_2(X_2) \underbrace{\int_{X_2} f(x) d\frac{\tilde{\lambda}_2(x)}{\tilde{\lambda}_2(X_2)}}_{\in X_2} \in co(X_1 \cup X_2).$$

□

Finally, we can present sets commonly used in financial mathematics that are measure-convex and thus suitable for applying robust Neyman-Pearson theory.

Corollary 3.3.21.
The following sets of measures are measure-convex:

(a) Singletons

$$\{Q\}.$$

This case corresponds to complete financial markets or cases where the usually large set of price measures is reduced to one measure using statistical or other methods.

(b) The convex hull of a finite number of measures

$$co(\{Q_1 \cup ... \cup Q_n\}).$$

When using a finite set of pricing measures, this is the smallest convex set, that contains $Q_1, ..., Q_n$.

(c) The set of all martingale measures equivalent to a market measure $\mathbb{P}$

$$\mathcal{Q}^{\mathbb{P}}.$$

This case corresponds to general incomplete financial markets.

(d) The set of all equivalent martingale measures, such that the claim remains under a specific price constraint $c \in \mathbb{R}$:

$$\{Q \in \mathcal{Q}^{\mathbb{P}} : \mathbb{E}^{Q}[C] \leq c\}.$$

For incomplete markets with very high prices, such as stochastic volatility models, this case can be useful.

(e) In general, any convex and closed set of probability measures, especially any convex and closed subset of $\mathcal{Q}^{\mathbb{P}}$.

(f) The inner convex hull of all equivalent martingale measures in a finite robust market model $\mathcal{P} = \{\mathbb{P}_1, ..., \mathbb{P}_n\}$

$$\left\{ \sum_{i=1}^{n} \lambda_i Q_i : Q_i \in \mathcal{Q}^{\mathbb{P}_i},\ \sum_{i=1}^{n} \lambda_i = 1 \text{ with } \lambda_i > 0 \right\}.$$

Proof.

a) Every singleton is compact and therefore complete, bounded and convex, see Lemma 3.3.11.

b) The convex hull of a finite number of points is compact and therefore complete, bounded and of course convex, which suffices for measure-convexity with Lemma 3.3.11.

c) Let λ be a finite measure on $\mathcal{Q}^{\mathbb{P}}$ with $\lambda(\mathcal{Q}^{\mathbb{P}}) = 1$. First of all, the expression

$$\int_{\mathcal{Q}^{\mathbb{P}}} Q d\lambda(Q)$$

is again a probability measure, since the set of all probability measures is measure-convex by Theorem 3.3.17.

Also $\int_{\mathcal{Q}^{\mathbb{P}}} Q d\lambda(Q)$ remains an equivalent measure. If it were not equivalent, there would either be a measurable set $A \in \mathcal{F}$ with

$$\mathbb{P}(A) > 0 \text{ and } \int_{\mathcal{Q}^{\mathbb{P}}} Q(A) d\lambda(Q) = 0,$$

which is a contradiction to $\lambda(\mathcal{Q}^{\mathbb{P}}) = 1$, since $Q(A) > 0$ for every $Q \in \mathcal{Q}$, or on the other hand

$$\mathbb{P}(A) = 0 \text{ and } \int_{\mathcal{Q}^{\mathbb{P}}} Q(A) d\lambda(Q) > 0,$$

which again is impossible, since in this case $Q(A) = 0$ for every $Q \in \mathcal{Q}^{\mathbb{P}}$, because $Q \sim \mathbb{P}$ for every $Q \in \mathcal{Q}^{\mathbb{P}}$.

And finally the martingale property remains preserved. Let $\tilde{Q} := \int_{\mathcal{Q}^{\mathbb{P}}} Q d\lambda(Q)$.

$$\begin{aligned} \mathbb{E}^{\tilde{Q}}[S_t] &= \int_{\Omega} S_t d\tilde{Q} = \int_{\Omega} S_t d \int_{\mathcal{Q}^{\mathbb{P}}} Q d\lambda(Q) = \int_{\mathcal{Q}^{\mathbb{P}}} \int_{\Omega} S_t dQ d\lambda(Q) \\ &= \int_{\mathcal{Q}^{\mathbb{P}}} \int_{\Omega} S_t dQ d\lambda(Q) = \int_{\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{Q}[S_t] d\lambda(Q) = \int_{\mathcal{Q}^{\mathbb{P}}} S_0 d\lambda(Q) = S_0. \end{aligned}$$

Note that these expressions must be finite by the definition of $\mathbb{P}$ and by the fact that $\int_{\mathcal{Q}^{\mathbb{P}}} Q d\lambda(Q)$ is already known to be a probability measure.

d) The space of all probability measures such that the expectation of the claim is below $c \in \mathbb{R}$

$$\{\mathbb{P} : \mathbb{E}^{\mathbb{P}}[C] \leq c\}$$

is closed. To see this, take a convergent series of measures $(\mathbb{P}_n) \subseteq A$ with $\mathbb{P}_n \to \mathbb{P}$. Then we have

$$\mathbb{E}^{\mathbb{P}_n}[C] = \int_\Omega C d\mathbb{P}_n \to \int_\Omega C d\mathbb{P} = \mathbb{E}^{\mathbb{P}}[C],$$

where the left hand side is bounded by $c \in \mathbb{R}$, such that $\mathbb{E}^{\mathbb{P}}[C] \leq c$.

As a closed subset of a complete space, the space is again complete and thus measure-convex as a closed, bounded and convex space with Lemma 3.3.11. The intersection of measure-convex sets remains measure-convex and thus $\{\mathbb{P} : \mathbb{E}^{\mathbb{P}}[C] \leq c\} \cap \mathcal{Q}^{\mathbb{P}}$ is measure-convex.

e) This is just a reformulation of Corollary 3.3.18.

f) Note that the set can be rewritten as

$$co\left(\bigcup_{i=1}^{n} \mathcal{Q}^{\mathbb{P}_i}\right) \cap \mathcal{Q}^{\mathbb{P}_\lambda},$$

where $\mathbb{P}_\lambda = \sum_{i=1}^n \lambda_i \mathbb{P}_i$ is a convex mixture of all market measures with $\sum_{i=1}^n \lambda_i = 1$ and $\lambda_i > 0$ for all $i = 1, ..., n$. Using Theorem 3.3.20, the set is an intersection of two measure-convex sets, which remains measure-convex by Lemma 3.3.19.

□

Chapter 4.

Optimality of hedging strategies

4.1. Optimal hedging strategies for single market models

If we assume that we are in a complete market, then we can find an exact hedging strategy for every claim, i.e. we can neutralise any risk. In incomplete markets, the same applies to superhedging strategies as we introduced in the previous section. Since superhedging prices are far too high from a practical point of view, it is reasonable that an investor will usually not be willing to invest the full amount of money needed to superhedge a claim. On the other hand, it makes sense for an investor to take some risk to make money. In these cases, we face the problem of finding an investment strategy that does not exceed a given constraint on investment capital, but also maximises the probability of hedging successfully or minimises some other risk measure.

To solve this problem, we construct a knockout barrier for the claim that is supposed to be hedged. With this knockout barrier, we can construct a new modified and cheaper claim that can be replicated or superhedged without exceeding the investment constraint. This approach was first introduced in Föllmer and Leukert [23], from which the proofs and results presented in this section are largely inspired and summarised.

We define $\tilde{V}_0 > 0$ as a capital constraint that cannot be exceeded by an investor, and we will try to find the optimal hedging strategy under this constraint regarding different risk measures.

4.1.1. Value at risk

One of the most widely used risk measures is the value at risk. Despite many disadvantages, not only mathematical but also practical, this risk measure is one used by regulators in the European Union to determine the solvency capital of banks and insurance companies, see Solvency 2 [20] for insurance regulation and Basel 3 [3] for banking regulation.

Definition 4.1.1. *Value at risk*
Let X be a random variable determining a loss and let $\alpha \in (0,1)$ be a security level. The value at risk of X with respect to level α represents the smallest constant c, such that $X - c$ is positive with a maximum probability of α. In mathematical terms:

$$VaR_\alpha(X) = \inf\{c \in \mathbb{R}_+ : \mathbb{P}(X - c \geq 0) \leq \alpha\} = F_X^{-1}(1-\alpha).$$

Due to the widespread use of the value at risk in risk management and its application in regulation, we will first aim to minimise the risk of a hedging strategy in terms of the value at risk. Mathematically speaking we will try to minimise the value at risk of the random variable $(C - V_T^{H,x})$, which leads to the next definition of a success set that describes all events where the hedging strategy is successful.

Definition 4.1.2. *Success sets*
Let C be a $\mathcal{F}_T$-measurable claim, let H be an admissible investment strategy and let $x > 0$ be the initial investment. We define the success set of the hedging strategy H with initial investment x as

$$A_{H,x} := \{x + \int_0^T H dS \geq C\},$$

which defines the set of all $\omega \in \Omega$ for which the strategy H successfully (super)hedges the claim C. The probability of not hedging successfully, $1 - \mathbb{P}(A_{H,x})$, is called the shortfall probability. If we do not know the corresponding hedging strategy or the initial capital, we omit the indices.

We define the set A pathwise with the idea that, for a given success set, we can construct a strategy (H, x) that exactly recreates the success set.

If we define the probability of making a loss according to the physical measure $\mathbb{P}(\Omega \setminus A)$ as a risk measure that defines how well an investment strategy hedges a claim, then we can rewrite our problem of finding the best hedging strategy under a market model $\mathbb{P}$ as follows:

Find a strategy (H, x) that solves

$$\begin{aligned} &\text{maximise } \mathbb{P}(A_{H,x}) \\ &\text{under the constraint } x \leq \tilde{V}_0. \end{aligned} \tag{4.1.1}$$

The corresponding solution $\tilde{A} = A_{\tilde{H},\tilde{x}}$ to this problem defines a new claim $\tilde{C} := \mathbb{1}_{\tilde{A}}C$, which can be superhedged with only using no more than $\tilde{V}_0$ initial investment, which is ensured by the constraint and the pricing-hedging duality for single market models, i.e. Theorem 3.2.15 with $\mathcal{P} = \{\mathbb{P}\}$.

Theorem 4.1.3. *Optimal hedging strategy under value at risk*
Let $\mathbb{P}$ be a time-continuous market measure and let C be a $\mathcal{F}_T$-measurable claim, such that C is not a modification of $\alpha\frac{d\mathbb{P}}{dQ}$ for every $\alpha \in \mathbb{R}$ and every $Q \in \mathcal{Q}^{\mathbb{P}}$. The optimisation problem (4.1.1) has a solution, which is given by the superhedging strategy for the knockout option

$$\tilde{C} = C\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}} > aC\}},$$

where the measure $\tilde{Q} \in \mathcal{Q}^{\mathbb{P}}$ defines a worst-case equivalent martingale measure and the constant a is chosen to satisfy the constraint $\tilde{V}_0$.

Proof. The proof of this result follows from the work of Föllmer and Leukert [23]. We will prove this theorem in three steps that summarise some statements and proofs in the said paper. In these steps we will first formulate a dual problem to (4.1.1) and show the uniqueness and existence of the solution. In the last step the explicit representation follows.

1) For every set A we can construct a knockout option $\tilde{C} = C\mathbb{1}_A$. From the pricing hedging duality, Theorem 3.2.15, we know, that there exists a strategy (H, x) that superhedges the claim $\tilde{C}$ with $x = \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[\tilde{C}]$, i.e. $V_T^{H,x} \geq \tilde{C}$. For a solution $(\tilde{H}, \tilde{x})$ of (4.1.1) this means that $\mathbb{P}(V_T^{\tilde{H},\tilde{x}} \geq C) \geq \mathbb{P}(V_T^{H,x} \geq C)$ for every other admissible strategy (H, x) with $x \leq \tilde{V}_0$ and we can rewrite (4.1.1) as

$$\begin{aligned} &\text{maximise } \mathbb{P}(A) \\ &\text{under the constraint } \mathbb{E}^Q[C\mathbb{1}_A] \leq \tilde{V}_0 \ \forall Q \in \mathcal{Q}^{\mathbb{P}}. \end{aligned} \tag{4.1.2}$$

2) The next step is to show, for a success set $\tilde{A}$ which solves (4.1.2), that the corresponding superhedging strategy $(\tilde{H}, \tilde{x})$ that superhedges $C\mathbb{1}_A$ generates the same success set, i.e. $\tilde{A} = A_{\tilde{H},\tilde{x}}$ and thus (4.1.2) is indeed equivalent to (4.1.1).

Suppose A is a solution of (4.1.2) and (H, x) is the cheapest superhedging strategy to $C\mathbb{1}_A$ that exists according to the pricing-hedging duality. First, note that the stochastic integral $V_t^{H,x} = x + (H \cdot S)_t$ is a local martingale under every local martingale measure $Q \in \mathcal{Q}^{\mathbb{P}}$. This means that $x + (H \cdot S)_t$ is a supermartingale and we get the inequality

$$\mathbb{E}^Q[C\mathbb{1}_{A_{H,x}}] \leq \mathbb{E}^Q[x + (H \cdot S)_T] \leq x \ \forall Q \in \mathcal{Q},$$

which means that (H, x) satisfies the constraints in (4.1.2). Since A is a solution of (4.1.2), it is also clear that

$$\mathbb{P}(A) \geq \mathbb{P}(A_{H,x}). \tag{4.1.3}$$

On the other hand we can see that

$$A \subseteq \{\omega \in \Omega : C\mathbb{1}_A \geq C\} \subseteq A_{H,x}. \tag{4.1.4}$$

With (4.1.3) and (4.1.4) it follows that $A = A_{H,x}$, i.e. the superhedging strategy of a solution to (4.1.2) indeed solves the optimisation problem (4.1.1).

3) To find a more explicit representation of the set $\tilde{A}$, we introduce a new set of equivalent measures $\{R(Q)|Q \in \mathcal{Q}\}$ via the equation

$$\frac{dR(Q)}{dQ} = \frac{C}{\mathbb{E}^Q[C]}.$$

With these measures we can rewrite the constraint in (4.1.2) as

$$\mathbb{E}^Q[C\mathbb{1}_A] = \mathbb{E}^R\left[C\mathbb{1}_A\frac{E^Q[C]}{C}\right] = \mathbb{E}^Q[C]E^R[\mathbb{1}_A] = \mathbb{E}^Q[C]R(A).$$

This gives us the new optimisation problem

$$\text{maximise } \mathbb{P}(A) \tag{4.1.5}$$

$$\text{under the constraint } R(A_{H,x}) \leq \frac{\tilde{V}_0}{\mathbb{E}^Q[C]} =: \alpha \; \forall Q \in \mathcal{Q}. \tag{4.1.6}$$

This optimisation problem is in the form of testing a hypothesis $\mathbb{P}$ against hypotheses $\{R(Q)|Q \in \mathcal{Q}\}$ with a given maximum error of the first kind $\alpha := \frac{\tilde{V}_0}{E^Q[C]}$. Note that the assumption that C is not a modification of $\alpha\frac{d\mathbb{P}}{dQ}$ implies that

$$\mathbb{P} \notin \{R(Q) : Q \in \mathcal{Q}\},$$

which is important, since otherwise $\frac{dR}{dQ} = \frac{C}{\mathbb{E}^Q[C]} = \frac{\alpha\frac{d\mathbb{P}}{dQ}}{\mathbb{E}^Q[C]}$ implies $\frac{dR}{d\mathbb{P}} \equiv 1$ for a $R \in \{R(Q) : Q \in \mathcal{Q}\}$. This fact allows us to use the robust Neyman-Pearson Theorem 3.3.14 and Corollary 3.3.21, which means that we know that there exists a worst case equivalent martingale measure $\tilde{Q}$ and a corresponding measure $\tilde{R} = R(\tilde{Q})$.

With these measures we can construct the solution to Problem (4.1.5) using

$$\tilde{a} := \inf\left\{a : \tilde{R}\left(\frac{d\mathbb{P}}{dQ} > aC\right) \leq \alpha\right\}, \tag{4.1.7}$$

and we can define the optimal success set with

$$\tilde{A} = \left\{ \frac{d\mathbb{P}}{d\tilde{Q}} > \tilde{a}C \right\}.$$

As already mentioned in the first step, we can find a superhedging strategy $(\tilde{H}, \tilde{x})$ for the option $C\mathbb{1}_{\tilde{A}}$. This strategy solves the original optimisation problem (4.1.1) as shown in step 2. □

Remark 4.1.4.

As long as we work with time-continuous, non-deterministic models, it holds that $\mathbb{P}(\{\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C\}) = 0$ and the above optimal solutions exist. Otherwise, the optimal Neyman-Pearson test could be a randomised test leading to a knockout option that provides a payout in the cases $\{\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C\}$ that depends on an external, independent source of randomness, such as a coin flip or, more mathematically, a Bernoulli distribution. More precisely, if $\{\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C\}$ is not a null set, the transformed claim $\tilde{C}$ is generally of the form

$$\tilde{C} = (\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}} > \tilde{a}C\}} + B\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C\}})C,$$

where $B \sim Ber(p)$ is a Bernoulli distributed random variable with

$$p = \frac{\tilde{V}_0 - \sup_{Q \in \mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}} > \tilde{a}C\}} C]}{\frac{1}{\tilde{a}} \mathbb{P}(\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C)}.$$

However, this setup is not only very unusual from a practical point of view, it is also mathematically unfavourable because without adjustments this claim is not well defined, since the sample space Ω, i.e. the Skorokhod space in this case, does not allow $C : \Omega \to \mathbb{R}_+$ to be defined by an additional random variable B. Therefore, the pricing-hedging duality does not ensure the existence of a corresponding superhedging strategy. On the contrary, there are several simple counterexamples, where the pricing-hedging duality does not hold for additional random sources.

The issue of a potential non-null set $\{\frac{d\mathbb{P}}{d\tilde{Q}} = \tilde{a}C\}$ can be solved by using a risk measure that rewards not only successful hedges, but also scenarios where the shortfall is reduced. One example of such a risk measure is the success ratio, which also takes into account the ratio of claim to hedge in cases where the hedging strategy fails.

It should be noted that in the following work we always assume time-continuous market models and therefore do not need to consider the case of Remark 4.1.4.

4.1.2. Optimal hedging strategies under different risk measures

The previously used proof techniques in Theorem 4.1.3 can be applied not only to the value at risk, but also to many other risk measures. On the one hand, the previous risk measure can be generalised to the success ratio taking into account the shortfall in the event of losses, which solves the problem of randomisation in discrete-time models. On the other hand one can use the expected shortfall, another frequently used risk measure in regulation, see for example the Swiss solvency test [64], the swiss counterpart to European Solvency 2. It is then shown that even for arbitrary coherent risk measures, the optimal hedging strategy can be determined using the robust Neyman-Pearson lemma.

4.1.2.1. Success ratio

First, we consider another risk measure that is broadly consistent with the value at risk, but takes into account the ratio of claim to hedging strategy in cases where a loss is incurred and penalises large losses. Interestingly, the optimal hedging strategy remains in a 0-1 form, with the exception of randomisation, which mostly occurs on a null-set.

Definition 4.1.5. *Success ratio*
For a hedging strategy (H, x) and a claim $C : \Omega \to \mathbb{R}_+$, the success ratio is defined as

$$\varphi(H, x) = \mathbb{1}_{\{V_T^{H,x}\} \geq C\}} + \frac{V_T^{H,x}}{C} \mathbb{1}_{\{V_T^{H,x} < C\}}.$$

Theorem 4.1.6. *Optimal hedging strategy under success ratio*
Let $\mathbb{P}$ be a market measure and let C be a $\mathcal{F}_T$-measurable claim, such that C is not a modification of $\alpha \frac{d\mathbb{P}}{dQ}$ for every $\alpha \in \mathbb{R}$ and every $Q \in \mathcal{Q}^{\mathbb{P}}$. The hedging strategy (H, x) that solves

$$\begin{aligned} &\text{maximise } \mathbb{E}^{\mathbb{P}}[\varphi(H, x)] \\ &\text{under the constraint } x \leq \tilde{V}_0 \end{aligned}$$

is the superhedging strategy for the claim

$$\tilde{C} = \left(\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > aC\}} + \gamma \mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} = aC\}} \right) C,$$

where a is defined as in Equation (4.1.7) and $\gamma = \frac{V_0 - \sup_{Q \in \mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q\left[\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > aC\}} C\right]}{\mathbb{E}^Q\left[\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} = aC\}} C\right]}$.

Proof. The results follow analogously to the previous proof of Theorem 4.1.3. □

Remark 4.1.7.
If $\mathbb{P}(\frac{d\mathbb{P}}{d\tilde{Q}} = aC) = 0$ for every constant a, which is the case for every time continuous diffusion model, then the hedging strategy that optimises shortfall probability and the strategy that optimises the success ratio do coincide.

Since we are working in time-continuous market models, we use the shortfall probability due to easier notation. On the other hand, using the success ratio has the advantage that it is much easier to prove concavity of the optimal success ratio as a function of the capital constraint, see for example Figure 5.2.3.

Corollary 4.1.8.
The optimal expected success ratio and the optimal shortfall probability as a function of the initial investment are continuous, non-decreasing and concave.

Proof. Using ideas in the proof of Lemma 3.1 in Bayraktar and Wang [4] we define optimal hedging strategies (H_1, x_1) and (H_2, x_2) with $x_1 < x_2 \leq \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[C]$. First of all, due to optimality it holds that

$$\mathbb{E}^{\mathbb{P}}[\varphi(H_1, x_1)] < \mathbb{E}^{\mathbb{P}}[\varphi(H_2, x_2)],$$

as H_1 is admissible under the initial investment x_2 as well, therefore it holds that $V_T^{H_1,x_2} > V_T^{H_1,x_1}$, which already leads to $\mathbb{E}^{\mathbb{P}}[\varphi(H_2, x_2)] \geq \mathbb{E}^{\mathbb{P}}[\varphi(H_1, x_2)] > \mathbb{E}^{\mathbb{P}}[\varphi(H_1, x_1)]$. In addition, it can be seen that a convex combination of success ratios leads to a new success ratio that can be superhedged with less or the same initial investment:

$$\begin{aligned}\sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[(\lambda\varphi(H_1, x_1) + (1-\lambda)\varphi(h_2, x_2))C] \leq& \lambda \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[\varphi(H_1, x_1)C] \\ &+ (1-\lambda) \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[\varphi(H_2, x_2)C] \\ =& \lambda x_1 + (1-\lambda)x_2.\end{aligned}$$

In a final step, the expected success ratio is additive:

$$\mathbb{E}^{\mathbb{P}}[\lambda\varphi(H_1, x_1) + (1-\lambda)\varphi(H_2, x_2)] = \lambda\mathbb{E}^{\mathbb{P}}[\varphi(H_1, x_1)] + (1-\lambda)\mathbb{E}^{\mathbb{P}}[\varphi(H_2, x_2)].$$

Finally, it follows that there exists an optimal strategy H such that

$$\mathbb{E}^{\mathbb{P}}[\varphi(H, \lambda x_1 + (1-\lambda)x_2)] > \lambda\mathbb{E}^{\mathbb{P}}[\varphi(H_1, x_1)] + (1-\lambda)\mathbb{E}^{\mathbb{P}}[\varphi(H_2, x_2)],$$

where we can use that the superhedging strategy to $(\lambda\varphi(H_1, x_1) + (1 - \lambda)\varphi(h_2, x_2))C$ uses an initial investment of $\lambda x_1 + (1 - \lambda)x_2$ but is not optimal, since it is not of 0-1-form. This means that the expected optimal success ratio is non-decreasing and concave as a function of the initial investment. Continuity follows because every non-decreasing concave function is also continuous. For values larger than $\sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^Q[C]$ the function is constant 1.

The same arguments hold for the value at risk, since both risk measures coincide in time-continuous markets. □

4.1.2.2. Expected shortfall

Another reasonable and regularly used risk measure is the so called expected shortfall. Like the value at risk, the expected shortfall is a widely used risk measure in regulation and risk management. While in the EU, for example, solvency 2 prescribes the value at risk to determine the risk capital of insurance companies, the Swiss solvency test uses the expected shortfall, also known as the average value at risk. This setup is also examined in Föllmer [22] and the problem of finding optimal hedging strategies regarding the expected shortfall can be solved using the same arguments as before. The solution is different in this case, but is still in a 0-1 form, which follows from Neyman-Pearson theory.

Definition 4.1.9.
Let C be a $\mathcal{F}_T$-measurable claim, let H be an admissible investment strategy and let $x \in \mathbb{R}$ be the initial investment. The expected shortfall of (H, x) is defined as

$$\mathbb{E}^{\mathbb{P}}[(C - V_T^{H,x})^+].$$

A major advantage over the shortfall probability is that the expected shortfall also takes into account the loss that occurs in the cases where the hedging strategy fails.

Using the same argumentation as in Theorem 4.1.3 one can again find a unique hedging strategy that minimises the expected shortfall under a capital constraint $x \leq \tilde{V}_0$.

Theorem 4.1.10. *Optimal hedging strategy under expected shortfall*
Let $\mathbb{P}$ be a market model, such that $\mathbb{P} \notin \mathcal{Q}^{\mathbb{P}}$, and let C be a $\mathcal{F}_T$-measurable claim. The optimisation problem

$$\text{minimise } \mathbb{E}^{\mathbb{P}}[(C - V_T^{H,x})^+]$$

$$\text{under the constraint } x \leq \tilde{V}_0$$

has a unique solution, which is given by the superhedging strategy to the option

$$\tilde{C} = C\left(\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}}>a\}} + \gamma\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}}=a\}}\right).$$

The measure $\tilde{Q} \in \mathcal{Q}^{\mathbb{P}}$ defines a worst-case equivalent martingale measure.

Proof. The proof works almost analogously to the proof of Theorem 4.1.3 and therefore follows ideas in Föllmer [22]. Note that we are looking for a hedging strategy that is the superhedging strategy for a claim $\tilde{C} = \varphi C$, where $\varphi : \Omega \to [0,1]$ is a test function. Using this equation and thinking of φ as a hedging ratio, we can rewrite the minimisation problem as

$$\begin{aligned} &\text{minimise } \mathbb{E}^{\mathbb{P}}[(1-\varphi)C] \\ &\text{under the constraint } \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{Q}[\varphi C] \leq \tilde{V}_0, \end{aligned} \tag{4.1.8}$$

which rewrites into the maximization problem

$$\begin{aligned} &\text{maximise } \mathbb{E}^{\mathbb{P}}[\varphi C] \\ &\text{under the constraint } \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{Q}[\varphi C] \leq \tilde{V}_0. \end{aligned}$$

With a new set of measures $\{R : \frac{dR}{d\mathbb{P}} = \frac{C}{\mathbb{E}^{\mathbb{P}}[C]}\}$ and $\{\tilde{R} : \frac{d\tilde{R}}{dQ} = \frac{C}{\mathbb{E}^{Q}[C]}\}$ we can formulate the optimisation problem as follows, so that it is solved by Neyman-Pearson theory as shown in Theorem 3.3.14.

$$\begin{aligned} &\text{maximise } \mathbb{E}^{R}[\varphi] \\ &\text{under the constraint } \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{\tilde{R}}[\varphi] \leq \frac{\tilde{V}_0}{\mathbb{E}^{Q}[C]} := \alpha. \end{aligned}$$

This gives the optimal solution

$$\tilde{\varphi} = \mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}}>a\}} + \gamma\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}}=a\}}.$$

Finally, it only remains to show that the superhedging strategy corresponding to a solution $\tilde{\varphi}$ of (4.1.8) is indeed the optimal hedging strategy. Let $(\tilde{H}, \tilde{x})$ be the superhedging strategy to $\tilde{\varphi}C$ and let $\varphi(\tilde{H}, \tilde{x})$ be the success ratio of $(\tilde{H}, \tilde{x})$. Using the pricing-hedging duality it is clear that $\tilde{x} \leq \tilde{V}_0$, such that the superhedging strategy satisfies the capital constraint. Then it follows:

$$\mathbb{E}^{\mathbb{P}}[(1-\varphi(\tilde{H},\tilde{x}))C] \geq \mathbb{E}^{\mathbb{P}}[(1-\tilde{\varphi})C],$$

since $\tilde{\varphi}$ solves (4.1.8) and

$$\varphi(\tilde{H},\tilde{x})C = \min\{V_T^{\tilde{H},\tilde{x}}, C\} \geq \tilde{\varphi}C,$$

because $(\tilde{H},\tilde{x})$ is the superhedging strategy to $\tilde{\varphi}C$. This finally leads to $\varphi(\tilde{H},\tilde{x})C = \tilde{\varphi}C$. □

Even more generally, the l weighted expected shortfall can be defined to model risk aversion or affinity. The optimisation problem is still solvable, but becomes much more complex in its general form. This result is not proved here, but is presented as an outlook on how general the class of risk measures can be chosen using these proof techniques.

Definition 4.1.11.
Let again C be a $\mathcal{F}_T$-measurable claim, let H be an admissible investment strategy and let $x \in \mathbb{R}$ be the initial investment. In addition, let $l : [0,\infty) \to [0,\infty)$ be an increasing convex function with $l(0) = 0$. The l weighted expected shortfall of (H,x) is defined as

$$\mathbb{E}^{\mathbb{P}}[l((C - V_T^{(H,x)})^+)].$$

Theorem 4.1.12. *Optimal hedging strategy regarding expected shortfall*
Take again a market model $\mathbb{P}$ with $\mathbb{P} \notin \mathcal{Q}^{\mathbb{P}}$ and a $\mathcal{F}_T$-measurable claim C and let l be a strictly convex function with $l(0) = 0$, $l'(0+) = 0$ and $l'(\infty) = \infty$. Let furthermore $\mathbb{P}$ be a complete market, i.e. $\mathcal{Q} = \{Q\}$. The optimisation problem

$$\begin{aligned} &\text{minimise } \mathbb{E}^{\mathbb{P}}[l((C - V_T^{(H,x)})^+)] \\ &\text{under the constraint } x \leq \tilde{V}_0 \end{aligned}$$

has a unique solution, which is given by the superhedging strategy to the option

$$\tilde{C} = C\left(1 - \left(\frac{(l')^{-1}\left(a\frac{dQ}{d\mathbb{P}}\right)}{C} \wedge 1\right)\right),$$

with a suitable constant a such that the constraint is satisfied.

Proof. The result is a combination of Theorem 5.1 in Föllmer and Leukert [22] and the pricing-hedging duality 3.2.15. □

4.1.2.3. Coherent risk measures

As shown in Huber [34], coherent risk measures allow for a representation as a supremum of expectations. With robust Neyman-Pearson theory, or more specifically Theorem 3.3.14, this allows to find optimal hedging strategies even for general the general class of coherent risk measures.

Definition 4.1.13. *Coherent risk measures*
A risk measure $\rho : \mathcal{L} \to \mathbb{R} \cup \{\infty\}$ is called coherent if for every $X_1, X_2 \in \mathcal{L}$:

(a) $\rho(0) = 0$, (normalization)

(b) If $X_1 \leq X_2$, then $\rho(X_1) \leq \rho(X_2)$, (monotonicity)

(c) $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$, (subadditivity)

(d) $\rho(\alpha X_1) = \alpha\rho(X_1)$ for any $\alpha \geq 0$, (positive homogeneity)

(e) $\rho(X_1 + a) = \rho(X_1) + a$ for any $a \in \mathbb{R}$ or any random variable that equals a almost surely. (translation invariance)

Lemma 4.1.14.
A risk measure ρ is a lower semi-continuous, coherent risk measure if and only if there is a dominated set of measures $\mathcal{R}$ such that the set $\{\frac{dR}{d\mathbb{P}} | R \in \mathcal{R}\}$ is weak*-closed, convex and it holds that

$$\rho(X) = \sup_{R \in \mathcal{R}} \mathbb{E}^R[-X].$$

Proof. This statement is a well known fact in the risk theory. It was first established in Proposition 10.1 in chapter 10 of Huber [34], whose proof we will adapt to our notation.

It is straightforward to show that for any $\mathcal{R}$ the expression $\sup_{R \in \mathcal{R}} E^R[-X]$ satisfies the properties of a coherent risk measure. It only remains to show that for any coherent risk measure ρ there exists a set $\mathcal{R}$ representing ρ. It suffices to show that for any random variable X_0 there exists a measure R such that $\mathbb{E}^R[X] \leq \rho(X)$ and $\mathbb{E}^R[X_0] = \rho(X_0)$. Then the set of all these measures is the set representing ρ.

In a first step one can show the existence of a suitable function f that can define the desired measure R, in a second step we need to prove that this f in fact has the properties to define a probability measure.

1) Let X_0 be a random variable such that without loss of generality $\rho(X_0) = 1$. Otherwise we can just consider $\frac{X}{\rho(X)}$. We can define the open and convex set

$$U := \{X|\rho(X) < 1\}.$$

The set is convex because of the subadditivity of ρ and open because for every $X \in U$ it holds $\rho(X + \epsilon) = \rho(X) + \epsilon < 1$ for every $\epsilon < 1 - \rho(X)$.

Since U is open and convex and $X_0 \notin U$, the Hahn Banach separation theorem implies the existence of a linear functional $\tilde{f}$ with

$$\tilde{f}(X) < \tilde{f}(X_0) \tag{4.1.9}$$

for every $X \in U$. Using $0 = \tilde{f}(0) < \tilde{f}(X_0)$ one can, without loss of generality, choose $f(X) = \frac{\tilde{f}(X)}{\tilde{f}(X_0)}$, such that $f(X_0) = 1$. This functional will define the desired measure R via $f(X) = \mathbb{E}^R[X]$. It remains that on the one hand we need $f(X) < \rho(X)$ and on the other hand the measure defined by f must be a probability measure.

2) For any X with $\rho(X) < 1$ it follows $X \in U$ by the definition of U, which implies with (4.1.9) that $f(X) < \rho(X_0) = 1$ or in short:

$$\rho(X) < 1 \quad \Rightarrow \quad f(X) < 1. \tag{4.1.10}$$

Finally, for any $c > 0$ and $X \geq 0$, it follows from $\rho(-cX) \leq 0 < 1$ (which holds due to monotonicity of ρ with $X_2 = 0$) that $cf(X) = -f(-cX) > -1$. We can see that $f(X) > -\frac{1}{c}$ for any $c > 0$, which implies that f is a positive functional.

Let $c < 1$: By (4.1.10) it directly follows that $cf(1) = f(c) < 1$ for every $c < 1$, implying $f(1) \leq 1$.

Let $c > 1$: Now we can use $\rho(2X_0 - c) = 2\rho(X_0) - c = 2 - c < 1$ to conclude again with (4.1.10) that $2f(X_0) - c = f(2X_0 - c) < 1$, implying $f(1) > \frac{1}{c}$ for every $c > 1$, such that $f(1) \geq 1$. Combined we have

$$f(1) = 1 \tag{4.1.11}$$

In a last step, $\rho(X) < c$ implies $\rho(\frac{X}{c}) < 1$, which means, using again (4.1.10), that $f(\frac{X}{c}) < 1$ and finally $f(X) < c$, in short

$$\rho(X) < c \quad \Rightarrow f(X) < c. \tag{4.1.12}$$

Now (4.1.12) shows that $f(X) < \rho(X)$. This can be shown by taking an arbitrary $c > 1$ and using $\rho(\frac{X}{\rho(X)}) = 1 < c$, which means $f(\frac{X}{\rho(X)}) < c$. This implies $f(X) < c\rho(X)$ for every $c > 1$. In addition (4.1.11), together with the linearity properties ensures that $R(A) = f(\mathbb{1}_A)$ is in fact a probability measure. Note that $f(1) = 1$ provides $R(\Omega) = 1$. This means for any X_0 we can find the desired measure R, which proves the claim.

Note that this argumentation is limited to finite Ω but is generalised in Föllmer and Schied [24]. □

Using the above property for coherent risk measures ρ with a set of equivalent martingale measures $\mathcal{Q}$ allows to solve the optimal hedging problem

$$\min_{(H,x)} \rho\left((C - V_T^{(H,x)})^+\right) \tag{4.1.13}$$
$$\text{under the constraint } x \leq \tilde{V}_0$$

by representing it as the optimisation problem

$$\min_{\varphi} \sup_{R \in \mathcal{R}} \mathbb{E}^R[(1-\varphi)C]$$
$$\text{under the constraint } \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[\varphi C].$$

Again, as in Rudloff [58] without the assumption of measure-convexity, one can show that the solution to this optimisation problem will be in the typical 0-1 structure and very similar to the structure we already could see for the other risk measures.

Theorem 4.1.15.
The optimal hedging problem (4.1.13) is solved by the superhedging strategy (H, x) to the claim

$$\tilde{C} = \mathbb{1}_{\{\frac{d\tilde{R}}{d\tilde{Q}} > aC\}} C$$

with worst-case measures $\tilde{R} \in \mathcal{R}$ and $\tilde{Q} \in \mathcal{Q}$.

Proof. Since the set $\mathcal{R}$ is weak*-closed and $\mathcal{Q}$ is measure-convex, because it is assumed to be a set of equivalent martingale measures we can apply Theorem 3.3.14. □

4.2. Optimal Hedging Strategies for robust market models

In the preceding section, optimal hedging strategies were determined in relation to a single, clearly defined market model. Since not only Neyman Pearson theory, which is fundamental to the construction of optimal hedging strategies, can be extended to sets of measures (see

Section 3.3), but also the pricing-hedging duality can be extended (see Section 3.2), it is reasonable that the construction of optimal hedging strategies, more precisely Theorem 4.1.3 and the following corresponding theorems, can be extended to the robust market models $\mathcal{P}$ defined in Section 3.1.

The presented theory of robust modelling allows for two different approaches. First, we can consider Convex sets, which mostly can describe small deviation from a specific market model. Second, we can consider a countable set of measures, i.e. explicitly not convex sets of measures, which is more suitable to describe different scenarios $\{\mathbb{P}_1, \mathbb{P}_2, ...\}$ that can differ greatly from each other. In that case these different scenarios, that all could occur in reality, represent another kind of model risk, that cannot be captured by allowing small deviations from a single market measure.

From the perspective of Theorem 3.3.14 and Assumption 3.3.3 it is useful to take convex and closed sets of market measures. Indeed, a very common form of robust modelling is allowing for arbitrary deviation of a model regarding a specific metric.

Example 4.2.1.
Let $\tilde{\mathbb{P}}$ be a fixed probability measure on $(\Omega, \mathcal{F})$ that resembles the chosen physical measure. The following sets are regularly used kinds of robustness regarding a measure.

- ϵ-contamination:

$$\mathcal{P} = \{\mathbb{P} : (1-\epsilon)\tilde{\mathbb{P}} + \epsilon H,\ H \text{ a probability measure on } \Omega\} \text{ or}$$
$$\mathcal{P} = \{\mathbb{P} : (1-\epsilon)\tilde{\mathbb{P}} + \epsilon H,\ H \in \mathcal{H}\} \text{ with } \mathcal{H} \text{ a closed set of measures.}$$

- total variation norm:

$$\mathcal{P} = \{\mathbb{P} : |\mathbb{P}(A) - \tilde{\mathbb{P}}(A)| \leq \epsilon \ \ \forall A \in \mathcal{F}\}.$$

- Prokhorov-distance:

$$\mathcal{P} = \{\mathbb{P} : \mathbb{P}(A) \leq \tilde{\mathbb{P}}(A^\delta) + \epsilon\} \text{ with } A^\delta \text{ the closed } \delta\text{-neighbourhood of } A.$$

Theorem 4.2.2.
The robust market models in Example 4.2.1 are convex and closed, which means they suffice the Assumption 3.3.3, as long as $\mathcal{P} \cap \mathcal{Q}^{\tilde{\mathbb{P}}} = \emptyset$

Proof. Let $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}$ and let $(\mathbb{P}_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}$ be a sequence of measures with existing limit.

(a) ϵ-contamination:

- Convexity:

$$\begin{aligned}\lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2 &= \lambda((1-\epsilon)\tilde{\mathbb{P}} + \epsilon H_1) + (1-\lambda)((1-\epsilon)\tilde{\mathbb{P}} + \epsilon H_2)\\ &= (1-\epsilon)\tilde{\mathbb{P}} + \epsilon(\lambda H_1 + (1-\lambda)H_2),\end{aligned}$$

  which is sufficient for convexity if $\mathcal{H}$ is convex.

- Closedness:

$$\lim_{n\to\infty}\mathbb{P}_n = \lim_{n\to\infty}(1-\epsilon)\tilde{\mathbb{P}} + \epsilon H_n = (1-\epsilon)\tilde{\mathbb{P}} + \epsilon\lim_{n\to\infty}H_n,$$

  which proves closedness if and only if $\mathcal{H}$ is closed.

(b) Total Variation distance:

- Convexity:

$$\begin{aligned}|\lambda\mathbb{P}_1(A) + (1-\lambda)\mathbb{P}_2(A) - \tilde{\mathbb{P}}(A)| &= \lambda(\mathbb{P}_1(A) - \tilde{\mathbb{P}}(A)) + (1-\lambda)(\mathbb{P}_2(A) - \tilde{\mathbb{P}}(A))|\\ &\le \lambda|\mathbb{P}_1(A) - \tilde{\mathbb{P}}(A)| + (1-\lambda)|\mathbb{P}_2(A) - \tilde{\mathbb{P}}(A)| \le \epsilon\end{aligned}$$

- Closedness: The closedness follows directly by using the total variation distance as metric on $\mathcal{P}$.

(c) Prohorov distance:

- Convexity:

$$\lambda\mathbb{P}_1(A) + (1-\lambda)\mathbb{P}_2(A) \le \lambda(\tilde{\mathbb{P}}(A^\delta) + \epsilon) + (1-\lambda)(\tilde{\mathbb{P}}(A^\delta) + \epsilon) = \tilde{\mathbb{P}}(A^\delta) + \epsilon$$

- Closedness: The Prokhorov metric metrizes weak convergence, see Theorem 6.8 in Billingsley [6], which is sufficient to show weak closedness.

□

These models consist of infinitely many, often uncountably many measures, which only allows a solution in the minimax form. This minimax optimising approach is a very well researched with plenty of literature, see for example Kirch [38] for a financial context or Huber [34] for robust statistics. Therefore, in this thesis we will specifically consider robust models containing only finitely many market measures.

The assumptions for the robust form of the Neyman Pearson theorem to apply include a convex set of physical measures $\mathcal{P}$, see Assumption 3.3.3. However, if a finite or discrete set of measures is used, the theorem cannot be easily applied. In this section, we show that for a finite set of measures there is a way to work without this assumption. In addition, using only finitely many models allows a much more detailed construction of optimal hedging strategies and, accordingly, allows for additional optimal solutions besides the already known minimax solution.

In order to construct optimal hedging strategies in this case, we need to give a new definition for the optimality of a hedging strategy. For this section if not stated otherwise, let $\mathcal{P}$ be a finite set of measures.

$$\mathcal{P} = \{\mathbb{P}_1, ..., \mathbb{P}_n\}.$$

4.2.1. Value at risk

As in the previous section on single market models, we also start here with the value at risk, i.e. shortfall probability, and then generalise the statements to other risk measures.

Definition 4.2.3. *Robust optimal hedging strategies regarding shortfall probability*
A hedging strategy $\tilde{H}$ with initial investment $\tilde{x} \leq \tilde{V}_0$ on an option C is said to be robust optimal with respect to the shortfall probability if for every other admissible strategy (H, x) with $x \leq \tilde{V}_0$ and

$$\mathbb{P}_i(V_T^{H,x} \geq C) > \mathbb{P}_i(V_T^{\tilde{H},\tilde{x}} \geq C)$$

there exists $j \neq i$ with

$$\mathbb{P}_j(V_T^{\tilde{H},\tilde{x}} \geq C) > \mathbb{P}_j(V_T^{H,x} \geq C).$$

This means that any strategy that leads to a higher probability of success under $\mathbb{P}_i$ must have a worse success probability under at least one other measure. In the following, the term "optimal strategy" generally means "robust optimal with respect to shortfall probability".

If we reduce the set of measures $\mathcal{P}$ to just two measures, namely $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$, it follows directly from the previous chapter that we can obtain optimal strategies and optimal success sets A_1 and A_2 that are optimal under $\mathbb{P}_1$ and $\mathbb{P}_2$ respectively, without considering the other measure. In general, these sets will not coincide and will not be optimal under the other measure, i.e. $\mathbb{P}_1(A_2) < \mathbb{P}_1(A_1)$ and $\mathbb{P}_2(A_1) < \mathbb{P}_2(A_2)$. The question is whether there are optimal solutions between these extremes, and if so, is there a way to find them and their probabilities of success? These are the two questions we address in the following section.

The main idea in this chapter is to define a convex combination of measures for which we write

$$\mathbb{P}_\lambda := \sum_{i\in I} \lambda_i \mathbb{P}_i \text{ with } \sum_{i\in I} \lambda_i = 1,\ \lambda_i > 0 \text{ and } \mathbb{P}_i \in \mathcal{P}$$

and apply the previous theory for single market models to these new market measures. For a simpler notation, in the proofs of the following main results we reduce the set of market measures to $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$, which means that $\mathbb{P}_\lambda$ is of the form

$$\mathbb{P}_\lambda := \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$$

with $\lambda \in (0,1)$. It will be seen that we can extend this construction to arbitrary finite sets of measures, which generalises the following proofs to finite sets for $\mathcal{P}$. In Section 3.3 we already proved that the convex combination of a countable set of probability measures is again a probability measure.

The first step in this section is to show in Theorem 4.2.8 that with this idea we can create optimal solutions according to the previous Definition 4.2.3. The second statement in the Theorem 4.2.9 shows that not only can we create optimal solutions with this approach but we can find every optimal solution. Finally, in Theorem 4.2.16 we will examine further properties of these optimal solutions, which means we will show that the set of all hedging strategies is strictly convex in terms of success probabilities.

Definition 4.2.4.
For a robust market model $\mathcal{P} = \bigcup_{i\in I}\{\mathbb{P}_i\}$, the set $\mathcal{Q}_\lambda$ is defined as

$$\mathcal{Q}_\lambda := \left\{ \sum_{Q_i \in \mathcal{Q}^{\mathbb{P}_i}} \lambda_i Q_i : \sum_{i\in I} \lambda_i = 1, \lambda_i > 0 \right\},$$

i.e. $\mathcal{Q}_\lambda$ is defined as the set of convex combinations of equivalent martingale measures.

Defining this set is necessary because every measure contained in $\mathcal{Q}_\lambda$ is equivalent to every convex combination $\mathbb{P}_\lambda$ of measures in $\mathcal{P}$, which does not necessarily hold for the equivalent measures in $\mathcal{Q}^{\mathbb{P}_i}$.

Lemma 4.2.5.
Let $\mathcal{P}$ be a countable robust market model and I an arbitrary index set on $\mathcal{P}$. Using the

set $\mathcal{Q}_\lambda$ for pricing under $\mathbb{P}_\lambda$ does not change the supremum of all arbitrage free prices. This means

$$\sup_{Q \in \mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_A C] = \sup_{Q \in \bigcup_{i \in I} \mathcal{Q}^{\mathbb{P}_i}} \mathbb{E}^Q[\mathbb{1}_A C]$$

for every $A \in \mathcal{F}$.

Proof. Since the expectation is a linear functional, the lemma follows by

$$\begin{aligned}
\sup_{Q \in \mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_A C] &= \sup_{\substack{Q_i \in \mathcal{Q}^{\mathbb{P}_i} \\ \lambda_i \in (0,1) \\ \sum \lambda_i = 1}} \mathbb{E}^{\sum_{i \in I} \lambda_i Q_i}[\mathbb{1}_A C] \\
&= \sup_{\substack{Q_i \in \mathcal{Q}^{\mathbb{P}_i} \\ \lambda_i \in (0,1) \\ \sum \lambda_i = 1}} \sum_{i \in I} \lambda_i \mathbb{E}^{Q_i}[\mathbb{1}_A C] \\
&= \sup_{\substack{i \in I \\ Q_i \in \mathcal{Q}^{\mathbb{P}_i}}} \mathbb{E}^{Q_i}[\mathbb{1}_A C] \\
&= \sup_{Q \in \bigcup_{i \in I} \mathcal{Q}^{\mathbb{P}_i}} \mathbb{E}^Q[\mathbb{1}_A C].
\end{aligned}$$

The third equation holds, since the expectations are deterministic values, where the largest of these values determines the supremum. □

This lemma shows that considering $\mathcal{Q}_\lambda$ as a new set of equivalent martingale measures does not produce prices that exceed the previous upper bound, which is important since we are creating a new option that is supposed to be superhedged in $\mathcal{P}$. However, $\mathcal{Q}_\lambda$ only contains measures that are equivalent to $\mathbb{P}_\lambda$ which is important in the proof of Theorem 4.2.9.

Remark 4.2.6.

As an interesting side note, $\mathcal{Q}_\lambda$ does not necessarily include every equivalent martingale measure of $\mathbb{P}_\lambda$, and it is possible that $\mathbb{P}_\lambda$, as a new market measure, would lead to higher prices than any single market measure in $\mathcal{P}$.

As a simple example, consider two binomial models with

- $\mathbb{P}_1$ with $u_1 = 4$ and $d_1 = \frac{1}{4}$,
- $\mathbb{P}_2$ with $u_2 = 2$ and $d_2 = \frac{1}{2}$,

- A claim $C : \Omega = \{u_1, u_2, d_1, d_2\} \to \mathbb{R}$ with $C(u_1) = 1$, $C(u_2) = 2$, $C(d_1) = C(d_2) = 0$.

Under $\mathbb{P}_1$ we have the equivalent martingale measure $Q_1(u_1) = \frac{1}{5}, Q_1(d_1) = \frac{4}{5}$ which leads to $\mathbb{E}^{Q_1}[C] = \frac{1}{5}$.

Under $\mathbb{P}_2$ we have the equivalent martingale measure $Q_2(u_2) = \frac{1}{3}, Q_2(d_2) = \frac{2}{3}$ which leads to $\mathbb{E}^{Q_2}[C] = \frac{2}{3}$.

Finally, under any $\mathbb{P}_\lambda = \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$ the martingale measure $Q(u_1) = 0$, $Q(u_2) = \frac{3}{7}$, $Q(d_1) = \frac{4}{7}$ and $Q(d_2) = 0$ can be approximated by equivalent martingale measures leading to $\sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}^Q[C] = \frac{6}{7} > \frac{2}{3}$.

On the other hand there are cases where every equivalent martingale measure of $\mathbb{P}_\lambda$ is contained in $\mathcal{Q}_\lambda$, as the following proposition states.

Proposition 4.2.7.
If for every $\mathbb{P}_i, \mathbb{P}_j \in \mathcal{P}$ either $\mathbb{P}_i \sim \mathbb{P}_j$ or $\mathbb{P}_i \perp \mathbb{P}_j$, where $\perp$ denotes the singularity of measures, and if $supp(\mathbb{P}_i) \in \mathcal{F}_0$ for every $\mathbb{P}_i \in \mathcal{P}$, then the set $\mathcal{Q}_\lambda$ contains every equivalent martingale measure of $\mathbb{P}_\lambda$.

Proof. The above statement is a more general form of Proposition 5.13 in Kirch [38], whose proof we loosely follow.

It is clear by definition that every measure Q in $\mathcal{Q}_\lambda$ is equivalent to $\mathbb{P}_\lambda$ and is a martingale measure, so we only need to show that every martingale measure equivalent to $\mathbb{P}_\lambda$ can be written in the form

$$Q = \sum_{\substack{Q_i\in\mathcal{Q}^{\mathbb{P}_i}\\ i\in I}} \lambda_i Q_i.$$

Let Q be a martingale measure equivalent to $\mathbb{P}_\lambda$. We can say without loss of generality that all measures in $\mathcal{P}$ are singular, because two equivalent measures have the same equivalent martingale measures.

Now we can define the support of the measure $\mathbb{P}_i$ as Ω_i and define new measures

$$\tilde{Q}_i(A) := Q(A|\Omega_i) = \frac{Q(A\cap\Omega_i)}{Q(\Omega_i)}.$$

which leads to a representation of Q as

$$Q(A) = \sum_{i=1}^{n} \tilde{Q}_i(A)Q(\Omega_i), \tag{4.2.1}$$

Since $Q \sim \mathbb{P}_\lambda$ and thus $\mathbb{P}_i \ll Q$ and $Q_i(\Omega_i) = 1$, it is clear that $\tilde{Q}_i \sim \mathbb{P}_i$, so it only remains to show that $\tilde{Q}_i$ is a martingale measure on Ω_i.

First, we can represent conditional expectation under Q by conditional expectation under the new measure $\tilde{Q}$ using Proposition 2.4.6:

$$\begin{aligned}
\mathbb{E}^Q[S_t|\mathcal{F}_s] &= \frac{1}{\mathbb{E}^{\mathbb{P}_\lambda}[\frac{dQ}{d\mathbb{P}_\lambda}|\mathcal{F}_s]}\mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{dQ}{d\mathbb{P}_\lambda}S_t\middle|\mathcal{F}_s\right] \\
&= \frac{1}{\mathbb{E}^{\mathbb{P}_\lambda}[\frac{dQ}{d\mathbb{P}_\lambda}|\mathcal{F}_s]}\mathbb{E}^{\mathbb{P}_\lambda}\left[\sum_{i=1}^{\infty} Q(\Omega_i)\frac{d\tilde{Q}_i}{d\mathbb{P}_\lambda}S_t\middle|\mathcal{F}_s\right] \\
&= \frac{1}{\mathbb{E}^{\mathbb{P}_\lambda}[\frac{dQ}{d\mathbb{P}_\lambda}|\mathcal{F}_s]}\sum_{i=1}^{\infty} Q(\Omega_i)\mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{d\tilde{Q}_i}{d\mathbb{P}_\lambda}S_t\middle|\mathcal{F}_s\right] \\
&= \frac{1}{\mathbb{E}^{\mathbb{P}_\lambda}[\frac{dQ}{d\mathbb{P}_\lambda}|\mathcal{F}_s]}\sum_{i=1}^{\infty} Q(\Omega_i)\mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{d\tilde{Q}_i}{d\mathbb{P}_\lambda}\middle|\mathcal{F}_s\right]\mathbb{E}^{\tilde{Q}_i}\left[S_t|\mathcal{F}_s\right]. \qquad (4.2.2)
\end{aligned}$$

The second equation follows due to Equation (4.2.1), while the third equation holds with monotone convergence theorem. Now we can make use of the assumption that $\Omega_i \in \mathcal{F}_0$. This allows for

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{dQ}{d\mathbb{P}_\lambda}\middle|\mathcal{F}_s\right]\mathbb{1}_{\Omega_i} &= \mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{dQ}{d\mathbb{P}_\lambda}\mathbb{1}_{\Omega_i}\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}^{\mathbb{P}_\lambda}\left[Q(\Omega_i)\frac{d\tilde{Q}_i}{d\mathbb{P}_\lambda}\middle|\mathcal{F}_s\right] \\
&= Q(\Omega_i)\mathbb{E}^{\mathbb{P}_\lambda}\left[\frac{d\tilde{Q}_i}{d\mathbb{P}_\lambda}\middle|\mathcal{F}_s\right],
\end{aligned}$$

where we can use measurability of Ω_i in the first equation and Equation (4.2.1) in the second equation. Inserting this into Equation (4.2.2) shows that on Ω_i it holds that

$$\mathbb{E}^Q[S_t|\mathcal{F}_s]\mathbb{1}_{\Omega_i} = \mathbb{E}^{\tilde{Q}_i}[S_t|\mathcal{F}_s]. \qquad (4.2.3)$$

Now we can finally show for any $A \in \mathcal{F}_s$ and any $i \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{E}^{\tilde{Q}_i}[S_s\mathbb{1}_A] &= \mathbb{E}^{\tilde{Q}_i}[S_s\mathbb{1}_{A\cap\Omega_i}] = \frac{1}{Q(\Omega_i)}\mathbb{E}^Q[S_s\mathbb{1}_{A\cap\Omega_i}] \\
&= \frac{1}{Q(\Omega_i)}\mathbb{E}^Q\left[\mathbb{E}^Q[S_t|\mathcal{F}_s]\mathbb{1}_{A\cap\Omega_i}\right] = \frac{1}{Q(\Omega_i)}\mathbb{E}^Q\left[\mathbb{E}^{\tilde{Q}_i}[S_t|\mathcal{F}_s]\mathbb{1}_{A\cap\Omega_i}\right] \\
&= \mathbb{E}^{\tilde{Q}_i}\left[\mathbb{E}^{\tilde{Q}_i}[S_t|\mathcal{F}_s]\mathbb{1}_A\right],
\end{aligned}$$

where the fourth equation follows with Equation (4.2.3). Since A can be chosen arbitrary this finally shows $\mathbb{E}^{\tilde{Q}_i}[S_t|\mathcal{F}_s] = S_s$ and proves that $\tilde{Q}_i$ is indeed a martingale measure on Ω_i and $\tilde{Q}_i \in \mathcal{Q}^{\mathbb{P}_i}$. □

After introducing the necessary notation and sets, one of the main results of this section is to show that the approach using $\mathbb{P}_\lambda$ as a new single measure leads to optimal solutions.

Theorem 4.2.8.
Let $\mathcal{P}$ be a market that satisfies the assumptions 3.2.8, 3.2.11 and 3.2.12, and let $\mathcal{P}$ be a finite robust market model. For every fixed $\lambda \in [0,1]^n$ with $\sum_{i=1}^n \lambda_i = 1$ there exists a superhedging strategy for the claim $\tilde{C} = C\mathbb{1}_{A_{H,x}}$, where $A_{H,x}$ solves the following optimisation problem

$$\begin{aligned} &\text{maximise } \mathbb{P}_\lambda(A_{H,x}) \\ &\text{under the constraint } \sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_{A_{H,x}}C] \leq \tilde{V}_0. \end{aligned} \tag{4.2.4}$$

This strategy is optimal regarding $\mathcal{P}$.

Proof. Let $\tilde{A}$ be a solution to this optimisation problem. Since the market satisfies the assumptions 3.2.8, 3.2.11 and 3.2.12, the previous Lemma 4.2.5 ensures that we can apply Theorem 3.2.15, which means that we can indeed find a hedging portfolio $(\tilde{H},\tilde{x})$ with

$$\begin{aligned} &\mathbb{P}_1\left(V_T^{\tilde{H},\tilde{x}} \geq \mathbb{1}_{\tilde{A}}C\right) = 1, \\ &\mathbb{P}_2\left(V_T^{\tilde{H},\tilde{x}} \geq \mathbb{1}_{\tilde{A}}C\right) = 1, \\ &\sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_{\tilde{A}}C] = \tilde{x} \leq \tilde{V}_0. \end{aligned}$$

It remains to show that this strategy is indeed an optimal strategy in view of Definition 4.2.3. If we assume that $\tilde{A}$ does not lead to an optimal strategy, then there exists a set $A \in \mathcal{F}$ with $\sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_A C] \leq \tilde{V}_0$ and

$$\mathbb{P}_1(A) > \mathbb{P}_1(\tilde{A}) \quad \text{but} \quad \mathbb{P}_2(A) \geq \mathbb{P}_2(\tilde{A})$$

This means we can get the following inequality:

$$\begin{aligned} \mathbb{P}_\lambda(A) &= \lambda\mathbb{P}_1(A) + (1-\lambda)\mathbb{P}_2(A) \\ &> \lambda\mathbb{P}_1(\tilde{A}) + (1-\lambda)\mathbb{P}_2(\tilde{A}) \\ &= \mathbb{P}_\lambda(\tilde{A}). \end{aligned}$$

This is a contradiction to the assumption, that $\tilde{A}$ is a solution to the optimisation problem (4.2.4).

□

This theorem shows us that we can indeed find optimal solutions with this new convex combination of measures. The subsequent theorem will demonstrate that this procedure can actually lead to every optimal solution.

Theorem 4.2.9.
Let $\mathcal{P}$ be a market that fulfills the assumptions 3.2.8, 3.2.11 and 3.2.12 and let $\mathcal{P}$ be a finite set of measures. Then every optimal strategy according to Definition 4.2.3 is the solution of a optimisation problem

$$\begin{aligned} &\text{maximise } \mathbb{P}_\lambda(A_{H,x}) \\ &\text{under the constraint } \sup_{Q \in \mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_{A_{H,x}} C] \leq \tilde{V}_0, \end{aligned} \tag{4.2.5}$$

for a weight vector $\lambda \in [0,1]^n$ with $\sum_{i=1}^n \lambda_i = 1$.

In view of Theorem 4.1.3 this can be rewritten to: Every Optimal Strategy is the super-hedging strategy to an option

$$\tilde{C} = C \mathbb{1}_{\left\{\frac{d\mathbb{P}_\lambda}{dQ_\lambda} > kC\right\}},$$

where $\tilde{Q} \in \mathcal{Q}_\lambda$ is a worst case measure and k is chosen such that the constraint is fulfilled.

Proof. First, we assume that we have a set $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$. As shown in Theorem 4.1.3, it is possible to find the optimal strategy for both $\mathbb{P}_1$ and $\mathbb{P}_2$. Consequently, for $\lambda = 1$ and $\lambda = 0$ the optimal sets A_1 and A_2 can be determined as

$$\begin{aligned} A_1 &= \left\{\frac{d\mathbb{P}_1}{d\tilde{Q}_1} > k_1 \, C\right\} \\ A_2 &= \left\{\frac{d\mathbb{P}_2}{d\tilde{Q}_2} > k_2 \, C\right\}, \end{aligned}$$

where $k_1, k_2 \in \mathbb{R}_+$ are constants, that can be determined by the investment constraint, and $\tilde{Q}_1, \tilde{Q}_2 \in \mathcal{Q}_\lambda$ are worst-case measures, that exist according to Neyman-Pearson theory, see Theorem 3.3.14.

Now assume $\lambda \in (0,1)$. With the argumentation in Theorem 4.1.3 we know that the solution to the optimisation problem (4.2.5) is of the form

$$A_\lambda = \left\{\frac{d\mathbb{P}_\lambda}{d\tilde{Q}_\lambda} > k_\lambda C\right\}.$$

Since $\mathbb{P}_1$ and $\mathbb{P}_2$ are absolutely continuous with respect to Q for any $Q \in \mathcal{Q}_\lambda$, i.e. $\mathbb{P}_1 \ll Q$ and $\mathbb{P}_2 \ll Q$, we can rewrite the set A_λ as

$$A_\lambda = \left\{\lambda \frac{d\mathbb{P}_1}{d\tilde{Q}_\lambda} + (1-\lambda)\frac{d\mathbb{P}_2}{d\tilde{Q}_\lambda} > k_\lambda C\right\}.$$

To finish the proof we need to show that the functions

$$p_i : [0,1] \to [0,1],\ \lambda \mapsto \mathbb{P}_i\left(\lambda\frac{d\mathbb{P}_1}{d\tilde{Q}} + (1-\lambda)\frac{d\mathbb{P}_2}{d\tilde{Q}} > k_\lambda C\right) \tag{4.2.6}$$

are continuous. This function represents the success probability under $\mathbb{P}_i$ for the success set that is optimal under $\mathbb{P}_\lambda$. Note that p_1 always measures the success probability of optimal solutions, i.e. the measured set in (4.2.6) solves

$$\begin{aligned}&\text{maximise } \mathbb{P}_\lambda(A)\\&\text{under the constraint } \sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}[\mathbb{1}_A C] \le \tilde{V}_0.\end{aligned}$$

First of all we will show that the function

$$p_{\lambda_1}(\lambda_2) :=\mathbb{P}_{\lambda_1}\left(\lambda_2\frac{d\mathbb{P}_1}{d\tilde{Q}_{\lambda_2}} + (1-\lambda_2)\frac{d\mathbb{P}_2}{d\tilde{Q}_{\lambda_2}} > k_{\lambda_2} C\right) \tag{4.2.7}$$

is continuous in λ_1 and λ_2. Since λ_1 only gives the weights in the convex combination of $\mathbb{P}_\lambda$ the function is continuous in λ_1. This means we have the following properties:

$$\begin{aligned}&0 \le p_{\lambda_1}(\lambda_2) \le 1\\&p_{\lambda_1}(\lambda_2) \text{ is continuous in } \lambda_1\\&\text{for a fixed } \lambda_1 \text{ we have } p_{\lambda_1}(\lambda_1) = \max_{\lambda_2\in[0,1]} p_{\lambda_1}(\lambda_2),\end{aligned} \tag{4.2.8}$$

where property (4.2.8) follows due to Theorem 4.1.3. These properties are sufficient to show continuity in λ_2:

Suppose that $p_{\lambda_1}(x)$ is not continuous in x. Continuity in λ_1 together with boundedness makes clear that if there are any discontinuities those cannot be limited to a single point, furthermore this set of discontinuities must be parallel to the λ_1 axis, due to continuity in this variable, which means there must be a tupel $(\lambda_0, \lambda_0) \in (0,1)\times(0,1)$ where it holds

$$p_{\lambda_0}(\lambda_0+\epsilon) \nrightarrow p_{\lambda_0}(\lambda_0).$$

This means we get the following chain of inequalities:

$$p_{\lambda_0}(\lambda_0) > p_{\lambda_0}(\lambda_0+\epsilon) = p_{\lambda_0+\epsilon}(\lambda_0+\epsilon) + \delta(\epsilon) \ge p_{\lambda_0+\epsilon}(\lambda_0) + \delta(\epsilon) = p_{\lambda_0}(\lambda_0) + \delta(\epsilon) + \tilde{\delta}(\epsilon), \tag{4.2.9}$$

for all $\epsilon > 0$ small enough, where $\delta(\epsilon), \tilde{\delta}(\epsilon) \to 0$ for $\epsilon \to 0$. The inequalities are following from the maximum property (4.2.8) and the δ and $\tilde{\delta}$ follow from continuity in λ_1. Considering the limit for $\epsilon \to 0$ we get

$$p_{\lambda_0}(\lambda_0) > \lim_{\epsilon\to 0} p_{\lambda_0}(\lambda_0+\epsilon) = \lim_{\epsilon\to 0} p_{\lambda_0+\epsilon}(\lambda_0+\epsilon) \ge p_{\lambda_0+\epsilon}(\lambda_0) = p_{\lambda_0}(\lambda_0),$$

which is a contradiction and verifies that $p_{\lambda_1}(\lambda_2)$ has to be continuous in both variables.

In the last step it remains to show that not only $p_\lambda = \lambda p_1 + (1-\lambda)p_2$ is continuous but also p_1 and p_2. So let us assume again that p_1 is not continuous at position λ_0, i.e. $\lim_{\epsilon\to 0} p_1(\lambda_0+\epsilon) \neq p_1(\lambda_0)$. As we know that $p_{\lambda_0} := \lambda_0 p_1 + (1-\lambda_0)p_2$ is a continuous function, we conclude that p_2 is also discontinuous with

$$\begin{aligned}\lim_{\epsilon\to 0} p_1(\lambda_0+\epsilon) - p_1(\lambda_0) =& \lim_{\epsilon\to 0}\left(\frac{1}{\lambda_0}p_{\lambda_0}(\lambda_0+\epsilon) - \frac{\lambda_0}{1-\lambda_0}p_2(\lambda_0+\epsilon)\right)\\ &- \left(\frac{1}{\lambda_0}p_{\lambda_0}(\lambda_0) - \frac{\lambda_0}{1-\lambda_0}p_2(\lambda_0)\right)\\ =& -\frac{\lambda_0}{1-\lambda_0}\left(\lim_{\epsilon\to 0} p_2(\lambda_0+\epsilon) - p_2(\lambda_0)\right).\end{aligned}$$

The second equality holds with continuity of $p_{\lambda_1}(\lambda_2)$ that we already showed.

On the other hand we also know that p_λ is continuous for every fixed λ in the subscript which leads to

$$\lim_{\epsilon\to 0} p_1(\lambda_0+\epsilon) - p_1(\lambda_0) = -\frac{\lambda}{1-\lambda}\left(\lim_{\epsilon\to 0} p_2(\lambda_0+\epsilon) - p_2(\lambda_0)\right)$$

for every $\lambda \in (0,1)$. The above equation can only hold for every λ if both sides are 0. Thus we finally conclude that p_1 and p_2 have to be continuous in λ_0, which is a contradiction.

This means we have continuous functions p_1 and p_2 and see that for every success probability $\tilde{p} \in [p_1(0), p_1(1)]$, there has to be a weight $\lambda \in [0,1]$ such that $p_1(\lambda) = \tilde{p}$. The same holds for p_2, which means we can indeed create every optimal success set and find every optimal success probability. The previous arguments can also be continued recursively to any finite set $\mathcal{P} = \{\mathbb{P}_1, ..., \mathbb{P}_n\}$. □

Corollary 4.2.10.

For a fixed index i and a capital constraint $\tilde{V}_0$ let $\overline{p}$ be the maximal success probability an optimal strategy can have under $\mathbb{P}_i$ and let $\underline{p}$ be the minimal success probability an optimal strategy can have under $\mathbb{P}_i$. Theorem 4.2.9 implies that for every $p \in [\underline{p}, \overline{p}]$ there exists a weight vector λ such that the optimisation problem

$$\begin{aligned}&\text{maximise } \mathbb{P}_\lambda(A_{H,x})\\ &\text{under the constraint } \sup_{Q\in\mathcal{Q}_\lambda} \mathbb{E}^Q[\mathbb{1}_{A_{H,x}}C] \leq \tilde{V}_0,\end{aligned}$$

leads to a strategy (H,x) with $\mathbb{P}_i(A_{H,x}) = p$.

Remark 4.2.11.

Theorem 4.2.9 implies the existence of a solution to the minimax problem we were solving in the case of convex sets of physical measures.

4.2.2. Extension to additional risk measures and concavity

In Section 4.1.2.1 we showed that in single market models all results can also be applied in a very similar way to the expected shortfall or the success ratio as a risk measure. This leads to the question of whether the results for robust markets can also be generalised to these risk measures, which will turn out to be correct.

Definition 4.2.12. *Robust optimal hedging strategy regarding success ratio*
A hedging strategy $(\tilde{H}, \tilde{x})$ with initial investment $\tilde{x} \leq \tilde{V}_0$ is optimal with respect to the success ratio, defined in Definition 4.1.5, if for every other hedging strategy (H, x) with $x \leq \tilde{V}_0$ and

$$\mathbb{E}^{\mathbb{P}_i}[\varphi(\tilde{H}, \tilde{x})] < \mathbb{E}^{\mathbb{P}_i}[\varphi(H, x)]$$

there exists $j \neq i$ with

$$\mathbb{E}^{\mathbb{P}_j}[\varphi(\tilde{H}, \tilde{x})] > \mathbb{E}^{\mathbb{P}_j}[\varphi(H, x)]$$

Definition 4.2.13. *Robust optimal hedging strategies regarding expected shortfall*
Corresponding to Definition 4.2.3 a hedging strategy $\tilde{H}$ with initial investment $\tilde{x} \leq \tilde{V}_0$ to a claim C is called robust optimal with respect to the expected shortfall if for every other admissible hedging strategy (H, x) with $x \leq \tilde{V}_0$ and

$$\mathbb{E}^{\mathbb{P}_i}[(C - V_T^{\tilde{H},x})^+] > \mathbb{E}^{\mathbb{P}_i}[(C - V_T^{H,x})^+]$$

there exists a $j \neq i$ with

$$\mathbb{E}^{\mathbb{P}_j}[(C - V_T^{\tilde{H},x})^+] < \mathbb{E}^{\mathbb{P}_j}[(C - V_T^{H,x})^+].$$

Theorem 4.2.14.
Theorem 4.2.8 and 4.2.9 are also valid under optimality regarding success ratio and expected shortfall.

Proof. The proof of Theorem 4.2.8 can be applied analogously.

The proof of Theorem 4.2.9 can be applied analogously again. Note that we used the continuity of $\mathbb{P}$ which also holds for $\mathbb{E}^{\mathbb{P}}$. This means that the arguments used there also apply to this definition of optimality. □

Using success ratios instead of success probabilities again has advantages when examining the remaining risk as a function of initial investment. Considering optimal strategies that maximise the minimum probability of success (Minimax optimisation) or maximise the sum of all success probabilities can again be proven to be concave as function of initial investment.

Corollary 4.2.15.
The optimal expected success ratio that maximises the minimal success ratio or the sum of success ratios depending on capital constraint is non decreasing and concave.

Proof. Let $x_1 > x_2$ and let H_1 and H_2 be the optimal strategies that maximise $\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(H,x)]$ depending on x. As in Corollary 4.1.8 one can again set up a hedging strategy that has a success ratio equal to a convex combination of both previous optimal success ratios for less or equal capital

$$\begin{aligned}\sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{Q}[(\lambda\varphi(H_1,x_1)+(1-\lambda)\varphi(h_2,x_2))C] \leq &\lambda \sup_{Q\in\mathcal{Q}^{\mathbb{P}}} \mathbb{E}^{Q}[\varphi(H_1,x_1)C] \\ &+(1-\lambda)\sup_{Q\in\mathcal{Q}^{\mathbb{P}}} E^{Q}[\varphi(H_2,x_2)C] \\ &=\lambda x_1+(1-\lambda)x_2.\end{aligned}$$

On the other hand the linearity of the summed success ratio

$$\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\lambda\varphi(H_1,x_1)+(1-\lambda)\varphi(H_2,x_2)] = \lambda\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(H_1,x_1)]+(1-\lambda)\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(H_2,x_2)]$$

is sufficient to proof the existence of an optimal hedging strategy $\tilde{H}$ with

$$\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(\tilde{H},\lambda x_1+(1-\lambda)x_2)] \geq \lambda\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(H_1,x_1)]+(1-\lambda)\sum_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[\varphi(H_2,x_2)].$$

While for the maximisation of the minimal success ratio we have

$$\begin{aligned}&\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\lambda\varphi(H_1,x_1)+(1-\lambda)\varphi(H_2,x_2)]\right)\\ &\quad\geq\lambda\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\varphi(H_1,x_1)]\right)+(1-\lambda)\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\varphi(H_2,x_2)]\right),\end{aligned}$$

which again proofs existence of $\tilde{H}$ with

$$\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\varphi(\tilde{H},\lambda x_1+(1-\lambda)x_2)]\right) \geq \lambda\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\varphi(H_1,x_1)]\right)+(1-\lambda)\min_{\mathbb{P}\in\mathcal{P}}\left(\mathbb{E}^{\mathbb{P}}[\varphi(H_2,x_2)]\right).$$

□

Theorem 4.2.16.
Let $\overline{p_i} \neq \underline{p_i}$. The set of all success probabilities of hedging strategies that satisfy the constraints is strictly convex (see for example Figure 5.5).

If we assume $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ and define $\overline{p_i}$ and $\underline{p_i}$ to largest and smallest possible success probability an optimal hedging strategy can have under $\mathbb{P}_i$, then the above theorem means that the optimal solutions we can find according to Theorem 4.2.9 form a strictly concave curve between $(\overline{p_1}, \underline{p_2})$ and $(\underline{p_1}, \overline{p_2})$, which means that these solutions have larger success probabilities than the straight line between $(\overline{p_1}, \underline{p_2})$ and $(\underline{p_1}, \overline{p_2})$.

Proof. We will again use the hedging ratio to prove this statement, but we can use the fact that the shortfall probability and hedging ratio coincide whenever the optimal solution for the shortfall probability exists. Fix $\lambda_1, \lambda_2 \in [0,1]$ with $\lambda_1 \neq \lambda_2$ and let $(H_1, \tilde{V}_0)$ be the optimal hedging strategy regarding $\mathbb{P}_{\lambda_1}$ and $(H_2, \tilde{V}_0)$ be the optimal hedging strategy regarding $\mathbb{P}_{\lambda_2}$. For any convex combination of these hedging strategies we have

$$\begin{aligned}\mathbb{E}^{\mathbb{P}_1}[\varphi(\lambda H_1 + (1-\lambda)H_2, \tilde{V}_0)] &= \mathbb{E}^{\mathbb{P}_1}[\lambda\varphi(H_1, \tilde{V}_0) + (1-\lambda)\varphi(H_2, \tilde{V}_0)] \\ &= \lambda\mathbb{E}^{\mathbb{P}_1}[\varphi(H_1, \tilde{V}_0)] + (1-\lambda)\mathbb{E}^{\mathbb{P}_1}[\varphi(H_2, \tilde{V}_0)].\end{aligned}$$

This proves that the optimal success ratio under $\mathbb{P}_1$ to $\lambda \in (\lambda_1, \lambda_2)$ is at least linear between the two optimal solutions generated by λ_1 and λ_2. If we assume $\overline{p_i} \neq \underline{p_i}$ the set $\{\varphi(H_1, \tilde{V}_0) \neq \varphi(H_2, \tilde{V}_0)\}$ is not a null set and thus $\varphi(\lambda H_1 + (1-\lambda)H_2, \tilde{V}_0)$ is not of 0-1-form. Therefore, there exists a strictly better hedging strategy than the convex combination $\lambda H_1 + (1-\lambda)H_2$. The same arguments hold for $\mathbb{P}_2$ as well. Since the optimal strategies for the shortfall probability and the hedging ratio coincide and since we have already proved the existence of the optimal strategies regarding the shortfall probability this suffices to prove the statement. Since λ_1 and λ_2 were chosen arbitrarily the indifferent curve of all optimal hedging strategies is strictly concave. If $\overline{p_i} = \underline{p_i}$, then the set of all optimal strategies is a singleton. □

The above Theorem 4.2.16 gives an idea of why the theory of robust superhedging can find useful applications. Since it is impossible, or at least very difficult, to determine the model selection risk, it is quite difficult to find an appropriate weight for each model in the definition of $\mathbb{P}_\lambda$. Theorem 4.2.16 shows that if there is no preference for any of the models considered in $\mathcal{P}$, then looking at more than one model can reward an investor with a higher overall probability of success while reducing model selection risk. For example, one can aim to maximise the sum of all success probabilities, which, as shown in the previous corollary, is higher than for optimal strategies in single market models. An example of this consideration is shown in Figure 5.6 in the next section, where calculations have been made for two Black-Scholes models with different drift terms. As another example, one can choose a minimum probability of success for certain models depending on outside considerations, while maximising it for other models. One can think of modelling high and low interest rates, which may require different levels of collateral or one can think of scenario-based approaches in general, which we often see in practice and as regulatory requirements. As a third example, one can maximise the minimum success probability across all models, this approach is explored in Kirch [38] in a more general setting.

4.2.3. Worst-case measures $\tilde{Q}$ for different weightings

The worst-case measure for an incomplete market model $\mathbb{P}$ is usually very difficult or almost impossible to find analytically. This circumstance motivates a grid search over all possible worst-case measures in $\mathcal{Q}^{\mathbb{P}}$, which is very expensive in terms of computational time. In the case of robust market modelling, a new measure $\mathbb{P}_\lambda$ is constructed from the measures in $\mathcal{P}$, which is a convex combination of all measures in $\mathcal{P}$ with a weight vector λ. Finding the worst case measure for each possible weights λ via a new grid search each time is therefore undesirable from a computational point of view.

The question therefore is whether the worst-case measures might have some continuity properties with respect to the weights λ. Intuitively, a small change in λ should result in only a small change in the corresponding worst-case measure, which will be mathematically proven in the following chapter. This means that changing the weights does not require a completely new global grid search, but only a local search around the previous worst-case measure.

Definition 4.2.17.
A sequence of probability measures (μ_n) converges in total variation to a probability measure μ if for every ϵ there exists a n_0 such that for every $n \geq n_0$ it holds that

$$||\mu_n - \mu||_{TV} < \epsilon,$$

where the total variation norm for probability measures is defined as

$$||\mu_n - \mu||_{TV} := \sup_{A \in \mathcal{F}} |\mu_n(A) - \mu(A)|.$$

Remark 4.2.18.
The total variation norm is generally defined for measures μ, ν on a measurable space

$(X, \mathcal{F})$ as

$$||\mu - \nu||_{TV} = \sup_f \int_X f d\mu - \int_X f d\nu,$$

where f is a measurable function from X to $[-1, 1]$. This definition simplifies to the definition in 4.2.17 in cases of probability measures.

Definition 4.2.19.

A sequence of random variables (X_n) converges in measure regarding μ to a random variable X if

$$\lim_{n\to\infty} \mu(\{w \in \Omega : |X_n - X| > \epsilon\}) = 0 \quad \forall \epsilon > 0.$$

Before proving that worst-case measures behave continuous under changes in the weighting term we need two auxiliary results. The following lemma is a special case of Scheffe's theorem, see Scheffe [61]. For the sake of completeness, a proof is given.

Lemma 4.2.20.

Let $(\Omega, \mathcal{F})$ be a measurable space with (μ_n) a sequence of probability measures absolutely continuous with respect to a probability measure μ. The convergence $\frac{d\mu_n}{d\mu} \to 1$ in measure is equivalent to the convergence $\mu_n \to \mu$ in total variation.

Proof. Let $C \in \mathcal{F}$ be an arbitrary event. For this take the sets

$$A_n := \Big\{\omega \in C \Big| \frac{d\mu_n}{d\mu}(\omega) \leq 1\Big\}$$

and

$$B_n := \Big\{\omega \in C \Big| \frac{d\mu_n}{d\mu}(\omega) > 1\Big\}.$$

First, it holds that

$$\mu(C) = \int_C d\mu_n = \int_C \frac{d\mu_n}{d\mu} d\mu = \int_{A_n} \frac{d\mu_n}{d\mu} d\mu + \int_{B_n} \frac{d\mu_n}{d\mu} d\mu.$$

This in turn gives

$$\begin{aligned}\int_{B_n} \frac{d\mu_n}{d\mu} d\mu &= \int_C d\mu - \int_{A_n} \frac{d\mu_n}{d\mu} d\mu = \int_{B_n} d\mu + \int_{A_n} (1 - \frac{d\mu_n}{d\mu}) d\mu \\ &= \mu(B_n) + \int_{A_n} (1 - \frac{d\mu_n}{d\mu}) d\mu.\end{aligned}$$

For any arbitrary set $C \in \mathcal{F}$, there is

$$\begin{aligned}||\mu_n - \mu||_{TV} =& |\int\limits_C d(\mu_n - \mu)| = |\int\limits_C (\frac{d\mu_n}{d\mu} - 1)d\mu| = \int\limits_{A_n} (1 - \frac{d\mu_n}{d\mu})d\mu + \int\limits_{B_n} (\frac{d\mu_n}{d\mu} - 1)d\mu \\ =& \int\limits_{A_n} (1 - \frac{d\mu_n}{d\mu})d\mu + \int\limits_{B_n} \frac{d\mu_n}{d\mu}d\mu - \mu(B_n) \\ =& \int\limits_{A_n} (1 - \frac{d\mu_n}{d\mu})d\mu + \mu(B_n) + \int\limits_{A_n} (1 - \frac{d\mu_n}{d\mu})d\mu - \mu(B_n) \\ =& 2\int\limits_{A_n} (1 - \frac{d\mu_n}{d\mu})d\mu \to 0.\end{aligned}$$

Here, convergence in the last step follows by the dominated convergence theorem, since $(1 - \frac{d\mu_n}{d\mu})$ is bounded by 1 and is positive. The convergence in total variation thus follows from the convergence $\frac{d\mu_n}{d\mu} \to 1$ in measure with respect to μ.

The backward direction, on the other hand, follows by Markov inequality

$$\mu(\{\omega \in \Omega : |\frac{d\mu_n}{d\mu} - 1| > \epsilon\}) \leq \frac{\mathbb{E}^\mu[|\frac{d\mu_n}{d\mu} - 1|]}{\epsilon} = \frac{1}{\epsilon} \int\limits_\Omega |\frac{d\mu_n}{d\mu} - 1| d\mu \to 0 \quad \forall \epsilon > 0.$$

Note here that the condition of convergence implies $\frac{d\mu_n}{d\mu} \to 1$ in $L^1(\mu)$. □

Lemma 4.2.21.
Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables with $X_n \geq 0$ and $\mathbb{E}[X_n] = 1$ for all $n \in \mathbb{N}$. Let $k \in \mathbb{R}$ be constant. It holds that

$$\mathbb{P}(\lim_{n\to\infty} Y + \frac{1}{n}X_n > k) = \mathbb{P}(Y > k).$$

Proof. Assuming that the assertion does not hold, then $\mathbb{P}(\{\lim_{n\to\infty} \frac{1}{n}X_n > 0\}) \neq 0$ or $\mathbb{P}(X_n > nk) \nrightarrow 0$ holds. Thus

$$1 = \mathbb{E}[X_n] \geq \mathbb{E}[X_n \mathbb{1}_{\{X_n > n\}}] \geq n\mathbb{P}(X_n > n) \to \infty.$$

□

These two results will be used to proof the main theorem of this section:

Theorem 4.2.22.
Let $\mathcal{P} = \{\mathbb{P}_1, ..., \mathbb{P}_n\}$ be a robust market model and λ a weight vector. Let $f : (0,1)^n \to \mathcal{Q}$ be a function with $f(\lambda) = Q_\lambda$, where Q_λ is a worst-case measure regarding $\mathbb{P}_\lambda$. The functional f is continuous with respect to the total variation norm.

Proof. It is to be shown that for any sequence $(\lambda_n) \to \lambda^*$ there is also $\lim_{n\to\infty} f(\lambda_n) = f(\lambda^*)$. Let without restriction $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$. For a sequence $\lambda + \frac{1}{n}$ the following holds:

$f(\lambda)$ solves the following optimisation problem in Q:

$$\max_{Q\in\mathcal{Q}} \mathbb{P}_\lambda \left(\frac{d\mathbb{P}_\lambda}{dQ} > k\ C \right) \tag{4.2.10}$$
$$\text{under the constraint } \sup_{Q\in\mathcal{Q}} E^Q[\mathbb{1}_{\left\{\frac{d\mathbb{P}_\lambda}{dQ} > kC\right\}} C] \leq \tilde{V}_0.$$

Additionally, for all $n \in \mathbb{N}$, $f(\lambda + \frac{1}{n})$ solves:

$$\max_{Q\in\mathcal{Q}} \mathbb{P}_{\lambda+\frac{1}{n}} \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{dQ} > kC \right)$$
$$\text{under the constraint } \sup_{Q\in\mathcal{Q}} E^Q[\mathbb{1}_{\left\{\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{dQ} > k\ C\right\}} C] \leq \tilde{V}_0.$$

When considering the limits for $n \to \infty$, it can be used that for any sequence (A_n):

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{P}(A_n) = 0, \tag{4.2.11}$$

since the measure $\mathbb{P}$ is bounded by 1.

Furthermore, for an increasing or decreasing sequence of sets $(A_1 \subseteq A_2 \subseteq A_3 ...)$ or $(A_1 \supseteq A_2 \supseteq A_3 ...)$, the monotone convergence theorem states

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\lim_{n\to\infty} A_n). \tag{4.2.12}$$

For the optimal success set with respect to $\lambda + \frac{1}{n}$, by definition.

$$\left\{ \frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right\}$$
$$= \left\{ \lambda \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda) \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} + \frac{1}{n} \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} - \frac{1}{n} \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right\}.$$

Note that the above sets would be monotonically increasing (decreasing) if the two posterior terms are both added (subtracted), since the Radon-Nikodym derivative is always non-negative.

Looking at the limit for $n \to \infty$ it can be seen that

$$\lim_{n\to\infty} \mathbb{P}_{\lambda+\frac{1}{n}} \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right) = \lim_{n\to\infty} \mathbb{P}_\lambda \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right)$$
$$\geq \lim_{n\to\infty} \mathbb{P}_\lambda \left(\lambda \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda) \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} - \frac{1}{n} \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} - \frac{1}{n} \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right)$$

$$
\begin{aligned}
&= \mathbb{P}_\lambda \left(\lim_{n\to\infty} \lambda \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda)\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} - \frac{1}{n}\frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} - \frac{1}{n}\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&= \mathbb{P}_\lambda \left(\lambda \lim_{n\to\infty} \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda) \lim_{n\to\infty} \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&= \mathbb{P}_\lambda \left(\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} > k\ C \right),
\end{aligned}
$$

while at the same time

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{P}_{\lambda+\frac{1}{n}} \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right) &= \lim_{n\to\infty} \mathbb{P}_\lambda \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&\le \lim_{n\to\infty} \mathbb{P}_\lambda \left(\lambda \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda)\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} + \frac{1}{n}\frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + \frac{1}{n}\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&= \mathbb{P}_\lambda \left(\lim_{n\to\infty} \lambda \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda)\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} + \frac{1}{n}\frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + \frac{1}{n}\frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&= \mathbb{P}_\lambda \left(\lambda \lim_{n\to\infty} \frac{d\mathbb{P}_1}{df(\lambda+\frac{1}{n})} + (1-\lambda) \lim_{n\to\infty} \frac{d\mathbb{P}_2}{df(\lambda+\frac{1}{n})} > k\ C \right) \\
&= \mathbb{P}_\lambda \left(\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} \right).
\end{aligned}
$$

The second last equation in each case holds with Lemma 4.2.21. Combining these two inequalities we get

$$
\mathbb{P}_\lambda \left(\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} \right) \le \lim_{n\to\infty} \mathbb{P}_{\lambda+\frac{1}{n}} \left(\frac{d\mathbb{P}_{\lambda+\frac{1}{n}}}{df(\lambda+\frac{1}{n})} > k\ C \right) \le \mathbb{P}_\lambda \left(\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} \right).
$$

That is, the set $\{\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} > k\ C\}$ also solves the first optimisation problem (4.2.10) with λ instead of $\lambda + \frac{1}{n}$. The uniqueness of the solution then provides that both sets must be equal except for null sets, i.e.

$$
\left\{ \lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} > k\ C \right\} = \left\{ \frac{d\mathbb{P}_\lambda}{df(\lambda)} > k\ C \right\} \quad \mathbb{P} - a.s.
$$

As f only generates measures equivalent to $\mathbb{P}_\lambda$, continuity of f now follows by

$$
\begin{aligned}
\lim_{n\to\infty} \frac{d\mathbb{P}_\lambda}{df(\lambda+\frac{1}{n})} &= \frac{d\mathbb{P}_\lambda}{df(\lambda)} \quad \mathbb{P}_\lambda\text{-a.s.} \\
\Rightarrow \lim_{n\to\infty} \frac{df(\lambda+\frac{1}{n})}{df(\lambda)} &= 1 \quad \mathbb{P}_\lambda\text{-a.s.}
\end{aligned}
$$

This is sufficient with Lemma 4.2.20 to proof convergence in the total variation norm. Iteratively we can generalise this theorem to finite sets $\mathcal{P}$. □

As already mentioned above, this continuity property of worst-case measures can significantly reduce the computational effort by allowing a local grid search instead of a global one. By exploiting this property, we can focus on relevant regions of the parameter space, improving efficiency. The numerical examples in the following chapter illustrate the computational effort rquierd for a grid search.

Chapter 5.

Application and examples

The expression $C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ}>kC\}}$ is very unintuitive and not directly applicable for real-world problems. Therefore, it is interesting to have a more detailed look at this expression and its application in specific models. Beginning with single market models we will present an analytic solution in Black-Scholes models to call and put options as well as binary options. The example of call and put options was already examined in Föllmer [23] but is reintroduced here to extend it to robust market models. To get an analytic solution, we need many properties of the Black-Scholes model that usually do not hold for more complex market models, which leads to the question of a more general applicable way to treat these problems. Therefore, we will show that one can use Monte Carlo simulations, that are generally applicable without strong assumptions, to find optimal success sets and probabilities.

First, we will present a short introduction into Monte-Carlo methods and numerical solutions of stochastic differential equations. This only serves as a justification that the following numerical approaches are mathematically correct. After that we will first present solutions in the single market, i.e. analytic solutions to call and binary options in the Black-Scholes model and numerical approaches in jump models. Then we will generalise these examples to robust models.

Note that all numerical calculations were made with the software R.

5.1. Monte-Carlo methods and approximation of SDE-solutions

One of the main problems we are facing when trying to find optimal hedging strategies for a given market model is to calculate expectations, $\mathbb{E}[f(S_T)]$, either under a physical measure to determine the risk or under an equivalent martingale measure to find arbitrage free prices. For most commonly used market models, i.e. jump-diffusion models, we have information about $(S_t)_{t\in\mathbb{R}_+}$ in the form of

$$dS_t = \mu(S_t,t)dt + \sigma(S_t,t)dW_t + c(S_{t-},t)dJ_t, \tag{5.1.1}$$

where W represents a Brownian motion and $J_t = \sum_{i=1}^{N_t} X_i$ represents the jump part with independent and identically distributed random variables X_i and a Poisson process N.

The problems we are facing can be reduced to finding a solution to the stochastic differential equation and then calculating expectations depending on the simulated paths of S.

If we know an explicit solution to the underlying differential equation, like in the Black-Scholes model or the Merton-Jump model, simulation of paths of S and especially of marginal distributions S_t is straightforward. However, we are often not able to determine the distribution of S_T explicitly if it is only described by a stochastic differential equation. In these cases one can try to approximate S_T using a time discrete approximation scheme. In this section we are describing a simple method to approximate distributions for S_T in a general set-up with corresponding convergence order.

First, most of the time one can guarantee the existence of a solution to the stochastic differential equation by restricting to lipschitz functionals for μ, σ and c.

Lemma 5.1.1.
A stochastic differential equation of form (5.1.1) has a unique and strong solution if μ, σ and c are lipschitz. Uniqueness is to be understood in the sense that two solutions of the stochastic differential equation are indistinguishable.

Proof. Chapter V, Theorem 7 in Protter [56]. □

As long as there exists a solution we can use a simple stochastic version of the Euler method.

Definition 5.1.2. *Euler-Maruyama approximation*
One can approximate the solution of a stochastic differential equation, for example (5.1.1), using an Euler scheme to simulate realisations according to the stochastic differential equation.

$$S_{t+\Delta t} = S_t + \mu(S_t,t)\Delta t + \sigma(S_t,t)\sqrt{\Delta t}\cdot W_t + c(S_t,t)\Delta J_t,$$

where W_t is standard normally distributed and $\Delta J_t = \sum_{i=N_{t-\Delta t}}^{N_t} X_i$ a compound Poisson process. This procedure can be used recursively on an equidistant grid $[0, T]$ using $\frac{T}{\Delta t}$ steps. This method is named Euler-Maruyama approximation.

Definition 5.1.3. *Weak and strong convergence*
Let S be the solution to a stochastic differential equation and let $\hat{S}^{\Delta t}$ be a time discrete approximation with step size Δt.

We define that $\hat{S}^{\Delta t}$ converges to S weakly with respect to a class of functions $\mathfrak{F}$ if

$$\lim_{\Delta t \to 0} |\mathbb{E}[f(S_t)] - \mathbb{E}[f(\hat{S}_t^{\Delta t})]| = 0 \text{ for any function } f \in \mathfrak{F}.$$

We say that $\hat{S}^{\Delta t}$ converges to S strongly if

$$\lim_{\Delta t \to 0} \mathbb{E}[|S_t - \hat{S}_t^{\Delta t}|] = 0.$$

In general, the more regular the functions in $\mathfrak{F}$ are, the better is the weak convergence rate. Using polynomials for $\mathfrak{F}$, which is still a comparably weak choice, is sufficient for convergence of all moments. This is also enough for the purpose of this work. As we are using the approximation of stochastic processes only for pricing, i.e. for determining expectations, weak convergence is sufficient, which is why we are only introducing this kind of convergence.

Definition 5.1.4. *Convergence rate*
Let $\hat{S}^{\Delta t}$ be a time discrete approximation of S. The approximation $\hat{S}^{\Delta t}$ converges with convergence rate q if there is a constant $C \in \mathbb{R}$ such that

$$|\mathbb{E}[f(S_t)] - \mathbb{E}[f(\hat{S}_t^{\Delta t})]| \leq C(\Delta t)^q.$$

Theorem 5.1.5.
If $\mu(x, t)$ and $\sigma(x, t)$ in (5.1.1) are four times continuously differentiable functions with polynomial growth and uniformly bounded derivatives and if the approximation is weakly consistent, then the Euler-Maruyama approximation converges weakly with order $\frac{1}{2}$.

Proof. Theorem 9.7.4 in Glasserman [26]. □

After finding a numerical approximation to the solution of stochastic differential equations it remains to determine expectations. The concept of Monte-Carlo simulation is well known. Roughly speaking, one can approximate expectations of random variables by simulating many independent realisations of this random variable and using the law of large numbers to get

$$\mathbb{E}[X] \approx \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Theorem 5.1.6.
Monte-Carlo simulation with n realisations converges with a rate of $\sqrt{n}$. More precisely if σ^2 denotes the variance of $X \in L^2(\Omega)$ then for the error $\epsilon_n = \mathbb{E}[X] - \frac{1}{n}\sum_{i=1}^{n} X_i$ it holds that $\sqrt{n}\epsilon_n$ is normally distributed with variance σ^2.

Proof. This is a direct consequence of the central limit theorem. □

These methods, both the monte carlo simulation and the Euler method aswell, are quite inefficient as there is only a convergence rate of $\frac{1}{2}$. There are way more efficient methods as the Milstein method with convergence rate 1, see for example Kloeden [39], that requires stronger assumptions. There is also a huge variety of more efficient variants of Monte-Carlo methods but as the focus in this work is not on numerical aspects these methods are sufficient to present applications of the previous results.

5.2. Single market models

5.2.1. Analytic solutions in the Black-Scholes model

First, we present some examples in the Black-Scholes model that can be solved analytically. The first example is a European call option followed by a binary option. The main feature of Black-Scholes models that allows us to find analytic solutions is the fact that the Radon-Nikodym derivative can be written solely in terms of the underlying itself, see Proposition 5.2.1.

5.2.1.1. Call options

The following idea of rewriting success sets was first formulated in Föllmer and Leukert [23]. Let $\mathbb{P}$ define a Black-Scholes market, i.e. the underlying S solves the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{5.2.1}$$

with a standard Brownian motion W, $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{>0}$. In this section we optimally hedge a vanilla call option with strike $K \in \mathbb{R}$:

$$C(S_T) = (S_T - K)^+$$

under capital constraint $\tilde{V}_0$.

First, we will present some well-known properties that hold in the Black-Scholes model:

Proposition 5.2.1.

(a) The solution of the stochastic differential equation (5.2.1) is given by

$$S_t = S_0 \exp\left(\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right). \tag{5.2.2}$$

(b) The Black-Scholes model is a complete market with a unique equivalent martingale measure. The change of measure to the equivalent martingale measure can be described by

$$\left.\frac{d\mathbb{P}}{dQ}\right|_{\mathcal{F}_t} = \exp\left(\frac{\mu - r}{\sigma} W_t + \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t\right). \tag{5.2.3}$$

Note that throughout this thesis it is assumed that $r = 0$.

(c) The Radon Nikodym derivative $\frac{d\mathbb{P}}{dQ}|_{\mathcal{F}_t}$ can be expressed in terms of S_t:

$$\left.\frac{d\mathbb{P}}{dQ}\right|_{\mathcal{F}_t} = c S_t^{\frac{\mu}{\sigma^2}} \tag{5.2.4}$$

with a constant $c \in \mathbb{R}$.

(d) The price of call options with strike K and maturity T in the Black-Scholes model can be calculated using the Black-Scholes pricing formula:

$$\mathbb{E}^Q[(S_T - K)^+)] = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2), \tag{5.2.5}$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$
$$d_2 = d_1 + \sigma\sqrt{T},$$

and Φ denotes the distribution function of a standard normal distribution.

Proof. Parts (a), (b) and (d) are standard results from financial mathematics, see Black and Scholes [8] for the pricing formula and Girsanov theory in Björk [7] for the Radon-Nokodym derivative. Part (c) follows with a straightforward calculation:

$$\begin{aligned}
\frac{d\mathbb{P}}{dQ} &= \exp\left(\frac{\mu}{\sigma}W_T + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 T\right)\\
&= \exp\left(\frac{\mu}{\sigma^2}(\sigma W_T + \frac{\mu}{2}T)\right)\\
&= S_0^{\frac{\mu}{\sigma^2}} \exp\left(\sigma W_T + (\mu - \frac{1}{2}\sigma^2)T\right)^{\frac{\mu}{\sigma^2}} S_0^{-\frac{\mu}{\sigma^2}} \exp\left(\frac{\mu^2}{2\sigma^2}T - \frac{\mu^2}{\sigma^2}T + \mu T\right)\\
&= cS_T^{\frac{\mu}{\sigma^2}}.
\end{aligned}$$

□

Corollary 5.2.2.
The optimal success set in Black-Scholes markets for call options is of the form

$$A = \{S_T < c_1\} \cup \{S_T > c_2\},$$

with $c_1 < c_2$ and $c_2 = \infty$ if $\mu < \sigma^2$. The optimal success probability is

$$\mathbb{P}(A) = \Phi\left(\frac{\ln\left(\frac{c_1}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu\right)T}{\sigma\sqrt{T}}\right) \qquad \text{if } \mu \leq \sigma^2,$$

$$\mathbb{P}(A) = \Phi\left(\frac{\ln\left(\frac{c_1}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu\right)T}{\sigma\sqrt{T}}\right) + \Phi\left(-\frac{\ln\left(\frac{c_2}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu\right)T}{\sigma\sqrt{T}}\right) \qquad \text{if } \mu > \sigma^2.$$

Proof. The main ideas of the following proof are already shown in Föllmer [23]. According to Theorem 4.1.3 the optimal success set is always of the form

$$A = \left\{\frac{d\mathbb{P}}{dQ} > \tilde{k}C\right\}.$$

Using (5.2.4) the optimal success set is of the form

$$A = \left\{ S_T^{\frac{\mu}{\sigma^2}} > \tilde{k}(S_T - K)^+ \right\}$$

with a constant $\tilde{k} \in \mathbb{R}$ determined by the capital constraint. Now one can see that the left term $S_T^{\frac{\mu}{\sigma^2}}$ is monotone in S_T. More importantly, this term is strictly convex for $\frac{\mu}{\sigma^2} > 1$ and concave for $\frac{\mu}{\sigma^2} \leq 1$. Since the right term $(S_T - K)^+$ is linear for $S_T > K$ and zero otherwise, this means that, with a proper choice of $\tilde{k}$, the equation $S_T^{\frac{\mu}{\sigma^2}} = \tilde{k}(S_T - K)^+$ has two solutions for $\mu > \sigma^2$ and one solution for $\mu \leq \sigma^2$. If the capital constraint $\tilde{V}_0$ is between zero and the price of the option, $\tilde{k}$ will be determined such that there are solutions to the equation. If c_1 and c_2 are solutions for S_T to $S_T^{\frac{\mu}{\sigma^2}} = \tilde{k}(S_T - K)^+$, then the success set is

$$A = \begin{cases} \{S_T > c_1\} \cap \{S_T < c_2\} & \text{for } \mu < \sigma^2 \\ \{S_T < c_1\} \cup \{S_T > c_2\} & \text{for } \mu > \sigma^2, \end{cases}$$

where $c_1 = 0$ in the first case.

Finding a more explicit representation of the optimal hedging strategy requires to determine the constants c_1 and c_2, or just c_2 in the case $\mu \leq \sigma^2$. This can be done using the Black-Scholes pricing formula (5.2.5) and adjusting it for the knockout option $C\mathbb{1}_A$.

First we can see that this option can be rewritten as a sum of call options and a binary option.

The case $\mu \leq \sigma^2$:

In the case $\mu \leq \sigma^2$ with $A = \{S_T < c\}$, this results in

$$C\mathbb{1}_A = (S_T - K)^+ - (S_T - c)^+ - (c - K)\mathbb{1}_{\{S_T > c\}}.$$

The first two summands can be priced using the Black-Scholes pricing formula, while the last summand is easy to price as a binary option. For an easier notation we define

$$b := \frac{ln\left(\frac{c}{S_0}\right) + \frac{\sigma^2}{2}T}{\sigma}$$

and get the following results:

$$\mathbb{E}^Q[(S_T - K)^+] = S_0\Phi(d_1) - K\Phi(d_2),$$

$$\begin{aligned}
\mathbb{E}^Q[(S_T-c)^+] &= S_0\Phi\left(\frac{ln\left(\frac{S_0}{c}\right)+\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) - c\Phi\left(\frac{ln\left(\frac{S_0}{c}\right)-\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) \\
&= S_0\Phi\left(-\frac{ln\left(\frac{c}{S_0}\right)+\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}+\sigma\sqrt{T}\right) - c\Phi\left(-\frac{ln\left(\frac{c}{S_0}\right)-\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) \\
&= S_0\Phi\left(\frac{-b+\sigma T}{\sqrt{T}}\right) - c\Phi\left(\frac{-b}{\sqrt{T}}\right), \\
\mathbb{E}^Q[(c-K)\mathbb{1}_{\{S_T>c\}}] &= (c-K)Q(S_T>c) \\
&= (c-K)Q\left(W_T^*>\frac{ln\left(\frac{c}{S_0}\right)+\frac{\sigma^2}{2}T}{\sigma}\right) \\
&= (c-K)\Phi\left(-\frac{b}{\sqrt{T}}\right).
\end{aligned}$$

Here we set W_t^* as a Q-Brownian motion. Using these three results we get the following price for $C\mathbb{1}_A$:

$$\mathbb{E}^Q[C\mathbb{1}_A] = S_0\Phi(d_1) - K\Phi(d_2) - S_0\Phi\left(\frac{-b+\sigma T}{\sqrt{T}}\right) + K\Phi\left(\frac{-b}{\sqrt{T}}\right).$$

Finally, the constant b can be found solving

$$\tilde{V}_0 = S_0\Phi(d_1) - K\Phi(d_2) - S_0\Phi\left(\frac{-b+\sigma T}{\sqrt{T}}\right) + K\Phi\left(\frac{-b}{\sqrt{T}}\right)$$

for $\mu<\sigma^2$ for a fixed capital constraint $\tilde{V}_0$, which requires numerical approaches.

After finding the constant b or c respectively one can find the optimal success probability using the lognormal distribution of S_T or the representation of S_T in (5.2.2).

$$\begin{aligned}
\mathbb{P}(A) = \mathbb{P}(S_T<c) &= \mathbb{P}\left(W_T<\frac{\ln\left(\frac{c}{S_0}\right)+\left(\frac{\sigma^2}{2}-\mu\right)T}{\sigma}\right) \\
&= \Phi\left(\frac{\ln\left(\frac{c}{S_0}\right)+\left(\frac{\sigma^2}{2}-\mu\right)T}{\sigma\sqrt{T}}\right) \\
&= \Phi\left(\frac{b-\frac{\mu}{\sigma}T}{\sqrt{T}}\right).
\end{aligned}$$

The case $\mu>\sigma^2$:

On the other hand, the case $\mu>\sigma^2$ is more complex, since we need to find two constants b_1 and b_2 which leads to an additional equation. We can use almost the same argumentation in this case. To price the corresponding knockout option $C\mathbb{1}_A$ with $A=\{S_T<c_1\}\cup\{S_T>c_2\}$ we define

$$b_1 := \frac{ln\left(\frac{c_1}{S_0}\right)+\frac{\sigma^2}{2}T}{\sigma}, \quad b_2 := \frac{ln\left(\frac{c_2}{S_0}\right)+\frac{\sigma^2}{2}T}{\sigma},$$

and rewrite the $C\mathbb{1}_A$ into

$$C\mathbb{1}_A = (S_T - K)^+ - (S_T - c_1)^+ - (c_1 - K)\mathbb{1}_{\{S_T > c_1\}} + (S_T - c_2)^+ + (c_2 - K)\mathbb{1}_{\{S_T > c_2\}},$$

which gives the pricing formula

$$\begin{aligned}\mathbb{E}^Q[C\mathbb{1}_A] =& S_0\Phi(d_1) - K\Phi(d_2) - S_0\Phi\left(\frac{-b_1 + \sigma T}{\sqrt{T}}\right) + K\Phi\left(\frac{-b_1}{\sqrt{T}}\right) \\ &+ S_0\Phi\left(\frac{-b_2 + \sigma T}{\sqrt{T}}\right) - K\Phi\left(\frac{-b_2}{\sqrt{T}}\right). \end{aligned} \tag{5.2.6}$$

Using the capital constraint we get the first equation to determine b_1 and b_2

$$\begin{aligned}\tilde{V}_0 =& S_0\Phi(d_1) - K\Phi(d_2) - S_0\Phi\left(\frac{-b_1 + \sigma T}{\sqrt{T}}\right) + K\Phi\left(\frac{-b_1}{\sqrt{T}}\right) \qquad (5.2.7) \\ &+ S_0\Phi\left(\frac{-b_2 + \sigma T}{\sqrt{T}}\right) - K\Phi\left(\frac{-b_2}{\sqrt{T}}\right). \end{aligned}$$

We have the additional condition that $S_T^{\frac{\mu}{\sigma^2}} = \tilde{k}(S_T - K)^+$, which means that c_1 and c_2 must solve the equation

$$x^{\frac{\mu}{\sigma^2}} = \tilde{k}(x - K)^+.$$

Since b_1 and b_2 are strictly monotone transformations of c_1 and c_2 these two equations are sufficient to uniquely determine c_1 and c_1 or b_1 and b_2 respectively.

Finally, we can again determine the success probability by using (5.2.2).

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(S_T \leq c_1, S_T \geq c_2) \\ &= \Phi\left(\frac{\ln\left(\frac{c_1}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu\right)T}{\sigma\sqrt{T}}\right) + \Phi\left(-\frac{\ln\left(\frac{c_2}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu\right)T}{\sigma\sqrt{T}}\right) \\ &= \Phi\left(\frac{b_1 - \frac{\mu}{\sigma}T}{\sqrt{T}}\right) + \Phi\left(-\frac{b_2 - \frac{\mu}{\sigma}T}{\sqrt{T}}\right). \end{aligned}$$

□

Example 5.2.3.

To use some concrete numbers, we will have a look at the following example: Let $\mathbb{P}$ denote a Black-Scholes market with $S_0 = 100$ and a vanilla call option $C(S_T) = (S_T - 100)^+$. Additionally we set $\mu = 0.1$, $\sigma = 0.5$ and $T = 1$.

Using the above formulas we can calculate success probabilities for different capital constraints, which gives us the following very typical concave curve for the success probabilities as a function of the capital constraint:

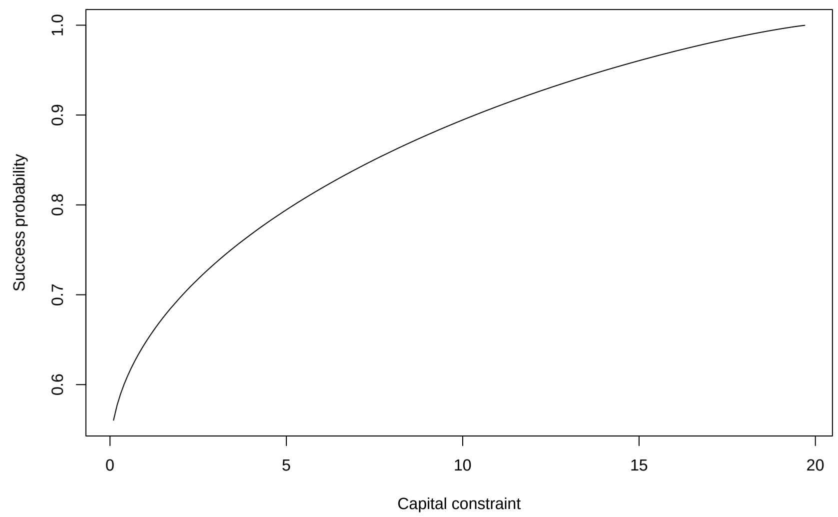

Figure 5.1.: Success probabilities of optimal hedging strategies for a call option in a Black-Scholes market for different capital constraints.

5.2.1.2. Binary option

Binary options, i.e. $C = \mathbb{1}_{\{S_T \in B\}}$ for a set $B \subseteq \mathbb{R}_+$, have the very useful property that prices can be derived using only the cumulative distribution function. In this case, this leads to prices and success probabilities that can be calculated analytically.

Corollary 5.2.4.
Let $p \in \mathbb{R}_+$ and consider a Black-Scholes model $\mathbb{P}$ with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$.

(a) For a binary option $C = \mathbb{1}_{\{S_T > p\}}$, the optimal success set is of the form

$$A = \{S_T \leq p\} \cup \{S_T > \tilde{k}\}, \text{ with } \tilde{k} = \exp\left(\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) - \frac{\sigma^2}{2}T\right).$$

The optimal success probability is

$$\mathbb{P}(A) = 1 - \Phi\left(\frac{\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) + \mu T}{\sqrt{T}\sigma}\right) + \Phi\left(\frac{\ln(p) - \mu T + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\right).$$

(b) For a binary options $C = \mathbb{1}_{\{S_T < p\}}$, the optimal success set is of the form

$$A = \{S_T < \tilde{k}\} \text{ with } \tilde{k} = \exp\left(\Phi^{-1}\left(\tilde{V}_0 + \Phi\left(\frac{\ln(p) + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\right)\right)\sqrt{T}\sigma - \frac{\sigma^2}{2}T\right).$$

The optimal success probability is

$$\mathbb{P}(A) = \Phi\left(\Phi^{-1}\Big(\tilde{V}_0 + \Phi\Big(\frac{\ln(p) + \frac{\sigma^2}{2}T}{\sqrt{T}}\Big)\Big) - \frac{\mu}{\sigma}\sqrt{T}\right).$$

Proof. Let $C = \mathbb{1}_{\{S_T > p\}}$ for a $p \in \mathbb{R}_+$. In this case we get

$$\begin{aligned}
\mathbb{E}^Q[C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > k\ C\}}] &= \mathbb{E}^Q[\mathbb{1}_{\{S_T > p\}}\mathbb{1}_{\{S_T^{\frac{\mu}{\sigma^2}} > k\ \mathbb{1}_{\{S_T > p\}}\}}] \\
&= \mathbb{E}^Q[\mathbb{1}_{\{S_T > p\} \cap \{S_T > \tilde{k}\ \mathbb{1}_{\{S_T > p\}}\}}] \\
&= \mathbb{E}^Q[\mathbb{1}_{\{S_T > p\} \cap \{S_T > \tilde{k}\}}] \\
&= Q(S_T > \tilde{k}),
\end{aligned}$$

where we use that, as long as $\tilde{V}_0 \leq \mathbb{E}^Q[C]$, it holds that $\tilde{k} \geq p$.

Now we can use that we know the distribution of $\frac{S_T}{S_0}$ under Q to be a lognormal distribution, $\frac{S_T}{S_0} \sim LN(-\frac{\sigma^2}{2}T, \sigma^2 T)$. This leaves us with

$$\tilde{V}_0 = \mathbb{E}^Q[C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > k\ C\}}] = Q(S_T > \tilde{k}) = 1 - \Phi\Big(\frac{\ln(\tilde{k}) + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\Big),$$

and we can determine $\tilde{k}$ as

$$\tilde{k} = \exp\left(\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) - \frac{\sigma^2}{2}T\right).$$

Finally, the optimal success probability can be determined as

$$\begin{aligned}
\mathbb{P}\Big(\frac{d\mathbb{P}}{dQ} > k\ C\Big) &= \mathbb{P}(S_T^{\frac{\mu}{\sigma^2}} > k\mathbb{1}_{\{S_T > p\}}) = \mathbb{P}(S_T > \tilde{k}\mathbb{1}_{\{S_T > p\}}) \\
&= \mathbb{P}(\{S_T > \tilde{k}\} \cup \{S_T \leq p\}) \\
&= \mathbb{P}\left(\left\{S_T > \exp\left(\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) - \frac{\sigma^2}{2}T\right)\right\} \cup \{S_T \leq p\}\right) \\
&= \mathbb{P}\left(S_T > \exp\left(\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) - \frac{\sigma^2}{2}T\right)\right) + \mathbb{P}(S_T \leq p),
\end{aligned}$$

where again we can use that $\tilde{k} \geq p$ as long as $\tilde{V} < \mathbb{E}^Q[C]$. Since under $\mathbb{P}$ the random variable $\frac{S_T}{S_0}$ is still lognormally distributed, with $\frac{S_T}{S_0} \sim LN(\mu T - \frac{\sigma^2}{2}T, \sigma^2 T)$, we can solve the last expression explicitly:

$$\begin{aligned}
\mathbb{P}\Big(\frac{d\mathbb{P}}{dQ} > k\ C\Big) &= \mathbb{P}\left(S_T > \exp\left(\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) - \frac{\sigma^2}{2}T\right)\right) + \mathbb{P}(S_T \leq p) \\
&= 1 - \Phi\Big(\frac{\sqrt{T}\sigma\Phi^{-1}(1 - \tilde{V}_0) + \mu T}{\sqrt{T}\sigma}\Big) + \Phi\Big(\frac{\ln(p) - \mu T + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\Big).
\end{aligned}$$

If instead we define $C = \mathbb{1}_{\{S_T<p\}}$, then we come to the conclusion that

$$\begin{aligned}\mathbb{E}^Q[C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ}>k\ C\}}] &= \mathbb{E}^Q[\mathbb{1}_{\{S_T<p\}\cap\{S_T>\tilde{k}\}}] = Q(p < S_T < \tilde{k}) \\ &= \Phi\Big(\frac{\ln(\tilde{k}) + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\Big) - \Phi\Big(\frac{\ln(p) + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\Big),\end{aligned}$$

which leads to

$$\tilde{k} = \exp\left(\Phi^{-1}\Big(\tilde{V}_0 + \Phi\Big(\frac{\ln(p) + \frac{\sigma^2}{2}T}{\sqrt{T}\sigma}\Big)\Big)\sqrt{T}\sigma - \frac{\sigma^2}{2}T\right)$$

and finally gives an optimal success probability of

$$\mathbb{P}\Big(\frac{d\mathbb{P}}{dQ} > k\ C\Big) = \mathbb{P}(S_T < \tilde{k}) = \Phi\left(\Phi^{-1}\Big(\tilde{V}_0 + \Phi\Big(\frac{\ln(p) + \frac{\sigma^2}{2}T}{\sqrt{T}}\Big)\Big) - \frac{\mu}{\sigma}\sqrt{T}\right).$$

□

5.2.2. Numerical approach for jump models

In the previous section we showed that in the simple case of a Black-Scholes model with path independent options it is possible to find an analytical representation of the optimal success sets and to determine the success probability analytically. To find these sets we make use of the very strong properties (5.2.4) and (5.2.5), i.e. an analytic solution to the stochastic differential equation, an analytic solution to the Radon-Nikodym derivative in terms of S itself and an analytic pricing formula. These strong properties are not fulfilled for many models.

This leads to the question of whether the theory presented for finding optimal hedging strategies in single market models is applicable to models that are more complex than a Black-Scholes model. In this section, we show how Monte Carlo simulation can be used to compute optimal success sets and probabilities in models that do not satisfy all of the above properties. Optimal success sets can also be determined implicitly using the same method.

The main advantage of the presented theory is that we already know the structure of the optimal hedging strategy, which means that we do not need to use a complex optimiser over numerous competing strategies. Instead we only need to find a way to find the worst case measure $\tilde{Q}$ and the constant k in the set $\{\frac{d\mathbb{P}}{dQ} > kC\}$ and determine the probability of this set. In complete market models, finding optimal hedging strategies and their success probabilities can be reduced to the following two tasks:

(a) Find $k \in \mathbb{R}$ that solves

$$\mathbb{E}^Q \left[C \mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > kC\}} \right] = \tilde{V}_0. \tag{5.2.8}$$

(b) Calculate success probability

$$\mathbb{P}\left(\frac{d\mathbb{P}}{dQ} > kC \right) = \mathbb{E}^{\mathbb{P}} \left[\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > kC\}} \right]. \tag{5.2.9}$$

As one can see, we are only interested in expected values of random variables, which motivates the usage of Monte Carlo simulation.

Note here that $\frac{d\mathbb{P}}{dQ}$ and C are random processes, which can be simulated, and note that $\frac{d\mathbb{P}}{dQ}$ already implies $\frac{dQ}{d\mathbb{P}}$.

Lemma 5.2.5.
Let μ and ν be σ-finite measures on a measurable space. If $\mu \sim \nu$, then for the Radon-Nikodym derivative $\frac{d\mu}{d\nu}$ it holds that

$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu} \right)^{-1}.$$

Proof. It follows by the Radon-Nikodym theorem that the expression $\frac{d\nu}{d\mu}$ which is defined as the solution for f in

$$\nu(A) = \int_A f d\mu$$

for every measurable set A is unique ν-almost surely. It follows directly from

$$\int_A d\nu = \nu(A) = \int_A \frac{d\nu}{\mu} d\mu = \int_A \frac{d\nu}{d\mu} \frac{d\mu}{d\nu} d\nu \quad \text{for every measurable } A,$$

that $\frac{d\nu}{d\mu} \frac{d\mu}{d\nu} = 1$ ν-almost surely. □

5.2.2.1. Numerical approach

In Monte Carlo simulation, we define discrete time steps $t_0, ..., t_m$ and simulate n paths $\omega_1, ..., \omega_n$ of S via its SDE. The Radon-Nikodym derivative can be determined pathwise. Since the claim C depends only on the simulated paths of S, we apply the law of large numbers to solve tasks (5.2.8) and (5.2.9).

Note that if the market is incomplete, the choice of the price measure $Q \in \mathcal{Q}^{\mathbb{P}}$ is not clear a priori. Thus, a worst-case choice can be found using a grid search on the set $\mathcal{Q}^{\mathbb{P}}$ of all possible choices for Q.

In short:

(a) Simulate n paths $\omega_1, ..., \omega_n$ under $(\Omega, \mathbb{P})$ using the physical measure $\mathbb{P}$

(b) If the market is incomplete fix a possible measure Q of a grid on $\mathcal{Q}^{\mathbb{P}}$.

(c) Calculate $\frac{dQ}{d\mathbb{P}}(\omega_i)$ for every ω_i.

(d) Find a constant $k \in \mathbb{R}$, such that the constraint is fulfilled, using the following approximation:

$$\begin{aligned}\tilde{V}_0 &= \mathbb{E}^Q\left[C\mathbb{1}_{\left\{\frac{d\mathbb{P}}{dQ}>kC\right\}}\right] \\ &= \mathbb{E}^{\mathbb{P}}\left[C\mathbb{1}_{\left\{\frac{d\mathbb{P}}{dQ}>kC\right\}}\frac{dQ}{d\mathbb{P}}\right] \\ &\approx \frac{1}{n}\sum_{i=1}^{n} C(\omega_i)\mathbb{1}_{\left\{\frac{d\mathbb{P}}{dQ}>kC\right\}}(\omega_i)\frac{dQ}{d\mathbb{P}}(\omega_i).\end{aligned}$$

Numerically one could for instance find the smallest $k \in \mathbb{R}$ such that

$$\frac{1}{n}\sum_{i=1}^{n} C(\omega_i)\mathbb{1}_{\left\{\frac{d\mathbb{P}}{dQ}>kC\right\}}(\omega_i)\frac{dQ}{d\mathbb{P}}(\omega_i) \geq \tilde{V}_0.$$

(e) Calculate the success probability via

$$\begin{aligned}\mathbb{P}\left(\frac{d\mathbb{P}}{dQ} > kC\right) &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{1}_{\left\{\frac{d\mathbb{P}}{dQ}>kC\right\}}\right] \\ &\approx \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\left\{(\frac{dQ}{d\mathbb{P}})^{-1}>kC\right\}}(\omega_i).\end{aligned}$$

(f) If the market is incomplete, return to step (c) and fix another possible measure Q or choose the worst case.

Note that the equation in step (e) has a unique solution in k because the sum is increasing with growing k as C and $\frac{dQ}{d\mathbb{P}}$ are positive processes.

Remark 5.2.6.

A major disadvantage of this approach is that we only get success probabilities but not the exact structure of the knockout option as in the analytic approach for the Black-Scholes model. But one can implicitly reconstruct the knockout option by plotting the hedged paths, which sometimes have an observable pattern.

5.2.2.2. Merton-Jump-Diffusion model

If we consider the Merton-Jump model, we are working with a typical jump diffusion model. The stochastic differential equation that the underlying is supposed to solve is the following:

$$dS_t = S_t(\mu - \lambda\kappa)dt + \sigma S_t dW_t + S_t(e^{J_t} - 1)dN_t.$$

In this model μ is a drift term, J_t affects the jumpsize at time t and is normally distributed and N is a poisson point process with rate λ, that defines the times at which a jump occurs. In addition κ is defined as $\kappa = \mathbb{E}[e^{J_t} - 1]$, such that $((e^{J_t} - 1)dN_t - \lambda\kappa dt)$ has a mean drift of zero, or in other words μ is indeed the drift of this process.

First of all, we will present some results concerning the Merton-Jump Model:

Lemma 5.2.7.
The stochastic differential equation of the Merton-Jump model has the following solution

$$S_t = \exp\left(\Big(\mu - \frac{\sigma^2}{2} - \lambda\kappa\Big)t + \sigma W_t + \sum_{n=1}^{N_t} J_n\right)$$

Proof. The generalized Itô formula for a C^2-function f, that can be found in Chapter II, Theorem 33 in Protter [56], gives us:

$$df(S_t) = \frac{\partial f(S_t)}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 f(S_t)}{\partial^2 S_t} d[S_t, S_t]^c + \left(f(S_t) - f(S_{t-}) - \frac{\partial f(S_t)}{\partial S_t}\Delta S_t\right) \tag{5.2.10}$$

We can use the fact that

$$\begin{aligned}\frac{\partial f(S_t)}{\partial S_t} dS_t - \frac{\partial f(S_t)}{\partial S_t}\Delta S_t =& \frac{\partial f(S_t)}{\partial S_t} dS_t^c \\ =& \frac{\partial f(S_t)}{\partial S_t}(S_t(\mu - \lambda\kappa)dt + \sigma S_t dW_t)\end{aligned}$$

to get the Itô formula for jump-diffusion processes. If we set $f(x) = \ln(x)$ and set $J_t = 0$ whenever there is not a jump in t, i.e. $\Delta N_t \neq 0$, (note that there are almost surely not two jumps at the same time) we get the following equations:

$$\begin{aligned}df(S_t) &= \frac{\partial f(S_t)}{\partial S_t} dS_t + \frac{1}{2}\frac{\partial^2 f(S_t)}{\partial^2 S_t} d[S_t, S_t]^c + \left(f(S_t) - f(S_{t-}) - \frac{\partial f(S_t)}{\partial S_t}\Delta S_t\right) \\ &= \frac{1}{S_t}(S_t(\mu - \lambda\kappa)dt + \sigma S_t dW_t) - \frac{\sigma^2 S_t^2}{2S_t^2}dt + ln(S_t) - ln(S_{t-}) \\ &= ((\mu - \lambda\kappa)dt + \sigma dW_t) - \frac{\sigma^2}{2}dt + ln(e^{J_t} S_{t-}) - ln(S_{t-})\end{aligned}$$

$$= ((\mu - \lambda\kappa)dt + \sigma dW_t) - \frac{\sigma^2}{2}dt + ln(e^{J_t}) + ln(S_{t-}) - ln(S_{t-})$$
$$= ((\mu - \lambda\kappa)dt + \sigma dW_t) - \frac{\sigma^2}{2}dt + ln(e^{J_t}).$$

Integrating both sides gives us

$$\ln(S_t) = \int_0^t \left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right) ds + \int_0^t \sigma dW_s + \sum_{n=1}^{N_t} J_n,$$

where we can apply the exponential function to get

$$S_t = \exp(\ln(S_t)) = \exp\left(\left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right) t + \sigma W_t + \sum_{n=1}^{N_t} J_n\right).$$

□

Remark 5.2.8.

As the Merton-Jump model contains three independent stochastic terms, it is not a complete model, which means that we can find more than one (even infinitely many) equivalent martingale measures. In his original publication [48] Merton proposed an equivalent martingale measure that leaves the jump heights and intensities unchanged. The idea behind this suggestion is that the risk of jumps might be fully diversifiable, so that there is no risk premium for jumps. However, it is known nowadays that jump risks are not purely unsystematic risks, which means we should be aware of every possible equivalent martingale measure.

To apply a Monte Carlo simulation, we need to find the Radon-Nikodym derivative of the physical measure $\mathbb{P}$ with respect to a worst case equivalent martingale measure Q. The next lemma gives a solution to the Radon-Nikodym derivative in the Merton-Jump model. The solution can be found in Cheang and Chiarella [12], but is provided with a more straightforward proof here. See Appendix A for a more detailed derivation for general jump diffusion model.

Corollary 5.2.9.

Let $\mathbb{P}$ be a physical measure with

$$dS_t = S_t(\mu - \lambda\kappa)dt + \sigma S_t dW_t + S_t(e^{J_t} - 1)dN_t,$$

where N is a Poisson process with rate λ and $(J_n)_{n\in\mathbb{N}}$ are normally distributed jumps with $\mathbb{E}^{\mathbb{P}}[e^{J_n} - 1] = \kappa$. The Radon-Nikodym Derivative of $\frac{d\mathbb{P}}{dQ}$ for an equivalent martingale measure Q is of the form

$$\left.\frac{d\mathbb{P}}{dQ}\right|_{\mathcal{F}_t} = \exp\left(\theta W_t + \frac{\theta^2}{2}t - \sum_{n=1}^{N_t}(\gamma J_n + \nu) + \lambda\tilde{\kappa}t\right). \tag{5.2.11}$$

The terms $\gamma \in \mathbb{R}$ and $\nu \in \mathbb{R}$ manipulate the arrival rate and jump sizes for the jump part under Q while $\theta = \frac{\mu-\lambda\kappa+\lambda^Q\kappa^Q}{\sigma}$ with $\lambda^Q = \lambda(1+\tilde{\kappa})$ and $\kappa^Q = \mathbb{E}^Q[e^J - 1]$ changes the drift of the Wiener process. In addition let $\tilde{\kappa} = e^{\nu}\mathbb{E}^{\mathbb{P}}[e^{\gamma J}] - 1$.

Proof. First, we can determine some of the variables and expressions explicitly. We Assume J_n to be normally distributed with mean μ_J and satndard deviation σ_J.

(a) $\kappa = \mathbb{E}^{\mathbb{P}}[e^J - 1] = \exp(\mu_J + \sigma_J^2/2) - 1$,

(b) $\tilde{\kappa} = \exp(\nu + \gamma\mu_J + \gamma^2\sigma_J^2/2) - 1$,

(c) $\kappa^Q = \mathbb{E}^Q[e^J - 1] = \exp(\mu_J + (2\gamma+1)\sigma_J^2/2) - 1$,

(d) $\mathbb{E}[\exp(aJ+b)] = \exp(a\mu + b + a^2\sigma^2/2)$ for $J \sim \mathcal{N}(\mu, \sigma^2)$,

(e) $\mathbb{E}[e^{xN}] = \exp(\lambda t(e^x - 1))$ for $N \sim Pois(\lambda)$.

Wit some calculations we can see that $\frac{dQ}{d\mathbb{P}}$ indeed is a martingale:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[\left.\frac{dQ}{d\mathbb{P}}\right|_{\mathcal{F}_t}\right] =& \mathbb{E}^{\mathbb{P}}\left[\exp(-\theta W_t)\exp\left(\frac{-\theta^2}{2}t\right)\exp\left(\sum_{n=0}^{N_t}\gamma J_n + \nu\right)\exp(-\lambda\tilde{\kappa}t)\right] \\
=& \mathbb{E}^{\mathbb{P}}\Big[\exp\left(-\theta W_t\right)\Big]\exp\left(\frac{-\theta^2}{2}t\right)\mathbb{E}^{\mathbb{P}}\left[\exp\left(\sum_{n=0}^{N_t}\gamma J_n + \nu\right)\right]\exp(-\lambda\tilde{\kappa}t) \\
=& \exp\left(\frac{\theta^2}{2}t\right)\exp\left(-\frac{\theta^2}{2}t\right)\mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\exp\left(\sum_{n=0}^{N_t}\gamma J_n + \nu\right)\Big| N_t = N\right]\right]\exp(-\lambda\tilde{\kappa}t) \\
\overset{4)}{=}& \mathbb{E}^{\mathbb{P}}\left[\exp\left(N(\gamma\mu_J + \nu) + \frac{1}{2}N\sigma_J^2\gamma^2\right)\right]\exp(-\lambda\tilde{\kappa}t) \quad \text{with } N \sim Pois(\lambda) \\
=& \mathbb{E}^{\mathbb{P}}\left[\exp\left(Nt\left(\gamma\mu_J + \nu + \frac{\sigma_J^2\gamma^2}{2}\right)\right)\right]\exp(-\lambda\tilde{\kappa}t) \\
\overset{5)}{=}& \exp\left(\lambda t\left(\exp\left(\gamma\mu_J + \nu + \frac{\sigma_J^2\gamma^2}{2}\right) - 1\right)\right)\exp(-\lambda\tilde{\kappa}t) \overset{2)}{=} 1.
\end{aligned}$$

Then, with some extensive calculations one can see that $\frac{dQ}{d\mathbb{P}}$ defines a martingale measure:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[\frac{dQ}{d\mathbb{P}}S_t\right] =& \mathbb{E}^{\mathbb{P}}\Bigg[\exp(-\theta W_t)\exp\left(\frac{-\theta^2}{2}t\right)\exp\left(\sum_{n=0}^{N_t}\gamma J_n + \nu\right)\exp(-\lambda\tilde{\kappa}t) \\
& S_0\exp\left(\left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)t + \sigma W_t + \sum_{n=1}^{N_t}J_n\right)\Bigg]
\end{aligned}$$

$$
\begin{aligned}
=&S_0\mathbb{E}^{\mathbb{P}}[(\sigma-\theta)W_t]\exp\left(\frac{-\theta^2}{2}t\right)\exp\left(\left(\mu-\lambda\kappa-\frac{\sigma^2}{2}-\lambda\tilde{\kappa}\right)t\right)\\
&\mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\exp\left(\sum_{n=1}^{N}(\gamma+1)J_n+\nu\right)\middle|N_t=N\right]\right]\\
\overset{4)}{=}&S_0\exp\left(\frac{(\sigma-\theta)^2}{2}t\right)\exp\left(\frac{-\theta^2}{2}t\right)\exp\left((\mu-\lambda\kappa-\frac{\sigma^2}{2}-\lambda\tilde{\kappa})t\right)\\
&\mathbb{E}^{\mathbb{P}}\left[\exp\left(\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)N\right)\right]\quad\text{with } N\sim Pois(\lambda)\\
\overset{5)}{=}&S_0\exp\left(\left(\frac{\sigma^2}{2}-\sigma\theta+\mu-\lambda\kappa-\frac{\sigma^2}{2}-\lambda\tilde{\kappa}\right)t\right)\\
&\exp\left(\lambda t\left(\exp\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)-1\right)\right)\\
\overset{3)}{=}&S_0\exp((-\lambda^Q\kappa^Q-\lambda\tilde{\kappa})t)\exp\left(\lambda t\left(\exp\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)-1\right)\right)\\
=&S_0\exp(-\lambda t\kappa^Q)\exp(-\lambda t\tilde{\kappa}(1+\kappa^Q))\\
&\exp\left(\lambda t\left(\exp\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)-1\right)\right)\\
=&S_0\exp(-\lambda t\kappa^Q)\exp\Big[\lambda t\Big(-(\exp(\nu+\gamma\mu_J+\gamma^2\sigma_J^2/2)-1)\\
&\exp(\mu_J+(2\gamma+1)\sigma_J^2/2)+\exp\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)-1\Big)\Big]\\
=&S_0\exp(-\lambda t\kappa^Q)\exp\Big[\lambda t\Big(-\exp((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2})\\
&+\exp(\mu_J+(2\gamma+1)\sigma_J^2/2)+\exp\left((\gamma+1)\mu_J+\nu+\frac{(\gamma+1)^2\sigma_J^2}{2}\right)-1\Big)\Big]\\
=&S_0\exp(-\lambda t\kappa^Q)\exp(\lambda t\underbrace{(\exp(\mu_J+(2\gamma+1)\sigma_J^2/2)-1}_{=\kappa^Q})=S_0
\end{aligned}
$$

□

Example 5.2.10. *Success probabilities for a single market Merton-Jump model*
In this example we will have a look on a Merton-Jump model with a vanilla put option.

(a) Underlying: $dS_t = S_t(\mu-\lambda\kappa)dt + \sigma S_t dW_t + S_t(e^{J_t}-1)dN_t$

(b) Parameter: $S_0 = 100, \mu = 0.1, \sigma = 0.2, \lambda = 2$

(c) Jump parameter: $J_t \sim \mathcal{N}(\mu_J, \sigma_J)$ with $\mu_J = -0.05, \sigma_J = 0.02$

(d) Option: $C(S_T) = (K - S_T)^+$ with $K = 100$

First, we face the problem of finding a worst-case measure, since there are infinitely many equivalent martingale measures. Since there is no analytic solution to this problem, we will use a grid search. We can see that by varying ν and γ we can construct any equivalent martingale measure. Thinking of the Neyman-Pearson lemma we are trying to find the measure that is as difficult to distinguish from the original physical measure $\mathbb{P}$ as possible, i.e. we are not looking for inconveniently large values for ν and γ, which would be difficult anyway due to the very large variance. Note that the grid in this case is only two-dimensional, which makes the example computationally solvable in a reasonable amount of time. The following success probabilities can be calculated for different choices of ν and γ, i.e. different choices of $Q \in \mathcal{Q}^{\mathbb{P}}$, for a capital constraint of $\tilde{V}_0 = 4$:

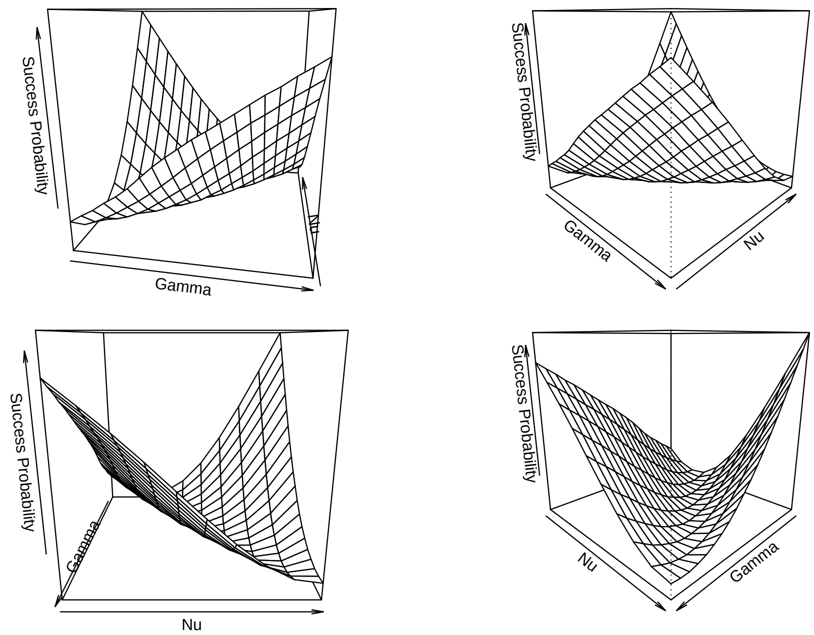

Figure 5.2.: Success probabilities for different values of ν and γ in the Merton-Jump model.

Then we can choose the values ν and γ such that success probability is minimal. Additionally calculating success probabilities for different capital constraints leaves us with the following very typical concave graph:

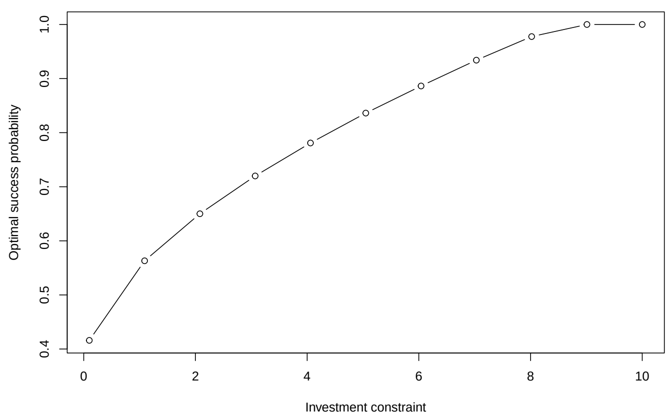

Figure 5.3.: Success Probabilities depending on the capital constraint in the Merton-Jump model.

All in all, the above presented way is very computationally intensive, as there needs to be a sufficiently large amount of simulations, since the variance in a jump model is quite large, while the grid for the worst-case martingale measure can be theoretically infinitely large if not restricted by external considerations. But on the other hand the used method can be easily applied to any market model as long as the Radon-Nikodym derivative for any measure change to an equivalent martingale measure is known.

5.3. Robust models

After presenting examples and numerical approaches for single markets we will have a look at robust markets. Again, as in the single market case, we will first present an analytical approach for the Black-Scholes model with parameter uncertainty. Then, because we again use a lot of properties that are not generally satisfied for most models, we will again have a look at possible numerical approaches.

In Theorem 4.2.9 it is already shown that finding optimal strategies and success sets can be reduced to the problem of constructing the set $\left\{\frac{d\mathbb{P}_\lambda}{dQ} > kC\right\}$. Before presenting an analytic solution as in Lemma 5.2.5 we can make use of some properties of the Radon-Nikodym derivative to make the construction of Radon-Nikodym derivatives easier.

Lemma 5.3.1.
Let μ, ν and ρ be σ-finite measures on a measurable space. The Radon-Nikodym derivative fulfils the following properties:

(a) If $\mu \ll \nu$ and $\rho \ll \nu$, it holds that $\frac{d(\mu+\rho)}{d\nu} = \frac{d\nu}{d\mu} + \frac{d\rho}{d\nu}$ ν-almost surely,

(b) If $\mu \ll \nu$ and $\lambda \in \mathbb{R}$ constant, it holds that $\frac{d(\lambda\mu)}{d\nu} = \lambda\frac{d\mu}{d\nu}$ ν-almost surely,

(c) If $\mu \ll \rho \ll \nu$, it holds that $\frac{d\mu}{d\nu} = \frac{d\mu}{d\rho}\frac{d\rho}{d\nu}$ ν-almost surely.

Proof. Note first that the Radon-Nikodym theorem states that Radon-Nikodym derivatives are uniquely defined, which is why the following equations are sufficient to prove the stated properties:

The first property follows with

$$\int\limits_A \frac{d(\mu+\rho)}{d\nu} d\nu = (\mu+\rho)(A) = \mu(A) + \rho(A) = \int\limits_A (\frac{d\mu}{d\nu} + \frac{d\rho}{d\nu}) d\nu.$$

The second property follows directly by

$$\int\limits_A \lambda\frac{d\mu}{d\nu} d\nu = \lambda \int\limits_A \frac{d\mu}{d\nu} d\nu = \lambda \cdot \mu(A) = (\lambda \cdot \mu)(A) = \int\limits_A \frac{d(\lambda\mu)}{d\nu} d\nu.$$

The last property follows by

$$\int\limits_A \frac{d\mu}{d\nu} d\nu = \mu(A) = \int\limits_A \frac{d\mu}{d\rho} d\rho = \int\limits_A \frac{d\mu}{d\rho}\frac{d\rho}{d\nu} d\nu.$$

□

These results can be used to work with the convex combinations of market measures $\mathbb{P}_\lambda = \sum\limits_{i=1}^{n} \lambda_i \mathbb{P}_i$ with $\lambda_i > 0$ and $\sum\limits_{i=1}^{n} \lambda_i = 1$. Since we are only interested in the Radon-Nikodym derivative $\frac{d\mathbb{P}_\lambda}{dQ}$, the above lemma allows to deal with this new model without the need of detailed understanding of these mixed measures. We can write

$$\frac{d\mathbb{P}_\lambda}{dQ} = \sum_{i=1}^{n} \lambda_i \frac{d\mathbb{P}_i}{dQ}.$$

For the case of only two models $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$, the Radon-Nikodym derivative is of the simple form

$$\frac{d\mathbb{P}_\lambda}{dQ} = \lambda\frac{d\mathbb{P}_1}{dQ} + (1-\lambda)\frac{d\mathbb{P}_2}{dQ}.$$

In both cases we do get the desired Radon-Nikodym derivative of $\mathbb{P}_\lambda$ with respect to Q only by adding already known and often simpler terms.

5.3.1. Analytic solution in Black-Scholes models with parameter uncertainty

In this section we will consider parameter uncertainty in Black-Scholes models, which is directly related to robust market modelling. We consider a robust market model that contains two models $\mathcal{P} = \{\mathbb{P}_1, \mathbb{P}_2\}$ such that under $\mathbb{P}_1$ we have

$$dS_t = \mu_1 S_t dt + \sigma S_t dW_t$$

and under $\mathbb{P}_2$ we have

$$dS_t = \mu_2 S_t dt + \sigma S_t dW_t.$$

This means that both measures are Black-Scholes models with different drift terms but the same variance, which has the advantage that both measures are equivalent and have the same equivalent martingale measure. Note that this is not strictly necessary for our calculations, but avoids problems we will examine in more detail in Chapter 6.

We already know from Corollary 5.2.2 that the optimal success sets in both cases are

$$A_i = \{S_T < c_1\} \cup \{S_T > c_2\} \text{ for } i \in \{1, 2\},$$

where $c_2 = \infty$ if $\mu_i < \sigma^2$.

In the case of $\mu_i > \sigma^2$ for both measures or $\mu_i < \sigma^2$ for both measures, we will get the same optimal success sets for both physical measures, because the prices of the options and the structure of the optimal sets coincide. But in the case $\mu_1 < \sigma^2 < \mu_2$, we are in the interesting case where the optimal success sets are different depending on the choice of the physical measure.

Example 5.3.2.
Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be Black-Scholes measures with $\mu_1 = 0.1$, $\mu_2 = 0.5$ and variance $\sigma^2 = 0.25$, such that we are in case of different success sets. Further let $S_0 = 100$ and $C(S_T) = (S_T - K)^+$ with a strike price of $K = 100$.

Using the solving strategy already presented in Section 5.2 for a capital constraint of $\tilde{V}_0 = 12$ we will get that

$$A_1 = \{S_T < 199.63\},$$
$$A_2 = \{S_T < 167.75\} \cup \{S_T > 247,59\}.$$

In other words, we can say that in the first case of $\mu_1 = 0.1$ we will hedge successfully whenever the value of the underlying is below 199.63 and in the second case we will hedge successfully whenever the value of the underlying is below 167.75 or larger than 247.75. As expected, these success sets are far from being optimal under the other measure, as we can see in the following probabilities

$$\mathbb{P}_1(A_1) = 0.9240, \quad \mathbb{P}_1(A_2) = 0.8922,$$
$$\mathbb{P}_2(A_1) = 0.7365, \quad \mathbb{P}_2(A_2) = 0.7559.$$

These success probabilities can be found using (5.2.2) which leads to

$$\mathbb{P}_i(A_j) = \Phi\left(\frac{\ln\left(\frac{c_1^j}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu_i\right)T}{\sigma}\right) + \Phi\left(-\frac{\ln\left(\frac{c_2^j}{S_0}\right) + \left(\frac{\sigma^2}{2} - \mu_i\right)T}{\sigma}\right).$$

It remains to find the success sets and success probabilities between these extreme sets, which can be done using the theory presented in Section 4.2. Due to Theorem 4.2.9 we only need to find the set $A_\lambda = \{\frac{d\mathbb{P}_\lambda}{dQ} > const \cdot C\}$, which does not require further theoretic work due to Lemma 5.3.1.

For every $\lambda \in [0, 1]$ we get new optimal sets

$$\begin{aligned} A_\lambda &= \left\{\frac{d\mathbb{P}_\lambda}{dQ} > k(S_T - K)^+\right\} \\ &= \left\{\lambda\frac{d\mathbb{P}_1}{dQ} + (1-\lambda)\frac{d\mathbb{P}_2}{dQ} > k(S_T - K)^+\right\} \\ &= \left\{\lambda k_1 S_T^{\frac{\mu_1}{\sigma^2}} + (1-\lambda)k_2 S_T^{\frac{\mu_2}{\sigma^2}} > k(S_T - K)^+\right\}, \end{aligned}$$

with suitable constants k_1, k_2 and k. For these sets we see that the equation $\frac{d\mathbb{P}_\lambda}{dQ} = k(S_T - K)^+$ still has two solutions in S_T in addition to $S_T = 0$. Note that k is always determined such that there are indeed two solutions, except for one of the cases $\lambda \in \{0, 1\}$. This means that the set A_λ is still of the form

$$A_\lambda = \left\{S_T < c_1^\lambda\right\} \cup \left\{S_T > c_2^\lambda\right\},$$

where c_1^λ and c_2^λ are constants depending on λ. Continuity as shown in Theorem 4.2.9 shows that c_1^λ and c_2^λ are continuous functions in λ that converge to the constants we have already determined in the extreme cases for $\lambda \to 0$ or $\lambda \to 1$. We can fix the lower constant c_1^λ and use the pricing formula for the knockout option (5.2.7) to determine c_2^λ. This leads to the following constants c_1 and c_2 for different values of λ.

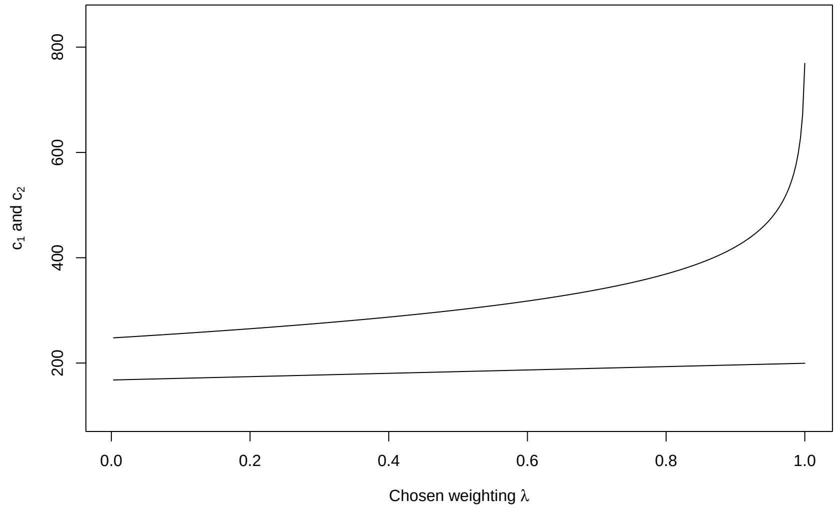

Figure 5.4.: Barrier c_1 and c_2 of the optimal knockout options for different weightings λ.

The upper bound represents the value c_2^λ, while the lower bound is c_1^λ. More interesting is the graph we get when we plot the success probabilities under both measures. In the following figure we can see, as already stated in Theorem 4.2.16, that the success probabilities of all hedging portfolios form a strictly convex set.

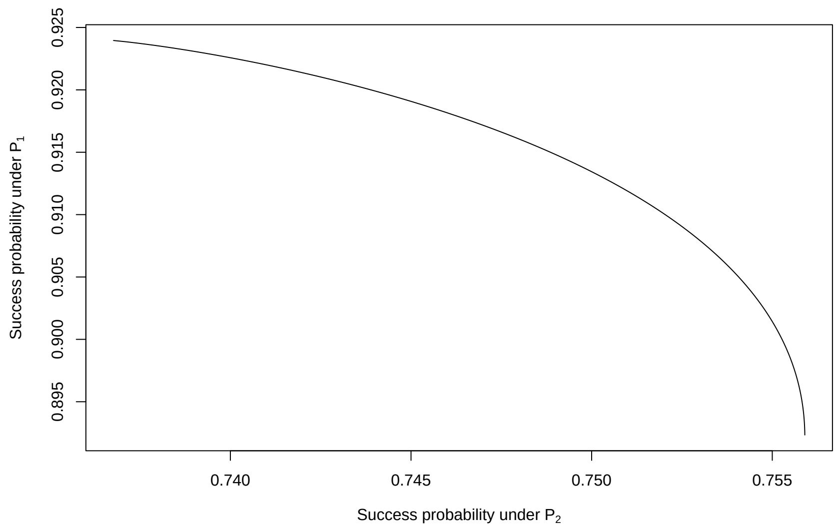

Figure 5.5.: Optimal success probabilities under $\mathbb{P}_1$ and $\mathbb{P}_2$ for different weightings λ.

As one can see in the previous figure, there are advantages of considering more than one model. If there is no preference for any model, one could aim to maximise the sum of all success probabilities which leads to an optimal strategy that never lies in the extreme cases $\lambda = 0$ or $\lambda = 1$. The following figure shows the sum of both success probabilities under different λ.

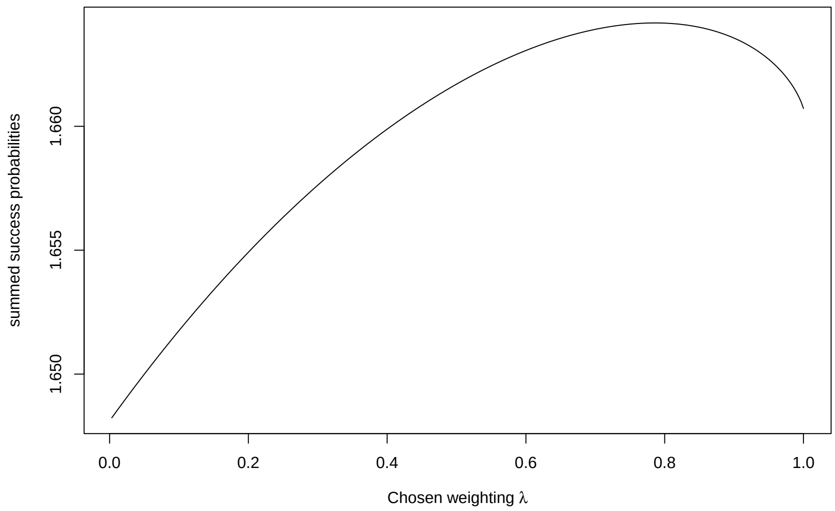

Figure 5.6.: Sum of success probabilities for different weightings λ.

5.3.2. Monte Carlo simulation

As we have already mentioned for the single market models, an analytic solution requires many properties that are lost when constructing models that are more complex than the Black-Scholes model. However, we can again make use of Monte Carlo simulation to find optimal success probabilities and optimal success sets. In this case we face the additional problem of more than one physical measure and more importantly a much larger set of possible worst case measures.

For every weight vector $(\lambda_1, ..., \lambda_m)$ on $\mathcal{P} = \{\mathbb{P}_1, ..., \mathbb{P}_m\}$ we can determine the corresponding optimal hedging strategy to $\mathbb{P}_\lambda$ as follows:

(a) Fix a possible worst case measure $Q \in \mathcal{Q}_\lambda$, i.e. fix a weighting vector $\mu = (\mu_1, ..., \mu_m)$.

(b) Simulate n paths $\omega_1, ..., \omega_n$ under measure Q.

(c) Calculate $\frac{d\mathbb{P}_j}{dQ}(\omega_i)$ for every ω_i and every j.

(d) Find a constant $k \in \mathbb{R}$ such that the constraint is fulfilled, e.g. using the approximation

$$\begin{aligned}\tilde{V}_0 &= \mathbb{E}^Q\left[C\mathbb{1}_{\left\{\frac{d\mathbb{P}_\lambda}{dQ}>kC\right\}}\right]\\ &\approx \frac{1}{n}\sum_{i=1}^{n} C(\omega_i)\mathbb{1}_{\left\{\sum\limits_{j=1}^{m}\frac{\lambda_j d\mathbb{P}_j}{dQ}>kC\right\}}(\omega_i).\end{aligned}$$

Numerically one could for instance find the smallest $k \in \mathbb{R}$ such that

$$\frac{1}{n}\sum_{i=1}^{n} C(\omega_i)\mathbb{1}_{\left\{\sum\limits_{j=1}^{m}\frac{\lambda_j d\mathbb{P}_j}{dQ}>kC\right\}}(\omega_i) \geq \tilde{V}_0.$$

(e) Calculate the success probabilities via

$$\begin{aligned}\mathbb{P}_j\left(\frac{d\mathbb{P}_\lambda}{dQ} > kC\right) &= \mathbb{E}^{\mathbb{P}_j}\left[\mathbb{1}_{\{\frac{d\mathbb{P}_\lambda}{dQ}>kC\}}\right] = \mathbb{E}^{\mathbb{P}_j}\left[\mathbb{1}_{\{\frac{d\mathbb{P}_\lambda}{dQ}>kC\}}\mathbb{1}_{supp(\mathbb{P}_j)}\right]\\ &= \mathbb{E}^Q\left[\mathbb{1}_{\{\frac{d\mathbb{P}_\lambda}{dQ}>kC\}}\mathbb{1}_{supp(\mathbb{P}_j)}\frac{dQ}{d\mathbb{P}_j}\right]\\ &\approx \sum_{i=1}^{\tilde{n}} \mathbb{1}_{\{\frac{d\mathbb{P}_\lambda}{dQ}>kC\}}(\omega_i)\frac{dQ}{d\mathbb{P}_j}(\omega_i),\end{aligned}$$

where the term $\frac{dQ}{d\mathbb{P}_j}$ only exists on the support of $\mathbb{P}_j$ and thus the last term only uses the $\tilde{n}$ paths on $supp(\mathbb{P}_j)$.

Finally, one can repeat the procedure on a grid of all possible choices for $Q \in \mathcal{Q}_\lambda$ to determine the worst case measure Q, which is computationally intensive.

Remark 5.3.3.
The previous pseudo code becomes much simpler when working with equivalent measures only. In this case the simulation can be done with respect to $\mathbb{P}_1$ and the fixing of an equivalent martingale measure can be done after the simulation of the paths, which allows to use the same simulation for all choices of Q and drastically reduces the computational effort. The necessity to restrict to the support of the measures when determining success probabilities leads to further problems, which we discuss in the next chapter, see Problem 6.1.1 and Problem 6.1.2.

Example 5.3.4.
Looking again at Example 5.3.2 we can validate the procedure of the Monte Carlo simulation. Again we set $S_0 = 100$, $C(S_T) = (S_T - K)^+$ with $K = 100$ and take Black-Scholes models $\mathbb{P}_1$ and $\mathbb{P}_2$ with parameters $\mu_1 = 0.1$, $\mu_2 = 0.5$ and $\sigma^2 = 0.25$.

In this case we can use the fact that in Black-Scholes models the Radon-Nikodym derivative, which changes the drift term from μ_1 under $\mathbb{P}_1$ to μ_2 under $\mathbb{P}_2$ is a well known as

$$\frac{d\mathbb{P}_2}{d\mathbb{P}_1} = \exp\left(-\frac{\mu_1 - \mu_2}{\sigma} W_t - \frac{1}{2}\left(\frac{\mu_2 - \mu_1}{\sigma}\right)^2 t\right),$$

and from the drift r under Q to μ_1 under $\mathbb{P}_1$ as

$$\frac{dQ}{d\mathbb{P}_1} = \exp\left(-\frac{r - \mu_2}{\sigma} W_t - \frac{1}{2}\left(\frac{r - \mu_1}{\sigma}\right)^2 t\right).$$

Note that these two Radon-Nikodym derivatives are sufficient to compute every other necessary Radon-Nikodym derivative and thus we do not need to simulate an additional Brownian motion $W^{\mathbb{P}_2}$ under $\mathbb{P}_2$ to find $\frac{dQ}{d\mathbb{P}_2}$, which is extremely useful since we already set $W^{\mathbb{P}_1}$ as a standard Brownian motion.

Following the presented algorithm, we obtain the same success probabilities as in Figure 5.5.

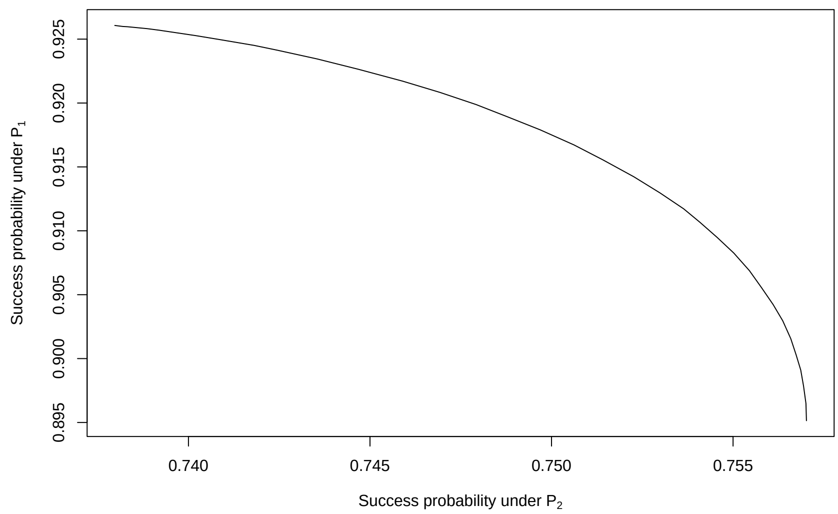

Figure 5.7.: Optimal success probabilities in robust Black-Scholes model using Monte Carlo simulation.

Chapter 6.

Analysis of the Skorokhod framework

As already mentioned in the examples, using the Skorokhod space as a sample space, while common in the literature, presents significant challenges, particularly in the context of robust market modeling. These issues become apparent when comparing theoretical frameworks with a real-world perspective. An alternative modeling approach can address these inconsistencies, providing a more coherent and reliable foundation for robust financial market analysis.

6.1. Issues of the Skorokhod space framework

In the previous examples, our focus was on markets that are equivalent. It remains to examine how we can deal with markets that are not equivalent. In mathematical terms: For the rest of this section we will assume a set $\mathcal{P}$ of singular measures. Let $\Omega_i \subseteq \Omega$ be the support of $\mathbb{P}_i$ for each measure $\mathbb{P}_i \in \mathcal{P}$. Due to singularity there is

$$\mathbb{P}_i(\Omega_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

We will see that this setup leads to unexpected results from a real-world perspective, as the following examples show.

Problem 6.1.1.

A major limitation of considering non-equivalent market measures is that the success probabilities of optimal hedging strategies under different measures do not align with the expected outcomes in real-world scenarios. Considering two very similar measures $\mathbb{P}_1 \perp \mathbb{P}_2$ (such as two Black-Scholes markets with slightly different volatility σ), the optimal success set $A_1 = \{\frac{d\mathbb{P}_1}{dQ_1} > const\ C\} \subseteq \Omega_1$ for $\mathbb{P}_1$ does have a success probability of $\mathbb{P}_2(A_1) = 0$, where we would expect a very similar success probability as the models are almost the same. These issues are caused by the pathwise construction.

Problem 6.1.2.
A second problem when working with singular market measures occurs when constructing robust optimal hedging strategies. For many examples of singular sets of market models, e.g. most uncertain variance models $\mathcal{P}$, where each $\mathbb{P} \in \mathcal{P}$ has a different but deterministic variance $(v_t)_{t\in[0,T]}$, there is $\Omega_i \in \mathcal{F}_{0+}$ for the respective supports. This leads to the problem that optimal hedging reduces to simply choosing the "correct" market measure, which is possible since the assumption $\Omega_i \in \mathcal{F}_{0+}$ allows to almost instantly reject every but one market measure, and setting up the corresponding simple optimal hedging strategy. In mathematical terms, we get the optimal hedging strategy

$$H_t = \sum_{i=1}^{n} H_t^i \mathbb{1}_{\Omega_i},$$

where H^i is the optimal hedging strategy regarding $\mathbb{P}_i$, see Corollary 6.1.3. In other words, the whole idea of robust modelling is entirely negated in this case.

It remains to prove the assertion of Problem 6.1.2, that optimal hedging strategies can be reduced to simply choosing one of the simple optimal hedging strategies.

Corollary 6.1.3.
If the measures in $\mathcal{P}$ are singular and for every corresponding support Ω_i there is $\Omega_i \in \mathcal{F}_{0+}$, then the robust optimal hedging strategy reduces to the strategy

$$H_t = \sum_{i=1}^{n} H_t^i \mathbb{1}_{\Omega_i} \tag{6.1.1}$$

with H_t^i being the optimal strategy under $\mathbb{P}_i$ and $\mathcal{Q}^{\mathbb{P}_i}$ corresponding to the capital constraint $\tilde{V}_0$ for every $i \in \mathbb{N}$.
Note that this is equivalent to the following statement: For any weights λ there exist a worst case measure $Q \in \mathcal{Q}$ such that it holds that $V_t^H = \mathbb{1}_A C$ with

$$A = \left\{ \sum_{i=1}^{n} \lambda_i \frac{d\mathbb{P}_i}{dQ} > kC \right\} \tag{6.1.2}$$

and with H of the form (6.1.1).

Proof. We will show the above statement by constructing the worst case measure and constant of set (6.1.2) explicitly depending on the worst case measures in the single market cases. Optimality in every single market case then ensures optimality of the constructed set A.

Let λ be an arbitrary but fixed vector that determines the weights of the convex combination $\mathbb{P}_\lambda$. For any $Q \in \mathcal{Q}$ there are weights $\tilde{\lambda}$ with

$$Q = \sum_{i=1}^{n} \tilde{\lambda}_i Q_i, \quad Q_i \in \mathcal{Q}^{\mathbb{P}_i}.$$

Due to Theorem 4.1.3, the optimal success set is of the form

$$A = \left\{ \frac{d\mathbb{P}_\lambda}{dQ} > kC \right\} = \left\{ \sum_{i=1}^{n} \frac{\lambda_i d\mathbb{P}_i}{dQ} > kC \right\}.$$

We can use that $\mathbb{P}_i \perp Q_j$ for all $i \in \mathbb{N}$ with $i \neq j$, to observe that

$$\frac{d\mathbb{P}_\lambda}{dQ} = \frac{\sum\limits_{i=1}^{n} \lambda_i d\mathbb{P}_i}{\sum\limits_{j=1}^{n} \tilde{\lambda}_j Q_j} = \sum_{i=1}^{n} \frac{\lambda_i d\mathbb{P}_i}{\tilde{\lambda}_i dQ_i}.$$

For any measure $\mathbb{P}_i$, we get success probabilities

$$\mathbb{P}_i(A) = \mathbb{P}_i\left(\frac{\lambda_i}{\tilde{\lambda}_i} \frac{d\mathbb{P}_i}{dQ_i} > kC \right).$$

Now let k_i be the constant in the optimal set $\{\frac{d\mathbb{P}_i}{dQ_i} > k_i\ C\}$ for any measure $\mathbb{P}_i$. For any $i \in \mathbb{N}$ there exists a $\tilde{\lambda}_i$ such that

$$\left\{ \frac{\lambda_i}{\tilde{\lambda}_i} \frac{d\mathbb{P}_i}{dQ_i} > C \right\} = \left\{ \frac{d\mathbb{P}_i}{dQ_i} > k_i\ C \right\}.$$

Since $k_i > 0$ it also holds that $\tilde{\lambda}_i > 0$. In a final step we can choose the constant as

$$k = \frac{1}{\sum\limits_{i=1}^{n} \tilde{\lambda}_i}.$$

Now $\lambda_i^* = k \cdot \tilde{\lambda}_i$ are the weights that construct the worst case measure $Q^* = \sum_{i=1}^{n} \lambda_i^* Q_i$ that leads to the optimal success set presented in the corollary. After all, if there were any better optimal success sets, this would contradict the optimality under single market measures. Finally, it remains to remark that $\Omega_i \in \mathcal{F}_{0+}$ for any $i \in \mathbb{N}$ keeps the hedging strategy well defined and the price for this claim is exactly $\tilde{V}_0$ for any $\mathbb{P}_i$. □

6.2. Push-Forward measures

The above problems 6.1.1 and 6.1.2 arise because of the pathwise construction, i.e. using the Skorokhod space as the underlying sample space Ω. Modelling the financial market in a different way could solve these problems. For path-independent options, using push-forward measures of S_T as an alternative way of modelling the financial market has many advantages. First, this approach solves the problem of non-equivalent measures since the push-forward measures of most models have support $\mathbb{R}_+$ and are therefore equivalent. Second, this approach does not have the problem of potential $\mathcal{F}_{0+}$-measurability, i.e. the impossibility of actually modelling robustness in a real-world sense. Third, we can still use the stochastic differential equation and Radon-Nikodym derivatives known from the Skorokhod approach, since the push-forward version can be derived analytically or can be obtained via simulation.

However, for a given market model, it is unclear a priori whether the two approaches lead to the same results. The following section shows that, for very simple models at least (models satisfying Assumption 6.2.5), both approaches are consistent. However, there are simple counter examples as well. Using the Skorokhod space will always lead to a less risky optimal hedging strategy.

Remark 6.2.1.
Many of the most commonly traded options are path-independent, i.e. they can be written as $C = f(S_T)$. For the sake of simpler notation, we will assume C to be path independent for the most of this section. The approach of using the push-forward measures is viable for a much larger amount of options.

In particular, the following results hold for any option $C : \Omega \to \mathbb{R}_+$, which can be written as $C(\omega) = f(X(\omega))$ with a continuous function $X : \Omega \to \mathbb{R}_+$ and a function $f : \mathbb{R}_+ \to \mathbb{R}_+$. In this case X defines the push-forward measures and f represents the payoff. For example, one can think of $X(\omega) = \max_{t\in[0,T]} S_t(\omega)$ for lookback options or $X(\omega) = \int_0^T S_t(\omega)dt$ for Asian options.

6.2.1. Theoretical results

Before proceeding with the rigorous mathematical notation and argumentation, we begin with a short heuristic and intuitive introduction to motivate the results from a more practical perspective.

Remark 6.2.2.
From a more practical point of view, the choice of the sample space also describes the possibilities to define a hedging strategy. If one chooses the Skorokhod space as the sample space, one can decide on a path-by-path basis whether to hedge or not, i.e. one can react to jumps or to other characteristics of the path. The less detailed the sample space is chosen, for example if only the distribution of the final value S_T on $\mathbb{R}_+$ is modelled, the less detailed the hedging strategy can be defined.

From a heuristic point of view, it should therefore be clear that the less detailed the sample space is, the worse the optimal hedging strategy will be. Therefore, even if it is impossible to actually implement a hedging strategy as detailed as the Skorokhod approach allows, one still obtains an upper bound on the success probabilities of hedging strategies and can thus decide whether a more detailed modelling of financial markets might be useful or not.

From a more mathematical point of view, every hedging strategy regarding push-forward measures is represented as a measurable subset of the positive real numbers. Furthermore each of these hedging strategies, denoted as $B \subseteq \mathbb{R}_+$, can of course be represented in the Skorohkod approach as $S_T^{-1}(B) \subseteq \Omega$, the set of all paths such that the final value of the underlying is in B.

The previous remark motivates that there is a close relationship between the optimal strategies under the original measures and the optimal strategies under the push-forward measures. In the following part of this section we will examine the connection between the set $B = \{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > kC^{S_T}\}$, where $\mathbb{P}^{S_T}$, Q^{S_T} and C^{S_T} are defined according to Definition 6.2.3, which represents the optimal solution under push-forward measures, and the set $A = \{\frac{d\mathbb{P}}{dQ} > kC\}$ representing original optimal solutions. If $S_T(A) = B$, i.e. if all paths ω in A lead to a final value ω_T lying in B, then both approaches lead to identical results.

Definition 6.2.3.
Push-forward measures regarding a random variable $X : \Omega \to \mathbb{R}_+$ are supposed to be understood in the usual way, i.e. for any $B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}^X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B)).$$

Additionally, for any function $f : \Omega \to \mathbb{R}_+$, we define the push-forward function $f^X : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$f^X(X(\omega)) = f(\omega) \quad \forall \omega \in \Omega.$$

Note here that if $X(\Omega) = \mathbb{R}_+$, the function f^X is defined uniquely.

First, it can be shown that the Radon-Nikodym derivatives of push-forward measures can be described as a conditional expectation of the original Radon-Nikodym derivatives.

Lemma 6.2.4.
Let $(\Omega, \mathcal{F})$ be a measurable space with equivalent measures $\mathbb{P}$ and Q. Let $X : \Omega \to \mathbb{R}$ be a random variable on Ω. For the Radon-Nikodym derivatives of push-forward measures $\mathbb{P}^X$ and Q^X generated by X it holds that

$$\int\limits_{X^{-1}(B)} \frac{d\mathbb{P}}{dQ} dQ = \int\limits_{X^{-1}(B)} \left(\frac{d\mathbb{P}^X}{dQ^X} \circ X \right) dQ.$$

Furthermore, this implies that

$$\frac{d\mathbb{P}^X}{dQ^X} = \mathbb{E}^Q \Big[\frac{d\mathbb{P}}{dQ} \Big| X \Big].$$

Proof. Similar to Li and Babu [42], Theorem 1.21, for any Borel set B we have:

$$\begin{aligned} \int\limits_{X^{-1}(B)} \frac{d\mathbb{P}^X}{dQ^X} \circ X dQ &= \int\limits_{\Omega} (\mathbb{1}_B \circ X)(\frac{d\mathbb{P}^X}{dQ^X} \circ X) dQ = \int\limits_{\Omega} (\mathbb{1}_B \frac{d\mathbb{P}^X}{dQ^X}) \circ X dQ \\ &= \int\limits_{\mathbb{R}_+} \mathbb{1}_B \frac{d\mathbb{P}^X}{dQ^X} dQ^X = \mathbb{P}^X(B) = \mathbb{P}(X^{-1}(B)) \\ &= \int\limits_{X^{-1}(B)} \frac{d\mathbb{P}}{dQ} dQ. \end{aligned}$$

□

It is important to note that in general this does not imply $\frac{d\mathbb{P}}{dQ}(\omega) = \frac{d\mathbb{P}^X}{dQ^X}(X(\omega))$.

It can be shown that, under the rather strict assumption 6.2.5, both approaches lead to exactly the same optimal hedging strategies and success probabilities.

Assumption 6.2.5.
Let $X : \Omega \to \mathbb{R}_+$ be a random variable defined on Ω and let $\mathbb{P}$ be a market measure with an equivalent martingale measure Q. We assume the implication:

$$X(\omega_1) = X(\omega_2) \Rightarrow \frac{d\mathbb{P}}{dQ}(\omega_1) = \frac{d\mathbb{P}}{dQ}(\omega_2).$$

In other words, we assume that the random variable X contains enough information to infer the equality of the change of measure from the equality of X, which depends not only on the choice of X, but also on the choice of the market measure.

Lemma 6.2.6.
If the Assumption 6.2.5 holds, then it holds that

$$\frac{d\mathbb{P}}{dQ}(\omega) = \frac{d\mathbb{P}^X}{dQ^X}(X(\omega)).$$

Proof. With Assumption 6.2.5 and Lemma 6.2.4 we get

$$\frac{d\mathbb{P}^X}{dQ^X}(X(\omega)) = \mathbb{E}^Q\left[\frac{d\mathbb{P}}{dQ}\Big|X = X(\omega)\right] = \frac{d\mathbb{P}}{dQ}(\omega).$$

The second equation follows by Assumption 6.2.5, because in this case, conditional on X, the Radon-Nikodym derivative becomes constant. □

Before formulating the main results regarding optimal hedging strategies using push-forward measures, we should recall the assumptions used for the robust version of the Neyman-Pearson Lemma 3.3.14. Assumption IV in 3.3.3 requires that the set of physical measures is closed and convex, while the set of pricing measures is assumed to be measure-convex. It turns out, as the next two lemmas show, that these properties transfer to the push-forward measures.

Lemma 6.2.7.
Let $\mathcal{P}$ be a closed set, regarding Prokhorov metric, of measures on $(\Omega, \mathcal{F})$ and let $X : \Omega \to \mathbb{R}_+$ be a continuous function on Ω. Then $\tilde{\mathcal{P}} = \{\mathbb{P}^X : \mathbb{P} \in \mathcal{P}\}$ is again a closed set of measures with respect to the Prokhorov metric.

Proof. Let $(\tilde{\mathbb{P}}_n)_{n\in\mathbb{N}} \subseteq \tilde{\mathcal{P}}$ be an arbitrary Cauchy sequence. It suffices to show that this sequence is convergent with limit in $\tilde{\mathcal{P}}$ to conclude that $\tilde{\mathcal{P}}$ is closed. Note that we again endow the set of probability measures on $\mathbb{R}_+$ with the Prokhorov metric equivalent to the set of probability measures on $(\Omega, \mathcal{B}(\Omega))$. Since $(\tilde{\mathbb{P}}_n)_{n\in\mathbb{N}}$ is supposed to be Cauchy we know that $d_P(\tilde{\mathbb{P}}_n, \tilde{\mathbb{P}}_{n+1}) \to 0$. Recall that the Prokhorov metric is defined as follows:

$$\begin{aligned} d_P(\tilde{\mathbb{P}}_n, \tilde{\mathbb{P}}_{n+1}) = \max\{ &\inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \tilde{\mathbb{P}}_n(B) \leq \tilde{\mathbb{P}}_{n+1}(B_\epsilon) + \epsilon\}, \\ &\inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \tilde{\mathbb{P}}_{n+1}(B) \leq \tilde{\mathbb{P}}_n(B_\epsilon) + \epsilon\}\}, \end{aligned}$$

where $B_\epsilon = \{x \in \mathbb{R}_+ : d(x, b) < \epsilon \text{ for a } b \in B\}$. Since $(\tilde{\mathbb{P}}_n)_{n\in\mathbb{N}} \subseteq \tilde{\mathcal{P}}$ there exists a sequence $(\mathbb{P}_n)_{n\in\mathbb{N}} \subseteq \mathcal{P}$ with $\tilde{\mathbb{P}}_n = \mathbb{P}_n^X$. Note that this sequence is not unique and it is not yet clear whether it is convergent. If we can show that this sequence converges, we can conclude that $\tilde{\mathcal{P}}$ must be closed, because

$$\lim_{n\to\infty} \tilde{\mathbb{P}}_n = (\lim_{n\to\infty} \mathbb{P}_n)^X \in \tilde{\mathcal{P}}.$$

As the functional X is continuous we know that $X^{-1}(B_\epsilon) \subseteq X^{-1}(B)_\epsilon$, and thus $\mathbb{P}_n(X^{-1}(B_\epsilon)) \leq \mathbb{P}_n(X^{-1}(B)_\epsilon)$. With

$$\begin{aligned}
&\max\{\inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \mathbb{P}_n(X^{-1}(B)) \leq \mathbb{P}_{n+1}(X^{-1}(B)_\epsilon) + \epsilon\},\\
&\qquad \inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \mathbb{P}_{n+1}(X^{-1}(B)) \leq \mathbb{P}_n(X^{-1}(B)_\epsilon) + \epsilon\}\}\\
&\leq \max\{\inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \mathbb{P}_n(X^{-1}(B)) \leq \mathbb{P}_{n+1}(X^{-1}(B_\epsilon)) + \epsilon\},\\
&\qquad \inf\{\epsilon > 0 | \forall B \in \mathcal{B}(\mathbb{R}_+) : \mathbb{P}_{n+1}(X^{-1}(B)) \leq \mathbb{P}_n(X^{-1}(B_\epsilon)) + \epsilon\}\}\\
&= d_P(\tilde{\mathbb{P}}_n, \tilde{\mathbb{P}}_{n+1}) \to 0,
\end{aligned}$$

at least on the sub-σ-algebra $X^{-1}(\mathcal{B}(\mathbb{R}_+))$, the sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is Cauchy as well. Since the metric space of all probability measures on $(\Omega, \mathcal{B}(\Omega))$ endowed with the Prokhorov metric is complete, because the Skorokhod space is complete and separable, this is sufficient for $(\mathbb{P}_n)_{n\in\mathbb{N}}$ to be convergent at least for preimages of X. Finally, we can see that each of the possible sequences $(\mathbb{P}_n)_{n\in\mathbb{N}}$ converges to the same limit on $X^{-1}(\mathcal{B}(\mathbb{R}_+))$ and thus for any measurable set that has an impact on the corresponding push-forward measure $\mathbb{P}^X$ we already have convergence. For any other sets $A \in \mathcal{B}(\Omega) \setminus X^{-1}(\mathcal{B}(\mathbb{R}_+))$, that do not affect $\mathbb{P}^X$, we can find a sequence with $d_P(\mathbb{P}_n, \mathbb{P}_{n+1}) \leq 2\epsilon$. Thus, we get a convergent sequence with $\lim_{n\to\infty} \mathbb{P}_n \in \mathcal{P}$, as $\mathcal{P}$ is assumed to be closed. For any $B \in \mathcal{B}(\mathbb{R}_+)$ we have

$$\lim_{n\to\infty} \tilde{\mathbb{P}}(B) = \lim_{n\to\infty} \mathbb{P}(X^{-1}(B)),$$

which means $\lim \tilde{\mathbb{P}}_n \in \tilde{\mathcal{P}}$. So every Cauchy series in $\tilde{\mathcal{P}}$ is convergent and $\tilde{\mathcal{P}}$ is closed. □

The measure-convexity of pricing measures should also be preserved by similar arguments, as shown in the next Lemma 6.2.8.

Lemma 6.2.8.
Let $\mathcal{Q}$ be a measure-convex set of measures on $(\Omega, \mathcal{F})$ and let $X : \Omega \to \mathbb{R}_+$ be a function on Ω. Then $\tilde{\mathcal{Q}} = \{Q^X : Q \in \mathcal{Q}\}$ is again a measure-convex set of measures.

Proof. Let $\mathcal{Q}$ be a measure-convex set. Let $\tilde{\lambda}$ be a probability measure on $\tilde{\mathcal{Q}}$. We want to show that $\int_{\tilde{\mathcal{Q}}} \tilde{Q} d\tilde{\lambda} \in \tilde{\mathcal{Q}}$. Now, per definition of $\tilde{\mathcal{Q}}$, there must be a probability measure λ on $\mathcal{Q}$ with

$$\tilde{Q}^* := \int_{\tilde{\mathcal{Q}}} \tilde{Q} d\tilde{\lambda} = \int_{\mathcal{Q}} Q^X d\lambda.$$

Since $\mathcal{Q}$ is measure-convex we know that $Q^* := \int_{\mathcal{Q}} Q d\lambda \in \mathcal{Q}$. Now it remains to show that $Q^{*X} = \tilde{Q}^*$.

For any Borel set $B \in \mathcal{B}(\mathbb{R}_+)$ there is

$$Q^{*X}(B) = \left(\int_{\mathcal{Q}} Q d\lambda\right)^X (B) = \left(\int_{\mathcal{Q}} Q d\lambda\right)(X^{-1}(B)) = \int_{\mathcal{Q}} Q(X^{-1}(B)) d\lambda$$
$$= \int_{\mathcal{Q}} Q^X(B) d\lambda = \left(\int_{\mathcal{Q}} Q^X d\lambda\right)(B) = \tilde{Q}^*(B),$$

which proves that $Q^{*X} = \tilde{Q}^*$ and thus $\int_{\tilde{\mathcal{Q}}} \tilde{Q} d\tilde{\lambda} \in \tilde{\mathcal{Q}}$. □

Together, Lemma 6.2.7 and Lemma 6.2.8 show that the assumptions made to apply the robust Neyman-Pearson Lemma 3.3.4 do indeed carry over to the corresponding set of push-forward measures.

Theorem 6.2.9.
If $\mathcal{P}$ is a robust market model satisfying Assumptions 3.3.3, $\mathcal{Q}$ is the convex hull of corresponding equivalent martingale measures, and $X : \Omega \to \mathbb{R}_+$ is a continuous function, then the set $\tilde{\mathcal{P}} = \{\mathbb{P}^X : \mathbb{P} \in \mathcal{P}\}$ satisfies Assumptions 3.3.3 and the set $\tilde{\mathcal{Q}} = \{Q^X : Q \in \mathcal{Q}\}$ is measure-convex. Therefore, the theory of optimal hedging strategies still applies to corresponding push-forward measures.

Proof. This theorem follows directly from Lemma 6.2.7 and 6.2.8. Note that convexity of $\tilde{\mathcal{P}}$ follows from straightforward construction of the corresponding measures in $\mathcal{P}$, analogous to the proof of Lemma 6.2.8. □

The choice of X depends on the options that is supposed to be hedged. While $X = S_T$ corresponds to the large amount of options, who are path independent, the case $X = \max\{S_t : t \in (0, T])\}$ corresponds to lookback options and $X = \frac{1}{T}\int S_t dt$ can be used to model Asian options.

Finally, we can come to the conclusion that if X satisfies the Assumption 6.2.5, both approaches, i.e. construction with Skorokhod space and push-forward measures, lead to the same optimal hedging strategies, success sets and success probabilities.

Theorem 6.2.10.
Let $(\Omega, \mathbb{F})$ be a probability space with Ω the Skorokhod space and let the market measure $\mathbb{P}$ and the function $X : \Omega \to \mathbb{R}_+$ satisfy Assumption 6.2.5 such that $X(\Omega) = \mathbb{R}_+$. Let $\mathbb{P}^X$ and Q^X be push-forward measures of market models $\mathbb{P} \sim Q$. Let $C : \Omega \to \mathbb{R}_+$ be a $\mathcal{F}_T$-measurable claim with $C^X : \mathbb{R}_+ \to \mathbb{R}_+$ such that $C(\omega) = C^X(X(\omega))$. Then

a) $X\Big(\Big\{\frac{d\mathbb{P}}{dQ} > kC\Big\}\Big) = \Big\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\Big\}$

b) $\mathbb{E}^Q\Big[C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > kC\}}\Big] = \mathbb{E}^{Q^X}\Big[C^X\mathbb{1}_{\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\}}\Big]$

c) $\mathbb{P}\Big(\frac{d\mathbb{P}}{dQ} > kC\Big) = \mathbb{P}^X\Big(\frac{d\mathbb{P}^X}{dQ^X} > kC^X\Big)$

This implies that the constant $k \in \mathbb{R}$, that is chosen such that the capital constraint is satisfied, remains unchanged when using the push-forward approach.

Proof.

a) The first statement follows using the Assumption 6.2.5:

$$\begin{aligned} X\Big(\Big\{\frac{d\mathbb{P}}{dQ} > kC\Big\}\Big) &= \Big\{X(\omega) : \frac{d\mathbb{P}}{dQ}(\omega) > kC(\omega)\Big\} \\ &= \Big\{X(\omega) : \frac{d\mathbb{P}^X}{dQ^X}(X(\omega)) > kC^X(X(\omega))\Big\} \\ &\supseteq \Big\{x \in \mathbb{R}_+ : \frac{d\mathbb{P}^X}{dQ^X}(x) > kC^X(x)\Big\} \\ &= \Big\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\Big\}, \end{aligned}$$

with equality in the third step if $X(\Omega) = \mathbb{R}_+$.

b) The second statement is a consequence of statement a).

$$\begin{aligned} \mathbb{E}^{Q^X}[C^X\mathbb{1}_{\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\}}] &= \int_{\mathbb{R}_+} C^X\mathbb{1}_{\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\}} dQ^X \\ &= \int_\Omega (C^X \circ X)(\mathbb{1}_{\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\}} \circ X) dQ \\ &= \int_\Omega C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > kC\}} dQ \\ &= \mathbb{E}^Q[C\mathbb{1}_{\{\frac{d\mathbb{P}}{dQ} > kC\}}]. \end{aligned}$$

c) The last statement again follows directly using the statement a). We obtain

$$\begin{aligned} \mathbb{P}^X\Big(\frac{d\mathbb{P}^X}{dQ^X} > kC^X\Big) &= \mathbb{P}\Big(X \in \Big\{\frac{d\mathbb{P}^X}{dQ^X} > kC^X\Big\}\Big) \\ &= \mathbb{P}\Big(X \in X\Big(\Big\{\frac{d\mathbb{P}}{dQ} > kC\Big\}\Big)\Big) \\ &= \mathbb{P}\Big(\Big\{\frac{d\mathbb{P}}{dQ} > kC\Big\}\Big). \end{aligned}$$

□

Corollary 6.2.11.
In the Black-Scholes model the function $S_T : \Omega \to \mathbb{R}_+$ satisfies the assumption 6.2.5.

Proof. As shown in Proposition 5.2.1 we can see that $\frac{d\mathbb{P}}{dQ}$ depends only on the underlying itself:

$$\begin{aligned}\frac{d\mathbb{P}}{dQ} &= \exp\left(\frac{\mu}{\sigma}W_T + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2 T\right)\\ &= \exp\left(\frac{\mu}{\sigma^2}(\sigma W_T + \frac{\mu}{2}T)\right)\\ &= S_0^{\frac{\mu}{\sigma^2}} \exp\left(\sigma W_T + (\mu - \frac{\sigma^2}{2})T\right)^{\frac{\mu}{\sigma^2}} S_0^{-\frac{\mu}{\sigma^2}} \exp\left(\frac{\mu^2}{2\sigma^2}T - \frac{\mu^2}{\sigma^2}T + \mu T\right)\\ &= S_T^{\frac{\mu}{\sigma^2}} \cdot c.\end{aligned}$$

□

Example 6.2.12.
Even in comparably simple complete markets, the Assumption 6.2.5 is not necessarily fulfilled for the function S_T: Consider a symmetric, two period binomial model with different physical up and down probabilities for each period. Let p_1 be the up-probability for the first period and p_2 be the up-probability for the second period. We assume $p_1 \neq p_2$. The martingale up-probability q for each period is constant due to symmetry. We can then look at the paths $\omega_1 = (up, down)$ and $\omega_2 = (down, up)$ to see that

$$S_{t_2}(\omega_1) = S_{t_2}(\omega_2)$$

and

$$\frac{d\mathbb{P}}{dQ}(\omega_1) = \frac{p_1}{q}\frac{1-p_2}{1-q} \neq \frac{1-p_1}{1-q}\frac{p_2}{q} = \frac{d\mathbb{P}}{dQ}(\omega_2)$$

The previous example also implies that, in general, this assumption should not be satisfied in all models with non-constant drift. One can even show that the assumption 6.2.5 is necessary for both approaches to coincide for any capital constraint $\tilde{V}_0$ (see Corollary 6.2.13).

From a more mathematical point of view, we can show that without Assumption 6.2.5, there must be a capital constraint $\tilde{V}_0$ such that there is a set $A \subseteq \Omega$ of paths that are not hedged under optimal hedging in Skorokhod modelling, while $A \subseteq S_T^{-1}(B)$ for a set $B \subseteq \mathbb{R}_+$ of realisations that are hedged under optimal hedging in push-forward modelling.

Corollary 6.2.13.
If the Assumption 6.2.5 does not hold, there exists $x \in \mathbb{R}_+$ and a set $A \subseteq S_T^{-1}(x)$ of non-null probability with

$$\frac{d\mathbb{P}}{dQ}(\omega) > \frac{d\mathbb{P}^{S_T}}{dQ^{S_T}}(S_T(\omega)) \;\; \forall \omega \in A.$$

Since $C(\omega) = C^{S_T}(x)$ for every $\omega \in A$, there will be a capital constraint $\tilde{V}_0$, i.e. a corresponding constant k, such that $A \subseteq \{\frac{d\mathbb{P}}{dQ} > kC\}$ but $x \notin \{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > kC^{S_T}\}$.

Proof. This is a direct consequence of the representation $\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} = \mathbb{E}^Q[\frac{d\mathbb{P}}{dQ}|S_T]$. If $\frac{d\mathbb{P}}{dQ}$ is not constant on the set $S_T^{-1}(x)$ there must be a set $A \subseteq S_T^{-1}(x)$ with

$$\mathbb{E}^Q\left[\frac{d\mathbb{P}}{dQ}\middle|A\right] > \mathbb{E}^Q\left[\frac{d\mathbb{P}}{dQ}\middle|S_T = x\right] > \mathbb{E}^Q\left[\frac{d\mathbb{P}}{dQ}\middle|S_T^{-1}(x) \setminus A\right].$$

Since for continuous measures $\mathbb{P}^{S_T}$ and Q^{S_T} the value k depends continuously on $\tilde{V}_0$, the second statement holds as well. □

The corollary does not imply that there must be a difference between the two approaches for every capital constraint, but it does imply that there is a specific capital constraint such that the success sets do not coincide.

In these cases, as already heuristically mentioned in remark 6.2.2, one can still show, without any additional assumptions, that the Skorokhod approach provides an upper bound on the optimal success probability under the push-forward measures approach. More specifically, the following lemma shows that any hedging strategy under push-forward measures can be represented using the Skorokhod modelling approach with the same amount of money.

Lemma 6.2.14.
Let $C : \Omega \to \mathbb{R}_+$ be a path independent option with corresponding option $C^{S_T} : \mathbb{R}_+ \to \mathbb{R}_+$ such that $C(\omega) = C^{S_T}(x)$ for every ω with $S_T(\omega) = x$ and let $B \subseteq \mathbb{R}_+$ define the success set of a hedging strategy. For every probability measure Q on the Skorokhod space $(\Omega, \mathcal{F})$ it holds that

$$\mathbb{E}^{Q^{S_T}}[C^{S_T}\mathbb{1}_B] = \mathbb{E}^Q[C\mathbb{1}_{S_T^{-1}(B)}].$$

Proof. For every $Q \in \mathcal{Q}^{\mathbb{P}}$ it holds that

$$\mathbb{E}^{Q^{S_T}}[C^{S_T}\mathbb{1}_B] = \int_{\Omega} C^{S_T}\mathbb{1}_B dQ^{S_T} = \int_{\mathbb{R}_+}(C^{S_T}\circ S_T)(\mathbb{1}_B\circ S_T)dQ = \int_{\mathbb{R}_+} C\mathbb{1}_{S_T^{-1}(B)}dQ$$
$$= \mathbb{E}^{Q}[C\mathbb{1}_{S_T^{-1}(B)}].$$

□

Theorem 6.2.15.
Let $\mathbb{P}$ be a market measure defined on the Skorokhod space for an underlying $(S_t)_{t\in[0,T]}$. For $\mathbb{P}$ and $Q \in \mathcal{Q}^{\mathbb{P}}$ let $\mathbb{P}^{S_T}$ and Q^{S_T} be push-forward measures. Let $C : \Omega \to \mathbb{R}_+$ be a $\mathcal{F}_T$-measurable claim with $C^{S_T} : \mathbb{R}_+ \to \mathbb{R}_+$ such that $C(\omega) = C^{S_T}(S_T(\omega))$. Then

$$\mathbb{P}\Big(\frac{d\mathbb{P}}{d\tilde{Q}} > kC\Big) \geq \mathbb{P}^{S_T}\Big(\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > k^{S_T}C^{S_T}\Big)$$

where $\tilde{Q} \in \mathcal{Q}^{\mathbb{P}}$ is the worst-case measure, while k and k^{S_T} are chosen such that $\sup_{Q\in\mathcal{Q}}\mathbb{E}^{Q}[C\mathbb{1}_{\{\frac{d\mathbb{P}}{d\tilde{Q}}>kC\}}] = \tilde{V}_0$ and $\sup_{Q\in\mathcal{Q}} E^{Q^{S_T}}[C^{S_T}\mathbb{1}_{\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}}>k^{S_T}C^{S_T}\}}] = \tilde{V}_0$.

Proof. For any Q^{S_T} the set $S_T^{-1}(\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > k^{S_T}C^{S_T}\}) \subseteq \Omega$ represents a hedging strategy under the Skorokhod approach. As shown in Lemma 6.2.14, prices do not change compared to the push-forward measure approach. This implies that also the suprema of all prices are equal, i.e.

$$\sup_{Q\in\mathcal{Q}}\mathbb{E}^{Q}[C\mathbb{1}_{S_T^{-1}(\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}}>k^{S_T}C^{S_T})}] = \sup_{Q\in\mathcal{Q}}\mathbb{E}^{Q^{S_T}}[C^{S_T}\mathbb{1}_{\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}}>k^{S_T}C^{S_T}\}}].$$

Finally, success probabilities must coincide due to the definition of push-forward measures, i.e.

$$\mathbb{P}\Big(S_T^{-1}\Big(\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > k^{S_T}C^{S_T}\Big)\Big) = \mathbb{P}^{S_T}\Big(\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > k^{S_T}C^{S_T}\Big)$$

This completes the proof, as there is a set implying a hedging strategy under the Skorokhod approach that is at least as good as the best hedging strategy under the push-forward approach. □

In summary, we have seen that the Skorokhod approach allows for a very detailed (pathwise) construction of hedging strategies, which is mostly too detailed for practical purposes. Nevertheless, the pathwise view can still reveal certain characteristics of the optimal success sets and hedging strategies.

6.2.2. Numerical implementation of the push-forward approach

From a numerical point of view, the implementation of the push-forward measure in the context of optimal hedging strategies is straightforward, using only Monte Carlo methods and utilising only the stochastic differential equation and the corresponding Radon-Nikodym derivative:

(a) Simulate n paths of S and $\frac{d\mathbb{P}}{dQ}$ under the physical measure $\mathbb{P}$ using the defining stochastic differential equation.

(b) Approximate the CDF of S_T under $\mathbb{P}$ as the empirical CDF of the simulated paths, which gives an approximation of $\mathbb{P}^{S_T}$.

(c) Determine the CDF of S_T under Q using

$$Q(S_T \leq k) = \mathbb{E}^Q[\mathbb{1}_{\{S_T \leq k\}}] = \mathbb{E}^{\mathbb{P}}\Big[\frac{dQ}{d\mathbb{P}}\mathbb{1}_{\{S_T \leq k\}}\Big],$$

which gives an approximation of Q^{S_T}.

(d) Determine the density functions $f^{\mathbb{P}}_{S_T}$ and $f^{Q}_{S_T}$, which can be done directly for $\mathbb{P}^{S_T}$ using kernel density estimation or similar methods. For Q^{S_T}, one can generate a random sample according to the CDF and then use kernel density estimation again.

(e) The optimal success set under this approach will be

$$B = \Big\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > kC^{S_T}\Big\} = \Big\{\frac{f^{\mathbb{P}}_{S_T}}{f^{Q}_{S_T}} > kC^{S_T}\Big\} \subseteq \mathbb{R}_+,$$

where $C^{S_T} : \mathbb{R}_+ \to \mathbb{R}_+$ is defined such that $C(\omega) = C^{S_T}(S_T(\omega))$ for every $\omega \in \Omega$. Remember that this is possible because C is assumed to be path independent.

(f) Find the value $k \in \mathbb{R}_+$, that ensures that the capital constraint is satisfied, i.e. that

$$\mathbb{E}^{Q^{S_T}}\Big[C^{S_T}\mathbb{1}_{\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > kC^{S_T}\}}\Big] = \tilde{V}_0.$$

For numerical approaches it can be useful to see that the constant k to be determined is bounded from above by $\frac{1}{\tilde{V}_0}$.

Corollary 6.2.16.
The constant k in the set $B = \Big\{\frac{d\mathbb{P}^{S_T}}{dQ^{S_T}} > kC^{S_T}\Big\}$ is bounded by $\frac{1}{\tilde{V}_0}$.

Proof. A straightforward calculation shows:

$$\begin{aligned}\tilde{V}_0 = \mathbb{E}^{Q^X}[\mathbb{1}_{\frac{f^{\mathbb{P}}}{f^Q} > kC^X} C] &= \int \mathbb{1}_{\{\frac{f^{\mathbb{P}}}{f^Q}(x) > kC^X(x)\}} C^X(x) f^Q(x) dx \\ &= \int \mathbb{1}_{\{\frac{1}{k} f^{\mathbb{P}}(x) > f^Q(x) C^X(x)\}} C^X(x) f^Q(x) dx \\ &\leq \frac{1}{k} \int \mathbb{1}_{\{\frac{1}{k} f^{\mathbb{P}}(x) > f^Q(x) C^X(x)\}} f^{\mathbb{P}} dx \\ &\leq \frac{1}{k} \int f^{\mathbb{P}} dx = \frac{1}{k}.\end{aligned}$$

□

6.3. Examples

To keep the notation and formulas clear, we again take a Black-Scholes model. In this case the previous approach proves especially useful for claims that are path dependent, as for example a fixed strike lookback option $C = (\max_{u \in [0,T]} S_u - K)^+$. Now, the previously useful property of the Black-Scholes model that we can represent the Radon-Nikodym derivative in terms of the final value S_T does not help find the optimal hedging strategy, because in the optimal success set

$$\left\{ \frac{d\mathbb{P}}{dQ} > k \left(\max_{u \in [0,T]} S_u - K \right)^+ \right\}$$

the right hand side does not depend solely on S_T.

However, using the reflection principle it is possible to determine the distribution of the maximum of a geometric Brownian motion. This allows to solve the problem of finding an optimal hedging strategy for C analytically, whereas solving the initial problem in the Skorokhod space is difficult, because we have to find the distribution of S_T conditional on M_T.

Lemma 6.3.1.
Let $(S_t)_{t \in [0,T]}$ be a geometric Brownian motion with parameter $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_+$. Let M_t be the maximum of S up to time $t \in [0,T]$, i.e. $M_t := \max_{u \in [0,t]} S_u$.

(a) The cumulative distribution of M_t is

$$\begin{aligned}F_{M_t}(x) = \mathbb{P}(M_t \leq x) = &\Phi\left(\frac{-(\mu - \sigma^2/2)t + \log(\frac{x}{S_0})}{\sigma\sqrt{t}} \right) \\ &- \left(\frac{S_0}{x} \right)^{1-2\mu/\sigma^2} \Phi\left(\frac{-(\mu - \sigma^2/2)t - \log(\frac{x}{S_0})}{\sigma\sqrt{t}} \right).\end{aligned}$$

(b) The density function of M_t is

$$f_{M_t}(x) = \frac{1}{\sigma x\sqrt{2\pi T}} \exp\left(-\frac{\left(-(\mu-\frac{\sigma^2}{2})T + \log\left(\frac{x}{S_0}\right)\right)^2}{2\sigma^2 T}\right)$$
$$+ \frac{1}{\sigma x\sqrt{2\pi T}} \left(\frac{S_0}{x}\right)^{1-\frac{2\mu}{\sigma^2}} \exp\left(-\frac{\left((\mu-\frac{\sigma^2}{2})T + \log\left(\frac{x}{S_0}\right)\right)^2}{2\sigma^2 T}\right)$$
$$+ \frac{1}{x}\left(1-\frac{2\mu}{\sigma^2}\right)\left(\frac{S_0}{x}\right)^{1-\frac{2\mu}{\sigma^2}} \Phi\left(\frac{-(\mu-\frac{\sigma^2}{2})T - \log\left(\frac{x}{S_0}\right)}{\sigma\sqrt{T}}\right). \tag{6.3.1}$$

Proof. The distribution of the running maximum of a geometric brownian motion is proved in Corollary 10.5 in chapter 10 of Privault [55]. The proof uses the reflection principle that determines the distribution of the running maximum of a standard Brownian motion

$$\mathbb{P}\left(\sup_{0\le s\le t} B_s \ge x\right) = 2\mathbb{P}(B_t \ge x).$$

Representing a geometric Brownian motion S as $S_t = \exp\left(\sigma B_t + (\mu - \frac{\sigma^2}{2})t\right)$ allows to calculate its distribution, as seen in detail in Privault [55]. ☐

Using the density function (6.3.1) it is possible to find the optimal success set analytically

$$A^{M_T} = \left\{x \in \mathbb{R}_+ : \frac{f_{M_T}(x;\mu)}{f_{M_T}(x;0)} > k(x-K)^+\right\}, \tag{6.3.2}$$

where $k \in \mathbb{R}$ is determined such that

$$\mathbb{E}^{Q^{M_T}}[\mathbb{1}_A(x-K)^+] = \int_K^\infty \mathbb{1}_{A^{M_T}(x)}(x-K)dQ^{M_T}(x)$$
$$= \int_K^\infty \mathbb{1}_{A^{M_T}(x)}(x-K)f_{M_T}(x;0)dx = \tilde{V}_0$$

Example 6.3.2.

Even though there is an analytic solution (6.3.2) it does not provide a good visualization, since the Radon-Nikodym derivative does not simplify. Still, this approach proves to be much simpler than using the Skorokhod space. Using again the example of a high volatility stock with $S_0 = 100, \mu = 0.5, \sigma = 0.5$ and $T = 1$ we can get the following density functions for M_T:

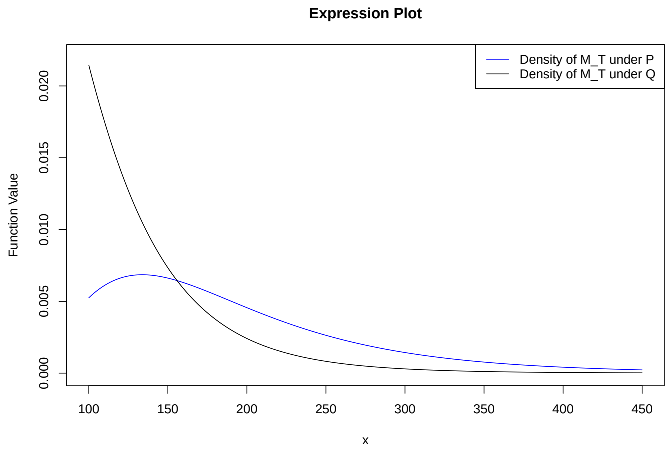

Figure 6.1.: Density functions of the maximum of a geometric Brownian motion up to $T = 1$ with $\sigma = 0,5$ and $\mu = 0.5$ under $\mathbb{P}$, $\mu = 0$ under Q.

Using a lookback option with fixed strike $K = 120$, which has a price of 30, calculating the the corresponding Radon Nikodym derivative, the constant k regarding the capital constraint $\tilde{V}_0 = 20$ and the respective transformed payout of the option leaves us with the following figure.

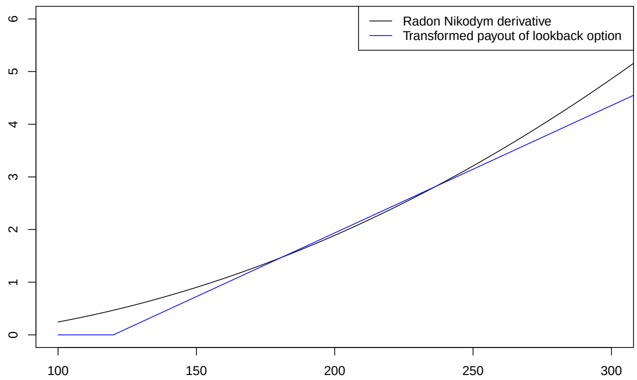

Figure 6.2.: Radon Nikodym derivative $\frac{d\mathbb{P}^{M_T}}{dQ^{M_T}}\Big|_{\mathcal{F}_T}$ of the maximum of a geometric Brownian motion and the transformed payout kC of a lookback option with strike $K = 120$ and capital constraint $\tilde{V}_0 = 20$.

Thus, an optimal hedge is the superhedging strategy $(H, \tilde{V}_0)$ to the option

$\tilde{C} = \mathbb{1}_{\{M_T \notin [181,236]\}} C$, which has a success probability $\mathbb{P}(V_T^{H,\tilde{V}_0} > C) = 74,71\%$ in this case.

Chapter 7.

Conclusion

In this thesis, we have investigated optimal hedging strategies under capital constraints in the context of robust market modelling. Our primary focus was on combining hedging methods, e.g. in Föllmer and Leukert [23], with a quasi-sure approach to account for model uncertainty, as introduced in Laurent and Martini [18]. Traditional hedging strategies often rely on assumptions about the market model, including the specification of a single reference probability measure. However, financial markets exhibit significant uncertainty, and relying on a single probability measure can result in misleading conclusions, as discussed in Knight [40]. By using the quasi-sure perspective of multiple valid market models, we have formulated a hedging strategy that remains valid across a wide range of market scenarios without being tied to a specific probabilistic assumption. To work in this quasi-sure set up, we have derived a Neyman-Pearson theorem that is more suitable for financial application, using the measure-convexity of the most relevant sets in financial mathematics.

One of the main results is the concept of an indifference curve of optimal hedging strategies, which illustrates the trade-off between different market models or scenarios. This indifference curve is not only constructed in a general way, but also explicitly determined for some examples. This perspective provides deeper insights into how different market assumptions and uncertainty influence optimal hedging decisions.

Another important result concerns continuity properties of the worst-case measure, that are usually difficult or even impossible to find analytically. We show that the worst-case measure has certain continuity properties across the indifferent curve of optimal hedging strategies. This allows to significantly reduce the numerical effort, since a global grid search across an arbitrarily large grid can be reduced to a local grid search.

In addition to these positive results, we have also presented issues when interpreting the mathematical results under a real-world perspective, especially when using non-equivalent or even singular measures. To overcome these problems, we presented an alternative to the common approach in the literature of using the Skorokhod space as a sample space. We have shown that many properties can be transferred to the push-forward measures of market measures without the problems we have regarding singular measures under the Skorokhod framework.

Chapter 8.

Outlook

The problem of optimal hedging can be divided into the representation problem, see Theorem 3.2.15, and the static optimisation problem, see Theorem 3.3.14. While the static optimisation problem is highlighted in great detail in this thesis the representation problem is reduced to an existence result. The field of stochastic analysis known as Malliavin calculus offers a way to examine the representation problem in more detail. Using these tools, allows to determine hedging strategies more explicitly, see for example Di Nunno et al. [19].

In addition, further research is needed to develop systematic methods to identify the worst-case measure. While we have proved continuity properties in this context, practical methods for determining the worst-case measure in general remain an open problem. A potential approach is inspired by mean-variance hedging, where the worst-case measure can be determined by minimising the relative entropy to the physical measure.

The mean-variance hedging problem, that minimises the expected quadratic hedging error, is typically solved by determining a minimum-variance measure, which plays a similar role to the worst-case measure in robust hedging. These very similar approaches motivate the idea of combining both approaches to optimise the success of hedging against the variance of the hedging error.

Finally, there is only little empirical research on the success of robust market modelling, see for example Guidolin [27] for an overview regarding Markov switching models. It may be of great interest to investigate robust market modelling and robust hedging strategies under empirical data, especially in view of financial crises.

So there are still many open questions in view of robust market modelling and hedging strategies under these models. Since financial markets will remain important and may become even more complex and unstable in future, this topic will continue to be of great interest.

Bibliography

[1] B. Acciaio et al. „A Model-Free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem“. eng. In: *Mathematical finance* 26.2 (2016), pp. 233–251.

[2] Erik M Alfsen. *Compact convex sets and boundary integrals*. Vol. 57. Springer Science & Business Media, 2012.

[3] Basel Committee on Banking Supervision. *Basel III: A global regulatory framework for more resilient banks and banking systems*. 2011.

[4] Erhan Bayraktar and Gu Wang. „Quantile Hedging in a semi-static market with model uncertainty“. eng. In: *Mathematical methods of operations research (Heidelberg, Germany)* 87.2 (2018), pp. 197–227.

[5] Sara Biagini et al. „Robust fundamental theorem for continuous processes“. In: *Mathematical Finance* 27.4 (2017), pp. 963–987.

[6] Patrick Billingsley. *Convergence of probability measures*. eng. Second edition. Wiley online library. 1999.

[7] Tomas Björk. *Arbitrage theory in continuous time*. eng. 1998.

[8] Fischer Black and Myron Scholes. „The Pricing of Options and Corporate Liabilities“. eng. In: *The Journal of political economy* 81.3 (1973), pp. 637–654.

[9] Bruno Bouchard and Marcel Nutz. „Arbitrage and duality in nondominated discrete-time models“. eng. In: *The Annals of applied probability* 25.2 (2015), pp. 823–859.

[10] Pierre Brémaud. *Point processes and queeues: martingale dynamics*. Vol. 50. Springer, 1981.

[11] Aleš Černý and Johannes Ruf. „Pure-jump semimartingales“. eng. In: *Bernoulli : official journal of the Bernoulli Society for Mathematical Statistics and Probability* 27.4 (2021).

[12] Gerald HL Cheang and Carl Chiarella. „A modern view on Merton’s jump-diffusion model“. In: *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J Elliott*. World Scientific, 2012, pp. 217–234.

[13] John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. „A Theory of the Term Structure of Interest Rates“. eng. In: *Econometrica* 53.2 (1985), pp. 385–407.

[14] Jakša Cvitanić and Ioannis Karatzas. „Generalized Neyman-Pearson lemma via convex duality“. In: *Bernoulli* (2001), pp. 79–97.

[15] Freddy Delbaen. „Representing Martingale Measures When Asset Prices Are Continuous And Bounded“. eng. In: *Mathematical finance* 2.2 (1992), pp. 107–130.

[16] Freddy Delbaen. *The Mathematics of Arbitrage*. eng. Springer Finance. 2006.

[17] Freddy Delbaen and Walter Schachermayer. „A general version of the fundamental theorem of asset pricing“. eng. In: *Mathematische annalen* 300.1 (1994), pp. 463–520.

[18] Laurent Denis and Claude Martini. „A Theoretical Framework for the Pricing of Contingent Claims in the Presence of Model Uncertainty“. eng. In: *The Annals of applied probability* 16.2 (2006), pp. 827–852.

[19] Giulia Di Nunno, Bernt K Øksendal, and Frank Proske. *Malliavin Calculus for Lévy Processes with Applications to Finance*. eng. Universitext. 2009.

[20] European Commission. „Commission delegated regulation (EU) 2015/35 of 10 October 2014 supplementing directive 2009/138/EC of the European parliament and of the council on the taking-up and pursuit of the business of insurance and reinsurance (Solvency II)“. In: *Official Journal of European Union* (2015).

[21] Hans Föllmer and Yuri M Kabanov. „Optional decomposition and Lagrange multipliers“. In: *Finance and Stochastics* 2.1 (1997), pp. 69–81.

[22] Hans Föllmer and Peter Leukert. „Efficient hedging: cost versus shortfall risk“. In: *Finance and Stochastics* 4.2 (2000), pp. 117–146.

[23] Hans Föllmer and Peter Leukert. „Quantile hedging“. In: *Finance and Stochastics* 3.3 (1999), pp. 251–273.

[24] Hans Föllmer and Alexander Schied. „Convex measures of risk and trading constraints“. eng. In: *Finance and stochastics* 6.4 (2002), pp. 429–447.

[25] Marie Gaudard and Donald Hadwin. „Sigma-Algebras on Spaces of Probability Measures“. eng. In: *Scandinavian journal of statistics* 16.2 (1989), pp. 169–175.

[26] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. eng. Stochastic Modelling and Applied Probability 53. 2003.

[27] Massimo Guidolin and David M. Drukker. *Markov Switching Models in Empirical Finance*. eng. Vol. 27 Part 2. aeco. Emerald Group Publishing Limited, 2011, pp. 1–86.

[28] J.Michael Harrison and David M Kreps. „Martingales and arbitrage in multiperiod securities markets“. eng. In: *Journal of economic theory* 20.3 (1979), pp. 381–408.

[29] Sheng-wu He, Jia-gang Wang, and Jia-an Yan. *Semimartingale theory and stochastic calculus*. Routledge, 2019.

[30] Steven L. Heston. „A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". eng. In: *The Review of financial studies* 6.2 (1993), pp. 327–343.

[31] David Hobson. „The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices". eng. In: *Paris-Princeton Lectures on Mathematical Finance 2010*. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 267–318.

[32] David G. Hobson. „Robust hedging of the lookback option". eng. In: *Finance and stochastics* 2.4 (1998), pp. 329–347.

[33] Zhaoxu Hou and Jan Obłój. „Robust pricing–hedging dualities in continuous time". eng. In: *Finance and stochastics* 22.3 (2018), pp. 511–567.

[34] Peter J Huber. *Robust statistics*. Vol. 523. John Wiley & Sons, 2004.

[35] Jean Jacod. *Calcul stochastique et problemes de martingales*. Vol. 714. Springer, 2006.

[36] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. eng. 1st ed. 1987. Grundlehren der mathematischen Wissenschaften, A Series of Comprehensive Studies in Mathematics, 288. 1987.

[37] Constantinos Kardaras. „Finitely additive probabilities and the fundamental theorem of asset pricing". In: *Contemporary quantitative finance*. Springer, 2010, pp. 19–34.

[38] Michael Kirch. „Efficient hedging in incomplete markets under model uncertainty". In: (2002).

[39] Peter E Kloeden. *Numerical solution of stochastic differential equations*. eng. 1999.

[40] Frank H. Knight. *Risk, Uncertainty and Profit*. Boston, MA: Houghton Mifflin Co, 1921.

[41] Thomas G Kurtz and Stewart N Ethier. „Markov Processes Characterization and Convergence". eng. In: *Markov Processes*. Vol. 282. United States: John Wiley & Sons, Incorporated, 1986.

[42] Bing Li and Gutti Jogesh Babu. *A graduate course on statistical inference*. Springer, 2019.

[43] R. Sh. Liptser and A. Shiryeav. *Statistics of random processes*. eng. 1st ed. 1977. Stochastic Modelling and Applied Probability, 5. 1977.

[44] George Lowther. *Almost Sure, Continuous Semimartingales*. `https://almostsuremath.com/2011/05/03/continuous-semimartingales/`. 2011.

[45] George Lowther. *Almost Sure, predictable FV processes.* `https://almostsuremath.com/2011/07/18/predictable-fv-processes/`. 2011.

[46] George Lowther. *Compactness of the set of densities of equivalent martingale measures.* MathOverflow. URL:https://mathoverflow.net/q/101784 (version: 2022-02-12). eprint: `https://mathoverflow.net/q/101784`.

[47] Reinhold Meise. *Introduction to functional analysis.* eng. 1997.

[48] Robert C Merton. „Option pricing when underlying stock returns are discontinuous". In: *Journal of financial economics* 3.1-2 (1976), pp. 125–144.

[49] Yumiharu Nakano. „Efficient hedging with coherent risk measure". eng. In: *Journal of mathematical analysis and applications* 293.1 (2004), pp. 345–354.

[50] Ariel Neufeld and Marcel Nutz. „Measurability of semimartingale characteristics with respect to the probability law". In: *Stochastic Processes and their Applications* 124.11 (2014), pp. 3819–3845.

[51] Marcel Nutz. „Robust superhedging with jumps and diffusion". In: *Stochastic Processes and their Applications* 125.12 (2015), pp. 4543–4555.

[52] Marcel Nutz and Ramon Van Handel. „Constructing sublinear expectations on path space". In: *Stochastic processes and their applications* 123.8 (2013), pp. 3100–3121.

[53] Nicolas Perkowski, Johannes Ruf, et al. „Supermartingales as Radon–Nikodym densities and related measure extensions". In: *Annals of Probability* 43.6 (2015), pp. 3133–3176.

[54] Robert R Phelps and Robert R. Phelps. *Lectures on Choquet's Theorem.* eng. Second Edition. Vol. 1757. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2001.

[55] Nicolas Privault. *Notes on stochastic finance.* Nanyang Technological University, Singapur, 2013.

[56] Philip E. Protter. *Stochastic integration and differential equations : a new approach.* eng. 1st ed. 1990. Stochastic Modelling and Applied Probability, 21. 1992.

[57] Daniel Revuz. *Continuous Martingales and Brownian Motion.* eng. 3rd ed. 1999. Grundlehren der mathematischen Wissenschaften, A Series of Comprehensive Studies in Mathematics, 293. 1999.

[58] Birgit Rudloff. „Coherent hedging in incomplete markets". In: *Quantitative Finance* 9.2 (2009), pp. 197–206.

[59] Wolfgang J. Runggaldier. „Chapter 5 - Jump-diffusion models". eng. In: *Handbook of Heavy Tailed Distributions in Finance.* North-Holland, 2003, pp. 169–209.

[60] Paul A Samuelson. „Rational Theory of Warrant Pricing". eng. In: *Industrial Management Review* 6.2 (1965), p. 13.

[61] Henry Scheffe. „A Useful Convergence Theorem for Probability Distributions". eng. In: *The Annals of mathematical statistics* 18.3 (1947), pp. 434–438.

[62] M. Schweizer. „A Guided Tour through Quadratic Hedging Approaches". eng. In: *Handbooks in Mathematical Finance.* Cambridge University Press, 2001, pp. 538–574.

[63] Daniel W Stroock. *Multidimensional Diffusion Processes.* eng. Classics in Mathematics / Grundlehren der mathematischen Wissenschaften, A Series of Comprehensive Studies in Mathematics. 1997.

[64] Swiss Financial Market Supervisory Authority (FINMA). *Swiss Solvency Test (SST).* `https://www.finma.ch/en/supervision/insurers/cross-sectoral-tools/swiss-solvency-test-sst/`.

[65] Onno Van Gaans. „Probability measures on metric spaces". In: *Lecture notes* (2003).

[66] Oldrich Vasicek. „An equilibrium characterization of the term structure". eng. In: *Journal of financial economics* 5.2 (1977), pp. 177–188.

[67] Gerhard Winkler. „Basic concepts from noncompact Choquet theory". In: *Choquet Order and Simplices: with Applications in Probabilistic Models* (2006), pp. 13–53.

[68] Gerhard Winkler. „Extreme Points of Moment Sets". eng. In: *Mathematics of operations research* 13.4 (1988), pp. 581–587.

[69] Hermann Witting. *Mathematische statistik I: parametrische verfahren bei festem stichprobenumfang.* Springer-Verlag, 2013.

Appendix A.

Martingale measures for jump diffusion model

In Corollary 5.2.9 we have presented a general form of the Radon-Nikodym derivative in the Merton-Jump model with a straightforward proof, for better readability and structure. Since this proof does not include a derivation of the Radon-Nikodym derivative, but only justifies the already known solution, a derivation of Radon-Nikodym derivatives with respect to equivalent martingale measures for jump diffusion models in general follows in this appendix.

Lemma A.0.1. *Novikov condition*
Let θ be an square integrable cadlag process such that for each $t \geq 0$ it holds that

$$\int_0^t \theta_s^2 ds < \infty,$$

then

$$Z_t = \exp\left(\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right)$$

defines a positive $\mathbb{P}$ martingale with $\mathbb{E}^{\mathbb{P}}[Z_t] = 1$.

Proof. For the simpler case of a deterministic θ we know that for every $t \in \mathbb{R}_+$ $\int_0^t \theta_s dW_s$ is a normally distributed random variable with mean zero and variance $\int_0^t \theta_s^2 ds$. Furthermore $\int_0^s \theta_s dW_s$ is $\mathcal{F}_s$ measurable while $\int_s^t \theta_s dW_s$ is independent of $\mathcal{F}_s$. This allows for the following calculation:

$$\begin{aligned}\mathbb{E}\left[\exp\left(\int_0^t \theta_u dW_u\right)\Big|\mathcal{F}_s\right] =&\mathbb{E}\left[\exp\left(\int_0^s \theta_u dW_u\right)\exp\left(\int_s^t \theta_u dW_u\right)\Big|\mathcal{F}_s\right]\\ =&\exp\left(\int_0^s \theta_u dW_u\right)\mathbb{E}\left[\exp\left(\int_s^t \theta_u dW_u\right)\right]\end{aligned}$$

$$= \exp\left(\int_0^s \theta_u dW_u\right) \exp\left(\frac{1}{2}\int_s^t \theta_u^2 du\right).$$

We can use this equation to conclude

$$\begin{aligned}
\mathbb{E}[Z_t|\mathcal{F}_s] =& \mathbb{E}\left[\exp\left(\int_0^t \theta_u dW_u\right) \exp\left(-\frac{1}{2}\int_0^t \theta_u^2 du\right) \Big| \mathcal{F}_s\right] \\
=& \exp\left(\int_0^s \theta_u dW_u\right) \exp\left(\frac{1}{2}\int_s^t \theta_u^2 du\right) \exp\left(-\frac{1}{2}\int_0^t \theta_u^2 du\right) \\
=& \exp\left(\int_0^s \theta_u dW_u - \frac{1}{2}\int_0^s \theta_u^2 du\right) = Z_s,
\end{aligned}$$

which means Z is a martingale. See Revuz [57] for a proof in the general case. □

Proposition A.0.2. *Girsanov measure transformation for Brownian motion*
Let S be a continuous process, that solves the stochastic differential equation

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t^{\mathbb{P}}$$

under a measure $\mathbb{P}$, where $W_t^{\mathbb{P}}$ is a standard Wiener process under $\mathbb{P}$. If $\int_0^t \theta_s^2 ds < \infty$, there exists an equivalent measure $Q \sim \mathbb{P}$ such that the process

$$W_t^Q = W_t^{\mathbb{P}} - \int_0^t \theta_s ds$$

is a standard Brownian motion under Q. This means that under Q the process S solves

$$dS_t = (\mu_t + \sigma_t\theta_t) S_t dt + \sigma_t dW_t^Q.$$

In addition we know that the Radon-Nikodym Derivative of $\mathbb{P}$ with respect to Q is of the form

$$\frac{d\mathbb{P}}{dQ} = \exp\left(\int_0^t \theta_s dW_s + \frac{1}{2}\int_0^t \theta_s^2 ds\right).$$

Proof. We denote a new measure Q by $dQ = Z_t d\mathbb{P}$ with

$$Z_t = \exp\left(\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right).$$

Since the exponential function is strictly positive, it is clear that $Q \sim \mathbb{P}$. First, we need to show that $W_t^Q = W_t^{\mathbb{P}} - \int\limits_0^t \theta_s ds$ defines a martingale under the measure Q. This means that it remains to show that W^Q has independent normally distributed increments with $W_t^Q - W_s^Q \sim \mathcal{N}(0, t-s)$, which can be shown via the moment generating function. The next equation shows that the moment generating function of W_t^Q is equal to the moment generating function of a normal distribution:

$$\begin{aligned}
\mathbb{E}^Q\left[\exp(aW_t^Q)\right] &= \mathbb{E}^Q\left[\exp(a(W_t^{\mathbb{P}} - \int_0^t \theta_u du))\right] = \mathbb{E}^{\mathbb{P}}\left[Z_t \exp(a(W_t^{\mathbb{P}} - \int\limits_0^t \theta_u du))\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(\int\limits_0^t \theta_s dW_s - \frac{1}{2}\int\limits_0^t \theta_s^2 ds + aW_t^{\mathbb{P}} - a\int\limits_0^t \theta_u du\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(\int\limits_0^t (a+\theta_s) dW_s - \frac{1}{2}\int\limits_0^t (2a\theta_s + \theta_s^2) ds\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(\int\limits_0^t (a+\theta_s) dW_s - \frac{1}{2}\int\limits_0^t (a+\theta_s)^2 ds\right)\right] \exp\left(\frac{a^2 t}{2}\right) \\
&= \exp\left(\frac{a^2 t}{2}\right).
\end{aligned}$$

The Nikov condition shows that the last expectation equals one. This shows that W_t^Q is normally distributed with variance t and it is easy to see that this can be extended to $W_t^Q - W_s^Q$ as well. The independency of the increments follows by construction, since $W^{\mathbb{P}}$ is assumed to be a standard Wiener process. □

The previous proposition can be extended in a general form for semimartingales. These results can be found in Chapter 3, Theorem 20 in Protter [56]. Since we only need the simpler version for a Brownian Motion, we omit this result and its proof here.

As we are trying to find the corresponding Radon-Nikodym derivative for jump diffusion models, we still need to investigate how a change in jump intensities and heights does affect the Radon-Nikodym derivative, which requires the following definitions and results.

Definition A.0.3. *E-marked point process*
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ a measurable space. Let N_t be a point process, where we define T_n as the time of the n-th jump. Let (J_n) be a sequence of random variables taking values in E. We call the double sequence (T_n, J_n) an E-marked

point process. We can define a corresponding counting measure $p(dt, dx)$ from $(\Omega, \mathcal{F})$ to $((0,\infty) \times E, (0,\infty) \otimes \mathcal{E})$ such that

$$\int_0^t \int_E H(s,x)p(ds \times dx) = \sum_{n=1}^{\infty} H(T_n, J_n)\mathbb{1}_{T_n<t}.$$

Definition A.0.4. *Local characteristics*
Let $p(dt, dx)$ be the measure of an E-marked point process. If $N_t(A)$ admits the $\mathcal{F}_t$-predictable intensity $\lambda_t(A)$ for every $A \in \mathcal{E}$ for a measure λ from $(\omega \times [0,\infty))$ to $(E, \mathcal{E})$, we say that $p(dt, dx)$ has an intensity kernel $\lambda_t(dx)$. If there exists a probability transition kernel $\Phi(dx)$ such that $\lambda_t(dx) = \lambda_t \Phi(dx)$ for a predictable process λ_t, the tupel $(\lambda_t, \Phi_t(dx))$ is called the local characteristics of (T_n, J_n).

Remark A.0.5.
To provide additional understanding of the previous definitions, note that for every non-negative E-marked process H the integral

$$\int_0^t \int_E H(p(ds \times dx) - \lambda_s(dx))$$

defines a local martingale, or in other words λ is a predictable compensator of p. In addition to the intuitive understanding one can find a proof of this statement in Chapter 8, Corollary C4 in Brémaud [10].

Lemma A.0.6.
Let a be an increasing right continuous function with $a(0) = 0$ and let h be a function with $\int_0^t |h(s)|da(s) < \infty$. The function f defined as

$$f(t) = f(0) \prod_{0<s\leq t} (1 + h(s)\Delta a(s)) \exp\left(\int_0^t h(s)da^c(s)\right)$$

with $a^c(s) = a(s) - \sum_{u\leq s} \Delta a(u)$ solves the integral equation

$$f(t) = f(0) + \int_0^t f(s-)h(s)da(s).$$

Proof. The lemma follows by application of the product formula to

$$f_1(t) = f(0) \prod_{0<s\leq t} (1 + h(s)\Delta a(s)),$$

$$f_2(t) = \exp\left(\int_0^t h(s)da^c(s)\right).$$

Using the product formula on $f(t) = f_1(t)f_2(t)$ results

$$\begin{aligned} f(t) &= f(0) + \int_0^t f_1(s-)df_2(s) + \int_0^t f_2(s)df_1(s) \\ &= f(0) + \int_0^t f_1(s-)f_2(s)h(s)da^c(s) + \sum_{0<s\leq t} f_1(s-)f_2(s)h(s)\Delta a(s) \\ &= f(0) + \int_0^t f(s-)h(s)da(s). \end{aligned}$$

The uniqueness of the solution follows by standard arguments, see the proof of Theorem T4 in Chapter A4 in Brémaud [10]. □

Theorem A.0.7. *Girasnov measure for pure jump processes*
Let $\mathbb{P}(dt, dx)$ be the measure of an E-marked point process with local characteristics $(\lambda_t, \Phi_t(dx))$. Let $h(t, x)$ be a predictable E-indexed non-negative process and μ_t be a predictable non-negative process with

$$\int_0^t \mu_s\lambda_s ds < \infty \quad \mathbb{P}\text{-a.s.},$$
$$\int_E h(t, x)\Phi_t(dx) = 1 \quad \mathbb{P}\text{-a.s.}$$

The process L_t defined as

$$L_t = \left(\prod_{n\geq 1} \mu_{T_n} h(T_n, J_n)\mathbb{1}_{T_n\leq t}\right) \exp\left(\int_0^t \int_E (1 - \mu_s h(s, x))\lambda_s\Phi_s(dx)ds\right) \tag{A.1}$$

is a $\mathbb{P}$ local martingale and for a measure Q with $\frac{dQ}{d\mathbb{P}}|_{\mathcal{F}_t} = L_t$ the process $p(dt, dx)$ has Q characteristics $(\mu_t\lambda_t, h(t, x)\Phi_t(dx))$.

Proof. The statement can be found as Theorem T10 in Chapter 8 in Brémaud [10].

□

In summary Theorem A.0.2 shows how to to modify the continuous part of a semi-martingale while Theorem A.0.7 shows how to modify the pure jump part of a semi-martingale when changing the measure. These two results can be combined to derive a general form for a Radon-Nikodym derivative. This further proves the result in Corollary 5.2.9.

Proof of Corollary 5.2.9. Similar to the proof of Theorem 2.5 in Runggaldier [59], which presents a more general form, we will first change the pure jump parts of the given semimartingale via a change of measure defined by $\frac{d\tilde{Q}}{d\mathbb{P}}$. Then we will change the continuous part with $\frac{dQ}{d\tilde{Q}}$. We will get the desired change of measure with $\frac{dQ}{d\mathbb{P}} = \frac{dQ}{d\tilde{Q}}\frac{d\tilde{Q}}{d\mathbb{P}}$.

Changing the pure jump parts leaves us with the result of Theorem A.0.7

$$\frac{d\tilde{Q}}{d\mathbb{P}} = \left(\prod_{n\geq 1} \mu_{T_n} h(T_n, \tilde{J}_n)\mathbb{1}_{\{T_n \leq t\}}\right) \exp\left(\int_0^t \int_{\mathbb{R}_+} (1 - \mu_s h(s,x))\lambda_s \Phi_s(dx)ds\right),$$

where we set J_n to be lognormally distributed and set $\tilde{J}_n = \exp(J_n)$. Additionally we can set $h(t,x) = \exp(tx)$ and $\mu_{T_n} = \exp(\nu)$, which leads to the desired form in (5.2.11) without the θ part (note that $\lambda_s = \lambda$ constant is given by the choice of S_t and $\Phi_s(\mathbb{R}_+) = 1$). For the continuous measure change we have to take care that S_t is supposed to form a martingale under the new measure Q, which, according to Theorem A.0.2, leads to the Radon-Nikodym derivative of the form

$$\frac{dQ}{d\tilde{Q}} = \theta W_t + \frac{1}{2}\theta^2 t,$$

with a θ that also compensates potential new drift of the measure change to the jump parts, i.e. $\theta = \frac{\mu - \lambda\kappa + \lambda^Q \kappa^Q}{\sigma}$, where $\kappa^Q = \mathbb{E}^Q[e^J - 1]$ is the new expected jump height under Q and $\lambda^Q = \lambda\mu_{T_n} h(T_n, \tilde{J}_n) = \lambda\tilde{\kappa}$ is the new arrival rate. The value κ^Q can be calculated using

$$\mathbb{E}^Q\left[e^{Jt}\right] = M_{Q,J}(t) = \frac{M_{\mathbb{P},J}(\gamma + t)}{M_{\mathbb{P},J}(\gamma)}$$

Multiplication of these two Radon-Nikodym derivatives gives exactly the desired form. □

Affidavit

(English)

I, Miguel Hinrichs, hereby declare that the submitted thesis is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I confirm that I followed the guidelines of good scientific practice of the Carl von Ossietzky University Oldenburg.
I further declare that neither the submitted thesis nor parts of it have been presented as part of an examination degree to any other university.

Eidesstattliche Erklärung

(German)

Hiermit versichere ich, Miguel Hinrichs, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Außerdem versichere ich, dass ich die allgemeinen Prinzipien wissenschaftlicher Arbeit und Veröffentlichung, wie sie in den Leitlinien guter wissenschaftlicher Praxis festgelegt sind, befolgt habe.
Ich versichere weiterhin, dass weder diese Arbeit noch Teile davon an einer anderen Universität eingereicht wurden.

Oldenburg, den 04.03.2025