OPTIMAL HEDGING IN THE PRESENCE OF SHORTFALL RISK

Yumiharu Nakano

Contents

Part I

Efficient hedging with coherent risk measure

1. Introduction

In a complete financial market, we can replicate a given contingent claim by a self-financing strategy. In an incomplete market, by using a "super-hedging" strategy, we can generate a final wealth that dominates the payoff of the contingent claim. If the seller of a contingent claim hopes to hedge the claim with a smaller initial amount of capital than that required by a perfect (or super-) hedging strategy, then the seller has to accept some risk. In such a situation, the seller seeks the optimal "partial" hedge that can be achieved with his/her initial amount. In Föllmer and Leukert [12], they introduced the strategy of "efficient hedging" that minimizes the *shortfall risk* under a capital constraint. They described the investor's attitude towards the shortfall in terms of a loss function, and defined the shortfall risk as the expectation of the shortfall weighted by the loss function. In other words, they used the expected loss functions as risk measures.

In this paper, we use *coherent measures of risk*, introduced by Artzner et al. [2] as risk measures. They are defined axiomatically by four desirable properties, that is, monotonicity, subadditivity, positive homogeneity, and translation invariance. In [2], they restrict themselves to finite probability spaces. Delbaen [9] extended the definition of coherent risk measures to general probability spaces (see also Kusuoka [19]). In [9], as the space of random variables, the space L^∞ of all essentially bounded random variables or the space L^0 of all random variables is adopted. We use intermediate space L^1 instead here. The space L^1 is large enough to be used in our hedging problem yet sufficiently small for nice properties to hold.

We show that, for a given contingent claim H, the optimal strategy consists in hedging a modified claim φH for some randomized test φ. This is an analogue of the results by [12].

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\mathcal{Q}$ be the set of all probability measures on $(\Omega, \mathcal{F})$ absolutely continuous with respect to P. We write L^1 and L^∞ for $L^1(\Omega, \mathcal{F}, P)$ and $L^\infty(\Omega, \mathcal{F}, P)$, respectively. For $Q \in \mathcal{Q}$, we denote expectation with respect to Q by E^Q and the Radon-Nykodim derivative dQ/dP by Z_Q. Following [2] and [9], we give the following definition.

Definition 1.1. We say that a function $\rho : L^1 \to \mathbf{R}$ is a coherent risk measure if the following are satisfied:

(1) For all $X \in L^1$ with $X \geq 0$, we have $\rho(X) \leq 0$.
(2) For all X and $Y \in L^1$, we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$.
(3) If $X \in L^1$ and $\lambda > 0$, then $\rho(\lambda X) = \lambda \rho(X)$.
(4) If $X \in L^1$ and $c \in \mathbf{R}$, then $\rho(X + c) = \rho(X) - c$.

We consider the coherent risk measures that are lower semi-continuous in the L^1-norm. We establish a representation theorem for them, which is a analogue of Proposition 4.1 in [2] and Theorem 2.3 in [9].

Theorem 1.2 ([21]). *For a function $\rho : L^1 \to \mathbf{R}$, the following are equivalent:*

(i) *The function ρ is a lower semi-continuous coherent risk measure.*

(ii) *There is a subset $\tilde{\mathcal{Q}}$ of $\mathcal{Q}$ such that*

$$\{Z_Q \mid Q \in \tilde{\mathcal{Q}}\} \text{ is a weak*-closed convex subset of } L^\infty, \tag{1.1}$$

$$\rho(X) = \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[-X] \qquad (X \in L^1). \tag{1.2}$$

The element of $\tilde{\mathcal{Q}}$ can be interpreted as a "scenario" (see [2] and [9]). We notice that, for ρ as in Theorem 1.2, the restriction of ρ on L^∞ satisfies the "Fatou property" defined in [9].

Let $(\mathcal{F}_t)_{0 \le t \le T}$ be a filtration on $(\Omega, \mathcal{F})$. For simplicity, we assume that $\mathcal{F}_0$ is trivial and $\mathcal{F}_T$ is equal to $\mathcal{F}$. The discounted price process of the underlying asset is described as a semimartingale $X = (X_t)_{0 \le t \le T}$ on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, P)$. Let $\mathcal{P}$ denote the set of all equivalent martingale measures. We assume absence of arbitrage in the sense that $\mathcal{P} \neq \emptyset$.

A self-financing strategy is described as a pair (V_0, ξ), where V_0 is an initial capital, and ξ is a predictable process such that the resulting value process

$$V_t = V_0 + \int_0^t \xi_s dX_s \qquad (t \in [0, T])$$

is well defined (see [12]). A self-financing strategy (V_0, ξ) is said to be *admissible* if the corresponding value process V satisfies

$$V_t \ge 0, \ \ \forall t \in [0, T], \ \ P\text{-a.s.}$$

We consider a contingent claim that is defined by a nonnegative random variable $H \in L^1$. We assume that

$$U_0 := \sup_{P^* \in \mathcal{P}} E^{P^*}[H] < \infty.$$

Let ρ be a coherent risk measure on L^1. The shortfall risk we consider here is given by $\rho((V_T - H) \wedge 0)$. For a given amount of initial capital $\tilde{V_0}$ which is smaller than U_0, we want to find an admissible strategy (V_0, ξ) that minimizes the shortfall risk $\rho((V_T - H) \wedge 0)$. Thus we consider the optimization problem

$$\rho\left((V_T - H) \wedge 0\right) = \rho\left(\left(V_0 + \int_0^T \xi_s dX_s - H\right) \wedge 0\right) = \min \tag{1.3}$$

under the constraint

$$V_0 \le \tilde{V_0}. \tag{1.4}$$

We take ρ from the class of lower semi-continuous coherent risk measures, and follow the method of [12]. We define the set

$$\mathcal{R} = \{\varphi : \Omega \to [0, 1] \mid \varphi \text{ is } \mathcal{F}\text{-measurable}\}$$

of "randomized tests" φ. We also define the constrained set

$$\mathcal{R}_0 = \left\{ \varphi \in \mathcal{R} \mid \sup_{P^* \in \mathcal{P}} E^{P^*}[\varphi H] \le \tilde{V}_0 \right\}.$$

We reduce our problem to the following proposition, which corresponding to Proposition 3.1 in [12].

Proposition 1.3 ([21]). *There exists $\tilde{\varphi} \in \mathcal{R}_0$ such that*

$$\inf_{\varphi \in \mathcal{R}_0} \rho(-(1-\varphi)H) = \rho(-(1-\tilde{\varphi})H). \tag{1.5}$$

Let $\tilde{\varphi}$ be the solution to the minimization problem defined by (1.5), and let $\tilde{U}$ be a right-continuous version of the process

$$\tilde{U}_t = \operatorname*{ess\,sup}_{P^* \in \mathcal{P}} E^{P^*}[\tilde{\varphi} H \mid \mathcal{F}_t].$$

The process $\tilde{U}$ is a $\mathcal{P}$-supermartingale, i.e., a supermartingale under any $P^* \in \mathcal{P}$. By the optional decomposition theorem (see [12]), there exists an admissible strategy $(\tilde{V}_0, \tilde{\xi})$ and an increasing optional process $\tilde{C}$ with $\tilde{C}_0 = 0$ such that

$$\tilde{U}_t = \tilde{V}_0 + \int_0^t \tilde{\xi}_s dX_s - \tilde{C}_t.$$

Following [12], we give the following definition.

Definition 1.4. For any admissible strategy (V_0, ξ) we define the corresponding success ratio as

$$\varphi_{(V_0,\xi)} = \mathbf{1}_{\{V_T \ge H\}} + \frac{V_T}{H} \mathbf{1}_{\{V_T < H\}}.$$

The next theorem corresponds to Theorem 3.2 in [12].

Theorem 1.5 ([21]). *Let $\tilde{\varphi}$ be a solution to the minimization problem* (1.5) *and let $(\tilde{V}_0, \tilde{\xi})$ be the admissible strategy determined by the optional decomposition of the claim $\tilde{\varphi} H$. Then the strategy $(\tilde{V}_0, \tilde{\xi})$ solves the optimization problem* (1.3) *and* (1.4).

We prove Theorem 1.2 in Section 2, and Proposition 1.3 and Theorem 1.5 in Section 3. In Section 4, we consider our hedging problem with some special coherent risk measures.

2. Proof of Theorem 1.2

Proof of Theorem 1.2. It is easy to prove the implication (2) $\Rightarrow$ (1). To prove the converse one (1) $\Rightarrow$ (2), we follow the method of proof of Theorem 2.3 in [9]. We put $\phi(X) = -\rho(X)$ and define the set $C = \{X \in L^1 \mid \phi(X) \ge 0\}$. Then since ϕ is upper semi-continuous, the set C is a convex and norm closed cone in

L^1. We regard L^∞ and L^1 as a duality pair associated with the nondegenerate bilinear form

$$L^1 \times L^\infty \ni \{X, Y\} \mapsto \langle X, Y\rangle = E[XY] \in \mathbf{R}.$$

Recall that the polar set C° of C is defined by

$$C^\circ = \{Y \in L^\infty \mid E[XY] \geq -1 \ \ (\forall X \in C)\}$$

(see [3, p. 30]). However, since C is cone, we have

$$C^\circ = \{Y \in L^\infty \mid E[XY] \geq 0 \ \ (\forall X \in C)\}.$$

This implies that C° is also a weak*-closed, convex cone in L^∞.

We put $\Phi = \{Y \in C^\circ \mid E[Y] = 1\}$. Then, it holds that

$$C^\circ = \cup_{\lambda \geq 0} \lambda\Phi. \tag{2.1}$$

Indeed, if $Y \in C^\circ$ with $E[Y] > 0$, then we have $Y = \lambda\tilde{Y}$, where $\tilde{Y} = Y/E[Y]$ and $\lambda = E[Y]$. Hence $Y \in \cup_{\lambda\geq 0}\lambda\Phi$. On the other hand, if $Y \in C^\circ$ with $E[Y] = 0$, then $Y = 0$ since $L^1_+ \subset C$. Hence $Y \in \cup_{\lambda\geq 0}\lambda\Phi$. Thus (2.1) follows. The bipolar theorem (see [3, p. 32]) then implies that

$$C = \left\{X \in L^1 \mid E[XY] \geq 0 \ \ (\forall Y \in \Phi)\right\}.$$

From this, we find that $\phi(X) \geq 0$ if and only if $E[XY] \geq 0$ for all $Y \in \Phi$. Since $\phi(X - \phi(X)) = 0$, we have that $E[(X - \phi(X))Y] \geq 0$ for all $Y \in \Phi$. Thus

$$\inf_{Y\in\Phi} E[XY] \geq \phi(X).$$

Now, for $\varepsilon > 0$, we have $\phi(X - \phi(X) - \epsilon) = -\varepsilon < 0$, so that there exists $Y \in \Phi$ such that $E[(X - \phi(X) - \varepsilon)Y] < 0$ or $E[XY] \leq \phi(X) + \varepsilon$. Since ε is arbitrary, we obtain

$$\inf_{Y\in\Phi} E[XY] \leq \phi(X),$$

hence

$$\inf_{Y\in\Phi} E[XY] = \phi(X). \tag{2.2}$$

If we put

$$\tilde{\mathcal{Q}} = \{Q \in \mathcal{Q} \mid Z_Q = Y \text{ for some } Y \in \Phi\},$$

then (2.2) implies (1.1). Since $\{Z_Q \mid Q \in \tilde{\mathcal{Q}}\} = \Phi$, we find that this is the desired representation for ρ. $\square$

3. Proofs of Proposition 1.3 and Theorem 1.5

Let $\rho : L^1 \to \mathbf{R}$ be a lower semi-continuous coherent risk measure. Then, by Theorem 1.2, there exists a subset $\tilde{\mathcal{Q}}$ of $\mathcal{Q}$ such that (1.1) and (1.2) hold. We use the representation (1.2) in the proofs of Proposition 1.3 and Theorem 1.5.

Proof of Proposition 1.3. First, we notice that $\mathcal{R}$ is weak*-compact, i.e., $\sigma(L^\infty, L^1)$-compact, in L^∞. Since the map

$$L^\infty \ni \varphi \mapsto \sup_{P^* \in \mathcal{P}} E^{P^*}[\varphi H] \in \mathbf{R}$$

is lower semi-continuous in the weak*-topology, the constrained set $\mathcal{R}_0$ is weak*-closed and so is weak*-compact. Since the map

$$L^\infty \ni \varphi \mapsto \sup_{Q \in \tilde{\mathcal{Q}}} E^{Q}[(1-\varphi) H] \in \mathbf{R}$$

is also lower semi-continuous in the weak*-topology, there exists $\tilde{\varphi} \in \mathcal{R}_0$ satisfying (1.5). □

Proof of Theorem 1.5. We consider an admissible strategy (V_0, ξ) with (1.4) and the corresponding success ratio φ. We have from $\varphi H = V_T \wedge H$ that

$$(V_T - H) \wedge 0 = -(H - V_T)_+ = -(H - V_T \wedge H) = -(1-\varphi)H.$$

Since the corresponding value process $(V_t)_{0 \le t \le T}$ is a $\mathcal{P}$-supermartingale, we obtain

$$E^{P^*}[\varphi H] \le E^{P^*}[V_T] \le V_0 \le \tilde{V}_0.$$

Thus the success ratio φ belongs to the constrained set $\mathcal{R}_0$ and so we have

$$\rho\left((V_T - H) \wedge 0\right) = \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\varphi)H] \ge \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\tilde{\varphi})H].$$

In particular, the success ratio $\varphi_{(\tilde{V}_0, \tilde{\xi})}$ satisfies

$$\sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\varphi_{(\tilde{V}_0,\tilde{\xi})})H] \ge \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\tilde{\varphi})H]. \tag{3.1}$$

On the other hand, we have

$$\varphi_{(\tilde{V}_0,\tilde{\xi})} H = \tilde{V}_T \wedge H \ge \tilde{\varphi} H, \quad P\text{-a.s.},$$

and so, for all $Q \in \tilde{\mathcal{Q}}$,

$$E^Q[(1-\varphi_{(\tilde{V}_0,\tilde{\xi})})H] \le E^Q[(1-\tilde{\varphi})H].$$

Hence we obtain from (3.1) that

$$\rho\left((\tilde{V}_T - H) \wedge 0\right) = \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\varphi_{(\tilde{V}_0,\tilde{\xi})})H] = \sup_{Q \in \tilde{\mathcal{Q}}} E^Q[(1-\tilde{\varphi})H],$$

which proves the theorem. □

4. Optimal hedging

In this section, we study our problem with two special coherent risk measures. The first one is the case of $\tilde{\mathcal{Q}}$ being a singleton, and the second one is the worst conditional expectation.

First, we take a singleton $\tilde{\mathcal{Q}} = \{Q\}$ with $Z_Q \in L^\infty$ as a scenario set. Then, the corresponding risk measure is

$$\rho(X) = E^Q[-X].$$

Thus we want to minimize the coherent risk measure

$$\rho\left((V_T - H) \wedge 0\right) = E^Q[-(V_T - H) \wedge 0] \tag{4.1}$$

under the constraint

$$V_0 \le \tilde{V}_0. \tag{4.2}$$

Theorem 1.5 shows that this is reduced to the optimization problem

$$E^Q[\varphi H] = \max \tag{4.3}$$

under the constraint that $\varphi \in \mathcal{R}$ satisfies

$$\sup_{P^* \in \mathcal{P}} E^{P^*}[\varphi H] \le \tilde{V}_0. \tag{4.4}$$

We assume that H is not trivial, i.e.,

$$E^Q[H] > 0.$$

Then the problem (4.3) and (4.4) can be reformulated as

$$E^R[\varphi] = \max \tag{4.5}$$

under the constraint

$$E^{P^*}[\varphi] \le \frac{\tilde{V}_0}{E^{P^*}[H]} \quad \forall P^* \in \mathcal{P}, \tag{4.6}$$

where the probability measures R and R^* are defined by

$$\frac{dR}{dQ} = \frac{H}{E^Q[H]}, \qquad \frac{dR^*}{dP^*} = \frac{H}{E^{P^*}[H]}.$$

In the terminology of the theory of hypothesis testing, the solution $\tilde{\varphi}_Q$ is identified as the most powerful test for the problem in which the null hypothesis is composite but the alternative simple.

In the complete case, by the fundamental lemma of Neyman and Pearson, we can solve the problem explicitly.

Proposition 4.1 ([21]). *Assume that $\mathcal{P} = \{P^*\}$. Then the most powerful test $\tilde{\varphi}_Q$ is given by*

$$\tilde{\varphi}_Q = \mathbf{1}_{\{Z_Q > \tilde{a} Z_{P^*}\}} + \gamma \mathbf{1}_{\{Z_Q = \tilde{a} Z_{P^*}\}},$$

where

$$\tilde{a} = \inf\{a \mid E^{P^*}[H\mathbf{1}_{\{Z_Q > a Z_{P^*}\}}] \leq \tilde{V}_0\}$$

and

$$\gamma = \begin{cases} \frac{\tilde{V}_0 - E^{P^*}[H\mathbf{1}_{\{Z_Q > \tilde{a} Z_{P^*}\}}]}{E^{P^*}[H\mathbf{1}_{\{Z_Q = \tilde{a} Z_{P^*}\}}]} & \text{if } P^*(\{Z_Q = \tilde{a} Z_{P^*}\} \cap \{H > 0\}) > 0, \\ \text{an arbitrary value from } [0,1] & \text{if } P^*(\{Z_Q = \tilde{a} Z_{P^*}\} \cap \{H > 0\}) = 0. \end{cases}$$

Remark 4.2 ([21]). When Q is equal to P, this proposition coincides with Proposition 4.1 in [12].

Proof of Proposition 4.1. From the Neyman-Pearson lemma (see Schmetterer [28, Chapter III, Section 3]) in terms of R and R^*, we obtain that

$$\tilde{\varphi}_Q = \mathbf{1}_{\{Z_R > \tilde{b} Z_{R^*}\}} + \beta \mathbf{1}_{\{Z_R = \tilde{b} Z_{R^*}\}},$$

where

$$\tilde{b} = \inf\left\{ b \mid P^*(Z_R > b Z_{R^*}) \leq \frac{\tilde{V}_0}{E^{P^*}[H]} \right\}$$

and

$$\beta = \begin{cases} \frac{\tilde{V}_0 / E^{P^*}[H] - R^*(Z_R > \tilde{b} Z_{R^*})}{R^*(Z_R = \tilde{b} Z_{R^*})} & \text{if } R^*(Z_R = \tilde{b} Z_{R^*}) > 0, \\ \text{an arbitrary value from } [0,1] & \text{if } R^*(Z_R = \tilde{b} Z_{R^*}) = 0. \end{cases}$$

We have

$$\{Z_R = b Z_{R^*}\} = \left\{ Z_Q = b Z_{P^*} \frac{E^Q[H]}{E^{P^*}[H]} \right\} \cap \{H > 0\}, \quad \frac{\tilde{b} E^Q[H]}{E^{P^*}[H]} = \tilde{a},$$

and $\gamma = \beta$. So

$$\tilde{\varphi}_Q \mathbf{1}_{\{H>0\}} = \mathbf{1}_{\{Z_Q > \tilde{a} Z_{P^*}\}} + \gamma \mathbf{1}_{\{Z_Q = \tilde{a} Z_{P^*}\}}.$$

Since

$$E^Q[\tilde{\varphi}_Q H] = E^Q[\tilde{\varphi}_Q H \mathbf{1}_{\{H>0\}}],$$

the proposition follows. □

Next we take the *worst conditional expectation* by Artzner et al. [2]. In our setting, this measure is given by

$$\mathrm{WCE}_\alpha(X) = \sup\left\{ E\left[(-X) \frac{\mathbf{1}_A}{P(A)} \right] \mid A \in \mathcal{F}, \ \ P(A) > \alpha \right\} \quad (X \in L^1),$$

where $\alpha \in (0,1)$. Now, for $\alpha \in (0,1]$, we define another coherent risk measure on L^1 as

$$\rho_\alpha(X) = \sup\left\{E[(-X)f] \mid f \in \Phi_\alpha\right\},$$

where

$$\Phi_\alpha = \left\{f \mid f \text{ is } \mathcal{F}\text{-measurable}, \;\; 0 \le f \le \alpha^{-1}, \;\; P\text{-a.s.}, \;\; E[f]=1\right\}.$$

For each $X \in L^1$ and $\alpha \in (0,1]$, both the coherent risk measures $\rho_\alpha(X)$ and $\mathrm{WCE}_\alpha(X)$ are bounded by $\alpha^{-1}\|X\|_1$. This implies that these coherent risk measures are continuous in the L^1-norm (see Inoue [13, Lemma 2.1]). As mentioned in [9, p. 12], if $(\Omega, \mathcal{F}, P)$ nonatomic, then $\rho_\alpha(X) = \mathrm{WCE}_\alpha(X)$ for $X \in L^\infty$. Since L^∞ is dense in L^1, we have that, for all $X \in L^1$,

$$\rho_\alpha(X) = \mathrm{WCE}_\alpha(X).$$

We consider our hedging problem with ρ_α ($\alpha \in (0,1]$) as a measure of risk. We do not need to assume that $(\Omega, \mathcal{F}, P)$ is nonatomic. Thus we consider the minimization problem of finding $\tilde{\varphi} \in \mathcal{R}_0$ such that

$$\rho_\alpha\left(-(1-\tilde{\varphi})H\right) = \inf_{\varphi \in \mathcal{R}_0} \rho_\alpha\left(-(1-\varphi)H\right). \tag{4.7}$$

Lemma 4.3 ([21])**.** *Let $X \in L^1$ such that $X \ge 0$. If $P(X>0) \le \alpha$, then we have*

$$\rho_\alpha(-X) = \frac{1}{\alpha} E^P[X]. \tag{4.8}$$

Proof. We fix $X \in L^1$ with $X \ge 0$. Then, by Theorem 1.2 in [13], we have

$$\rho_\alpha(-X) = \frac{1}{\alpha} E^P\left[X(\mathbf{1}_{\{X>k\}} + \beta \mathbf{1}_{\{X=k\}})\right], \tag{4.9}$$

where

$$k = \inf\left\{a \in \mathbf{R} \mid P(X>a) \le \alpha\right\}$$

and

$$\beta = \begin{cases} \frac{\alpha - P(X>k)}{P(X=k)} & \text{if } P(X=k)>0, \\ 0 & \text{if } P(X=k)=0. \end{cases}$$

If $P(X>0) \le \alpha$, then $k=0$ and hence (4.8) follows. □

For special H, our problem with ρ_α reduced to that with ρ_1 which has already been treated in [12] as $l(x)=x$.

Proposition 4.4 ([21])**.** *Suppose that $P(H>0) \le \alpha$. Then the solution to the minimization problem* (4.7) *is the most powerful test $\tilde{\varphi}_P$.*

Proof. By Lemma 4.3, we have

$$\rho_\alpha(-(1-\varphi)H) = \frac{1}{\alpha}E^P[(1-\varphi)H].$$

Therefore $\tilde{\varphi}_P$ minimizes $\rho_\alpha(-(1-\varphi)H)$ in $\mathcal{R}_0$. □

Example 4.5 ([21])**.** We consider the standard Black-Scholes model as in [12, Section 6.2]. Then, the discounted price process is given by

$$X_t = x_0 \exp\left(\sigma W_t + \left(m - \frac{\sigma^2}{2}\right)t\right),$$

where $m \in \mathbf{R}$, $\sigma > 0$, $x_0 > 0$, and W is a one-dimensional Wiener process on $(\Omega, \mathcal{F}, P)$. The unique equivalent martingale measure P^* is given by

$$\frac{dP^*}{dP} = \exp\left(-\frac{m}{\sigma}W_T - \frac{1}{2}\left(\frac{m}{\sigma}\right)^2 T\right) = \text{const } X_T^{-m/\sigma^2}.$$

We assume that $m > 0$, and consider an European call $H = (X_T - K)_+$ as in [12, Section 6.2]. The cost of replication of this claim is

$$U_0 = E^{P^*}[H] = x_0 N(d_+) - KN(d_-),$$

where

$$d_\pm = \frac{1}{\sigma\sqrt{T}}\log\left(\frac{x_0}{K}\right) \pm \frac{1}{2}\sigma\sqrt{T}$$

and N denotes the standard Gaussian cumulative distribution function. Let $\tilde{V}_0$ be a positive constant such that $\tilde{V}_0 \leq E^{P^*}[H]$. We assume

$$P(H > 0) = N(m\sqrt{T} + d_-) \leq \alpha.$$

Then by Proposition 4.4, the most powerful test $\tilde{\varphi}_P$ solves the minimization problem (4.7). By [12, Section 6.2], $\tilde{\varphi}_P$ is given by

$$\tilde{\varphi}_P = \mathbf{1}_{\{X_T > c\}}$$

where the constant c is determined by

$$\begin{aligned}\tilde{V}_0 &= E^{P^*}[H\mathbf{1}_{\{X_T > c\}}] \\ &= x_0 N\left(\frac{1}{\sigma\sqrt{T}}\log\left(\frac{x_0}{c}\right) + \frac{1}{2}\sigma\sqrt{T}\right) - KN\left(\frac{1}{\sigma\sqrt{T}}\log\left(\frac{x_0}{c}\right) - \frac{1}{2}\sigma\sqrt{T}\right).\end{aligned}$$

Part II

Minimizing coherent risk measures of shortfall in discrete-time models with cone constraints

5. Introduction

It is well-known that a frictionless, discrete-time and finite-horizon model of financial market is arbitrage-free if and only if there exists an equivalent martingale measure (the first fundamental theorem of asset pricing). Moreover, the arbitrage-free market is complete if and only if the equivalent martingale measure is unique (the second fundamental theorem of asset pricing) (Shiryaev [31]). In arbitrage-free and complete models, any contingent claim H is attainable, that is, there exists a trading strategy ξ^H such that the self-financed wealth process $V(x_0, \xi^H)$ is worth H at the maturity date T. The cost x_0 of replication is given by the expectation of H under the unique martingale measure.

Under the market incompleteness or in the presence of frictions, the standard no-arbitrage arguments are no longer available, and some contingent claims may not be attainable. However, even so, we can still *super-hedge* such claims: starting with enough initial wealth x_0, an agent can find a trading strategy ξ^H such that

$$V_T(x_0, \xi^H) \geq H, \quad \text{almost surely.}$$

The strategy ξ^H is called a *super-hedging strategy for H*. Then, the super-replication cost x_0 is given by the supremum of expectations of H over a suitable set of probability measures (El Karoui and Quenez [17], Föllmer and Kabanov [10], Karatzas [15], Schäl [27] and the references cited there). We define the *shortfall risk* by the following net profit of hedging loss:

$$-\left(H - V_T(x, \xi)\right)_+,$$

where $(a)_+ = \max(a, 0)$. For a super-hedging strategy ξ^H, the shortfall risk is equal to zero, that is,

$$-\left(H - V_T(x_0, \xi^H)\right)_+ = 0.$$

However, if an investor has only an initial wealth x less than x_0, then he or she cannot necessarily accomplish the super-hedging, that is, the shortfall risk may not be equal to zero.

If we are in a position of such an investor, then we have to accept the possibility of shortfall. As a result, we wish to measure the shortfall risk by some risk measure ρ, and try to find an optimal strategy that solves the minimization problem

$$\inf_{\xi} \rho\left(-\left(H - V_T(x, \xi)\right)_+\right).$$

See Cvitanić [6], Föllmer and Leukert [11, 12], and Pham [24]. These references adopted the $\rho(X) = E^P\left[l(-X)\right]$ as risk measures, where $l(\cdot)$ is a loss function and P is the objective probability. They studied the minimization problem

$$\inf_{\xi} E^P\left[l\left(\left(H - V_T(x, \xi)\right)_+\right)\right]$$

in complete or incomplete market models.

In this paper, we work in the general discrete-time models with convex cone-constrained trading strategies. Such models were studied in [24], and Pham and Touzi [25]. Under a suitable assumption, the market model is arbitrage-free if and only if the set $\mathcal{M}^e$ of all equivalent probability measures under which the discounted price process satisfies a generalized martingale property is not empty (see Theorem 2.3 below). Moreover, in such a market, super-hedging as stated above is possible. Theorem 2.4 below, due to [24], characterizes the super-replication cost of H by the supremum of expectations of H over $\mathcal{M}^e$. In particular, in the case of short-selling prohibition, the set $\mathcal{M}^e$ is given by the set of all "equivalent supermartingale measures".

As risk measures, we use the coherent risk measures introduced by Artzner, Delbaen, Eber, and Heath [2]. They are real-valued functions on a suitable space of random variables satisfying four desirable properties, that is, monotonicity, subadditivity, positive homogeneity, and translation invariance. We recall the precise definition in Section 3. Given a coherent risk measure ρ and a contingent claim H, we study the stochastic control problem

$$\inf_{\xi} \rho\left(-\left(H - V_T(x,\xi)\right)_+\right). \tag{5.1}$$

See Nakano [21], where a similar optimization problem for general frictionless continuous-time semimartingale models is studied.

It is known that a coherent risk measure ρ of a random variable X arises as the supremum of the expected negative of X over a set of "real-world" probability measures or "scenarios", that is,

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

for some set of probability measures $\mathcal{Q}$ (see Proposition 4.1 in [2]; Theorem 2.3 in Delbaen [9]; see also Theorem 1.2 in [21]; and Theorem 1.1 in Inoue [13]). Thus the problem (5.1) is equivalent to the minimization problem

$$\inf_{\xi} \sup_{Q\in\mathcal{Q}} E^Q\left[\left(H - V_T(x,\xi)\right)_+\right].$$

For a special scenario set $\mathcal{Q}$, problems of this type are studied in Cvitanić and Karatzas [7], and Sekine [30], essentially complete, continuous-time models. The problem (5.1) that consider in this paper is different from those in these references because of the presence of constraints on trading strategies (and the discrete-time setting).

In the proposed approach to the problem (5.1), we use the methods of convex duality and super-hedging. The supermartingale property of $V(x,\xi)$ and the super-hedging method enable one to reduce the dynamic problem (5.1) to the following static problem:

$$\inf_{X\in\mathcal{X}(x)} \rho\left(-(H - X)\right), \tag{5.2}$$

where the infimum is taken over a suitable set $\mathcal{X}(x)$ of random variables. To solve the static problem (5.2), we use the convex duality method (cf. [6]; Cvitanić and Karatzas [8]). After enlarging $\mathcal{Q}$ and $\mathcal{M}^e$ to suitable classes $\mathcal{Z}$ and $\mathcal{G}$, respectively, we define the auxiliary dual problem to (5.2) by

$$\sup\left\{E^P\left[H\left(Z\wedge yG\right)\right]-xy : Z\in\mathcal{Z},\ G\in\mathcal{G},\ y\geq 0\right\}, \tag{5.3}$$

where $a\wedge b=\min(a,b)$. Following [8], we prove the existence of a solution $(\hat{Z},\hat{G},\hat{y})$ to the dual problem (5.3). Using the triple, we show that there exists a solution to the problem (5.2) of the form:

$$\hat{X}=H\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+HB\mathbf{1}_{\{\hat{y}\hat{G}=\hat{Z}\}}, \tag{5.4}$$

where B is some random variable taking values in $[0,1]$. This result is similar to Theorem 4.1 in [8].

In conclusion, our approach may be summarized as follows: First, we solve the auxiliary dual problem (5.3) to obtain a triple $(\hat{Z},\hat{G},\hat{y})$. Next, we find a random variable B such that $\hat{X}$ of the form (5.4) is a solution to the static problem (5.2). Finally, we construct a super-hedging strategy $\hat{\xi}$ for $\hat{X}$. Then, the resulting strategy $\hat{\xi}$ solves (5.1).

This paper is organized as follows: In Section 6, we present the general framework and the basic results for the discrete-time models with cone-constrained trading strategies. We prepare coherent risk measures to be used in Section 7. In Section 8, we formulate the minimization problem (5.1), and reduce the dynamic problem (5.1) to the static one (5.2). Then we present the semi-closed form solution as in (5.4) to the problem (5.2) via the convex duality method. We also give a simple illustrating example. Section 9 is devoted to the proof of a key to the main theorem.

6. Discrete-time models with constraints

Let $T\in\mathbf{N}$. The discounted price process of d stocks is described as an $\mathbf{R}^d$-valued adapted stochastic process $S=\{S_t,\ t=0,\dots,T\}$ on some filtered probability space $((\Omega,\mathcal{F},P),(\mathcal{F}_t)_{t=0,\dots,T})$. We shall assume that $\mathcal{F}_0=\{\emptyset,\Omega\}$ and that $\mathcal{F}_T=\mathcal{F}$. For a probability measure Q on $(\Omega,\mathcal{F})$, we denote by E^Q the expectation with respect to Q. We write $L^1(P)$ for $L^1(\Omega,\mathcal{F},P)$. Here we consider only real-valued function. We assume that $S_t\in L^1(P)$ for $t=0,\dots,T$.

For $\mathbf{R}^d$-valued process $Y=(Y_t)_{t=0,\dots,T}$, we define $\Delta Y_t:=Y_t-Y_{t-1}$ for $t=1,\dots,T$. We write $a\cdot b$ for the inner product of $a,b\in\mathbf{R}^d$: $a\cdot b:=\sum_{k=1}^d a_kb_k$ for $a=(a_k)$ and $b=(b_k)$.

A trading strategy is an $\mathbf{R}^d$-valued predictable process $\xi=(\xi_t)_{t=1,\dots,T}$. Here, for $i=1,\dots,d$, ${\xi_t}^i$ represents the number of shares of the stock S^i held by the investor during $(t-1,t]$ for $t=1,\dots,T$. We denote by Ξ the set of all trading strategies. For $x \in \mathbf{R}$ and $\xi \in \Xi$, we define the discounted (self-financed) wealth process $V_t(x, \xi)$ by

$$
\begin{aligned}
V_t(x,\xi) &= x + \sum_{k=1}^{t} \xi_k \cdot \Delta S_k, \quad t = 1, \dots, T, \\
V_0(x,\xi) &= x.
\end{aligned}
$$

Let C be a nonempty closed convex cone in $\mathbf{R}^d$. A constrained trading strategy ξ is a trading strategy such that

$$
\xi_t \in C, \quad t = 1, \dots, T, \quad P \text{ a.s.}
$$

We denote by $\Xi(C)$ the set of all constrained trading strategies. Since C is a cone, the (negative) polar cone C° of C is given by

$$
C^\circ = \left\{ b \in \mathbf{R}^d : a \cdot b \le 0 \ \ (\forall a \in C) \right\}
$$

(see Section 5 of Chapter 1 in Aubin and Ekeland [4]).

Example 6.1. We have the following examples of constraint sets:

(1) *Unconstrained case*: $C = \mathbf{R}^d$. Then $C^\circ = \{0\}$.

(2) *Prohibition of short-selling of some stocks*:

$$
C = \left\{ a \in \mathbf{R}^d : a^i \ge 0 \ \ (\forall i \in I) \right\},
$$

where I is a subset of $\{1, \dots, d\}$. Then C° is given by

$$
C^\circ = \left\{ b \in \mathbf{R}^d : b^i \le 0 \ \ (i \in I), \ b^i = 0 \ \ (i \in \{1, \dots, d\} \setminus I) \right\}.
$$

We define the following convex cone associated with $\Xi(C)$:

$$
K = \left\{ V_T(0, \xi) : \xi \in \Xi(C) \right\}.
$$

Denote by $L^0_+(P)$ the space of all nonnegative random variables. We define the *no arbitrage condition* as follows:

Definition 6.2. We say that there is *no arbitrage opportunity* if

$$
\text{(NA)} \qquad K \cap L^0_+(P) = \{0\}.
$$

We define the subset $\Xi(x)$ of $\Xi(C)$ by

$$
\Xi(x) = \left\{ \xi \in \Xi(C) : V_T(x, \xi) \ge 0, \ P\text{–a.s.} \right\}.
$$

Thus we consider the admissibility condition that imposes the nonnegativity constraint only on the terminal wealth value. Denote $L^\infty(\Omega, \mathcal{F}, P)$ by $L^\infty(P)$, and let $\mathcal{P}$ be the set of all probability measures on $(\Omega, \mathcal{F})$ absolutely continuous with respect to P. As in [24], we introduce the following set of 'martingale measures':

$$
\begin{aligned}
\mathcal{M}(P) &= \left\{ Q \in \mathcal{P} : \begin{array}{l} dQ/dP \in L^\infty(P) \ \text{ and} \\ E^Q[\Delta S_t | \mathcal{F}_{t-1}] \in C^\circ, \ t = 1, \dots, T, \ \ Q\text{–a.s.} \end{array} \right\}, \\
\mathcal{M}^e(P) &= \left\{ Q \in \mathcal{M}(P) : Q \sim P \right\}.
\end{aligned}
$$

Writing $V_t(x,\xi) = V_{t-1}(x,\xi) + \xi_t \cdot \Delta S_t$, it is easily seen that for any $\xi \in \Xi(x)$, the process $V(x,\xi)$ is a supermartingale under any $Q \in \mathcal{M}(P)$. In the unconstrained case $C = \mathbf{R}^d$, the set $\mathcal{M}^e(P)$ is actually the set of equivalent probability measures with density in $L^\infty(P)$ under which S is a martingale. On the other hand, in the no short-selling constraints case $C = [0,\infty)^d$, the set $\mathcal{M}^e(P)$ is the set of equivalent probability measures with density in $L^\infty(P)$ under which S is a supermartingale.

Following [25] and [24], we shall make a nondegeneracy assumption on the price process. Set, for $t = 1, \dots, T$,

$$N(t-1) = \left\{ \eta \in L^{0,d}(\mathcal{F}_{t-1}, P) : \eta \cdot \Delta S_t = 0, \ \ P\text{−a.s.} \right\},$$

where $L^{0,d}(\mathcal{F}_t, P)$ denotes the space of all $\mathbf{R}^d$-valued $\mathcal{F}_t$-measurable random variables. We assume that

$$\text{(B)} \qquad N(t-1) = \{0\} \quad (t = 1, \dots, T).$$

The condition (B) is satisfied by most standard financial models. For example, the Cox-Ross-Rubinstein model satisfies (B) (see Sections 3 and 5 in [24]).

Under the condition (B), the no arbitrage condition (NA) implies the 'no free lunch' condition (see [25] and [24]). Further, we have the following extended version of first fundamental theorem of asset pricing:

Theorem 6.3 ([25])**.** *Assume* (B). *Then* (NA) *holds if and only if* $\mathcal{M}^e(P) \neq \emptyset$.

We define the super-replication cost x_0 of a nonnegative contingent claim $H \in L^1(P)$ by

$$x_0 = \inf \left\{ x \in \mathbf{R} : V_T(x,\xi) \geq H, \ P\text{−a.s., for some } \xi \in \Xi(x) \right\}.$$

In what follows, we write $\mathcal{M} = \mathcal{M}(P)$ and $\mathcal{M}^e = \mathcal{M}^e(P)$ for simplicity. The next theorem, due to [24], provides a duality result between the initial wealth and the expectation of the contingent claim under probability measures $\mathcal{M}^e$, within a general discrete-time framework with cone constraints.

Theorem 6.4 ([24])**.** *Assume* (NA) *and* (B). *Then the super-replication cost of a nonnegative contingent claim* $H \in L^1(P)$ *is given by*

$$\text{(6.1)} \qquad x_0 = \sup \left\{ E^Q[H] : Q \in \mathcal{M}^e \right\}.$$

In (6.1), *we may replace* $\mathcal{M}^e$ *by* $\mathcal{M}$. *Moreover, if* $\sup_{Q \in \mathcal{M}^e} E^Q[H] < \infty$, *then there exists* $\xi^H \in \Xi(x_0)$ *such that* $V_T(x_0, \xi^H) \geq H$, P−a.s. *The strategy* ξ^H *is called a super-hedging strategy for* H.

Remark 6.5. For a super-hedging strategy ξ^H in Theorem 6.4 and $x \geq x_0$, we easily see that $V_T(x, \xi^H) \geq H$.

7. Coherent risk measures

Let us consider a nonnegative contingent claim $H \in L^1(P)$. Following a trading strategy ξ and starting with an initial wealth x, an agent loses the resulting shortfall $(H - V_T(x,\xi))_+$ at time $t = T$. Thus his or her shortfall risk is $-(H - V_T(x,\xi))_+$. We are concerned with the problem of minimizing a coherent risk measure of this shortfall risk.

Definition 7.1 ([2] and [9])**.** We say that a functional $\rho : L^1(P) \to \mathbf{R}$ is a *coherent risk measure* if the following are satisfied:

(i) For $X \in L^1(P)$ with $X \geq 0$, we have $\rho(X) \leq 0$.
(ii) For X and $Y \in L^1(P)$, we have $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
(iii) If $X \in L^1(P)$ and $\lambda \in (0,\infty)$, then $\rho(\lambda X) = \lambda \rho(X)$.
(iv) If $X \in L^1(P)$ and $c \in \mathbf{R}$, then $\rho(X+c) = \rho(X) - c$.

Remark 7.2 ([22]). For interpretations of the above properties, we refer to [2], where they restrict themselves to finite probability spaces. Subsequently, in [9], the definition of coherent risk measures is extended to general probability spaces; as the space of random variables, the space $L^\infty(P)$ or the space $L^0(P)$ of all random variables is adopted. As in [21], we use the intermediate space $L^1(P)$ here. The space $L^1(P)$ is large enough to be used in our hedging problem since the payoff of, e.g., a European call option belongs to $L^1(P)$ since we have assumed that $S_T \in L^1(P)$.

Let $\rho : L^1(P) \to \mathbf{R}$ be a coherent risk measure that is continuous in the L^1-norm. Recall $\mathcal{P}$ from Section 6. Then, we have the following representation:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E^Q[-X] \qquad (X \in L^1(P)), \tag{7.1}$$

where

$$\mathcal{Q} = \left\{ Q \in \mathcal{P} : dQ/dP \in L^\infty(P),\ E^Q[-X] \leq \rho(X)\ \ (\forall X \in L^1(P)) \right\}. \tag{7.2}$$

See Theorem 1.2 in [21]; see also Proposition 4.1 in [2], Theorem 2.3 in [9], and Theorem 1.1 in [13]. So, every continuous coherent risk measure arises as the supremum of expected negatives of a random variable over a set $\mathcal{Q}$ of 'real-world' probability measures or 'scenarios'. Then, from the uniform boundedness theorem, we easily see that

$$\sup_{Q \in \mathcal{Q}} \|dQ/dP\|_\infty < \infty. \tag{7.3}$$

Remark 7.3 ([22]). In Theorem 1.2 in [21], it is proved that the representation (7.1) also holds for *lower semi-continuous* coherent risk measures ρ. However, even for such ρ, (7.3) still holds by the uniform boundedness theorem. This implies that every lower semi-continuous coherent risk measure turns out to be continuous (see Lemma 2.1 in [13]).

We need the next property of coherent risk measures.

Definition 7.4 ([2] and [9]). We say that a coherent risk measure $\rho : L^1(P) \to \mathbf{R}$ is *relevant* if, for every $A \in \mathcal{F}$ with $P(A) > 0$, we have $\rho(-\mathbf{1}_A) > 0$.

The next proposition, which is an analogue of Theorem 3.5 in [9], characterizes relevant coherent risk measures.

Proposition 7.5 ([22]). *Let $\rho : L^1(P) \to \mathbf{R}$ be a continuous coherent risk measure, and let $\mathcal{Q}$ be as in (7.2). Then the following are equivalent:*

(i) *ρ is relevant.*
(ii) *The set $\{Q \in \mathcal{Q} : Q \sim P\}$ is not empty.*

Proof. It is easy to prove the implication (ii) $\Rightarrow$ (i). To prove the converse (i) $\Rightarrow$ (ii), we follow the line of the proof of Theorem 3.5 in [9]. Set $\mathcal{Y} = \{dQ/dP : Q \in \mathcal{Q}\}$ and take a sequence $\{Y_n\}_{n=1}^\infty$ from $\mathcal{Y}$ so that

$$\lim_{n\to\infty} P(Y_n > 0) = \sup_{Y\in\mathcal{Y}} P(Y > 0).$$

We define the probability measure $\hat{Q} = \hat{Y}P$ by

$$\hat{Y} = \sum_{n=1}^\infty 2^{-n} Y_n.$$

Then, since $\{\hat{Y} > 0\} = \cup_{n=1}^\infty \{Y_n > 0\}$, we have

$$P(\hat{Y} > 0) \geq \sup_{Y\in\mathcal{Y}} P(Y > 0).$$

However, by (7.3), $\hat{Y} \in L^\infty(P)$, whence $\hat{Q} \in \mathcal{Q}$. Thus

$$P(\hat{Y} > 0) = \sup_{Y\in\mathcal{Y}} P(Y > 0). \tag{7.4}$$

Now suppose that $P(\hat{Y} = 0) > 0$. Since ρ is relevant, there exists $Y_0 \in \mathcal{Y}$ such that $E^P[Y_0 \mathbf{1}_{\{\hat{Y}=0\}}] > 0$. This implies

$$P(Y_0 > 0,\ \hat{Y} = 0) > 0.$$

However, if we put $Y' := (\hat{Y} + Y_0)/2 \in \mathcal{Y}$, then we find from (7.4) that

$$\begin{aligned} P(Y_0 > 0,\ \hat{Y} = 0) &= P(\hat{Y} = 0) - P(\hat{Y} = 0,\ Y_0 = 0) \\ &= P(\hat{Y} = 0) + P\left(\{\hat{Y} > 0\} \cup \{Y_0 > 0\}\right) - 1 \\ &= P(\hat{Y} = 0) + P(Y' > 0) - 1 \\ &= P(\hat{Y} = 0) + P(\hat{Y} > 0) - 1 = 0, \end{aligned}$$

which is a contradiction. Thus $\hat{Q}$ and P are equivalent. □

Remark 7.6 ([22]). The '*worst conditional expectation*' is a typical example of coherent risk measures. Given $\alpha \in (0,1)$, this measure is defined by, for $X \in L^1(P)$,

$$\mathrm{WCE}_\alpha(X) = \sup\left\{E^Q[-X] : Q(\cdot) = P(\cdot|A),\ P(A) > \alpha,\ A \in \mathcal{F}\right\}$$

(see Section 5 in [2] and Section 4 in [9]). WCE_α is the law-invariant, smallest coherent risk measure dominating the value at risk VaR_α, which is a popular risk measure, but not a coherent one. Now we define another coherent risk measure $\rho_\alpha(\cdot)$ on $L^1(P)$ by

$$\rho_\alpha(X) = \sup_{Q \in \mathcal{Q}} E^Q[-X] \qquad (X \in L^1(P)), \tag{7.5}$$

where

$$\mathcal{Q} = \{Q \in \mathcal{P} : dQ/dP \leq 1/\alpha\}.$$

As mentioned in Section 4 in [9], if the underlying probability space $(\Omega, \mathcal{F}, P)$ is nonatomic, then ρ_α coincides with WCE_α on $L^1(P)$ since the extreme points of $\mathcal{Q}$ are of the form $\mathbf{1}_A/P(A)$ with $P(A) = \alpha$ (see Lindenstrauss [20]). We see that ρ_α is relevant and continuous in the L^1-norm. For related work, we refer the reader to [1], [13], Kusuoka [19], and [30]. In particular, [30] studies the minimization problem of $\rho_\alpha\left(-(H - V_T(x,\xi))_+\right)$ in complete continuous-time market models. In many situations, it seems more convenient to use ρ_α than to use WCE_α itself.

8. The minimization

Let $\rho : L^1(P) \to \mathbf{R}$ be a relevant, continuous coherent risk measure. Then ρ has a representation of the form (7.1). We consider a nonnegative contingent claim $H \in L^1(P)$ that satisfies

$$x_0 := \sup_{Q \in \mathcal{M}} E^Q[H] < +\infty.$$

In this section, we assume the no arbitrage condition (NA) and the nondegeneracy condition (B). Then, it follows from Theorems 6.3 and 6.4 that $\mathcal{M}^e \neq \emptyset$ and that the super-replication cost of H is given by $\sup_{Q \in \mathcal{M}} E^Q[H]$.

Now let $x > 0$ and $\xi_1, \xi_2 \in \Xi(x)$. If

$$\rho\left(-(H - V_T(x,\xi_1))_+\right) \leq \rho\left(-(H - V_T(x,\xi_2))_+\right)$$

or, equivalently,

$$\sup_{Q \in \mathcal{Q}} E^Q\left[(H - V_T(x,\xi_1))_+\right] \leq \sup_{Q \in \mathcal{Q}} E^Q\left[(H - V_T(x,\xi_2))_+\right],$$

then we may regard the strategy ξ_1 as preferable to ξ_2. Therefore, an agent, who uses the coherent risk measure ρ as measure of shortfall risk, wishes to minimize $\rho(-(H - V_T(x,\xi))_+)$ for a given initial wealth x. Thus we consider the following optimization problem: for $x > 0$,

$$R(x) := \inf_{\xi \in \Xi(x)} \rho\left(-\left(H - V_T(x,\xi)\right)_+\right). \tag{8.1}$$

Remark 8.1 ([22]). For $x \geq x_0$, the super-hedging strategy ξ^H is obviously a solution to the problem (8.1) since $\left(H - V_T(x,\xi^H)\right)_+ = 0$.

We define the set $\mathcal{X}(x)$ by

$$\mathcal{X}(x) = \left\{X \in L^1(P) : 0 \leq X \leq H,\ P\text{−a.s.},\ E^P[GX] \leq x,\ (\forall G \in \mathcal{M})\right\}. \tag{8.2}$$

Here we identify each probability measure $Q \in \mathcal{M}$ with its Radon-Nikodým density $G = dQ/dP$. As in [12], [21], and [24], we reduce the dynamic problem (8.1) to a static one by the next proposition.

Proposition 8.2 ([22])**.** *Suppose that $\hat{X} \in \mathcal{X}(x)$ is a solution to the static problem*

$$\inf_{X \in \mathcal{X}(x)} \rho(-(H - X)). \tag{8.3}$$

Then there exists a super-hedging strategy $\hat{\xi} \in \Xi(x)$ for $\hat{X}$ that solves the dynamic problem (8.1)*. Moreover, we have*

$$R(x) = \inf_{X \in \mathcal{X}(x)} \rho(-(H - X)).$$

Proof. The proof is similar to that of Proposition 4.1 in [24]. Let $\xi \in \Xi(x)$ and $X := H - (H - V_T(x,\xi))_+$. Then $X \in L^1(P)$ and $X \leq V_T(x,\xi)$. Now, under every $Q \in \mathcal{M}$, the process $V(x,\xi)$ is a supermartingale, so that $X \in \mathcal{X}(x)$. So we obtain

$$\rho\left(-\left(H - V_T(x,\xi)\right)_+\right) = \rho(-(H - X)) \geq \inf_{X' \in \mathcal{X}(x)} \rho(-(H - X')).$$

Since ξ is an arbitrary element of $\Xi(x)$, we find that

$$R(x) \geq \inf_{X \in \mathcal{X}(x)} \rho(-(H - X)). \tag{8.4}$$

Conversely, suppose that $\hat{X} \in \mathcal{X}(x)$ solves the problem (8.3). We put

$$\hat{x} := \sup_{G \in \mathcal{M}} E^P[G\hat{X}].$$

Then $\hat{x} \leq x < +\infty$, and so Theorem 6.4 implies that there exists a super-hedging strategy $\hat{\xi} \in \Xi(x)$ for $\hat{X}$ such that

$$V_T(x,\hat{\xi}) \geq \hat{X}.$$

Since $-(H-\hat{X}) \leq -\left(H - V_T(x,\hat{\xi})\right)_+$, we have, from the monotonicity of ρ, that

$$R(x) \leq \rho\left(-\left(H - V_T(x,\hat{\xi})\right)_+\right) \leq \rho(-(H-\hat{X})). \tag{8.5}$$

The proposition follows from (8.4) and (8.5). □

In what follows, we assume

$$0 < x < x_0,$$

and study the static problem (8.3) for the initial wealth x. We follow the convex duality method as in [6] and [8].

We define

$$\mathcal{Z} = \left\{ Z \in L^\infty(P) : \begin{array}{l} Z \geq 0 \ (P\text{−a.s.}), \quad E^P[Z] \leq 1, \\ E^P[ZX] \leq \rho(-X) \quad (\forall X \in L^1_+(P)) \end{array} \right\},$$

where $L^1_+(P)$ denotes the space of all nonnegative random variables in $L^1(P)$. Since $\mathcal{Q} \subset \mathcal{Z}$, it holds that

$$\rho(-X) = \sup_{Z\in\mathcal{Z}} E^P[ZX] \qquad (X \in L^1_+(P)). \tag{8.6}$$

Remark 8.3 ([22]). In the case of ρ_α in (7.5) with (8.6), we can show that

$$\mathcal{Z} = \{Z \in L^\infty(P) : 0 \leq Z \leq 1/\alpha,\ E^P[Z] \leq 1\}.$$

The next proposition is needed to prove Lemmas 8.7 and 9.1.

Proposition 8.4 ([22])**.** *Let $\mathcal{Z}$ be as above.*

(i) *The set $\mathcal{Z}$ is convex and closed under P−a.s. convergence.*
(ii) *If a sequence $\{Z_n\}_{n=1}^\infty$ from $\mathcal{Z}$ converges to a random variable Z, P−a.s., on the set $\{H > 0\}$ as $n \to \infty$, and if $Z = 0$ on the set $\{H = 0\}$, then we have $Z \in \mathcal{Z}$.*

Proof. Suppose that a sequence $\{Z_n\}$ from $\mathcal{Z}$ converges to a random variable Z, P−a.s. Then, by Fatou's lemma, we have

$$E^P[ZX] \leq \liminf_{n\to\infty} E^P[Z_nX] \leq \rho(-X) < \infty \quad (X \in L^1_+(P)).$$

From this, it easily follows that $Z \in L^\infty(P)$. Similarly, by Fatou's lemma, we have $E^P[Z] \leq 1$. Thus $Z \in \mathcal{Z}$, and so (i) follows.

Let $\{Z_n\}$ and Z be as in (ii). By Fatou's lemma, we see that, for $X \in L^1_+(P)$,

$$\begin{aligned} E^P[ZX] &= E^P\left[ZX\mathbf{1}_{\{H>0\}}\right] \leq \liminf_{n\to\infty} E^P\left[Z_nX\mathbf{1}_{\{H>0\}}\right] \\ &\leq \liminf_{n\to\infty} E^P[Z_nX] \leq \rho(-X) < \infty. \end{aligned}$$

In particular, $Z \in L^\infty(P)$. Similarly, we have that $E^P[Z] \leq 1$. Thus $Z \in \mathcal{Z}$. This proves (ii). □

Following Section 3 in [6] and Section 3 in [8], we introduce the set

$$\mathcal{G} := \left\{ G \in L^1_+(P) : \begin{array}{l} E^P[G] \leq 1, \quad E^P[GH] \leq x_0, \\ E^P[GX] \leq x, \quad \forall x > 0, \quad \forall X \in \mathcal{X}(x) \end{array} \right\}$$

(recall that we have assumed that $x \in (0, x_0)$).

Proposition 8.5 ([22])**.** *We have the following:*

(i) *The set $\mathcal{G}$ is convex, closed under $P-$a.s. convergence, and bounded in $L^1(P)$, and it includes the convex hull of $\mathcal{M}$, that is,*

$$\operatorname{conv} \left\{ \frac{dQ}{dP} \right\}_{Q \in \mathcal{M}} \subset \mathcal{G}.$$

(ii) *If a sequence $\{G_n\}_{n=1}^{\infty}$ from $\mathcal{G}$ converges to some random variable G, $P-$a.s., on the set $\{H > 0\}$, and if $G = 0$ on the set $\{H = 0\}$, then we have $G \in \mathcal{G}$.*

Proof. As in the proof of Proposition 8.4, we obtain Proposition 8.5 using Fatou's lemma several times. □

Remark 8.6 ([22]). The properties (i) and (ii) in Proposition 8.5 are needed to prove Lemma 8.7 and Lemma 9.1, respectively. Therefore, for this purpose, we may replace $\mathcal{G}$ by another set $\mathcal{G}'$ if it satisfies (i) and (ii) in Proposition 8.5. For example, consider the one-period models with no-short-selling constraints. Then we know that

$$\mathcal{M} = \left\{ Q \in \mathcal{P} : E^Q[S_1] \leq S_0 \right\}. \tag{8.7}$$

We easily find that the set $\mathcal{G}'$ defined by

$$\mathcal{G}' := \left\{ G' \in L^1_+(P) : E^P[G'S_1] \leq S_0, \; E^P[G'] \leq 1 \right\}$$

satisfies the properties (i) and (ii) in Proposition 8.5 for every H, whence we may replace $\mathcal{G}$ by $\mathcal{G}'$.

Now, as in Section 3 in [8], we have the following important observation: for $Z \in \mathcal{Z}$, $G \in \mathcal{G}$, $y \geq 0$, and $X \in \mathcal{X}(x)$,

$$\begin{aligned} E^P\left[Z(H - X)\right] &= E^P[ZH] - E^P\left[X(Z - yG)\right] - yE^P[GX] \\ &\geq E^P[ZH] - E^P\left[H(Z - yG)_+\right] - xy \\ &= E^P\left[H(Z \wedge yG)\right] - xy. \end{aligned} \tag{8.8}$$

We define

$$f(y) = \sup_{Z \in \mathcal{Z}, G \in \mathcal{G}} E^P\left[H(Z \wedge yG)\right] \qquad (y \geq 0); \tag{8.9}$$

$$g(x) = \sup_{y \geq 0} \left(f(y) - xy\right). \tag{8.10}$$

Then we have

$$(8.11) \qquad g(x) \leq \inf_{X \in \mathcal{X}(x)} \sup_{Z \in \mathcal{Z}} E^P [Z(H - X)] = \inf_{X \in \mathcal{X}(x)} \rho(-(H - X)).$$

Lemma 8.7 ([22]). *For each $y \geq 0$, there exists a pair $(Z_y, G_y) \in \mathcal{Z} \times \mathcal{G}$ that attains the supremum in* (8.9).

Proof. Let $\{(Z_n, G_n)\}_{n=1}^{\infty}$ be a sequence from $\mathcal{Z} \times \mathcal{G}$ such that

$$\lim_{n \to \infty} E^P [H(Z_n \wedge yG_n)] = f(y).$$

Since the set $\mathcal{Z} \times \mathcal{G}$ is bounded in $L^1(P) \times L^1(P)$, the theorem of Komlós (see Komlós [16] and Schwartz [29]) implies that there exists a pair $(Z_y, G_y) \in L^1(P) \times L^1(P)$ and a relabeled subsequence $\left\{(Z'_j, G'_j)\right\}_{j=1}^{\infty}$ of $\{(Z_n, G_n)\}_{n=1}^{\infty}$ such that

$$\left(\frac{1}{k} \sum_{j=1}^{k} Z'_j, \frac{1}{k} \sum_{j=1}^{k} G'_j \right) \to (Z_y, G_y) \quad (k \to \infty), \quad P\text{-a.s.}$$

By the P-a.s. closedness of $\mathcal{Z}$ and $\mathcal{G}$, we have $(Z_y, G_y) \in \mathcal{Z} \times \mathcal{G}$. Since

$$|E^P[ZX]| \leq \rho(-|X|) < +\infty \quad (X \in L^1(P), Z \in \mathcal{Z}),$$

it follows from the uniform boundedness theorem that

$$(8.12) \qquad \sup_{Z \in \mathcal{Z}} \|Z\|_\infty < \infty.$$

From this, as well as Lebesgue's convergence theorem and the concavity of the function $(s, t) \mapsto s \wedge t$, we obtain

$$\begin{aligned} E^P [H(Z_y \wedge yG_y)] &= \lim_{k \to \infty} E^P \left[H \left\{ \left(\frac{1}{k} \sum_{j=1}^{k} Z'_j \right) \wedge y \left(\frac{1}{k} \sum_{j=1}^{k} G'_j \right) \right\} \right] \\ &\geq \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} E^P \left[H(Z'_j \wedge yG'_j) \right] \\ &= \lim_{j \to \infty} E^P \left[H(Z'_j \wedge yG'_j) \right] = f(y). \end{aligned}$$

Thus (Z_y, G_y) attains the supremum, as desired. □

Lemma 8.8 ([22]). *The function $f(\cdot)$ is concave.*

Proof. Let $y_1, y_2 \in [0, \infty)$ and $t \in (0, 1)$. If $y_1 = y_2 = 0$ then the concavity immediately follows. So we assume that $y_1 > 0$ or $y_2 > 0$. By Lemma 8.7, there exist $(Z_{y_j}, G_{y_j}) \in \mathcal{Z} \times \mathcal{G}$, $j = 1, 2$, such that

$$E^P \left[H(Z_{y_j} \wedge y_j G_{y_j}) \right] = f(y_j) \quad (j = 1, 2).$$

We put $G := (ty_1G_{y_1} + (1-t)y_2G_{y_2}) \,/\, (ty_1 + (1-t)y_2) \in \mathcal{G}$. Then it follows that

$$\begin{aligned}
&tf(y_1) + (1-t)f(y_2) \\
&= E^P\left[H\left\{t(Z_{y_1} \wedge y_1G_{y_1}) + (1-t)(Z_{y_2} \wedge y_2G_{y_2})\right\}\right] \\
&\le E^P\left[H\left\{(tZ_{y_1} + (1-t)Z_{y_2}) \wedge (ty_1G_{y_1} + (1-t)y_2G_{y_2})\right\}\right] \\
&= E^P\left[H\left\{(tZ_{y_1} + (1-t)Z_{y_2}) \wedge (ty_1 + (1-t)y_2)G\right\}\right] \\
&\le f\left(ty_1 + (1-t)y_2\right).
\end{aligned}$$

Thus we obtain the lemma. □

Lemma 8.9 ([22])**.** *There exists $\hat{y} \equiv \hat{y}_x > 0$ that attains the supremum in* (8.10)*.*

Proof. By Lemma 8.8, the function

$$h(y) := f(y) - xy \quad (y \ge 0)$$

is concave. From (8.12), we see that $h(+\infty) = -\infty$. Clearly, we have

$$h(0) = 0, \quad h(y) \ge -xy \;\; (y > 0).$$

We claim that there exists $y_0 > 0$ such that $h(y_0) > 0$. Suppose otherwise. Then $h(y) \le 0$ for $y > 0$. Since ρ is relevant, by Proposition 7.5 there exists $Z \in \mathcal{Z}$ such that $Z > 0$, $P-$a.s. We see that, for every $G \in \mathcal{G}$,

$$xy \ge f(y) \ge E^P\left[H(Z \wedge yG)\right] \quad (\forall y > 0).$$

Dividing by y and then letting $y \downarrow 0$, we have that, for any $G \in \mathcal{G}$,

$$E^P[HG] \le x.$$

However, this contradicts the assumption $x < x_0$. Thus the claim is proved. The lemma now follows from the concavity of $h(\cdot)$. □

Remark 8.10 ([22]). As mentioned in Remark 8.6, the necessity of considering $\mathcal{Z}$ and $\mathcal{G}$ rather than $\mathcal{Q}$ and $\mathcal{M}$, respectively, is to ensure the existence of a solution to the dual problem $g(x)$.

Now, here is our main theorem.

Theorem 8.11 ([22])**.** *Let $\hat{y} > 0$ be as in Lemma 8.9, and let $(\hat{Z}, \hat{G}) \equiv (Z_{\hat{y}}, G_{\hat{y}})$ be an optimal pair for the problem* (8.9) *with $y = \hat{y}$.*

(i) *There exists a $[0,1]$-valued random variable B such that the random variable*

$$\hat{X} := H\mathbf{1}_{\{\hat{y}\hat{G} < \hat{Z}\}} + HB\mathbf{1}_{\{\hat{y}\hat{G} = \hat{Z}\}}$$

is a solution to the static problem (8.3)*. Moreover, there is no "duality gap" in* (8.11)*, that is, $g(x) = R(x)$.*

(ii) *If $\hat{\xi}$ is a super-hedging strategy for $\hat{X}$, then $\hat{\xi}$ is a solution to* (8.1)*.*

The following proposition is a key to the proof of Theorem 8.11.

Proposition 8.12 ([22]). *There exists a* $[0,1]$*-valued random variable* B *such that* $\hat{X}$ *of the form in Theorem 8.11 satisfies* $\hat{X} \in \mathcal{X}(x)$ *as well as the conditions*

$$E^P\left[\hat{G}\hat{X}\right] = x, \tag{8.13}$$

$$\sup_{Z\in\mathcal{Z}} E^P\left[Z(H-\hat{X})\right] = E^P\left[\hat{Z}(H-\hat{X})\right]. \tag{8.14}$$

We prove this proposition in Section 9, following the method of [6] and [8]. In this method, some results from the nonsmooth convex analysis ([4]) are used.

Proof of Theorem 8.11. Using (8.13) and the fact $\hat{X}(\hat{Z}-\hat{y}\hat{G}) = H(\hat{Z}-\hat{y}\hat{G})_+$ in (8.8) with $Z=\hat{Z}$, $X=\hat{X}$, $G=\hat{G}$, and $y=\hat{y}$, we see that

$$\begin{aligned} E^P\left[\hat{Z}(\hat{H}-\hat{X})\right] &= E^P\left[H(\hat{Z}\wedge \hat{y}\hat{G})\right] - x\hat{y} \\ &= f(\hat{y}) - x\hat{y} = g(x), \end{aligned}$$

whence, by (8.14),

$$g(x) = \sup_{Z\in\mathcal{Z}} E^P\left[Z(H-\hat{X})\right] = \rho\left(-(H-\hat{X})\right).$$

Therefore, from (8.11), it follows from that

$$\rho\left(-(H-\hat{X})\right) \le \inf_{X\in\mathcal{X}(x)} \rho\left(-(H-X)\right).$$

However, $\hat{X}$ is in $\mathcal{X}(x)$, so that $\hat{X}$ attains the infimum above. In particular, by Proposition 8.2, we find that $g(x) = R(x)$ and that (ii) holds. Thus the theorem follows. $\square$

We illustrate Theorem 8.11 by using one-period binomial models.

Example 8.13 ([22]). Let $0 < d < 1 < u$ and $0 < \pi < 1$. We consider the probability space $\Omega = \{u, d\}$, $\mathcal{F} = 2^\Omega$, $P(\{u\}) = \pi = 1 - P(\{d\})$. Let $\{S_0, S_1\}$ be the discounted price process described as

$$S_0 = 1, \quad S_1 = \begin{cases} u & \text{if } \ \omega = u, \\ d & \text{if } \ \omega = d. \end{cases}$$

We consider the no-short-selling constraints. From (8.7), we see that $Q \in \mathcal{M}$ if and only if $Q(u) \le \hat{\pi} := (1-d)/(u-d)$.

Let $\alpha \in (0,1)$. We take ρ_α in (7.5) as the measure of risk here. We consider the European call $H = (S_1 - K)_+$ with $d < K < u$. Then it follows from Remark 8.3 that

$$\mathcal{Z} = \left\{Z : 0 \le Z \le \alpha^{-1}, \ \ \pi Z(u) + (1-\pi)Z(d) \le 1\right\}.$$

We put

$$\mathcal{G}' = \{G : 0 \le G(u) \le \hat{\pi}/\pi, \ \ \pi G(u) + (1-\pi)G(d) \le 1\}.$$

Then $\mathcal{G}'$ satisfies the properties (i) and (ii) in Proposition 8.5. Here, as mentioned in Remark 8.6, we consider $\mathcal{G}'$ rather than $\mathcal{G}$. The function $f(\cdot)$ in (8.9) with $\mathcal{G}$ replaced by $\mathcal{G}'$ is given by

$$f(y) = \sup_{Z\in\mathcal{Z}, G\in\mathcal{G}'} \pi(u-K)\left(Z(u) \wedge yG(u)\right).$$

We easily find that, for $y > 0$, this supremum is attained by, e.g.,

$$\hat{Z}(\omega) = \left(\frac{1}{\alpha} \wedge \frac{1}{\pi}\right)\mathbf{1}_{\{u\}}(\omega), \quad \hat{G}(\omega) = \hat{\pi}\mathbf{1}_{\{u\}}(\omega).$$

Then the dual problem (8.10) is written as

$$g(x) = \sup_{y\ge 0}\left[\pi(u-K)\left\{\left(\alpha^{-1}\wedge\pi^{-1}\right)\wedge y\hat{\pi}\right\} - xy\right].$$

Since $x < x_0 = \hat{\pi}(u-K)$, $\hat{y} = \pi(\alpha^{-1}\wedge\pi^{-1})/\hat{\pi}$ attains this supremum. Therefore, from Theorem 8.11, we deduce that the random variable

$$\begin{aligned}\hat{X} &:= H\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}} + HB\mathbf{1}_{\{\hat{y}\hat{G}=\hat{Z}\}} \\ &= (u-K)B\mathbf{1}_{\{u\}}\end{aligned}$$

is optimal for some $[0,1]$-valued random variable B. However, from the condition (8.13), we obtain

$$B(u) = \frac{x}{\hat{\pi}(u-K)},$$

whence

$$\hat{X} = \frac{x}{\pi}\mathbf{1}_{\{u\}} = \frac{x}{x_0}H.$$

9. Proof of Proposition 8.12

This section is devoted to the proof of Proposition 8.12. Following [6] and [8], we introduce the Banach space

$$\mathcal{K} := L^1(P)\times L^1(P)\times\mathbf{R}$$

with norm

$$\|(U,V,y)\|_{\mathcal{K}} := E^P\left[|U|+|V|\right] + |y|$$

and its subset

$$\mathcal{L} := \{(HZ, yHG, y)\in\mathcal{K} : Z\in\mathcal{Z}, \ G\in\mathcal{G}, \ y\ge 0\}.$$

Lemma 9.1 ([22])**.** *The set $\mathcal{L}$ is convex and closed in $\mathcal{K}$.*

Proof. Let $t \in (0,1)$ and $(HZ_j, y_jHG_j, y_j) \in \mathcal{L}$ for $j = 1, 2$. If $y_1 = y_2 = 0$, the convexity immediately follows. Assume that $y_1 > 0$ or $y_2 > 0$. Then we have

$$G := \frac{ty_1G_1 + (1-t)y_2G_2}{ty_1 + (1-t)y_2} \in \mathcal{G},$$

so that $(H(tZ_1 + (1-t)Z_2), H(ty_1G_1 + (1-t)y_2G_2), ty_1 + (1-t)y_2) \in \mathcal{L}$. Thus the convexity follows.

Next, let $(HZ_n, y_nHG_n, y_n) \in \mathcal{L}$ be a sequence that converges to some (U, V, y) in $\mathcal{K}$. Then $y_n \to y$, and $HZ_n \to U$, $P-$a.s. (possibly along a subsequence). Put

$$Z := \begin{cases} U/H & \text{on } \{H > 0\}, \\ 0 & \text{on } \{H = 0\}. \end{cases}$$

Then we have that $U = HZ$ on $\{H > 0\}$ and that $Z_n \to Z$ on $\{H > 0\}$, as $n \to \infty$, $P-$a.s. The property (ii) in Proposition 8.4 implies $Z \in \mathcal{Z}$.

On the other hand, we have

$$E^P\left[H|y_nG_n - yG_n|\right] \le |y_n - y|x_0 \to 0 \quad (n \to \infty),$$

hence

$$E^P\left[H|yG_n - W|\right] \to 0 \quad (n \to \infty),$$

where, we set $W = V/H$ on $\{H > 0\}$, and $= 0$ on $\{H = 0\}$. If $y = 0$, then $W = 0 = yHG_1$. If $y > 0$ and if we put $G := W/y$, then we have that $G_n \to G = W/y$, $P-$a.s., on the set $\{H > 0\}$ (possibly along a subsequence) and that $G = 0$ on the set $\{H = 0\}$. The property (ii) in Proposition 8.5 implies $G \in \mathcal{G}$. Thus the closedness of $\mathcal{L}$ follows. □

We now define a functional $\Phi : \mathcal{K} \to \mathbf{R}$ by

$$\Phi(U, V, y) = xy - E^P[U \wedge V].$$

Lemma 9.2 ([22])**.** *Let $\hat{Z}$, $\hat{G}$, and $\hat{y}$ be as in Section 8. The functional Φ is proper, convex and lower semi-continuous on $\mathcal{K}$ and attains its infimum over $\mathcal{L}$ at the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$.*

Proof. Since the proof of the properness, convexity, and lower semi-continuity are simple, we omit them. By Lemmas 8.7 and 8.9, we have

$$\begin{aligned} \Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) &= x\hat{y} - E^P\left[H(\hat{Z} \wedge \hat{y}\hat{G})\right] = x\hat{y} - f(\hat{y}) \\ &\le xy - f(y) \le xy - E^P\left[H(Z \wedge yG)\right] \\ &= \Phi(HZ, yHG, y), \qquad\qquad \forall (HZ, yHG, y) \in \mathcal{L}. \end{aligned}$$

Thus the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ is optimal. □

Proof of Proposition 4.8. The method of the proof is similar to that of Section 3 in [6] and Sections 4 and 6 in [8]. We now consider the dual space

$$\mathcal{K}^* := L^\infty(P) \times L^\infty(P) \times \mathbf{R}$$

of $\mathcal{K}$, and the normal cone (see Definition 4.1.3 and Proposition 4.1.4 in [4])

$$\mathcal{N}(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) := \left\{ (T, W, \zeta) \in \mathcal{K}^* : \begin{array}{l} E^P[H\hat{Z}T + \hat{y}H\hat{G}W] + \hat{y}\zeta \geq \\ E^P[UT + VW] + y\zeta, \quad (\forall (U, V, y) \in \mathcal{L}) \end{array} \right\}$$

to the set $\mathcal{L}$ at the point $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) \in \mathcal{L}$. We also consider the subdifferential at this point

$$\partial\Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) := \left\{ (T, W, \zeta) \in \mathcal{K}^* : \begin{array}{r} \Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y}) - \Phi(U, V, y) \leq \\ E^P\left[T(H\hat{Z} - U) + W(\hat{y}H\hat{G} - V)\right] \\ + \zeta(\hat{y} - y), \quad (\forall (U, V, y) \in \mathcal{K}) \end{array} \right\}.$$

Then, from Lemma 9.2, it follows that the triple $(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ solves the problem

$$\inf_{(U,V,y)\in\mathcal{M}} \Phi(U, V, y),$$

so that, by Corollary 4.6.3 in [4], there exists a triple $(\hat{T}, \hat{W}, \hat{\zeta}) \in \mathcal{K}^*$ such that $(\hat{T}, \hat{W}, \hat{\zeta}) \in \mathcal{N}(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$ and $(-\hat{T}, -\hat{W}, -\hat{\zeta}) \in \partial\Phi(H\hat{Z}, \hat{y}H\hat{G}, \hat{y})$. These are equivalent to the following, respectively:

$$E^P\left[H\hat{T}(\hat{Z} - Z) + H\hat{W}(\hat{y}\hat{G} - yG)\right] + \hat{\zeta}(\hat{y} - y) \geq 0, \qquad \forall (Z, G, y) \in \mathcal{Z} \times \mathcal{G} \times [0, +\infty), \tag{9.1}$$

$$E^P\left[\hat{T}(U - H\hat{Z}) + \hat{W}(V - \hat{y}H\hat{G}) + H(\hat{Z} \wedge \hat{y}\hat{G}) - U \wedge V\right] \geq (x + \hat{\zeta})(\hat{y} - y), \qquad \forall (U, V, y) \in \mathcal{K}. \tag{9.2}$$

By letting $y \to \pm\infty$, we see that (9.2) holds only if

$$\hat{\zeta} = -x. \tag{9.3}$$

From (9.1) with $\hat{\zeta} = -x$, $Z = \hat{Z}$, $G = \hat{G}$, and $y = \hat{y} \pm \delta$ ($\delta > 0$), we have

$$E^P[\hat{G}H\hat{W}] = x. \tag{9.4}$$

On the other hand, (9.1) with $Z = \hat{Z}$ and $y = \hat{y}$ implies

$$E^P[GH\hat{W}] \leq E^P[\hat{G}H\hat{W}], \quad \forall G \in \mathcal{G}, \tag{9.5}$$

and (9.1) with $G = \hat{G}$ and $y = \hat{y}$ implies

$$E^P[ZH\hat{T}] \leq E^P[\hat{Z}H\hat{T}], \quad \forall Z \in \mathcal{Z}. \tag{9.6}$$

Reading (9.2) with

$$U = H\hat{Z} - \mathbf{1}_{\{\hat{W}+\hat{T}>1\}}, \quad V = \hat{y}H\hat{G} - \mathbf{1}_{\{\hat{W}+\hat{T}>1\}}$$

and using (9.3), we have

$$E^P[(\hat{W}+\hat{T}-1)_+] \le 0.$$

Similarly we obtain $E^P\left[(1-\hat{W}-\hat{T})_+\right] \le 0$, whence

$$\hat{W}+\hat{T}=1. \tag{9.7}$$

Hence, the conditions (9.1) and (9.2) can be written as

$$E^P\left[H\hat{W}(\hat{y}\hat{G}-yG+Z-\hat{Z})+H(Z-\hat{Z})\right] \ge x(\hat{y}-y),$$
$$\forall (Z,G,y)\in \mathcal{Z}\times\mathcal{G}\times[0,+\infty),$$

$$E^P\left[\hat{W}(V-U+H\hat{Z}-\hat{y}H\hat{G})+U-H\hat{Z}+H(\hat{Z}\wedge\hat{y}\hat{G})-U\wedge V\right] \ge 0, \tag{9.8}$$
$$\forall (U,V)\in L^1(P)\times L^1(P).$$

Considering (9.8) for $U=H\hat{Z}$, $V=\hat{y}H\hat{G}+\mathbf{1}_A$ with arbitrary $A\in\mathcal{F}$, we see that

$$0\le E^P\left[\hat{W}\mathbf{1}_A+H(\hat{Z}\wedge\hat{y}\hat{G})-H\hat{Z}\wedge(\hat{y}H\hat{G}+\mathbf{1}_A)\right]\le E^P[\hat{W}\mathbf{1}_A],$$

so that $\hat{W}\ge 0$, $P-$a.s. Similarly we get

$$0\le E^P\left[(1-\hat{W})\mathbf{1}_A\right],$$

whence $\hat{W}\le 1$, $P-$a.s. Therefore

$$0\le\hat{W}\le 1 \qquad P\text{-a.s.} \tag{9.9}$$

The conditions (9.5) and (9.9) imply $H\hat{W}\in\mathcal{X}(x)$. The condition (9.8) for $U=V$ implies

$$E^P\left[H\hat{W}(\hat{Z}-\hat{y}\hat{G})\right]\ge E^P\left[H(\hat{Z}-\hat{y}\hat{G})_+\right]. \tag{9.10}$$

Hence, (9.9) and (9.10) lead to $H\hat{W}(\hat{Z}-\hat{y}\hat{G})=H(\hat{Z}-\hat{y}\hat{G})_+$, $P-$a.s., so that

$$\hat{W}=1 \quad\text{on}\quad \{\hat{y}\hat{G}<\hat{Z}\}\cap\{H>0\}, \quad =0 \quad\text{on}\quad \{\hat{y}\hat{G}>\hat{Z}\}\cap\{H>0\}.$$

Thus, there exist $[0,1]$-valued random variables B and J such that

$$\hat{W}=\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z},H>0\}}+B\mathbf{1}_{\{\hat{y}\hat{G}\ge\hat{Z},H>0\}}+J\mathbf{1}_{\{H=0\}}. \tag{9.11}$$

Reading (9.8) with $\hat{W}$ in (9.11), $U=HU'$, and $V=HV'$, we have

$$E^P\left[\left(\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+B\mathbf{1}_{\{\hat{y}\hat{G}\ge\hat{Z},H>0\}}\right)H(U-V+\hat{Z}-\hat{y}\hat{G})\right.$$
$$\left.+H(U-\hat{Z}+\hat{Z}\wedge\hat{y}\hat{G}-U\wedge V)\right]\ge 0, \quad \forall U,V\in L^\infty(P).$$

In particular, reading $U=\hat{Z}$, we have that, for $V\in L^\infty(P)$,

$$E^P\left[H\left(\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+B\mathbf{1}_{\{\hat{y}\hat{G}\ge\hat{Z},H>0\}}\right)(V-\hat{y}\hat{G})+H(\hat{Z}\wedge\hat{y}\hat{G}-\hat{Z}\wedge V)\right]\ge 0.$$

So, for $V \in L^\infty(P)$, we obtain

$$
\begin{aligned}
&E^P\left[BH\mathbf{1}_{\{\hat{y}\hat{G}\geq\hat{Z}\}}(\hat{y}\hat{G}-V)\right]\\
&\leq E^P\left[H(V-\hat{y}\hat{G})\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+H(\hat{Z}\wedge\hat{y}\hat{G}-\hat{Z}\wedge V)\right]\\
&= E^P\Big[H(V-\hat{y}\hat{G})\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+H(\hat{y}\hat{G}-V)\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z},V\leq\hat{Z}\}}\\
&\quad +H(\hat{y}\hat{G}-\hat{Z})\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}<V\}}+H(\hat{Z}-V)\mathbf{1}_{\{\hat{y}\hat{G}\geq\hat{Z}\geq V\}}\Big]\\
&= E^P\left[H(V-\hat{Z})\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}<V\}}+H(\hat{Z}-V)\mathbf{1}_{\{\hat{y}\hat{G}\geq\hat{Z}\geq V\}}\right].
\end{aligned}
\tag{9.12}
$$

From (9.12) with

$$
V=\begin{cases}\hat{Z}-\varepsilon & \text{on } \{\hat{y}\hat{G}<\hat{Z}\},\\ \hat{Z} & \text{on } \{\hat{y}\hat{G}\geq\hat{Z}\}\end{cases}
$$

for some $\varepsilon>0$, we have

$$
\{V<\hat{Z}\}=\{\hat{y}\hat{G}<\hat{Z}\},
$$

and so

$$
E^P\left[BH(\hat{y}\hat{G}-\hat{Z})_+\right]\leq E^P\left[H(-\varepsilon)\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}\right]\leq 0.
$$

This implies $B=0$ on $\{\hat{y}\hat{G}>\hat{Z},H>0\}$, $P-$a.s. Therefore we deduce that

$$
H\hat{W}=H(\mathbf{1}_{\{\hat{y}\hat{G}<\hat{Z}\}}+B\mathbf{1}_{\{\hat{y}\hat{G}=\hat{Z}\}})=\hat{X}\qquad P\text{-a.s.}
$$

On the other hand, the conditions (9.4), (9.6) and (9.7) imply that $\hat{X}$ satisfies (8.13) and (8.14). Thus the proposition follows. $\square$

Part III

Minimization of shortfall risk in a jump-diffusion model

10. Introduction

In this paper, we consider a frictionless, complete financial market consisting of one riskless bond and two risky assets S_i, $i = 1, 2$, that are traded up to a finite time horizon T. We suppose that the dynamics of S_i are described by jump-diffusion processes. Given a contingent claim H and an initial wealth x, we study the following optimization problem:

$$V(x) = \inf_{\pi \in \mathcal{A}} E\left[\ell_p\left((H - X^{x,\pi}(T))_+\right)\right], \tag{10.1}$$

where $\ell_p(x)$ is the power function x^p/p with $p > 1$, $X^{x,\pi}$ is the wealth process, and $\mathcal{A}$ is a set of admissible portfolios.

To explain the problem (10.1), we consider an investor with initial wealth x. If x is greater than the replication cost of H, say x_H, then the investor can hedge the contingent claim H without risk, by the completeness of the market. However, if x is strictly less than x_H, he/she faces the possibility of shortfall, i.e., for any portfolio π, the shortfall $(H - X^{x,\pi}(T))_+$ may be positive. In this situation, one method of hedging H is to follow a portfolio that minimizes the shortfall risk $E\left[\ell_p\left((H - X^{x,\pi}(T))_+\right)\right]$.

For work related to the problem (1), see, e.g., Föllmer and Leukert [12] and Pham [24]. See also Nakano [22]. In [12], general semimartingale models and general loss functions are considered. They impose the nonnegativity constraint on the wealth processes, and use the arguments involving Neyman-Pearson-type lemmas. This setting is crucial in solving their problem. In this paper, however, instead of the nonnegativity constraint, we impose only an integrability condition on the wealth processes (see Definition 11.2 below). Thus, our setting is similar to that of [24], except that we work in a jump-diffusion model of continuous-time markets. By requiring only the integrability condition on the wealth processes, we can obtain not only the optimal terminal wealth but also the optimal portfolio explicitly.

In Section 11, we explain the model and state the precise formulation of our problem. We then show that we can separate the problem into two problems, that is, the perfect hedging problem of H and the utility minimization problem

$$J(x_H - x) = \inf_{\pi \in \mathcal{A}_0(x_H - x)} E[\ell_p(X^{x_H - x,\pi}(T))], \tag{10.2}$$

where $A_0(x_H - x)$ is the set of portfolios. We prove that the optimal portfolio of (10.1) is represented as the difference between the perfect hedging portfolio of H and the optimal portfolio of the problem (10.2). As in the standard utility maximization problems (cf. Karatzas and Shreve [18, Chapter 3] and Jeanblanc-Picqué and Pontier [14]), we can solve the problem (10.2) by using the martingale method. In our main theorem, we give closed form expressions for the optimal portfolio and the value function $V(x)$. These results hold for every European-type contingent claim, such as, claims that can take negative values and path-dependent options. All the proofs of the results are given in Section 12.

11. The model and main results

We consider a frictionless financial market consisting of one riskless bond B and two risky asset S_i, $i = 1, 2$, that are traded up to a finite time horizon T. We suppose that B satisfies the equation

$$dB(t) = r(t)B(t)dt, \quad B(t) = 1.$$

We also suppose that S_i, $i = 1, 2$, satisfy the stochastic differential equations

$$\begin{cases} dS_i(t) = S_i(t-)(\mu_i(t)dt + \sigma_i(t)dW(t) + \gamma_i(t)dN(t)), \\ S_i(0) = s_i \in (0, \infty) \qquad\qquad\qquad\qquad (i = 1, 2), \end{cases} \tag{11.1}$$

where W is a one-dimensional standard Brownian motion on a complete probability space $(\Omega, \mathcal{F}, P)$, and the process N is a Poisson process with intensity $\lambda(\cdot)$, which is independent of W. The filtration $(\mathcal{F}_t)_{t\geq 0}$ is the P-augmentation of the natural filtration generated by W and N. Then W is a $(P, \mathcal{F}_t)$-Brownian motion, N is a $(P, \mathcal{F}_t)$-Poisson process with intensity $\lambda(\cdot)$, and the process $M(t) := N(t) - \int_0^t \lambda(s)ds$ is a P-martingale.

Assumption 11.1. For $i = 1, 2$, λ, r, μ_i, σ_i, and γ_i are bounded, measurable, deterministic functions on $[0, T]$ that satisfy the following conditions:

(i) $\lambda(t) > 0$, $r(t) \geq 0$, $\sigma_i(t) > 0$, $\gamma_i(t) > -1$, and $\gamma_i(t) \neq 0$ for $t \in [0, T]$ and $i = 1, 2$;

(ii) there exists $c_1 \in (0, \infty)$ such that, for $t \in [0, T]$,

$$|\sigma_1(t)\gamma_2(t) - \sigma_2(t)\gamma_1(t)| \geq c_1;$$

(iii) there exists $c_2 \in (0, \infty)$ such that, for $t \in [0, T]$,

$$\frac{(\mu_2(t) - r(t))\sigma_1(t) - (\mu_1(t) - r(t))\sigma_2(t)}{\lambda(t)(\sigma_2(t)\gamma_1(t) - \sigma_1(t)\gamma_2(t))} \geq c_2.$$

In this paper, we define the investor's wealth process $(X(t))_{0\leq t\leq T}$ in the standard self-financing way. Thus we assume that $X(\cdot) \equiv X^{x,\pi}(\cdot)$ satisfies

$$\begin{cases} dX^{x,\pi}(t) = r(t)X^{x,\pi}(t)dt \\ \qquad\qquad + \displaystyle\sum_{i=1}^{2} \pi_i(t) \left\{(\mu_i(t) - r(t))dt + \sigma_i(t)dW(t) + \gamma_i(t)dN(t)\right\}, \\ X^{x,\pi}(0) = x, \end{cases} \tag{11.2}$$

where $x \in \mathbf{R}$ is an initial wealth and the portfolio process $\pi(t) = (\pi_1(t), \pi_2(t))$ is an $\mathbf{R}^2$-valued $\mathcal{F}_t$-predictable process such that all the integrals in (11.2) are well-defined. The process $(\pi_1(t), \pi_2(t))$ represents the actual amounts of money invested in the risky assets $(S_1(t), S_2(t))$.

Throughout this paper, we fix $p \in (1, \infty)$, and consider the loss function ℓ_p defined by

$$\ell_p(x) = \frac{x^p}{p} \quad (x \geq 0).$$

We are concerned with the minimization of $E[\ell_p((H - X^{x,\pi}(T))_+)]$ over some suitable class of portfolios. To this end, we define the class of admissible portfolios as follows.

Definition 11.2. A portfolio process $(\pi(t))_{0 \leq t \leq T}$ is said to be *admissible* if

$$E\left[\sup_{0 \leq t \leq T} |X^{0,\pi}(t)|^p\right] < \infty.$$

We write $\mathcal{A}$ for the class of all such π.

Put, for $t \in [0, T]$,

$$\begin{aligned}
\theta(t) &:= \frac{(\mu_2(t) - r(t))\gamma_1(t) - (\mu_1(t) - r(t))\gamma_2(t)}{\sigma_2(t)\gamma_1(t) - \sigma_1(t)\gamma_2(t)}; \\
\beta(t) &:= \frac{(\mu_2(t) - r(t))\sigma_1(t) - (\mu_1(t) - r(t))\sigma_2(t)}{\lambda(t)(\sigma_2(t)\gamma_1(t) - \sigma_1(t)\gamma_2(t))}.
\end{aligned}$$

Then, by Assumption 11.1, the functions θ and β are bounded, and β is positive. Moreover we have

$$\mu_i(t) - r(t) - \sigma_i(t)\theta(t) + \lambda\gamma_i(t)\beta(t) = 0, \quad i = 1, 2. \tag{11.3}$$

We consider the exponential local martingale

$$\begin{aligned}
L(t) := \exp&\left(-\int_0^t \theta(s)dW(s) - \frac{1}{2}\int_0^t |\theta(s)|^2 ds\right) \\
&\times \exp\left(\int_0^t \log\beta(s)dN(s) + \int_0^t \lambda(s)(1 - \beta(s))ds\right).
\end{aligned}$$

From the boundedness of θ and β, the process $(L(t))_{0 \leq t \leq T}$ is a strictly positive P-martingale, and satisfies, for every $a \in \mathbf{R}$,

$$E[(L(t))^a] < \infty. \tag{11.4}$$

We consider the probability measure P_0 on $(\Omega, \mathcal{F}_T)$ defined by

$$\frac{dP_0}{dP} = L(T).$$

Then, the process

$$W_0(t) := W(t) + \int_0^t \theta(s)ds \quad (0 \leq t \leq T)$$

is a $(P_0, \mathcal{F}_t)$-Brownian motion, $(N(t))_{0\le t\le T}$ is a $(P_0, \mathcal{F}_t)$-Poisson process with intensity $\lambda(t)\beta(t)$, and the process

$$M_0(t) := N(t) - \int_0^t \lambda(s)\beta(s)ds \quad (0 \le t \le T)$$

is a P_0-martingale (cf. Brémaud [5]). Using (11.3), we have, for $i = 1, 2$,

$$d\tilde{S}_i(t) = \tilde{S}_i(t-)(\sigma_i(t)dW_0(t) + \gamma_i(t)dM_0(t)),$$

where $\tilde{S}_i(t) = S_i(t)/B(t)$. Thus, $(\tilde{S}_i(t))_{0\le t\le T}$, $i = 1, 2$, are also P_0-martingales.

In what follows, we use the following notation:

Notation. For a process $(Y(t))_{0\le t\le T}$, we denote by $\tilde{Y}(t)$ the discounted value of $Y(t)$, i.e., $\tilde{Y}(t) := Y(t)/B(t)$.

For $x \in \mathbf{R}$ and $\pi \in \mathcal{A}$, the discounted wealth process $\tilde{X}^{x,\pi}$ satisfies

$$\tilde{X}^{x,\pi}(t) = x + \sum_{i=1}^{2} \int_0^t \tilde{\pi}_i(s) \left\{\sigma_i(s)dW_0(s) + \gamma_i(s)dM_0(s)\right\}. \tag{11.5}$$

Thus, $\tilde{X}^{x,\pi}$ is a local P_0-martingale. However, since $\pi \in \mathcal{A}$, we have from Hölder's inequality and (11.4) that $E_0[\sup_{0\le t\le T} |\tilde{X}^{x,\pi}(t)|] < \infty$, where $E_0[\cdot]$ denotes the expectation with respect to P_0. This implies that $\tilde{X}^{x,\pi}$ is a P_0-martingale.

Now, the market above is complete in the following sense:

Proposition 11.3 ([23])**.** *Let $H \in L^{p+\varepsilon}(\Omega, \mathcal{F}_T, P)$ for some $\varepsilon > 0$, and put*

$$x_H := E_0\left[\frac{H}{B(T)}\right]. \tag{11.6}$$

Then there exists unique $\pi_H \in \mathcal{A}$ such that, for $t \in [0, T]$,

$$\tilde{X}_t^{x_H,\pi_H} = E_0\left[\left.\frac{H}{B(T)}\right| \mathcal{F}_t\right] \quad a.s. \tag{11.7}$$

Let $H \in L^{p+\varepsilon}(\Omega, \mathcal{F}_T, P)$. We interpret H as the investor's liability. By Proposition 11.3, starting with initial wealth $x_H := E_0[H/B(T)]$, the investor can find the replicating portfolio $\pi \in \mathcal{A}$ for H. However, if the initial wealth x is less than x_H, he/she faces the possibility of shortfall. In such a situation, one is naturally led to the minimization of shortfall in an adequate sense. Thus we consider the following stochastic control problem:

$$V(x) := \inf_{\pi\in\mathcal{A}} E\left[\ell_p((H - X^{x,\pi}(T))_+)\right], \quad x < x_H.$$

By definitions of H and $\mathcal{A}$, we easily find that $V(x) < \infty$.

As we stated in Section 10, we can separate the problem above into two problems, that is, the perfect hedging problem of H and a utility minimization problem. For $z > 0$, we denote by $\mathcal{A}_0(z)$ the set of portfolio processes $(\pi(t))_{0\le t\le T}$ satisfying

$$X^{z,\pi}(t) \ge 0 \quad t \in [0,T] \quad \text{a.s.},$$

and

$$E\left[\sup_{0\le t\le T} |X^{0,\pi}(t)|^p\right] < \infty.$$

We consider another optimization problem, that is,

$$J(z) := \inf_{\pi\in\mathcal{A}_0(z)} E\left[\ell_p(X^{z,\pi}(T))\right], \quad z > 0. \tag{11.8}$$

Proposition 11.4 ([23]). *For every $z \in (0,\infty)$, there exists $\pi_0 \in \mathcal{A}_0(z)$ that is optimal for the problem stated in* (11.8).

Define $q \in (0,\infty)$ by $(1/p) + (1/q) = 1$. The main theorem of this paper is stated in the following.

Theorem 11.5 ([23]). (i) *Let π_0 be as in Proposition 11.4 with $z = x_H - x$ and let π_H be as in Proposition 11.3. Then $\pi_H - \pi_0$ is optimal for the problem* (10.1).

(ii) *For $(t,u) \in [0,T]\times(0,\infty)$, let (Π_1,Π_2) be the unique solution to the linear system*

$$\begin{cases} \sigma_1(t)\Pi_1(t,u) + \sigma_2(t)\Pi_2(t,u) = -\dfrac{\theta(t)}{p-1}u, \\ \gamma_1(t)\Pi_1(t,u) + \gamma_2(t)\Pi_2(t,u) = (\beta(t))^{q-1}u. \end{cases}$$

Then the optimal portfolio $\pi_0 \in \mathcal{A}_0(z)$ of the problem (11.8) *is given by* $(\Pi_1(t, X^{z,\pi_0}(t-)), \Pi_2(t, X^{z,\pi_0}(t-)))$.

(iii) *The value function $V(x)$ in* (10.1) *is given by*

$$V(x) = \ell_p(x_H - x)\exp\left(-(p-1)\int_0^T a(s)ds\right) \quad (x < x_H),$$

where $a(\cdot)$ is defined by

$$a(s) = -qr(s) + \frac{1}{2}q(q-1)\theta^2(s) - \lambda(s)\left((q-1) - q\beta(s) + (\beta(s))^q\right). \tag{11.9}$$

(iv) *The optimal terminal wealth in* (10.1) *is given by*

$$\begin{aligned} &X^{x,\pi_H-\pi_0}(T) \\ &= H - (x_H - x)(L(T))^{q-1}\exp\left(-\int_0^T\left(a(s) + \frac{r(s)}{p-1}\right)ds\right). \end{aligned}$$

By Theorem 11.5, if we can obtain the hedging portfolio π_H, then we can minimize the shortfall risk by implementing the portfolio $\pi_H - \pi_0$. The problem (10.1) is thus reduced to the perfect hedging problem of H. In the case of a claim of the form $H = f(S_1(T), S_2(T))$, the hedging portfolio can be obtained as in the classical case of Black-Scholes model. See, e.g., Runggaldier [26], Section 6 in [14], and the references cited there.

Remark 11.6 ([23]). As in the most utility maximization problems, we can associate the problem (11.8) with a HJB equation. We define the wealth process $X^{t,z,\pi}$ with initial condition $(t,z) \in [0,T] \times (0,\infty)$ as in (11.2). We also define the class $\mathcal{A}_0(t,z)$ of portfolio processes in a way similar to the definition of $\mathcal{A}_0(z)$, and put $J(t,z) := \inf_{\pi \in \mathcal{A}_0(t,z)} E[\ell_p(X^{t,z,\pi}(T))]$. Then, as in Chapter 3 in [18] and Proposition 5.1 in [14], we can prove that the function $J(t,z)$ satisfies the following HJB equation:

$$\begin{cases} \dfrac{\partial J(t,z)}{\partial t} + \inf_{\pi \in \mathbf{R}^2} \mathcal{L}J(t,z) = 0, \\ J(T,z) = \ell_p(z), \qquad (t,z) \in [0,T) \times (0,\infty), \end{cases}$$

where

$$\begin{aligned} \mathcal{L}J(t,z) &= zr(t)\frac{\partial J}{\partial z}(t,z) + \sum_{i=1}^{2}(\mu_i(t) - r(t))\frac{\partial J}{\partial z}(t,z) \\ &\quad + \frac{1}{2}\left(\sum_{i=1}^{2}\pi_i\sigma_i(t)\right)^2 \frac{\partial^2 J}{\partial z^2}(t,z) + \lambda\left\{J\left(t, z + \sum_{i=1}^{2}\pi_i\gamma_i(t)\right) - J(t,z)\right\}. \end{aligned}$$

12. Proofs

12.1. Proof of Proposition 11.3.

Proof of Proposition 4. Let $H \in L^{p+\varepsilon}(\Omega, \mathcal{F}_T, P_0)$, and put $x_H := E_0[H/B(T)]$. Then, as in Proposition 2.1 in [14], there exists a portfolio process π_H such that (11.7) holds. We can easily show the admissibility of π_H using the P_0-martingale property of $\tilde{X}$, Hölder's inequality, (11.4), and Doob's maximal inequality. □

12.2. Proof of Proposition 11.4.

As in the references on the expected utility maximization such as Chapter 3 in [18], we use the martingale method.

First we write $I(\cdot)$ for the inverse function of $\ell_p'(\cdot)$, that is, $I(y) = y^{q-1}$ for $y > 0$, where $1/p + 1/q = 1$. Let U_p be the negative of the Legendre transform of ℓ_p:

$$\begin{aligned} U_p(y) &= -\sup_{x>0}(xy - \ell_p(x)) = \inf_{x>0}(\ell_p(x) - xy) \\ &= \ell_p(I(y)) - yI(y), \quad y > 0. \end{aligned} \tag{12.1}$$

The infimum in (12.1) is attained by $x = I(y)$.

For $(t, y) \in [0, T] \times (0, \infty)$, we define the process

$$\begin{aligned} Y^{t,y}(s) = y \exp\left(-\int_t^s r(u)du - \int_t^s \theta(s)dW(u) - \frac{1}{2}\int_t^s |\theta(s)|^2 du \right) \\ \times \exp\left(\int_t^s \log\beta(u)dN(u) + \int_t^s (1-\beta(u))\lambda du \right), \quad t \le s \le T. \end{aligned}$$

Then $Y^{t,y}$ satisfies

$$\begin{cases} dY^{t,y}(s) = Y^{t,y}(s-)\left(-r(s)ds - \theta(s)dW(s) + (\beta(s)-1)dM(s)\right), \\ Y^{t,y}(t) = y. \end{cases}$$

We put, for $(t, y) \in [0, T] \times (0, \infty)$,

$$\mathcal{X}(t, y) := E_0\left[e^{-\int_t^T r(s)ds} I(Y^{t,y}(T)) \right].$$

Then, we easily see that

$$\mathcal{X}(t, y) = y^{1/(p-1)} \exp\left(\int_t^T a(s)ds \right), \tag{12.2}$$

where $a(\cdot)$ is given by (11.9). We write $\mathcal{Y}(t, \cdot)$ for the inverse function of $\mathcal{X}(t, \cdot)$, that is,

$$\mathcal{Y}(t, z) = z^{p-1} \exp\left(-(p-1)\int_t^T a(s)ds \right). \tag{12.3}$$

Proof of Proposition 11.4. By (12.1), we find that, for $y > 0$ and $\pi \in \mathcal{A}_0(z)$,

$$\begin{aligned} &E[\ell_p(X^{z,\pi}(T))] \\ &= E[\ell_p(X^{z,\pi}(T)) - Y^{0,y}(T)X^{z,\pi}(T)] + E[Y^{0,y}(T)X^{z,\pi}(T)] \\ &= E[\ell_p(X^{z,\pi}(T)) - Y^{0,y}(T)X^{z,\pi}(T)] + yz \\ &\ge E[U_p(Y^{0,y}(T))] + yz. \end{aligned} \tag{12.4}$$

The equality in (12.4) holds if and only if

$$X^{z,\pi}(T) = I(Y^{0,y}(T)).$$

However, from Proposition 11.3, there exists $\pi_0 \in \mathcal{A}_0(z)$ such that

$$\tilde{X}^{z,\pi_0}(t) = E_0\left[e^{-\int_0^T r(s)ds} I(Y^{0,\mathcal{Y}(0,z)}(T)) \middle| \mathcal{F}_t \right]. \tag{12.5}$$

Then (12.4) implies that $E[\ell_p(X^{z,\pi_0}(T))] = J(z)$. $\square$

12.3. **Proof of Theorem 11.5.**

Proof of Theorem 11.5. We consider the stochastic Legendre transform $U(y,\omega)$ of $-\ell_p(H(\omega)-z)$: for $y>0$,

$$U(y,\omega) = \sup_{-\infty<z\le H(\omega)} \{-\ell_p(H(\omega)-z) - yz\} = \ell_q(y) - yH(\omega). \tag{12.6}$$

The supremum in (12.6) is attained by $H(\omega)-I(y)$.

From (11.7) and (12.5), we have

$$X^{x_H,\pi_H}(T) = H, \quad X^{x_H-x,\pi_0}(T) = I\left(Y^{0,y}(T)\right)$$

and

$$E_0\left[e^{-\int_0^T r(u)du} I\left(Y^{0,y}(T)\right)\right] = x_H - x,$$

where $y := \mathcal{Y}(0, x_H - x)$. It follows that

$$X^{x,\pi_H-\pi_0}(T) = X^{x_H,\pi_H}(T) - X^{x_H-x,\pi_0}(T) = H - I\left(Y^{0,y}(T)\right). \tag{12.7}$$

Now, by (12.6), we see that, for every $\pi \in \mathcal{A}$,

$$\begin{aligned}
&(x_H - x)\mathcal{Y}(0,x_H-x) - E\left[\ell_q\left(\mathcal{Y}(0,x_H-x)e^{-\int_0^T r(u)du}L(T)\right)\right] \\
&= E\left[Y^{0,y}(T)H - \ell_q\left(Y^{0,y}(T)\right)\right] - x\mathcal{Y}(0,x_H-x) \\
&\le E\left[Y^{0,y}(T)H - \ell_q\left(Y^{0,y}(T)\right) - Y^{0,y}(T)\left(H\wedge X^{x,\pi}(T)\right)\right] \\
&\le E\left[\ell_p(H - X^{x,\pi}(T)\wedge H)\right] = E\left[\ell_p((H-X^{x,\pi}(T))_+)\right].
\end{aligned}$$

Both equalities hold in the above inequalities if and only if

$$H \wedge X^{x,\pi}(T) = H - I\left(Y^{0,y}(T)\right). \tag{12.8}$$

However, (12.7) implies that the portfolio $\pi_H - \pi_0$ satisfies (12.8). Therefore,

(12.9)

$$\begin{aligned}
V(x) &= E\left[\ell_p((H - X^{x,\pi_H-\pi_0}(T))_+)\right] \\
&= (x_H - x)\mathcal{Y}(0,x_H-x) - E\left[\ell_q\left(\mathcal{Y}(0,x_H-x)e^{-\int_0^T r(u)du}L(T)\right)\right].
\end{aligned}$$

Thus Theorem 11.5 (i) follows.

From (12.5) and the Markov property of the process $Y^{t,y}$, we have, for $z>0$,

$$\begin{aligned}
X^{z,\pi_0}(t) &= E_0\left[e^{-\int_t^T r(s)ds} I\left(Y^{0,\mathcal{Y}(0,z)}(T)\right)\middle|\mathcal{F}_t\right] \\
&= \mathcal{X}\left(t, Y^{0,\mathcal{Y}(0,z)}(t)\right),
\end{aligned}$$

where π_0 is as in Proposition 11.4. Thus

$$Y^{0,\mathcal{Y}(0,z)}(t) = \mathcal{Y}(t, X^{z,\pi_0}(t)). \tag{12.10}$$

Itô formula and (12.2) imply that

$$d\left(e^{-\int_0^t r(s)ds}\mathcal{X}(t,Y^{0,y}(t))\right) = -\theta(t)e^{-\int_0^t r(s)ds}Y^{0,y}(t)\mathcal{X}_y(t,Y^{0,y}(t))dW_0(t) + \left\{\mathcal{X}(t,\beta(t)Y^{0,y}(t)) - \mathcal{X}(t,Y^{0,y}(t))\right\}dM_0(t).$$

From this and (12.10), we see that

$$\begin{aligned}
&\tilde{X}^{z,\pi_0}(t) - z \\
&= -\int_0^t \theta(s)e^{-\int_0^s r(u)du}\frac{\mathcal{Y}(s,X^{z,\pi_0}(s))}{\mathcal{Y}_z(s,X^{z,\pi_0}(s))}dW_0(s) \\
&\quad + \int_0^t e^{-\int_0^s r(u)du}\left\{\mathcal{X}(s,\beta(s)\mathcal{Y}(s,X^{z,\pi_0}(s-))) - X^{z,\pi_0}(s-)\right\}dM_0(s).
\end{aligned} \tag{12.11}$$

Therefore, by (12.3) we have

$$-\theta(t)\frac{\mathcal{Y}(t,u)}{\mathcal{Y}_z(t,u)} = -\frac{\theta(t)}{p-1}u = \sum_{i=1}^2 \sigma_i(t)\Pi_i(t,u),$$

$$\mathcal{X}(t,\beta(t)\mathcal{Y}(t,u)) - u = \beta(t)^{1/(p-1)}u = \sum_{i=1}^2 \gamma_i(t)\Pi_i(t,u).$$

Thus, by (12.11), Theorem 11.5 (ii) follows.

By (12.3), (12.9), and easy computation similar to that of (12.2), we have

$$V(x) = \ell_p(x_H - x)\exp\left(-(p-1)\int_0^T a(s)ds\right),$$

which proves Theorem 11.5 (iii). Finally, Theorem (11.5) (iv) follows immediately from (12.7). $\square$

References

[1] C. Acerbi and D. Tasche, Expected shortfall: a natural coherent alternative to Value at Risk, *Econ. Notes* **31** (2002), 379–388.

[2] Ph. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, *Math. Finance* **9** (1999), 203–208.

[3] J. P. Aubin, *Mathematical Methods of Game and Economic Theory*, revised ed., North-Holland, Amsterdam, 1982.

[4] J. Aubin and J. Ekeland, *Applied nonlinear analysis*, John Wiley & Sons, New York, 1984.

[5] P. Brémaud, *Point processes and queues: martingale dynamics*, Springer-Verlag, New York, 1981.

[6] J. Cvitanić, Minimizing expected loss of hedging in incomplete and constrained markets, *SIAM J. Control Optim.* **38** (2000), 1050–1066.

[7] J. Cvitanić and I. Karatzas, On dynamic measures of risk, *Finance Stoch.* **3** (1999), 904–950.

[8] J. Cvitanić and I. Karatzas, Generalized Neyman-Pearson lemma via convex duality, *Bernoulli* **7** (2001), 79–97.

[9] F. Delbaen, Coherent risk measures on general probability spaces, in K. Sondermann and P. Schönbucher eds., *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*, 1–37, Springer-Verlag, Berlin, 2002.

[10] H. Föllmer and Yu, M. Kabanov, Optional decompositions and Lagrange multipliers, *Finance Stoch.* **2** (1998), 69–81.

[11] H. Föllmer and P. Leukert, Quantile Hedging, *Finance Stoch.* **3** (1999), 251–273.

[12] H. Föllmer and P. Leukert, Efficient hedging: Cost versus shortfall risk, *Finance Stoch.* **4** (2000), 117–146.

[13] A. Inoue, On the worst conditional expectation, *J. Math. Anal. Appl.* **286** (2003), 237–247.

[14] M. Jeanblanc-Picqué and M. Pontier, Optimal portfolio for a small investor in a market model with discontinuous prices, *Appl. Math. Optim.* **22** (1990), 287–310.

[15] I. Karatzas, *Lectures on the Mathematics of Finance*, CRM Monographs series, American Mathematical Society, Providence, 1997.

[16] J. Komlós, A generalization of a problem of Steinhaus, *Acta Math. Acad. Sci. Hung.* **18** (1967), 217–229.

[17] N. El Karoui and M.-C. Quenez, Dynamic programming and pricing of contingent claims in an incomplete market, *SIAM J. Control Optim.* **33** (1995), 29–66.

[18] I. Karatzas and S. E. Shreve, *Methods of mathematical finance*, Springer-Verlag, New York, 1998.

[19] S. Kusuoka, On law invariant coherent risk measures, in S. Kusuoka and T. Maruyama eds., *Advances in Mathematical Economics* **3**, 83–95, Springer-Verlag, Tokyo, 2001.

[20] J. Lindenstrauss, A short proof of Lyapounoff's theorem, *J. Math. Mech.* **4** (1966), 971–972.

[21] Y. Nakano, Efficient hedging with coherent risk measure, *J. Math. Anal. Appl.* **293** (2004), 345–354.

[22] Y. Nakano, Minimizing coherent risk measures of shortfall in discrete-time models under cone constraints, *Appl. Math. Finance* **10** (2003), 163–181.

[23] Y. Nakano, Minimization of shortfall risk in a jump-diffusion model, *Statist. Probab. Lett.* **67** (2004), 87–95.

[24] H. Pham, Dynamic L^p-hedging in discrete time under cone constraints, *SIAM J. Control Optim.* **38** (2000), 665–682.

[25] H. Pham and N. Touzi, The fundamental theorem of asset pricing with cone constraints, *J. Math. Econom.* **31** (1999), 265–280.

[26] W. J. Runggaldier, Jump diffusion models, in: S. T. Rachev ed., *Heavy Tailed Distributions in Finance*, Handbooks in Finance, Book 1, pp.169–209, Elsevier, North-Holland, 2003.

[27] M. Schäl, Martingale measures and hedging for discrete time financial markets, *Math. Oper. Res.* **24** (1999), 509–528.

[28] L. Schmetterer, *Introduction to Mathematical Statistics*, Springer-Verlag, New York, 1974.

[29] M. Schwartz, New proofs of a theorem of Komlós, *Acta. Math. Hungar.* **47** (1986), 181–185.

[30] J. Sekine, Dynamic minimization of worst conditional expectation of shortfall, *Math. Finance* **14** (2004), 605–618.

[31] A. N. Shiryaev, *Essentials of Stochastic Finance, Facts, Models, Theory*, World Scientific, Singapore, 1999.