Noname manuscript No.
(will be inserted by the editor)

Risk Minimization under Transaction Costs

Paolo Guasoni

Bank of Italy
Research Department
Via Nazionale, 91
00184 Roma *
e-mail: guasoni@dm.unipi.it

The date of receipt and acceptance will be inserted by the editor

Abstract We study the general problem of an agent wishing to minimize the risk of a position at a fixed date. The agent trades in a market with a risky asset, with incomplete information, proportional transaction costs, and possibly constraints on strategies. In particular, this framework includes the problems of hedging contingent claims and maximizing utility from wealth.

We obtain a minimization problem on a space of predictable processes with finite variation. Borrowing a technique from Calculus of Variation, on this space we look for a convergence which makes minimizing sequences relatively compact, and risk lower semicontinuous.

For a class of convex decreasing risk functionals, we show the existence of optimal strategies. Examples include the problems of *shortfall minimization*, *utility maximization*, and minimization of *coherent risk measures*.

Key words transaction costs – incomplete markets – risk minimization – coherent risk measures – constraints
***JEL* Classification**: G12,G13
Mathematics Subject Classification (2000): 60H30, 62P05, 91B30, 26A45

---

* The first draft of this paper was completed while the author was affiliated to Scuola Normale Superiore. The author is grateful to Freddy Delbaen for stimulating discussions. Special thanks go to Maurizio Pratelli and Thorsten Rheinländer for helpful remarks, and for carefully reading earlier versions of this paper. Financial support from Scuola Normale Superiore and the Bank of Italy is gratefully acknowledged. The views expressed herein are those of the author and not necessarily those of the Bank of Italy.

1 Introduction

In a perfect market, asset pricing theory takes two very different forms for conventional stocks and bonds on the one hand, and for derivative securities on the other. The price of a conventional asset is determined in equilibrium by its risk/return properties, as wells as by the preferences of market participants. This is the classical portfolio theory, due to Markowitz, Sharpe, and Merton. By contrast, in perfect markets derivative contracts are redundant, since their payoff can be replicated exactly by a trading strategy in the underlying securities, as shown in the celebrated paper of Black and Scholes. As a result, derivatives are priced by pure arbitrage arguments.

This distinction ceases to exist when market frictions are taken into account. In an incomplete market, the risk associated to derivatives can only be partially hedged, and an optimal strategy involves the minimization of the residual risk. In other words, hedging a claim becomes equivalent to optimizing a portfolio including it.

This paper studies the general problem of minimizing the risk of a trading position at a fixed date, in a market with incomplete information, proportional transaction costs, and constraints on strategies. This framework encompasses different problems, such as hedging a contingent claim expiring at time T, or maximizing utility from terminal wealth.

In the last decade, these problems have raised considerable interest, and have been attacked with different approaches. Stochastic control theory has been employed by Davis and Norman [8] for the utility maximization problem under transaction costs and by Davis, Panas, and Zariphopoulou [9] for the option pricing problem in the Black-Scholes model with transaction costs. These techniques have the advantage that characterize value functions as weak solutions of PDE, but are applicable only when the risky assets follow Markov processes. Such limitation is overcome by the convex duality approach, introduced by Karatzas, Lehoczky and Shreve [19] in a complete market model, and progressively extended to more general settings. Assuming that risky assets follow Itô processes, Karatzas, Lehoczky, Shreve and Xu [20] solved the utility maximization problem under incomplete information, while the generalization to the semimartingale case is due to Kramkov and Schachermayer [22] and Schachermayer [24], who determined necessary and sufficient conditions on utility functions for the existence of an optimal solution. In this framework, the problem with transaction costs has been solved in the case of Itô processes by Cvitanic and Karatzas [5], assuming the existence of a solution to the dual problem, which has been recently proved by Cvitanic and Wang [7]. In the general semimartingale case, the problem has been addressed by Deelstra, Pham and Touzi [10], who show the existence of a solution and relax the regularity assumptions on utility functions.

Here we take a different approach, which avoids the dual formulation and deals with the original problem directly. We obtain a minimization problem of a convex functional over a class of predictable processes with finite variation. Borrowing an idea popular in Calculus of Variation, we look for a convergence which makes minimizing sequences relatively compact, and the functional lower semicontinuous. This provides the existence of a minimizer through the classical Weierstrass theorem.

Since we assume that the risky assets are continuous semimartingales, the scope of this paper goes beyond Itô processes, but does not reach the full semimartingale case. Also, we define the spaces of admissible strategies in terms of integrability conditions, along the lines of Delbaen and Schachermayer [14], as opposed to imposing the limited borrowing condition, generally used in most papers on convex duality (with the exception of Schachermayer [24]). Rather, we consider this condition among constraints on strategies.

One of the major benefits of our approach is that market frictions add little complexity to the problem, even when transaction costs and constraints are stochastic. Also, the risk functionals considered here are general enough to allow both the maximization of expected utility, and the minimization of *coherent risk measures* (see Artzner, Delbaen, Eber and Heath [2] for details). Utility functions will need minimal regularity, and can be defined on the whole real line. As a byproduct, the *shortfall minimization* problem can be easily handled, as it is equivalent to the maximization of a utility function consisting in the infimum of two lines.

The paper is organized as follows: in section 2 we describe in details our model for a market with transaction costs, and show that the formulation given here is substantially equivalent to others in the literature. Section 3 is devoted to the spaces of admissible strategies, where we look for a convergence providing both sufficiently many compact sets, and reasonable (semi)continuity properties for the portfolio value. Section 4 contains the main results on the existence of optimal strategies in markets with incomplete information and transaction costs, when X is a continuous martingale. Section 5 handles the constrained case, while the last section is devoted to the continuous semimartingale case.

In the appendix we recall a few not-so-popular results in functional analysis and stochastic integration which are used extensively in the present paper.

2 The Model

We consider a market with a risky asset X and a riskless asset B. X is a continuous semimartingale defined on a probability space $(\Omega, \mathcal{F}, P)$, adapted with respect to a given filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, which satisfies the usual hypotheses, and such that $\mathcal{F} = \mathcal{F}_T$. $\mathcal{F}_t$ represents the information available to the agent at time t, which includes the observation of the risky asset value X_t. We assume that B is deterministic, hence without loss of generality we can set $B = 1$, since it is the same as replacing X by $\frac{X}{B}$.

An agent starts with some initial capital c, and faces some contingent liability $H = (H_X, H_B)$ at time T, which requires the payment of H_X

shares of the risky asset, and H_B units of the numeraire. Her goal is to set up a portfolio which minimizes the total risk at time T. The self-financing condition implies that a trading strategy is uniquely determined by the number of shares θ_t invested in the risky asset at time t.

In this market continuous trade is allowed, but proportional transaction costs are present. Denoting by L_t, M_t respectively the cumulative number of shares purchased and sold at time t, and assuming that a cost of k_t is incurred for each share traded, the total cost of a strategy is given by:

$$C_t(\theta) = \int_0^T k_t(dL_t + dM_t)$$

where the processes k_t, L_t and M_t are adapted to $\mathcal{F}_t$, and k_t determines the cost scheme. For example, costs are proportional to the number of shares if k_t is a constant, or to the amount traded if $k_t = kX_t$, for some constant k. We obviously have:

$$\theta_t = L_t - M_t$$

and L_t, M_t are increasing processes, therefore θ_t is a process of bounded variation. Viceversa, any function of bounded variation can be represented as a difference between two increasing functions, given by:

$$L_t = D\theta^+([0,t]) \quad \text{and} \quad M_t = D\theta^-([0,t])$$

where $D\theta^+$ and $D\theta^-$ are respectively the positive and negative parts of the Radon measure $D\theta$. The increasing processes L_t and M_t are uniquely determined under the assumption that they do not simultaneously increase; see for instance Ambrosio, Fusco and Pallara (henceforth AFP) [1] for details. From the financial point of view, this is a natural condition, since it prevents opposite transactions from taking place at the same time.

With the convention that $\theta_t = 0$ for $t < 0$ and $\theta_T = \lim_{t\uparrow T} \theta_t$, the transaction cost up to time t is equal to:

$$C_t(\theta) = \int_{[0,t]} k_s d|D\theta_s|$$

and the market value of the portfolio at time t is given by the initial capital, plus the trading gain, minus the transaction cost, namely:

$$V_t^c(\theta) = c + G_t(\theta) - C_t(\theta) \qquad \text{where} \qquad G_t(\theta) = \int_0^t \theta_s dX_s$$

At the terminal date T, the payment of the liability H and the liquidation of the remaining portfolio will give a payoff equal to:

$$V_T^c(\theta) - k_T|\theta_T - H_X| - X_T H_X - H_B$$

which must be evaluated according to the agent preferences. We shall consider risk functionals $\rho : L^p \mapsto \mathbb{R}$, where ρ satisfies the following properties:

i) ρ is convex;
ii) if $X \leq Y$ a.e., then $\rho(X) \geq \rho(Y)$ (ρ is decreasing);
iii) ρ has the Fatou property. Namely, if $X_n \to X$ a.e., then

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

In particular, this class includes the *shortfall*, for $\rho(X) = E[X^-]$, and a class of σ-additive *coherent risk measures* (see Delbaen [11] for details), for $\rho(X) = \sup_{P\in\mathcal{P}} E_P[-X]$, where $\mathcal{P}$ is a set of probabilities. *Utility maximization* can also be embedded in this framework, minimizing the functional $\rho(X) = -E[U(X)]$, where U is the utility function.

Remark 2.1 In our framework, the costs associated to the purchase and sale of the risky asset are equal. This is not the case in several papers dealing with proportional transaction costs, where the costs for buying and selling one share of the risky asset are respectively λX_t and μX_t (for example, see Davis, Panas, and Zariphopoulou [9] and Cvitanić and Karatzas [5]). This feature can be embedded in our model via the alternative definition of $C_T(\theta)$:

$$C_t^{\lambda,\mu}(\theta) = \int_{[0,t]} k_t d(\lambda D\theta^+ + \mu D\theta^-)$$

Since we have that $|D\theta| = D\theta^+ + D\theta^-$, for $\lambda = \mu$ the above definition boils down to (2), up to a rescaling of k. In fact, the quantity $\lambda D\theta^+ + \mu D\theta^-$ shares all the relevant properties of $|D\theta|$, such as weak lower semicontinuity, and all the discussion that follows on $|D\theta|$ should be easy to adapt to the case $\lambda \neq \mu$.

In practice, given two processes for the bid and the ask prices, one can always obtain a model where $\lambda = \mu$, choosing X_t as the midpoint between the bid and the ask, and $2k_t$ equal to the bid-ask spread. Since this convention greatly simplifies notation, throughout the paper we shall only consider the case $\lambda = \mu$.

Remark 2.2 The pointwise definition of variation can be modified into the following (much less intuitive), which is invariant up to sets of Lebesgue measure zero:

$$|D\theta|(\omega) = \sup_{\substack{\phi\in C_c^1(0,T)\\ \|\phi\|_\infty\leq 1}} \int_{[0,T]} \theta_s(\omega)\phi'(s)ds$$

In fact, it can be shown that for each θ there exists a representative such that the expression above coincides with the (generally higher) pointwise variation. For details, see AFP [1].

3 Strategies with Transaction costs

In the literature on markets with incomplete information, the following spaces of strategies are often considered, especially for $p = 2$:

$$\Theta^p = \{\theta : \theta\ \mathcal{F}_t\text{-predictable},\ G_T(\theta) \in L^p(P)\}$$

Θ^p can be endowed with the topology induced by the map $G_T : \theta \mapsto \int_0^T \theta_t dX_t$. It is clear that $G_T(\Theta^p)$ is a linear subspace of L^p: if X is a continuous martingale, it turns out that it is also closed, as it follows from Theorems 6.9 and 6.10, in the cases $p > 1$ and $p = 1$ respectively.

The presence of transaction costs in fact forces a much narrower set of admissible strategies than Θ^p. As we argued in the previous section, in a market with proportional costs we should only consider strategies with finite variation. This leads us to define the following spaces:

$$\Theta^p_C = \{\theta \in \Theta^p,\ C_T(\theta) \in L^p(P)\}$$

endowed with the norm:

$$\|\cdot\|_{\Theta^p_C} :\ \theta \to \left(\left\| \int_0^T \theta_t dX_t \right\|_p^p + \left\| \int_{[0,T]} k_t d|D\theta| \right\|_p^p \right)^{\frac{1}{p}} =$$
$$= \left(\|G_T(\theta)\|_p^p + \|C_T(\theta)\|_p^p \right)^{\frac{1}{p}}$$

We begin our discussion with the following result:

Proposition 3.1 *Let X be a continuous local martingale, and k a continuous, adapted process, such that $\tilde{k}(\omega) = \min_{t\in[0,T]} k_t(\omega) > 0$ for a.e. ω. Then Θ^p_C is a Banach space for all $p \geq 1$.*

Proof Θ^p_C is a linear subspace of Θ^p, therefore $\|G_T(\theta)\|_p$ is a norm. Hence it is sufficient to prove that $\|C_T(\theta)\|_p$ is also a norm, and that the space is complete. Trivially, $\|C_T(\lambda\theta)\|_p = |\lambda| \|C_T(\theta)\|_p$. Let $\theta, \eta \in \Theta^p_C$. We have:

$$\|C_T(\theta + \eta)\|_p = \left\| \int_{[0,T]} k_t d|D(\theta + \eta)|_t \right\|_p \leq$$
$$\leq \left\| \int_{[0,T]} k_t d|D\theta|_t + \int_{[0,T]} k_t d|D\eta|_t \right\|_p \leq \|C_T(\theta)\|_p + \|C_T(\eta)\|_p$$

It remains to show that Θ^p_C is complete. Let θ^n be a Cauchy sequence for Θ^p_C: since it is also Cauchy in Θ^p and X is continuous, it follows that Θ^p is complete by Theorems 6.9 and 6.10, and we can assume that $\theta^n \to \theta$ in Θ^p.

We now show that convergence holds in the $\|C_T(\theta)\|_p$ norm. Through a standard Borel-Cantelli argument (see for instance Shiriayev [26], page 257), we obtain a strategy θ' and a subsequence n_k such that $C_T(\theta^{n_k} - \theta') \to 0$

almost surely. Since θ^n is a Cauchy sequence in $\|C_T(\theta)\|_p$, by the Fatou's Lemma we have

$$E\left[C_T(\theta^m-\theta')^p\right]=E\left[\liminf_{k\to\infty} C_T(\theta^m-\theta^{n_k})^p\right]\leq$$
$$\leq \liminf_{k\to\infty} E\left[C_T(\theta^m-\theta^{n_k})^p\right]<\varepsilon$$

which provides the desired convergence.

□

Remark 3.2 Θ^p_C is not a Hilbert space even for $p=2$. In fact, it is easily checked that the equality $|\theta+\eta|^2+|\theta-\eta|^2=2|\theta|^2+2|\eta|^2$, which is valid in any Hilbert space, is not satisfied from the deterministic strategies $\theta_t=1_{\{t<\frac{1}{3}T\}}$ and $\eta_t=1_{\{t\geq\frac{2}{3}T\}}$.

Remark 3.3 Θ^p_C is generally not *separable*. To see this, observe that the set of deterministic strategies $\{\theta^x\}_{x\in[0,T]}$, where $\theta^x_t=1_{\{t\geq x\}}$, is uncountable, and $\|\theta^x-\theta^y\|_{\Theta^p_C}\geq k_x(\omega)+k_y(\omega)$ for all $x\neq y$ and for all $p\geq 1$. If k is uniformly bounded away from zero (for instance, if k is constant), it follows that $\|\theta^x-\theta^y\|_{\Theta^p_C}\geq c$ for some positive c, which proves the claim.

The following inequality states a continuous immersion of Θ^r_C into Θ^p, for $r>p$, provided that $\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}}$ is sufficiently integrable.

Proposition 3.4 *Let X be a continuous local martingale. For any $p,q\geq 1$, we have:*

$$\|G_T(\theta)\|_p\leq\|C_T(\theta)\|_{pq}\left\|\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}}\right\|_{pq'}$$

where $q'=\frac{q}{q-1}$.

Proof For any p, we have, by the Burkholder-Davis-Gundy inequality:

$$E\left[|G_T(\theta)|^p\right]\leq E\left[\left(\int_0^T\theta_t^2 d\langle X\rangle_t\right)^{\frac{p}{2}}\right]\leq E\left[\left(\sup_{t\leq T}|\theta_t|\right)^p\langle X\rangle_T^{\frac{p}{2}}\right]\leq$$
$$\leq E\left[|D\theta|([0,T])^p\langle X\rangle_T^{\frac{p}{2}}\right]\leq E\left[C_T(\theta)^p\left(\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}(\omega)}\right)^p\right]\leq$$
$$\leq E\left[C_T(\theta)^{pq}\right]^{\frac{1}{q}}E\left[\left(\frac{\sqrt{\langle X\rangle_T}}{\tilde{k}(\omega)}\right)^{pq'}\right]^{\frac{1}{q'}}$$

and, raising both sides to the power $\frac{1}{p}$, the thesis follows.

□

Remark 3.5 Proposition 3.4 admits a simple financial interpretation. The transaction cost needed for a gain with a high moment of order p is bounded if the asset itself has a sufficiently high moment of the same order (a buy and hold strategy does the job). Otherwise, the strategy itself must amplify the swings of the market. In this case, the less the market is volatile, the higher the moment of the strategy.

Remark 3.6 Denoting the $\mathcal{H}^p$ norm of a martingale by $\|M\|_{\mathcal{H}^p} = E\left[\langle X\rangle_T^{\frac{p}{2}}\right]$, it is clear from the proof of Proposition 3.4 that we also have:

$$\|G_T(\theta)|_{\mathcal{H}^p} \leq \|C_T(\theta)\|_{pq} \left\| \frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}} \right\|_{pq'}$$

This is trivial for $p > 1$, as the $\mathcal{H}^p$ norm is equivalent to the L^p norm. On the other hand, the space $\mathcal{H}^1$ is strictly smaller than L^1. In this case, the last observation states that the gain $G_T(\theta)$ belongs to $\mathcal{H}^1$ and, *a fortiori*, is a uniformly integrable martingale.

Remark 3.7 The integrability condition for $\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}}$ in Proposition 3.4 seems not too restrictive. For example, if k is a constant, it boils down to an integrability condition for X. In most models of financial markets, the asset X has finite moments of any order, therefore is is automatically satisfied.

On the other hand, if $k_t = kX_t$, the moments of $\tilde{k}$ can be computed in terms of X_t, and integrability can be obtained via the Hölder inequality.

Remark 3.3 suggests that the norm topology in Θ_C^p is too restrictive to provide sufficient compactness on the space of strategies. The following lemma provides a more reasonable alternative:

Lemma 3.8 *Let X be a continuous local martingale, and $G_T(\theta^n) \to G_T(\theta)$ in the L^p-norm. Then:*

i) if $p > 1$, up to a subsequence $\theta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$;
ii) if $p = 1$, there exists some η^n, convex combinations of stoppings of θ^n, such that $\eta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$.

Proof i) Let τ_k be a reducing sequence of stopping times for the local martingale X_t. For any k, the Burkholder-Davis-Gundy inequality yields:

$$E\left[|G_{T\wedge\tau_k}(\theta^n) - G_{T\wedge\tau_k}(\theta)|^p\right] \geq c_p E\left[\left|\int_0^{T\wedge\tau_k} (\theta_t^n - \theta_t)^2 d\langle X\rangle_t\right|^{\frac{p}{2}}\right] \tag{1}$$

for some positive constant c_p. Since the left-hand side converges to zero, it follows that $\int_0^{T\wedge\tau_k}(\theta_t^n - \theta_t)^2 d\langle X\rangle_t$ also converges to zero in probability, and $\theta^n \to \theta$ in the measure $d\langle X\rangle_t(\omega)P(d\omega)$. Up to a subsequence, convergence holds a.e., and since $d\langle X^{\tau_k}\rangle_t P(d\omega)$ is a sequence of measures increasing to $d\langle X\rangle_t P(d\omega)$, we conclude that $\theta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$.

ii) The situation is more delicate here, because (1) is not true in L^1. Denote by $\tau_k = \inf\{t : |G_t(\theta)| \geq k\}$. The stopped martingales $G_{t\wedge\tau_k}(\theta)$ clearly converge to $G_t(\theta)$ almost surely. For each k, $G_{t\wedge\tau_k}(\theta) \in \mathcal{H}^1$ and we can apply Corollary 6.13, obtaining that for some stopping times $T_{n,k}$, $G_{t\wedge T_{n,k}}(\theta^n) \rightharpoonup G_{t\wedge\tau_k}(\theta)$ in $\sigma(\mathcal{H}^1, BMO)$. Up to a sequence of convex combinations $\xi_t^{n,k} = \sum_{j=n}^{M_n} \beta_j^{n,k} G_{t\wedge T_{j,k}}(\theta^j)$, we can assume that $\xi_t^{n,k} \to G_{t\wedge\tau_k}(\theta)$ in the strong topology of $\mathcal{H}^1$. Observe also that $\xi_t^{n,k} = G_t(\eta^{n,k})$, where $\eta_t^{n,k} = \sum_{j=n}^{M_n} \beta_j^{n,k}\theta_t^j 1_{\{t<T_{j,k}\}}$, and that if $k' < k$, then $(\xi^{n,k})^{\tau_{k'}} \to G_{t\wedge\tau_{k'}}(\theta)$. Hence, by a diagonalization argument, we consider the sequence $\eta^{n,n}$, which satisfies the condition $G_{t\wedge\tau_k}(\eta^{n,n}) \to G_{t\wedge\tau_k}(\theta)$ in the $\mathcal{H}^1$ norm for all k.

In other words:

$$\lim_{n\to\infty} E\left[\left(\int_0^{T\wedge\tau_k} (\eta_t^{n,n} - \theta_t)^2 d\langle X\rangle_t\right)^{\frac{1}{2}}\right] = 0$$

This means that $\eta^{n,n} \to \theta$ in the measure $d\langle X^{\tau_k}\rangle_t P(d\omega)$, and up to a subsequence, a.e. As in *i*), we conclude that $\eta^{n,n} \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$.

□

The next Proposition provides the lower semicontinuity of the cost process, with respect to the convergence in $dtP(d\omega)$. Intuitively, this means that taking limits can only reduce transaction costs, because in the limit strategy some transactions may cancel out, while no new ones can arise. As we shall see later, this property has interesting consequences on the type of risk functions and constrained problems that we are able to solve.

Proposition 3.9 *If θ^n is bounded in Θ_C^p, and $\theta_t^n \to \theta_t$ a.e. in $dtP(d\omega)$ then:*

$$C_T(\theta) \leq \liminf_{n\to\infty} C_T(\theta^n) \qquad \text{for a.e. } \omega \tag{2}$$

and

$$\|C_T(\theta)\|_p \leq \liminf_{n\to\infty} \|C_T(\theta^n)\|_p \tag{3}$$

for all $p \geq 1$.

The proof requires a few lemmas:

Lemma 3.10 *For a fixed ω, let $\theta^n(\omega)_t \to \theta(\omega)_t$ for a.e. t, and $|D\theta(\omega)^n|([0,T]) < C$ uniformly in n. Then $D\theta(\omega)^n \rightharpoonup D\theta(\omega)$ in the weak star topology of Radon measures.*

Proof For all $\phi \in C_c^\infty[0,T]$, we have:

$$\nu(\phi) = \lim_{n\to\infty}\int_{[0,T]} \phi_t dD\theta_t^n = -\lim_{n\to\infty}\int_{[0,T]} \phi_t'\theta_t^n dt = -\int_{[0,T]} \phi_t'\theta_t dt = D\theta(\phi)$$

It remains to show that the distribution $D\theta$ is in fact a Radon measure, and this follows from the inequality:

$$D\theta(\phi) \leq \sup_{t\in[0,T]} |\phi(t)| \limsup_{n\to\infty} |D\theta^n|([0,T]) \leq C \sup_{t\in[0,T]} |\phi(t)|$$

which completes the proof.

□

The following is a standard result in measure theory (see for instance AFP [1]):

Lemma 3.11 *Let $\mu^n \rightharpoonup \mu$, where μ^n, μ are Radon measures on I, and convergence is meant in the weak star sense. Then $|\mu| \leq \liminf_{n\to\infty} |\mu^n|$.*

Proof of Proposition 3.9 By assumption, for a.e. ω, $\theta_t^n(\omega) \to \theta_t(\omega)$ for a.e. t. To prove (2), we show that for all subsequences n_j for which $C_T(\theta^{n_j}(\omega))$ converges, we have:

$$C_T(\theta(\omega)) \leq \lim_{j\to\infty} C_T(\theta^{n_j}(\omega)) \tag{4}$$

If $C_T(\theta^{n_j}(\omega)) \to \infty$, then (4) is trivial. If not, then $C_T(\theta^{n_j}(\omega)) < M(\omega)$ for all j and hence

$$|D\theta^{n_j}(\omega)|([0,T]) < \frac{M(\omega)}{\tilde{k}(\omega)}$$

Lemma 3.10 implies that $D\theta^{n_j}(\omega) \rightharpoonup D\theta(\omega)$. By Lemma 3.11, we obtain:

$$C_T(\theta(\omega)) = \int_{[0,T]} k_t d|D\theta(\omega)| \leq \lim_{j\to\infty} \int_{[0,T]} k_t d|D\theta^{n_j}(\omega)| = \lim_{j\to\infty} C_T(\theta^{n_j}(\omega))$$

and (2) follows. For (3), notice that:

$$\|C_T(\theta(\omega))\|_p^p = E\left[\left(\int_{[0,T]} k_t d|D\theta(\omega)|\right)^p\right] \leq E\left[\liminf_{n\to\infty}\left(\int_{[0,T]} k_t d|D\theta^n(\omega)|\right)^p\right] \leq$$
$$\leq \liminf_{n\to\infty} E\left[\left(\int_{[0,T]} k_t d|D\theta^n(\omega)|\right)^p\right] = \liminf_{n\to\infty} \|C_T(\theta^n(\omega))\|_p^p < \infty$$

where the last inequality follows from the uniform boundedness of θ^n in Θ_C^p, and the previous one holds by Fatou's Lemma.

□

4 Existence of Optimal Strategies

This section contains the main existence results for optimal hedging strategies in unconstrained incomplete markets with proportional transaction costs, in the local martingale case.

In general, the existence of a minimum requires two basic ingredients: relative compactness of minimizing sequences (up to some transformation which leaves them minimizing), and lower semicontinuity of the functional.

Compactness is obviously much easier in L^p spaces with $p > 1$, since it coincides with boundedness, but this kind of information is rare to obtain in applications. On the other hand, some measures of risk (maximization of utility and minimization of shortfall) seem to provide natural bounds on the L^1 norms of optimizing portfolios at expiration.

Moreover, in the next section we shall see that if X is a semimartingale, then a minimization problem can be reduced through a change of measure to a problem in L^1. This shows that the L^1 case is both mathematically more challenging, and the most relevant in applications.

We also need the functional $F : \theta \mapsto \rho(V_T^c(\theta) - H)$ to be lower semicontinuous (shortly l.s.c.). Proposition 3.9 shows that in general $V_T^c(\cdot)$ is upper semicontinuous with respect to a.s. convergence in $dtP(d\omega)$, but not necessarily continuous. This means that we need a decreasing ρ to ensure the semicontinuity of F. Also, we are going to take convex combinations of minimizing strategies, and we need a convex ρ to leave them minimizing. This leads to the following

Definition 4.1 We define a *convex decreasing risk functional* as a function $\rho : L^p \mapsto \mathbb{R} \cup \{+\infty\}$, satisfying the following properties:

i) ρ is convex;
ii) if $X(\omega) \leq Y(\omega)$ for a.e. ω, then $\rho(X) \geq \rho(Y)$ (ρ is decreasing);
iii) ρ has the Fatou property. Namely, if $X_n \to X$ a.e., then

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

Remark 4.2 In the definition above, the critical point is that we require the Fatou property for *any* sequence of random variables representing the terminal wealth of a trading strategy. On the contrary, in the definition of σ-additive coherent risk measure (see Delbaen [11]) the same property is required only for sequences bounded in L^∞.

For $\rho(X) = E[X^-]$ (i.e. minimizing the *shortfall*), *iii*) follows from a straightforward application of Fatou's Lemma. On the other hand, this property does not hold for coherent risk measures, unless additional restrictions on trading strategies are made.

As we shall see in the next section, this difficulty disappears in presence of margin requirements, which force all strategies to be bounded.

We easily see that the above definition provides the desired properties of semicontinuity and convexity:

Lemma 4.3 *Let ρ be a convex decreasing functional, and $c > 0$. Denoting by $H(\theta) = k_T|\theta_T - H_X| + X_T H_X + H_B$ and $F : \theta \mapsto \rho\,(V_T^c(\theta) - H(\theta))$, if $\theta^n \to \theta$ a.e. in $dtP(d\omega)$, we have:*

i) F is convex;
ii) F is l.s.c. with respect to a.s. convergence in $dtP(d\omega)$.

Proof $i)$ Since ρ is convex decreasing, and $V_T^c - H$ is concave, it follows that $F = \rho \circ (V_T^c - H)$ is convex.

$ii)$ By Proposition 3.9, we have:

$$V_T^c(\theta) \geq \limsup_{n\to\infty} V_T^c(\theta^n)$$

Since ρ is decreasing, and $H(\theta)$ is continuous by definition:

$$\rho(V_T^c(\theta) - H(\theta)) \leq \rho\big(\limsup_{n\to\infty}(V_T^c(\theta^n) - H(\theta^n))\big)$$

and finally, by the Fatou property of ρ:

$$\rho(V_T^c(\theta) - H(\theta)) \leq \liminf_{n\to\infty} \rho(V_T^c(\theta^n) - H(\theta^n))$$

□

For convex decreasing functionals we are going to prove an existence result on bounded sets of Θ_C^p. Examples include σ-additive *coherent risk measures* (see Delbaen [11] for details).

A special class of these functionals consists of those which can be written as $\rho(X) = E\,[\nu(X)]$, where $\nu : \mathbb{R} \mapsto \mathbb{R}$ is a convex decreasing function. In this case, we show that an optimal strategy exists in the whole space Θ_C^1, since minimizing sequences are automatically bounded. Both the problems of *shortfall* minimization and *utility* maximization belong to this class.

Throughout this section, we make the following:

Assumption 4.4 The measures $d\langle X\rangle_t(\omega)P(d\omega)$ and $dtP(d\omega)$ are equivalent.

This assumption implies that X cannot have intervals of constancy, and that its bracket process $\langle X\rangle_t$ cannot exhibit a Cantor-ladder behavior. It is necessary to draw inference on $C_T(\theta^n)$, which depends on convergence with respect to the measure $dtP(d\omega)$, from the convergence of $G_T(\theta^n)$, which provides information in the measure $d\langle X\rangle_t P(d\omega)$. In practice, it is satisfied by all diffusion models, even with Hölder coefficients or volatility jumps.

We start with risk minimization in Θ_C^p, with $p > 1$. In this case, we prove the existence of optimal strategies among those with a moment of order p not exceeding M. As a result, the minimum will generally depend on the particular bound considered.

The next lemma provides a class of weakly compact sets:

Lemma 4.5 *For $C \in \mathbb{R}^+$ and $p > 1$, the set*

$$B_{C,D} = \{\theta : \|G_T(\theta)\|_p \leq C, \|C_T(\theta)\|_p \leq D\}$$

is Θ^p-weakly compact for $D \in \mathbb{R}^+ \cup \{+\infty\}$.

Proof Let $\theta^n \in B_{C,D}$. For $p > 1$, Θ^p is a reflexive Banach space (it is isometric to a closed subspace of L^p, which is reflexive). Hence, the set

$$B_C = \{\theta : \|G_T(\theta)\|_p \leq C\}$$

is weakly compact in Θ^p, and up to a subsequence $G_T(\theta^n) \rightharpoonup G_T(\theta) \in B_C$. Since $B_{C,D}$ is convex, by Theorem 6.7 there exists a sequence $\eta^n \in B_{C,D}$ of convex combinations of θ^n, such that $G_T(\eta^n) \to G_T(\theta)$ in L^p. By Lemma 3.8, it follows that, up to a subsequence, $\eta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$, and by Lemma 3.9, we conclude that $\theta \in B_{C,D}$.

□

Proposition 4.6 *Let ρ be a convex decreasing functional, $c > 0$ and $(H_B + X_T H_X, k_T H_X) \in L^p(\Omega, \mathcal{F}_T, P)$, with $p > 1$. For any $M > 0$ let us denote*

$$\Theta^p_{C,M} = \{\theta \in \Theta^p_C,\ \|G_T(\theta)\|_p \leq M\}$$

Then the following minimum problem admits a solution:

$$\min_{\theta \in \Theta^p_{C,M}} \rho\left(V^c_T(\theta) - H(\theta)\right)$$

Proof Let θ^n be a minimizing sequence, so that $F(\theta^n) \to \inf_{\theta \in \Theta^p_{C,M}} F(\theta)$. Since $\Theta^p_{C,M}$ is weakly compact by Lemma 4.5, up to a subsequence we can assume that $\theta^n \rightharpoonup \theta \in \Theta^p_{C,M}$. Then, by Theorem 6.7, there exists a sequence of convex combinations $\eta^n = \sum_{k=n}^{\infty} \alpha^n_k \theta^k$, such that $\eta^n \to \theta$ in the strong topology. By Lemma 3.8, we can assume up to a subsequence that $\eta^n \to \theta$ in the $d\langle X\rangle_t P(d\omega)$-a.s. convergence, and hence $dtP(d\omega)$-a.e. by Assumption 4.4.

From the semicontinuity of F (Lemma 4.3), and standard convexity argument (see for instance Ekeland and Temam [16]), it follows that θ is a minimizer.

□

We now turn to risk minimization in Θ^1_C. As mentioned before, this case has the advantage that minimizing sequences are bounded for some problems considered in applications. On the other hand, a few mathematical issues arise: a bounded sequence in L^1 does not necessarily converge, even in a weak sense, and the L^1 norm of a uniformly integrable martingale is not equivalent to the $\mathcal{H}^1$ norm.

It turns out that the first problem can be overcome through a result of Komlós [21] (see also Schwartz [25], for a shorter proof). We can circumvent the latter at the price of using stopping as a further transformation on minimizing sequences, besides extracting subsequences and taking convex combinations.

We start with the existence result for general convex decreasing risks:

Proposition 4.7 *Let ρ be a convex decreasing functional, $c > 0$ and $(H_B + X_T H_X, k_T H_X) \in L^1(\Omega, \mathcal{F}_T, P)$. For any $M > 0$ the following minimum problem admits a solution:*

$$\min_{\theta \in \Theta^1_{C,M}} \rho\left(V^c_T(\theta) - H(\theta)\right)$$

Proof Let θ^n be a minimizing sequence. By Komlós' Theorem (6.8), up to a subsequence of convex combinations $\eta^n = \sum_{k=n}^{M_n} \alpha^n_k \theta^k$ we can assume that $G_T(\eta^n) \to \gamma$ a.e. and in L^1, and by Yor's Theorem 6.10, there exists some $\theta \in \Theta^1$ such that $\gamma = G_T(\theta)$. To see that $\theta \in \Theta^1_C$, first we apply Lemma 3.8, to obtain a sequence $\zeta^n = \sum_{j=n}^{M_n} \beta^n_j (\eta^j)^{T_{j,n}}$, such that $\zeta^n \to \theta$ a.e. Then Lemma 3.9 implies that $C_T(\theta) \in L^1(P)$, as required.

As in the previous proof, the thesis follows from a standard convexity argument (see Ekeland and Temam [16]).

□

In the special case of ρ being the expectation of a convex decreasing function ν, it turns out that the optimal strategy in Θ^1_C coincides with that in $\Theta^1_{C,M}$, for some value of M. This is shown in the following

Proposition 4.8 *Let $\nu : \mathbb{R} \mapsto \mathbb{R}$ a strictly convex (at least in one point) decreasing function, $c > 0$ and $(H_B + X_T H_X, k_T H_X) \in L^1(\Omega, \mathcal{F}_T, P)$. Then the problem*

$$\min_{\theta \in \Theta^1_C} E\left[\nu\left(V^c_T(\theta) - H(\theta)\right)\right]$$

admits a solution.

Proof Let θ^n be a minimizing sequence, so that $F(\theta^n) \to \inf_{\theta \in \Theta^1_C} F(\theta)$. Since ν is strictly convex, we have that:

$$\nu(x) \geq a + bx^- - (b - \varepsilon)x^+ \tag{5}$$

which implies

$$E\left[Y^+\right] \leq \frac{1}{\varepsilon}\left(E\left[\nu(Y)\right] - a + bE\left[Y\right]\right)$$

for any integrable random variable Y. Substituting $Y = V^c_T(\theta^n) - H(\theta^n)$, we get:

$$E\left[(V^c_T(\theta^n) - H(\theta^n))^+\right] \leq \frac{1}{\varepsilon}\left(E\left[\nu(V^c_T(\theta^n) - H(\theta^n))\right] - a + bE\left[V^c_T(\theta^n) - H(\theta^n)\right]\right)$$

The first term in the right-hand side is bounded by assumption, since θ^n is a minimizing sequence. The second term is also bounded, because $E\left[V^c_T(\theta^n)\right] = c - E\left[C_T(\theta^n)\right] \leq c$ and $H(\theta^n)$ is integrable. Therefore $E\left[(V^c_T(\theta^n) - H(\theta^n))^+\right]$ is bounded and the inequality

$$E\left[V^c_T(\theta^n)^+\right] \leq E\left[(V^c_T(\theta^n) - H(\theta^n))^+\right] + E\left[H(\theta^n)^+\right]$$

implies that $E\left[V_T^c(\theta^n)^+\right]$ is bounded. In a similar fashion, (5) yields:

$$E\left[Y^-\right] \leq \frac{1}{b}(E\left[\nu(Y)\right] + (b-\varepsilon)E\left[Y^+\right] - a)$$

Substituting again $Y = V_T^c(\theta^n) - H(\theta^n)$, and using the foregoing result, we conclude that $\sup_n E\left|V_T^c(\theta^n)\right| < \infty$. Also:

$$E\left|V_T^c(\theta^n)\right| \geq E\left[-V_T^c(\theta^n)\right] = -c + E\left[C_T(\theta^n)\right]$$

which implies that $\sup_n E\left[C_T(\theta^n)\right] < \infty$, and hence $\sup_n E\left[\left|G_T(\theta^n)\right|\right] < \infty$. Now that boundedness is shown, the thesis follows from Proposition 4.7 for a suitable M.

□

Example 4.9 (Shortfall risk) In Proposition 4.8, choosing $\nu(x) = x^-$, we obtain the existence of a *shortfall minimizing* strategy, that is a solution of the problem

$$\max_{\theta \in \Theta_C^1} E\left[(H(\theta) - V_T^c(\theta))^+\right]$$

Without transaction costs, but assuming that X is only a semimartingale, this problem has been solved for European options by Cvitanic and Karatzas [6] in a complete market, and by Cvitanic [4] in incomplete and constrained markets. In both cases, they use the duality approach, as opposed to the Neyman-Pearson lemma approach, employed by Föllmer and Leukert [17] to solve the same problem in an unconstrained incomplete market.

Choosing $\nu(x) = (x^-)^p$, with $p > 1$, one obtains a solution for the problem studied by Pham [23] in discrete time.

Example 4.10 (Utility maximization) Let U be a concave increasing function. The *utility maximization* problem

$$\max_{\theta \in \Theta_C^1} E\left[U\left(V_T^c(\theta) - H(\theta)\right)\right]$$

admits a solution. In fact, apply Proposition 4.8, with $\nu(x) = -U(x)$. This problem has been studied for European options in a Markovian model by Hodges and Neuberger [18] and developed more rigorously by Davis, Panas, and Zariphopoulou [9]: in both papers, a stochastic control problem is considered, and the assumptions on the model lead to a Hamilton-Jacobi-Bellmann equation which can be solved in a weak sense. In more general models, the same problem has been studied by Cvitanic and Karatzas [5] and by Kramkov and Schachermayer [22] with the convex duality approach.

Remark 4.11 The variable H needs not be a function of X_T alone: in fact we only require that it is $\mathcal{F}$-measurable. This means that the existence result is valid for a general path-dependent option, as long as its exercise is fixed at time T. This excludes American-type options.

5 Problems with Constraints

In this section we study the problem of hedging with constraints on the space of strategies. Essentially, we consider two types of constraints: those on the position in the risky asset (such as limits on short-selling), and those on the portfolio value (such as margin requirements).

The existence of a constrained minimum depends on two conditions: the stability of the restricted set of admissible strategies under the transformations used on minimizing sequences, and its closedness in the topology where the risk functional is lower semicontinuous.

In this setting, it becomes evident that the more transformations are used in the proof of the unconstrained problem, the smaller is the set of tractable constraints. Since in the case of L^p we only take convex combinations of strategies, it follows that we can obtain an existence result for constraints of the type $\theta_t \in A_t$, where $A_t(\omega)$ is a closed convex subset of $\mathbb{R}$.

On the contrary, in the L^1 case we also use stopped strategies, hence $A_t(\omega)$ will have to be a closed convex containing zero. At any rate, it seems that these conditions are not restrictive for most applications.

Since a convex closed set in $\mathbb{R}$ is indeed a (possibly unbounded) interval, we shall describe a constraint by means of two processes m_t and M_t, taking values in $\mathbb{R} \cup \{-\infty\}$ and $\mathbb{R} \cup \{+\infty\}$ respectively, and representing the lower and upper bounds of the strategy at time t. Of course, both m_t and M_t must be predictable with respect to $\mathcal{F}_t$.

The next proposition establishes the existence of a minimum for a constraint of the type $\theta_t \in A_t$.

Proposition 5.1 *Let m_t and M_t two predictable processes, such that $m_t \leq \theta_t \leq M_t$. Denoting by*

$$\Gamma^p(m, M) = \{\theta_t \in \Theta^p_C : m_t \leq \theta \leq M_t \textit{ for } dtP(d\omega)\textit{-a.e.}\}$$

if $\Gamma^p(m, M)$ is not empty, we have that:

i) if $p > 1$, then for all $K > 0$ the minimum problem

$$\min_{\theta \in \Theta^p_{C,K} \cap \Gamma^p(m,M)} \rho\left(V^c_T(\theta) - H(\theta)\right)$$

admits a solution.

ii) if $p = 1$, and $m_t \leq 0 \leq M_t$ $dtP(d\omega)$-a.e., then for all $K > 0$ the minimum problem

$$\min_{\theta \in \Theta^1_{C,K} \cap \Gamma^1(m,M)} \rho\left(V^c_T(\theta) - H(\theta)\right)$$

admits a solution.

Proof Reread the proof of Propositions 4.6 and 4.7, observing that convex combinations of strategies in $\Gamma^p(m, M)$ remain in $\Gamma^p(m, M)$ and that (for 4.7) a strategy in $\Gamma^1(m, M)$, if stopped, remains in $\Gamma^1(m, M)$. Finally, if $\theta^n \to \theta$ a.e. in $dtP(d\omega)$, and $\theta^n \in \Gamma^p(m, M)$, then $\theta \in \Gamma^p(m, M)$.

□

Example 5.2 (Short-selling) For $m \equiv 0$ and $M \equiv +\infty$, the constraint above becomes $\theta_t \geq 0$ for all t, which amounts to forbid the short sale of X. Notice that in this case the constraint is deterministic.

Example 5.3 (Margin requirements) Despite the generality of Proposition 5.1, there are some relevant constraints of financial markets which look awkward to embed in the framework above. For instance, consider a solvability condition, which in our notation looks like:

$$V_t^c(\theta) - k_t|\theta_t| \geq -l$$

where l is the maximum credit line available at time t. This constraint is clearly stable with respect to the operations considered, but it does not fit well in Proposition 5.1, where the processes m and M should be written explicitly. Indeed, in this example it can be shown that:

$$(m_t, M_t) = \begin{cases} (-\infty, +\infty) & \text{if } t < \tau \\ (0, 0) & \text{if } t \geq \tau \end{cases}$$

where τ is defined as:

$$\tau = \inf\{t: \ V_t^c(\theta) - k_t|\theta_t| = -l\}$$

In other words, the agent is unconstrained until the solvency limit is hit, then the position must be closed for the rest of the period.

In these cases, where the processes m and M are defined implicitly by an inequality, it makes more sense to check directly that the constraint is stable under convex combinations, limits a.e. in $dtP(d\omega)$ and, in the L^1 case, under stopping.

As mentioned in Remark 4.2, the presence of margin requirements allows to consider a larger class of risk functionals. In fact we have the following:

Theorem 5.4 *Let ρ be a sigma-additive coherent risk measure. In other words, $\rho(X) = \lim_{k\to\infty} \sup_{P\in\mathcal{P}} E_P[-(X \wedge k)]$, where $\mathcal{P}$ is a set of probabilities, all absolutely continuous with respect to P.*

If m and M are defined as in Example 5.3, H is bounded, and $\mathcal{P}$ is weakly relatively compact, the same minimum problems as in Proposition 5.1 admit a solution.

We recall the following Lemma from Delbaen [11]:

Lemma 5.5 *Let $\mathcal{P}$ be a weakly relatively compact set of absolutely continuous probabilities. If X_n is uniformly bounded in $L^\infty(\Omega)$ and $X_n \to X$ a.e., then $\lim_{n\to\infty} \rho(X_n) = \rho(X)$.*

Proof of Theorem 5.4 A coherent risk measure has all the properties of a convex risk functional, except that the Fatou property is satisfied only by bounded sequences of random variables. As a result, the semicontinuity Lemma 4.3 generally fails. We now show that, given a minimizing sequence θ^n, under the additional assumptions above we can obtain another minimizing sequence η^n such that $F : \theta \mapsto \rho(V^c_T(\theta) - H(\theta))$ is lower semicontinuous with respect to η^n.

Let θ^n be a minimizing sequence, and denote by $\overline{\rho} = \inf_n F(\theta^n)$. As in the proofs of Propositions 4.6 and 4.7, up to a subsequence of convex combinations we can assume that $V^c_T(\theta^n) \to V^c_T(\theta)$ a.e. for some admissible strategy θ. Also, by definition of ρ, for each ε there exists some n and k_n such that $\rho((V^c_T(\theta^n) - H(\theta^n)) \wedge k) < \overline{\rho} + \varepsilon$ for all $k > k_n$.

By the assumptions on m and M, $V^c_T(\theta^n) > -l$ for all n. Define now the stopping times $\tau_k = \inf\{t : \ V^c_t(\theta) \geq k\}$. By construction, for the stopped strategies $\eta^{n,k}_t = \theta^n_{t\wedge\tau_k}$ and $\eta^k_t = \theta_{t\wedge\tau_k}$, we have that $V^c_T(\eta^{n,k}) < k$. Now, the sequence $V^c_T(\eta^{n,k})$ is uniformly bounded in L^∞, $\mathcal{P}$ is weakly relatively compact, and $\lim_{n\to\infty} V^c_T(\eta^{n,k})) = V^c_T(\eta^k)) = V^c_T(\theta) \wedge k$ a.e. By the above lemma we have that

$$\lim_{n\to\infty} \rho(V^c_T(\eta^{n,k}) - H(\eta^{n,k})) = \rho(V^c_T(\eta^k) - H(\eta^k)) = \rho(V^c_T(\theta) \wedge k - H(\eta^k))$$

and, since H is bounded,

$$\lim_{k\to\infty} \rho(V^c_T(\eta^k) - H(\eta^k)) = \lim_{k\to\infty} \rho(V^c_T(\theta) \wedge k - H(\eta^k)) = \\ = \lim_{k\to\infty} \rho((V^c_T(\theta) - H(\eta^k)) \wedge k) = \rho(V^c_T(\theta) - H(\eta^k))$$

Therefore η^k is a minimizing sequence, and F is continuous with respect to it.

□

6 The Semimartingale Case

In this section we discuss the problems arising in the more realistic case where X is a semimartingale, and give a partial extension of the foregoing results to this setting.

It is well-known that in a frictionless incomplete market the absence of arbitrage (and thus the well-posedness of a hedging problem) is equivalent to the existence of an equivalent (local) martingale measure (see Delbaen and Schachermayer [13] for a general version of this result). This is no longer true in presence of transaction costs, which may add downside risk to potential arbitrage opportunities, effectively excluding most of them.

In fact, the existence of a martingale measure is a stronger condition than the absence of arbitrage, since it implies that, removing frictions, the same market remains arbitrage-free. Nevertheless, most market models considered in the literature do admit equivalent martingale measures, and here we shall make this additional assumption.

Definition 6.1 We define the following sets of martingale measures:

$$\mathcal{M}_q^e(P) = \{Q \sim P : \frac{dQ}{dP} \in L^q(P), X \text{ is a } Q\text{-local martingale}\}$$

The second issue deals with the space of admissible strategies: if X is a semimartingale, the space:

$$G_T(\Theta^p) = \left\{ \theta : \int_0^T \theta_t dX_t \ \in L^p(P) \right\}$$

is generally not closed, unless additional assumptions are made on X. For $p = 2$, Delbaen, Monat, Schachermayer, Schweizer and Stricker [12] established a necessary and sufficient condition for the closedness of $G_T(\Theta^p)$. Unfortunately, this condition fails to hold for some stochastic volatility models (see for instance Biagini, Guasoni and Pratelli [3] for an example), suggesting that the choice of the space $G_T(\Theta^p)$ may not be satisfactory.

In this spirit, Delbaen and Schachermayer [14] have proposed a different space K_p with better closure properties. We briefly summarize a few facts:

Definition 6.2 Let K_p^s be the set of bounded simple integrals with respect to X (as defined in section 2). K_p denotes the closure of K_p^s in the norm topology of $L^p(P)$.

Remark 6.3 If X is a martingale, it is easily seen that $K_p = G_T(\Theta^p)$. This allows to generalize the definition of Θ_C^p to the semimartingale case.

A short version of the main result of Delbaen and Schachermayer [14] sounds as follows:

Theorem 6.4 *Let $1 \leq p \leq \infty$, and $p' = \frac{p}{p-1}$. If X is a continuous semimartingale locally in $L^p(P)$ such that $\mathcal{M}_{p'}^e(P) \neq \emptyset$, and $f \in L^p(P)$, the following conditions are equivalent:*

i) $f \in K_p$;
ii) There exists a X-integrable predictable process θ such that $G_t(\theta)$ is a uniformly integrable Q-martingale for each $Q \in \mathcal{M}_{p'}^e(P)$, and $G_t(\theta)$ converges to f in the $L^1(Q)$ norm (as t converges to T);
iii) $E_Q[f] = 0$ for each $Q \in \mathcal{M}_{p'}^e(P)$.

Since K_p replaces $G_T(\Theta^p)$ when X is a semimartingale, the definition of Θ_C^p can be extended as follows:

$$\Theta_C^p(P) = \{\theta : G_T(\theta) \in K_p, C_T(\theta) \in L^p(P)\}$$

We now see when a minimization problem of the type:

$$\min_{\theta \in \Theta_C^p(P)} \rho(V_T^c(\theta) - H(\theta))$$

fits into the framework outlined in the previous sections for the martingale case. The idea is to find an auxiliary problem under Q, equivalent to the above problem under P. We have the following:

Proposition 6.5 *Let $1 \leq p \leq \infty$, X a continuous semimartingale locally in $L^p(P)$, and $Q \in \mathcal{M}^e_{p'}(P)$. Then the problem:*

$$\min_{\theta \in \Theta^p_{C,M}(P)} \rho(V^c_T(\theta) - H(\theta)) \tag{S}$$

admits a solution for any $M > 0$.

Proof Since the set $\{\frac{dQ}{dP} : Q \in \mathcal{M}^e_{p'}(P)\}$ is convex and closed in $L^{p'}(P)$, it follows that there exists a countable set of martingale measures $\{Q_i\}_i$, such that $\left\{\frac{dQ_i}{dP}\right\}_i$ is dense in $\{\frac{dQ}{dP} : Q \in \mathcal{M}^e_{p'}(P)\}$ in the $L^{p'}(P)$ norm.

By the Hölder inequality, the identity map $Id : K_p(P) \mapsto G_T(\Theta^1(Q_i))$ is a continuous operator for all i. Note also that if $C_T(\theta) \in L^p(P)$, then $C_T(\theta) \in L^1(Q_i)$.

Let θ^n be a minimizing sequence for (S). Since the set of simple strategies is dense both in $K_p(P)$ and $\Theta^1_{C,M'}(Q_i)$, it follows that θ^n is a minimizing sequence for the problem:

$$\min_{\theta \in \Theta^1_{C,M'}(Q_i)} \rho(V^c_T(\theta) - H(\theta)) \tag{M}$$

for a suitable M'. By Proposition 4.7 we can extract a minimizing sequence of convex combinations of θ^n converging to a minimizer θ of (M). Since minimizing sequences are stable under convex combinations, we can take further subsequences of convex combinations converging to θ in $L^1(Q_i)$ for any finite set of i. By a diagonalization argument, we obtain a η^n such that $\eta^n \to \theta$ in $L^1(Q_i)$ for all i.

If $Q \in \mathcal{M}^e_{p'}(P)$, we can assume up to a subsequence that $\frac{dQ_i}{dP} \to \frac{dQ}{dP}$ in $L^{p'}(P)$. By construction, for all i we have that:

$$E_{Q_i}\left[G_T(\theta)\right] = 0$$

and as $i \to \infty$, we obtain that $E_Q\left[G_T(\theta)\right] = 0$ for any $Q \in \mathcal{M}^e_{p'}(P)$. By Theorem 6.4, it follows that $\theta \in K_p$. Also, Lemma 3.9 implies that $C_T(\theta) \in L^p(P)$. Therefore $\theta \in \Theta^p_C(P)$, and the proof is complete.

□

Remark 6.6 The diagonalization procedure in Proposition 6.5 is necessary since the map $T : K_p(P) \mapsto G_T(\Theta^1(Q))$ is generally not onto (see Delbaen and Schachermayer [14], Remark 2.2 c) for details).

Appendix

We recall here a few results in functional analysis and stochastic integration that we use in the main text. Unlike the previous sections, here T denotes a generic stopping time.

This result dates back to Banach, and is well-known:

Theorem 6.7 *Let x_n be a relatively weakly compact sequence in a Banach space V. Then there exists a sequence of convex combinations $y_n = \sum_{i=n}^{\infty} \alpha_i^n x_n$ such that y_n converges in the norm topology of V.*

Bounded sets are relatively compact in L^p spaces for $p > 1$. For L^1 this is not the case, since its weak star closure leads to the space of Radon measures. Nonetheless, relative compactness can be recovered through convex combinations, as shown by the following:

Theorem 6.8 (Komlós) *Let X_n be a sequence of random variables, such that $\sup_n E|X_n| < \infty$. Then there exists a subsequence X'_n and a random variable $X \in L^1$ such that, for each subsequence X''_n of X'_n,*

$$\frac{1}{n}\sum_{i=1}^{n} X''_n \to X \ \ a.e.$$

The next type of results states when a sequence of stochastic integrals converges to a stochastic integral. Again, the situation is different for $p > 1$ and $p = 1$. The first case is a classic result of stochastic integration, and dates back to Kunita-Watanabe for $p = 2$:

Theorem 6.9 (Kunita-Watanabe) *Let X be a continuous local martingale, θ^n a sequence of X-integrable predictable stochastic processes such that each $\int_0^t \theta_s^n dX_s$ is a L^p-bounded martingale, and such that the sequence $\int_0^\infty \theta_s^n dX_s$ converges to a random variable G in the norm topology of L^p.*

Then there is an $\mathcal{F}^X$-predictable stochastic process θ such that $\int_0^t \theta_s dX_s$ is an L^p-bounded martingale, and such that $\int_0^t \theta_s dX_s = G$.

The case $p = 1$ is due to Yor [27]:

Theorem 6.10 (Yor) *Let X be a continuous local martingale, θ^n a sequence of X-integrable predictable stochastic processes such that each $\int_0^t \theta_s^n dX_s$ is a uniformly integrable martingale, and such that the sequence $\int_0^\infty \theta_s^n dX_s$ converges to a random variable G in the norm topology of L^1 (or even in the $\sigma(L^1, L^\infty)$ topology).*

Then there is an $\mathcal{F}^X$-predictable stochastic process θ such that $\int_0^t \theta_s dX_s$ is a uniformly integrable martingale, and such that $\int_0^t \theta_s dX_s = G$.

The main difference between $p > 1$ and $p = 1$ is that in the latter case the norms $L^p : M \mapsto E\left[M_\infty^p\right]$ and $\mathcal{H}^p : M \mapsto \langle M \rangle_\infty^{\frac{p}{2}}$ are no longer equivalent. Yor's idea is to reduce the L^1 case to $\mathcal{H}^1$ by stopping arguments.

In fact, Theorem 6.10 is a consequence of a more general result (see Yor [27], Theorem 2.4 page 277), which gives further information on converging sequences of uniformly integrable martingales. We report this result with its proof, and an easy corollary used in this paper.

Note that a different proof of 6.10 can be found in Delbaen and Schachermayer [15], with an excellent exposition of the properties of the space $\mathcal{H}^1$.

Theorem 6.11 (Yor) *Let A be a set of uniformly integrable martingales, and denote*

$$\Phi(A) = \{Y^T : \ Y \in A, \ Y^T \in \mathcal{H}^1\}$$

Let Y be a uniformly integrable martingale, $Y^n \in A$ and $Y^n \to Y$ in L^1 (or even weakly in $\sigma(L^1, L^\infty)$). Then, for all stopping times T such that $Y^T \in \mathcal{H}^1$, we have that Y^T belongs to the closure of $\Phi(A)$ in the weak topology $\sigma(H^1, BMO)$.

Proof We separate the proof into three steps.

Step 1: we first show that if the theorem holds for $Y \in \mathcal{H}^1$ and $T = \infty$, then it holds in general. If $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$, then $(Y^n)^T \rightharpoonup Y^T$, and we can apply the theorem to $(Y^n)^T, Y^T$ (since $Y^T \in \mathcal{H}^1$ by assumption) obtaining that Y^T belongs to the closure of $\Phi(\{(Y^n)^T\}_n)$, which is smaller than $\Phi(A)$.

Step 2: we further reduce the proof to the case of $Y^n \in \mathcal{H}^1$. For all n, there exists a sequence of stopping times $S_k^n \to \infty$ such that $(Y^n)^{S_k^n} \in \mathcal{H}^1$. Denoting by $Z_n = (Y^n)^{S_k^n}$, by martingale convergence there exists some $k = k_n$ such that $\|Y_\infty^n - Z_\infty^n\|_{L^1} = \|Y_\infty^n - E\left[Y_\infty^n | \mathcal{F}_{S_k^n}\right]\|_{L^1} \leq \frac{1}{n}$. Therefore, $Z^n \in \mathcal{H}^1$ for all n, and we have, for all $g \in L^\infty$:

$$|E\left[(Z_\infty^n - Y_\infty)g\right]| \leq \|Z_\infty^n - Y_\infty^n\|_{L^1}\|g\|_{L^\infty} + |E\left[(Y_\infty^n - Y_\infty)g\right]|$$

therefore Z_∞^n converges weakly in $\sigma(L^1, L^\infty)$ to Y_∞. Also, the set $\Phi(\{Z^n\}_n)$ is smaller then $\Phi(A)$.

Step 3: we now give the proof under the assumptions $Y^n, Y \in \mathcal{H}^1$, $T = \infty$. It is sufficient to show, for any finite subset $(U^1, \ldots, U^d) \in BMO$, that there exists some stopping time T such that:

$$|E\left[[(Y^n)^T - Y, U^i]_\infty\right]| < \varepsilon \quad \text{for all } i$$

We take $T = \inf\{t : \sum_{i=1}^d |U^i| \geq k\}$, choosing k such that:

$$E\left[\int_{(T,\infty]} |d[Y, U^i]_s|\right] < \frac{\varepsilon}{2} \quad \text{for all } i$$

which is always possible by the Fefferman inequality, since $[Y, U^i]$ has integrable variation, $Y \in \mathcal{H}^1$, and $U^i \in BMO$. Therefore it remains to show that, for some fixed T, and for all $i = 1, \ldots, d$:

$$\lim_{n\to\infty} E\left[[(Y^n)^T, U^i]_\infty\right] = \lim_{n\to\infty} E\left[[Y^n, (U^i)^T]_\infty\right] = E\left[[Y, (U^i)^T]_\infty\right]$$

U^i is bounded in $[0, T)$, but it belongs to BMO, it has bounded jumps, hence it is also bounded on $[0, T]$. As a result, $(U^i)^T$ is bounded. The local martingale $[Y^n, (U^i)^T] - Y^n(U^i)^T$ hence belongs to the class D, and we have:

$$E\left[[Y^n, (U^i)^T]_\infty\right] = E\left[Y_\infty^n U_T^i\right] \quad \text{and} \quad E\left[[Y, (U^i)^T]_\infty\right] = E\left[Y_\infty U_T^i\right]$$

Finally, $U_T^i \in L^\infty$, and the thesis follows from the assumption $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$.

□

Remark 6.12 The statement of Theorem 6.11 with $A = \{Y^n\}_n$ says that, for any $Y^\tau \in \mathcal{H}^1$, there exists a sequence of indices n_k, and a sequence of stopping times τ_k such that $Y^{\tau_k}_{n_k} \rightharpoonup Y^\tau$ in $\sigma(H^1, BMO)$. However, *a priori* the sequence n_k may not tend to infinity, and the the stopping times τ_k may not converge almost surely to τ.

The proof provides more information on these issues. For the first, note that in fact $n_k = k$. For the latter, let us look more closely to the three steps.

Step 1 simply shows that τ_k may be chosen such that $\tau_k \leq \tau$ a.e.

In Step 2, we have $S^n_k \to \infty$, and k_n must be sufficiently high. Therefore we can replace it with some higher k_n such that the condition $P(S^n_{k_n} < n) < \frac{1}{n}$ is satisfied as well.

Likewise, in Step 3 the stopping time T needs to be sufficiently high, hence it may be chosen to satisfy the condition $P(T < n) < \frac{1}{n}$. By a diagonalization argument, we can select a sequence $(Y^n)^{T_n}$ such that $(Y^n)^{T_n} \rightharpoonup Y$ in $\sigma(H^1, BMO)$, and $T_n \to \infty$ a.e.

The previous Remark provides the following:

Corollary 6.13 *If $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$, and $Y^\tau \in \mathcal{H}^1$ for some stopping time τ, then there exists a subsequence n_k and a sequence of stopping times $\tau_k \to \tau$ a.e. such that $(Y^{n_k})^{\tau_k} \rightharpoonup Y^\tau$ in $\sigma(\mathcal{H}^1, BMO)$.*

References

1. L. Ambrosio, N. Fusco, and D. Pallara, *Special functions of bounded variation and free discontinuity problems*, Oxford University Press, 1999.
2. P. Artzner, F. Delbaen, J.M. Eber, and D. Heath, *Coherent measures of risk*, Mathematical Finance **9** (1999), no. 3, 203–228.
3. F. Biagini, P. Guasoni, and M. Pratelli, *Mean-variance hedging for stochastic volatility models*, Mathematical Finance **10** (2000), no. 2, 109–123.
4. J. Cvitanic, *Minimizing expected loss of hedging in incomplete and constrained markets*, SIAM Journal of Control and Optimization (1999), To appear.
5. J. Cvitanic and I. Karatzas, *Hedging and portfolio optimization under transaction costs: a martingale approach*, Mathematical Finance **6** (1996), 133–165.
6. ———, *On dynamic measures of risk*, Finance and Stochastics **3** (1999), no. 4.
7. J. Cvitanic and H. Wang, *On optimal terminal wealth under transaction costs*, Journal of Mathematical Economics (1999), To appear.
8. M.H.A. Davis and A.R. Norman, *Portfolio selection with transaction costs*, Mathematics for Operations Research **15** (1990), 676–713.
9. M.H.A. Davis, V. Panas, and T. Zariphopoulou, *European option pricing with transaction costs*, SIAM Journal of Control and Optimization **31** (1993), no. 2, 470–493.
10. G. Deelstra, H. Pham, and N. Touzi, *Dual formulation of the utility maximization problem under transaction costs*, Preprint, 2000.
11. F. Delbaen, *Coherent risk measures on general probability spaces*, Preprint, March 2000.

12. F. Delbaen, P. Monat, Schachermayer W., M. Schweizer, and C. Stricker, *Weighted norm inequalities and hedging in incomplete markets*, Finance and Stochastics **1** (1997), no. 3, 181–227.
13. F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Mathematischen Annalen **300** (1994), 463–520.
14. ______, *Attainable claims with p-th moments*, Annales Institute H. Poincaré de Probabilites and Statistiques **32** (1996), no. 6, 743–763.
15. ______, *A compactness principle for bounded sequences of martingales with applications*, Preprint, March 1998.
16. I. Ekeland and R. Temam, *Convex analysis and variational problems*, SIAM, 1999.
17. H. Föllmer and P. Leukert, *Efficient hedging: Cost versus shortfall risk*, Finance and Stochastics **4** (2000), no. 2, 117–146.
18. Hodges and Neuberger, *Optimal replication of contingent claims under transaction costs*, Review of Futures Markets **8** (1989), 222–239.
19. I. Karatzas, J.P. Lehoczky, and S.E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM Journal of Control and Optimization **25** (1987), 1557–1586.
20. I. Karatzas, J.P. Lehoczky, S.E. Shreve, and G.L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM Journal of Control and Optimization **29** (1991), no. 3, 702–730.
21. J. Komlós, *A generalization of a problem of steinhaus*, Acta Mathematica Academiae Scientiarum Hungaricae **18** (1967), no. 1-2, 217–229.
22. D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, Annals of Applied Probability **9** (1999), 904–950.
23. H. Pham, *Dynamic l^p hedging in discrete time under cone constraints*, Preprint, 1998.
24. W. Schachermayer, *Optimal investment in incomplete markets when wealth may become negative*, Working paper no. 65 Vienna University of Technology, January 2000.
25. M. Schwartz, *New proofs of a theorem of komlós*, Acta Mathematica Hungarica **47** (1986), no. 1-2, 181–185.
26. A.N. Shiryaev, *Probability*, Springer-Verlag, 1984.
27. M. Yor, *Sous-espaces denses dans l^1 ou h^1 et reprèsentation des martingales*, Séminaire de Probabilité XII, Lecture Notes in Mathematics, vol. 649, Springer-Verlag, 1978, pp. 265–309.