HEDGING BY SEQUENTIAL REGRESSIONS REVISITED

ALEŠ ČERNÝ AND JAN KALLSEN

City University London and CAU Kiel

Abstract. Almost 20 years ago Föllmer and Schweizer (1989) suggested a simple and influential scheme for the computation of hedging strategies in an incomplete market. Their approach of *local* risk minimization results in a sequence of one-period least squares regressions running recursively backwards in time. In the meantime there have been significant developments in the *global* risk minimization theory for semimartingale price processes. In this paper we revisit hedging by sequential regression in the context of global risk minimization, in the light of recent results obtained by Černý and Kallsen (2007). A number of illustrative numerical examples is given.

1. Introduction

It is hard to imagine modern financial markets without derivative securities. From equity and currency options, to interest rate swaptions, to exotics, derivatives are an important tool for financial risk sharing. From the buyer's point of view derivative securities provide significant reduction in risk with a relatively small initial outlay. The issuer, on the other hand, can take comfort in the result of Black and Scholes (1973) who show that issuer's exposure can be offset by frequent enough trading in the underlying asset. In practice, however, the Black–Scholes result cannot and should not be taken literally, because it is well documented that most asset price dynamics are inconsistent with a pure diffusion process. It may be the case that frequent hedging of a derivative security removes most of issuer's risk, but such conclusion can only be reached after a detailed investigation of the hedging error *outside* of the Black–Scholes model in an environment which allows for price jumps. To succeed in this task one needs to understand how to compute good hedging strategies in a situation where perfect replication is *not* guaranteed a-priori.

In an influential article Föllmer and Schweizer (1989) argued that the Black–Scholes result can be viewed as a special case of a sequential least squares regression whose aim is to minimize recursively the one-step expected squared hedging error. This procedure is easy to visualize and implement in practice and it will recover the per-

---

Date: February 4, 2007.

2000 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.

Key words and phrases. option, hedging, CAPM, sequential regression, opportunity-neutral measure.

We would like to thank two anonymous referees for their thoughtful comments.

fect replicating portfolio when such a portfolio exists. In subsequent work by Föllmer and Schweizer (1991) and Schweizer (1991) it is shown that the concept of sequential regressions can be extended to semimartingale models using the notion of local risk minimization in which the minimal martingale measure and the Föllmer–Schweizer (F–S) decomposition play a central role. But while in a finite state model the minimal martingale measure always exists, it may happen that in a perfectly reasonable arbitrage-free model with continuous asset prices it does not, see Delbaen and Schachermayer (1998).

As an alternative to local risk minimization one may consider the minimization of the *unconditional* expected squared hedging error

$$\inf_{\vartheta} E((v + \vartheta \bullet S_T - H)^2), \tag{1.1}$$

where v is an (admissible) initial endowment, ϑ is an (admissible) trading strategy, S is a stock price, H is a contingent claim to be hedged and $\vartheta \bullet S_T$ represents gains from trading in the time interval $[0, T]$. Here v, S, H are expressed in terms of an appropriate numeraire, most commonly the risk-free bank account. For ease of exposition we only consider one risky asset in the main body of the paper, relegating the multivariate case to section 8.

Criterion (1.1) is more appealing than local risk minimization, because after all one cares about the total hedging error and not the daily profit–loss ratios. The reason why global risk minimization has not been used more ubiquitously up until now is that its solution is generally considered more involved compared to the approach of Föllmer and Schweizer. The mathematical history of the solution to (1.1) is traced in Pham (2000) and Schweizer (2001). Černý and Kallsen (2007) show that (1.1) admits a solution in a very general class of arbitrage-free semimartingale models where local risk minimization may fail to be well defined.

The purpose of this paper is threefold. Firstly, we will demonstrate that in discrete time the solution of (1.1) is as simple as the solution to local risk minimization of Föllmer and Schweizer and can be implemented by means of a sequential regression. We show that there are two differences between the F–S decomposition and the globally optimal regression: the former uses two explanatory variables (safe and risky returns) and it is always performed under the objective measure P; the latter uses only one explanatory variable (risky return) and is performed under so-called opportunity-neutral measure $P^{\star}$ which may or may not coincide with P. Crucially, the global risk minimizing strategy is always well defined and hence it provides a more robust theoretical concept compared to the local risk minimization. Secondly, we highlight the link between (1.1) and globally mean–variance efficient portfolios which simplifies and extends the analysis of Li and Ng (2000) and Leippold et al. (2004). Finally, we translate the general semimartingale setup of Černý and Kallsen (2007) into discrete time and draw comparison with the existing literature.

The paper is organized as follows. Section 2 introduces notation and assumptions. In section 3 we study hedging in a one-period model, establishing basic properties of least squares coefficients and some connections to the Capital Asset Pricing Model.

Section 4 considers a multiperiod model with IID stock returns and examines, in turn, the Föllmer–Schweizer sequential regression, its relation to the F–S decomposition, and the globally optimal sequential regression. Section 5 considers a model with non-IID returns and constructs the corresponding opportunity process and the opportunity-neutral measure $P^\star$. In section 6 we provide interpretation of the opportunity process in terms of unconditional Sharpe ratios, and compute the globally mean–variance efficient portfolio. Section 7 contrasts our approach with the Gourieroux–Laurent–Pham numeraire method. Most proofs and technicalities are deferred to section 8.

2. Notation and assumptions

Consider a time horizon $T \in \mathbb{N}$ and the set of trading dates $\mathcal{T} := \{0, 1, \ldots, T\}$. We fix a probability space $(\Omega, P, \mathcal{F})$, a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathcal{T}}$, $\mathcal{F}_T = \mathcal{F}$, and an $\mathcal{F}_T$-measurable contingent claim $H \in L^2(P)$. We introduce the following notation for conditional expectations,

$$\begin{aligned} E_t(X) &:= E(X|\mathcal{F}_t), \\ \mathrm{Var}_t(X) &:= E_t(X^2) - (E_t(X))^2 . \end{aligned}$$

The discounted stock price process $\{S_t\}_{t\in\mathcal{T}}$ is adapted to $\mathbb{F}$ and we assume throughout that S is locally square-integrable, i.e. for $\Delta S_{t+1} := S_{t+1} - S_t$ we have

$$E_t((\Delta S_{t+1})^2) < \infty \text{ for } t < T.$$

This assumption is weaker than the commonly encountered requirement $S_t \in L^2(P)$ for $t \in \mathcal{T}$, cf. Hipp (1993), Melnikov and Nechaev (1999), Schäl (1994), Schweizer (1995).

Definition 2.1. *We say that process S admits no arbitrage, if for all $t \in \mathcal{T} \setminus \{0\}$ and all $\mathcal{F}_{t-1}$-measurable portfolios ϑ_t we have that $\vartheta_t \Delta S_t \geq 0$ a.s. implies $\vartheta_t \Delta S_t = 0$ a.s.*

We assume that S is arbitrage-free in the sense of the above definition. Strictly speaking one can define a solution of (1.1) *without* the no-arbitrage requirement (cf. Melnikov and Nechaev 1999) but such extension, while mathematically elegant, does not bring additional economic insight.

Definition 2.2. *We say that (v, ϑ) is an* admissible *endowment–strategy pair if and only if v is $\mathcal{F}_0$-measurable, $\vartheta = \{\vartheta_t\}_{t\in\mathcal{T}\setminus\{0\}}$ is predictable, meaning that ϑ_t is $\mathcal{F}_{t-1}$-measurable, and*

$$v + \vartheta \bullet S_T := v + \sum_{t=1}^{T} \vartheta_t \Delta S_t \in L^2(P).$$

The set of admissible trading strategies with initial endowment v is denoted $\overline{\Theta}(v)$. We write $\overline{\Theta}$ as a shorthand for $\overline{\Theta}(0)$.

From now we take $\mathcal{F}_0$ trivial to simplify the exposition. We will revert to general $\mathcal{F}_0$ in section 8.

3. Mean–variance hedging and the CAPM model

Consider $T = 1$ and a contingent claim $H \in L^2(P)$. Set $V_1 := H$ and define

$$\{V_0, \xi_1\} := \arg\min_{v,\vartheta_1} E((v + \vartheta_1 \Delta S_1 - V_1)^2).$$

By standard least squares arguments we have

$$\begin{aligned} \xi_1 &= \frac{\mathrm{Cov}(V_1, \Delta S_1)}{\mathrm{Var}(\Delta S_1)}, && (3.1)\\ V_0 &= E(V_1) - \xi_1 E(\Delta S_1). && (3.2) \end{aligned}$$

Next we will provide alternative expressions for V_0 and ξ_1 which, although less immediately obvious, are more useful than (3.1), (3.2).

Consider an auxiliary regression of the constant onto the explanatory variable ΔS_1,

$$\tilde{\lambda}_1 := \arg\min_{\vartheta_1 \in \mathbb{R}} E((\vartheta_1 \Delta S_1 - 1)^2) = \frac{E(\Delta S_1)}{E((\Delta S_1)^2)},$$

and denote by $\Delta \tilde{K}_1$ the sum of explained squares from this auxiliary regression,

$$1 - \Delta \tilde{K}_1 := \min_{\vartheta_1 \in \mathbb{R}} E((\vartheta_1 \Delta S_1 - 1)^2) = 1 - \tilde{\lambda}_1 E(\Delta S_1) = 1 - \frac{(E(\Delta S_1))^2}{E((\Delta S_1)^2)}.$$

Then using the Frisch–Waugh–Lovell theorem (cf. Davidson and MacKinnon 1993, p.20) we can obtain V_0 from the regression of V_1 onto the residuals from the auxiliary regression,

$$V_0 = \arg\min_{v} E((v(1 - \tilde{\lambda}_1 \Delta S_1) - V_1)^2) = E\left(\frac{1 - \tilde{\lambda}_1 \Delta S_1}{1 - \Delta \tilde{K}_1} V_1\right). \tag{3.3}$$

With V_0 known one can recover ξ_1 from a univariate regression,

$$\xi_1 = \arg\min_{\vartheta_1} E((V_0 + \vartheta_1 \Delta S_1 - V_1)^2) = \frac{E((V_1 - V_0)\Delta S_1)}{E((\Delta S_1)^2)}. \tag{3.4}$$

Remark 3.1. *An easy calculation shows that equation (3.3) is equivalent to the CAPM pricing formula for derivative asset H with S_1/S_0 being the return on the market portfolio and 1 the risk-free return. Indeed on dividing equation (3.2) by V_0 and substituting for ξ_1 from (3.1) we obtain*

$$\begin{aligned} E\left(V_1/V_0\right) &= 1 + \frac{\mathrm{Cov}\left(V_1/V_0, \Delta S_1\right)}{\mathrm{Var}\left(\Delta S_1\right)} E(\Delta S_1) \\ &= 1 + \frac{\mathrm{Cov}\left(V_1/V_0, S_1/S_0\right)}{\mathrm{Var}\left(S_1/S_0\right)} E(S_1/S_0 - 1), \end{aligned}$$

which yields the desired CAPM formula.

Remark 3.1 highlights that the (possibly signed) measure Q defined by

$$\frac{dQ}{dP} := \frac{1 - \tilde{\lambda}_1 \Delta S_1}{1 - \Delta \tilde{K}_1}$$

is a martingale measure. To see this mathematically we observe,

$$
\begin{aligned}
E\left(\frac{dQ}{dP}\right) &= E\left(\frac{1-\tilde{\lambda}_1\Delta S_1}{1-\Delta\tilde{K}_1}\right) = 1,\\
E^Q\left(\Delta S_1\right) &= E\left(\frac{1-\tilde{\lambda}_1\Delta S_1}{1-\Delta\tilde{K}_1}\Delta S_1\right) = \frac{E\left(\Delta S_1\right)-\tilde{\lambda}_1 E((\Delta S_1)^2)}{1-\Delta\tilde{K}_1} = 0.
\end{aligned}
$$

The first result states that Q has total mass 1, whereas the second asserts that the stock is priced correctly by Q.

The quantity

$$
\Delta\hat{K}_1 := \frac{(E(\Delta S_1))^2}{\mathrm{Var}\left(\Delta S_1\right)} = \frac{(E(\Delta S_1))^2}{E((\Delta S_1)^2)-(E(\Delta S_1))^2} = \frac{\Delta\tilde{K}_1}{1-\Delta\tilde{K}_1}
$$

is the squared market Sharpe ratio of the stock return. The portfolio weight $\tilde{\lambda}_1$ represents the optimal number of shares bought by a quadratic utility investor with unit initial wealth and unit relative risk aversion, cf. Černý (2004b, Chapter 3).

Example 3.2. *Consider a one-period model where $S_0 > 0$ and the one-period return S_1/S_0 takes three values, $(0.9\ 1.2\ 1.6)$, with probabilities $p = \left(\frac{2}{3}\,\frac{1}{12}\,\frac{1}{4}\right)$. Denote the excess return $X := S_1/S_0 - 1$. Then we have $E(X) = E(X^2) = 0.1$, and*

$$
\begin{aligned}
\tilde{\lambda}_1 &= \frac{E(X)}{E(X^2)S_0} = \frac{1}{S_0},\\
\Delta\tilde{K}_1 &= \frac{(E(X))^2}{E(X^2)} = 0.1.
\end{aligned}
$$

Denoting by Q the CAPM risk-neutral probability measure we have,

$$
\frac{dQ}{dP} = \frac{1-XE(X)/E(X^2)}{1-\Delta\tilde{K}_1} = \left(\frac{11}{9}\ \frac{8}{9}\ \frac{4}{9}\right),
$$

and the CAPM risk-neutral probabilities read $q = \frac{dQ}{dP}p = \left(\frac{22}{27}\ \frac{2}{27}\ \frac{1}{9}\right)$. The squared market Sharpe ratio equals $\Delta\hat{K}_1 = \Delta\tilde{K}_1/(1-\Delta\tilde{K}_1) = 1/9$.

4. Model with IID stock returns

Suppose now that $S_0 > 0$ and $\{R_t\}_{t\in\mathcal{T}}$ is a collection of IID random variables with finite second moment such that $R_t > 0$ almost surely. Define

$$
\begin{aligned}
S_t &:= S_0\prod_{j=1}^{t} R_j \text{ for } t \geq 1,\\
\mu &:= E(R_t),\\
\sigma^2 &:= \mathrm{Var}(R_t).
\end{aligned}
$$

4.1. **Local risk minimization by sequential regression.** In a dynamic model one can perform the least squares regressions outlined in the previous section recursively, by defining

$$\begin{aligned}\{V_{t-1},\xi_t\} &:= \arg\min_{v_{t-1},\vartheta_t}\left\{E_{t-1}((v_{t-1}+\vartheta_t\Delta S_t-V_t)^2) : v_{t-1},\vartheta_t \text{ are } \mathcal{F}_{t-1}\text{-measurable}\right\},\\ V_T &:= H. \end{aligned} \tag{4.1}$$

This approach is taken in Föllmer and Schweizer (1989). Assuming, for the time being, that the process V thus obtained is well defined we have that

$$V_{t-1}=E_{t-1}\left(\frac{1-\tilde{\lambda}_t\Delta S_t}{1-\Delta\tilde{K}_t}V_t\right)=:E^Q_{t-1}(V_t), \tag{4.2}$$

where Q is the so-called minimal martingale measure,

$$\frac{dQ}{dP}:=\prod_{t=1}^T\frac{1-\tilde{\lambda}_t\Delta S_t}{1-\Delta\tilde{K}_t},$$

and the quantities

$$\tilde{\lambda}_t := \frac{E_{t-1}(\Delta S_t)}{E_{t-1}((\Delta S_t)^2)}, \tag{4.3}$$

$$\Delta\tilde{K}_t := \frac{(E_{t-1}(\Delta S_t))^2}{E_{t-1}((\Delta S_t)^2)}=1-E_{t-1}((1-\tilde{\lambda}_t\Delta S_t)^2), \tag{4.4}$$

reflect the amount and performance of myopic one-period investment in the stock. For future reference we denote the one-period realized locally optimal hedging error by e_t,

$$e_t := V_{t-1}+\xi_t\Delta S_t-V_t. \tag{4.5}$$

The locally optimal hedging coefficient ξ is obtained from a conditional version of (3.1, 3.4)

$$\xi_t=\frac{\mathrm{Cov}_{t-1}(V_t,\Delta S_t)}{\mathrm{Var}_{t-1}(\Delta S_t)}=\frac{E_{t-1}((V_t-V_{t-1})\,\Delta S_t)}{E_{t-1}((\Delta S_t)^2)}. \tag{4.6}$$

Throughout in the notation we suppress the explicit dependence of V,ξ,e on the contingent claim H.

Extending the analysis of Föllmer and Schweizer (1989) one can now ask what is the *unconditional* hedging error of a self-financing strategy starting with capital v and with ξ_t shares bought at time $t-1$. Denote the value of this portfolio at time t by $G^{v,\xi}_t := v+\xi\bullet S_t$ and set $V_T=H$. By the law of iterated expectations and the self-financing property, $G^{v,\xi}_T=G^{v,\xi}_{T-1}+\xi_T\Delta S_T$, we have

$$\begin{aligned}E((G^{v,\xi}_T-V_T)^2) &= E(E_{T-1}((G^{v,\xi}_T-V_T)^2))\\ &= E(E_{T-1}((G^{v,\xi}_{T-1}-V_{T-1}+V_{T-1}+\xi_T\Delta S_T-V_T)^2)). \end{aligned} \tag{4.7}$$

Since V_{T-1} and ξ_T are the least squares coefficients from local risk minimization at time $T-1$ the realized locally optimal hedging error $e_T = V_{T-1} + \xi_T \Delta S_T - V_T$ must be orthogonal to 1 implying $E_{T-1}(e_T) = 0$. Consequently (4.7) yields

$$E((G_T^{v,\xi} - V_T)^2) = E((G_{T-1}^{v,\xi} - V_{T-1})^2 + \psi_T), \tag{4.8}$$

where we have defined

$$\psi_t := E_{t-1}(e_t^2) = \mathrm{Var}_{t-1}(V_t) - \xi_t \mathrm{Cov}_{t-1}(\Delta S_t, V_t). \tag{4.9}$$

After recursive application of (4.8) one obtains

$$E((G_T^{v,\xi} - V_T)^2) = (v - V_0)^2 + \sum_{t=1}^{T} E(\psi_t). \tag{4.10}$$

In words, the unconditional expected squared hedging error equals the sum of one-period expected squared hedging errors plus the square of initial misalignment $v - V_0$.

Example 4.1. *Consider a two-period ($T = 2$) trinomial model where the one-period returns take three values, $S_t/S_{t-1} = (0.9\ 1.2\ 1.6)$, with conditional probabilities $p_t = \left(\frac{2}{3}\,\frac{1}{12}\,\frac{1}{4}\right)$ as in Example 3.2. Denote by X_t the excess return $X_t := S_t/S_{t-1} - 1$. Then we have*

$$\begin{aligned}
E_{t-1}(X_t) &= E_{t-1}(X_t^2) = 0.1,\\
\tilde{\lambda}_t &= \frac{E_{t-1}(X_t)}{E_{t-1}(X_t^2)S_{t-1}} = \frac{1}{S_{t-1}},\\
\Delta\tilde{K}_t &= \frac{(E_{t-1}(X_t))^2}{E_{t-1}(X_t^2)} = 0.1,\ \textit{for } t = 1, 2.
\end{aligned}$$

The one-period change of measure reads

$$\frac{1 - \tilde{\lambda}_t \Delta S_t}{1 - \Delta\tilde{K}_t} = \frac{1 - \frac{E_{t-1}(X_t)}{E_{t-1}(X_t^2)} X_t}{1 - \Delta\tilde{K}_t} = \begin{pmatrix} \frac{11}{9} & \frac{8}{9} & \frac{4}{9} \end{pmatrix},$$

and the conditional risk-neutral probabilities of the minimal martingale measure are

$$q_t = \frac{1 - \tilde{\lambda}_t \Delta S_t}{1 - \Delta\tilde{K}_t} p_t = \begin{pmatrix} \frac{22}{27} & \frac{2}{27} & \frac{1}{9} \end{pmatrix}.$$

As a consequence the stock returns remain IID under Q.

Consider a European call option with strike $K = 108$ expiring at $T = 2$. The intercepts V arising from the local risk minimization are computed recursively from $T = 2$ using (4.2). For example, the value of V in the middle node at time 1 is given by

$$(V_1 | S_1 = 120) = E^Q(V_2 | S_1 = 120) = 12.$$

Both the stock price lattice and the corresponding mean value process V are given in Figure 1. Next we will evaluate the locally optimal hedging coefficient ξ using the

FIGURE 1. Stock price and mean value process.

S

t=0	t=1	t=2
		256
	160	192
100	120	144
	90	108
		81

V

t=0	t=1	t=2
		148
	52	84
268/27	12	36
	4	0
		0

second expression in (4.6). For example the value of ξ_2 in the lowest node at time $t = 1$ equals,

$$(\xi_2|S_1 = 90) = \frac{E((V_2 - V_1)\,\Delta S_2|S_1 = 90)}{E((\Delta S_2)^2|S_1 = 90)} = \frac{5}{9}.$$

The locally optimal hedge ξ is depicted in Figure 2.

FIGURE 2. The locally optimal hedge ξ and the one-period expecteed squared hedging error of an initially perfectly balanced hedging portfolio ψ.

xi

t=1	t=2
	1
91/135	1
	5/9

psi

t=1	t=2
	0
968/81	0
	18

Now one can evaluate the conditional expected one-period squared hedging error of a perfectly balanced initial position (one where $G_{t-1} = V_{t-1}$). For example, in the lowest node at $t = 1$ the one-step locally optimal realized hedging error equals

$$(e_2|S_1 = 90) = (V_2 - V_1 - \xi_2\Delta S_2|S_1 = 90) = \begin{pmatrix} 1 & -14 & 2 \end{pmatrix},$$

and the conditional expected squared hedging error is therefore

$$(\psi_2|S_1 = 90) = E((V_2 - \xi_2\Delta S_2 - V_1)^2|S_1 = 90) = 18.$$

Values of ψ are depicted in Figure 2.

Finally, using equation (4.10), we evaluate the unconditional expected squared hedging error of the hedging strategy ξ, assuming that $v = V_0$,

$$E((V_0 + \xi \bullet S_2 - H)^2) = \psi_1 + E(\psi_2) = 23\frac{77}{81}. \tag{4.11}$$

4.2. **Föllmer–Schweizer decomposition.**

Definition 4.2.

(1) *We say that a contingent claim* $H \in L^2(P)$ *has an* extended *F–S decomposition if there is* $v \in \mathbb{R}$, *a predictable process* ϑ *and a square-integrable martingale* N *starting at* 0 *such that* $H = v + \vartheta \bullet S_T + N_T$ *and* $\mathrm{Cov}_{t-1}(\Delta S_t, \Delta N_t) = 0$ *for all* $t \in \mathcal{T} \setminus \{0\}$.

(2) *We say that* H *has a standard F–S decomposition if in addition* $\vartheta \bullet S_t \in L^2(P)$ *for all* $t \in \mathcal{T}$.

We will now relate the sequential regressions of the previous section to the F–S decomposition. If the process V defined in (4.2) is square-integrable then

$$H = V_0 + \xi \bullet S_T + N_T,$$

with ξ given in (4.6) and $\Delta N_t = e_t = V_{t-1} + \xi_t \Delta S_t - V_t$, is the standard F–S decomposition of the contingent claim H. The unconditional squared hedging error of a self-financing strategy (V_0, ξ) can be expressed in terms of the process N,

$$E((V_0 + \xi \bullet S_T - H)^2) = \sum_{t=1}^{T} E(\psi_t) = E(N_T^2).$$

In example 8.6 we will construct a price process S and a contingent claim H which does not have the standard F–S decomposition but admits the extended decomposition. In example 8.9 we will exhibit a two-period model and a contingent claim $H \in L^2(P)$ for which even the extended decomposition fails to exist and the local risk minimization is no longer well defined.

4.3. **Global risk minimization by sequential regression.** Let us now examine the solution to the global risk minimization

$$\min_{\vartheta} E((G_T^{v,\vartheta} - V_T)^2), \qquad V_T = H,$$

with the initial wealth v fixed. To facilitate the exposition we assume that the globally optimal strategy exists, deferring the proof of its existence to section 8.2. The optimal strategy is denoted by $\varphi(v)$ and we again suppress its explicit dependence on H in the notation. Using the law of iterated expectations, the definition of a self-financing strategy, and the optimality of φ we obtain

$$E((G_T^{v,\varphi(v)} - V_T)^2) = E(\min_{\vartheta_T} E_{T-1}((G_{T-1}^{v,\varphi(v)} + \vartheta_T \Delta S_T - V_T)^2)).$$

Recall the locally optimal strategy ξ defined in (4.1, 4.6). It transpires that ξ_T *cannot* be globally optimal in general, because by its construction it will solve

$$\min_{\vartheta_T} E_{T-1}((G_{T-1}^{v,\varphi(v)} + \vartheta_T \Delta S_T - V_T)^2) \tag{4.12}$$

only if

$$G_{T-1}^{v,\vartheta} = V_{T-1}, \text{ or} \tag{4.13}$$

$$E_{T-1}(\Delta S_T) = 0. \tag{4.14}$$

Unless the contingent claim is perfectly replicable one cannot rely on the special case (4.13). This, incidentally, explains why the hedging strategy in Föllmer and Schweizer (1989) is *not* self-financing but only "mean self-financing", i.e. G_t equals V_t *on average* if one starts with $G_{t-1} = V_{t-1}$. In an incomplete market the hedging portfolio G_{T-1} may frequently undershoot or overshoot the target value V_{T-1} and the globally optimal hedge φ must take this fact into account. The second special case corresponding to (4.14) was discussed already by Föllmer and Sondermann (1986) for a general square-integrable martingale S.

From (4.12) the globally optimal hedging strategy in the final period, $t = T$, reads

$$\varphi_t(v) = \arg\min_{\vartheta_t} E_{t-1}((G_{t-1}^{v,\varphi(v)} + \vartheta_t \Delta S_t - V_t)^2), \tag{4.15}$$

which once again represents a least squares regression, but this time *without* an intercept because the value of $G_{t-1}^{v,\varphi(v)}$ is given by past trading performance and the constraint of self-financing strategies prevents the hedger to add or withdraw funds along the way. Thus in the globally optimal regression the dependent variable is $V_t - G_{t-1}^{v,\varphi(v)}$, the explanatory variable equals ΔS_t and no intercept is present. By standard univariate regression we obtain

$$\begin{aligned} \varphi_t(v) &= \frac{E_{t-1}((V_t - G_{t-1}^{v,\varphi(v)})\Delta S_t)}{E_{t-1}((\Delta S_t)^2)} \\ &= \frac{E_{t-1}((V_t - V_{t-1})\Delta S_t)}{E_{t-1}((\Delta S_t)^2)} + \tilde{\lambda}_t (V_{t-1} - G_{t-1}^{v,\varphi(v)}) \\ &= \xi_t + \tilde{\lambda}_t (V_{t-1} - G_{t-1}^{v,\varphi(v)}), \end{aligned} \tag{4.16}$$

where the last line follows from (4.6).

To evaluate the hedging error of strategy $\varphi(v)$ we substitute (4.16) back into the right hand side of (4.15), adding and subtracting V_{t-1},

$$\begin{aligned} G_{t-1}^{v,\varphi(v)} + \varphi_t(v)\Delta S_t - V_t &= (G_{t-1}^{v,\varphi(v)} - V_{t-1})(1 - \tilde{\lambda}_t \Delta S_t) \\ &\quad + V_{t-1} + \xi_t \Delta S_t - V_t. \end{aligned} \tag{4.17}$$

By construction of V_{t-1} and ξ_t the realized locally optimal hedging error is orthogonal to 1 and ΔS_t, implying

$$E_{t-1}((1 - \tilde{\lambda}_t \Delta S_t)\,(V_{t-1} + \xi_t \Delta S_t - V_t)) = 0. \tag{4.18}$$

Equations (4.17) and (4.18) yield

$$E_{t-1}((G_{t-1}^{v,\varphi(v)} + \varphi_t(v)\Delta S_t - V_t)^2) = (1 - \Delta\tilde{K}_t)(G_{t-1}^{v,\varphi(v)} - V_{t-1})^2 + \psi_t, \tag{4.19}$$

with ψ defined in equation (4.9).

Most importantly, in an IID case the quantity $L_{T-1} := 1 - \Delta\tilde{K}_T$ is deterministic and therefore the optimization at time $t = T-1$ is essentially the same as the optimization at T, i.e.

$$\begin{aligned}\varphi_t(v) &= \arg\min_{\vartheta_t} E_{t-1}(L_t(G_{t-1}^{v,\varphi(v)} + \vartheta_t \Delta S_t - V_t)^2 + \psi_{t+1}) \\ &= \arg\min_{\vartheta_t} L_t E_{t-1}((G_{t-1}^{v,\varphi(v)} + \vartheta_t \Delta S_t - V_t)^2) \\ &= \arg\min_{\vartheta_t} E_{t-1}((G_{t-1}^{v,\varphi(v)} + \vartheta_t \Delta S_t - V_t)^2)\end{aligned}$$

because ψ_{t+1} does not depend on ϑ and L_t is deterministic. On defining

$$L_t = \prod_{j=t+1}^{T} (1 - \Delta \tilde{K}_j), \qquad L_T = 1, \tag{4.20}$$

and after recursive application of (4.19) we obtain

$$E((G_T^{v,\varphi(v)} - V_T)^2) = L_0 (v - V_0)^2 + \sum_{t=1}^{T} E(L_t \psi_t). \tag{4.21}$$

Example 4.3. *Consider the setup of Example 4.1. We have* $L_t = 0.9^{T-t}$ *for* $t = 0, 1, 2$. *Assuming* $v = V_0$ *we find* $\varphi_1(V_0) = \xi_1 = 55/81$,

$$G_1^{V_0,\varphi(V_0)} = V_0 + \xi_1 \Delta S_1 = \begin{pmatrix} 1\frac{44}{81} & 21\frac{74}{81} & 49\frac{2}{27} \end{pmatrix},$$

and therefore

$$\varphi_2(V_0) = \xi_2 + \tilde{\lambda}_2(V_1 - G_1^{V_0,\varphi(V_0)}) = \begin{pmatrix} \frac{4249}{7290} & \frac{8917}{9720} & \frac{4399}{4320} \end{pmatrix} \approx \begin{pmatrix} 0.583 & 0.917 & 1.018 \end{pmatrix}.$$

Referring to (4.21) the unconditional expected squared hedging error of the globally optimal strategy equals

$$E((G_2^{V_0,\varphi(V_0)} - H)^2) = L_1 \psi_1 + E(L_2 \psi_2) = 22\frac{34}{45},$$

which is indeed less than the corresponding value for the locally optimal strategy ξ *in equation (4.11).*

Further numerical examples are available in Černý (2004b, Chapter 12).

5. Stochastic opportunity set

In previous sections we have considered a multi-period stock price model with IID returns and reviewed the computation of hedging strategies by *sequential regressions* due to Föllmer and Schweizer (1989). We have noted that the hedging strategy resulting from the Föllmer–Schweizer sequential regression will not minimize the unconditional hedging risk, because it chooses at each node a specific value for the intercept (intercept being the value of the replicating portfolio), whereas in reality this value is given by the past trading performance. We have modified the Föllmer–Schweizer sequential regression by removing the intercept to obtain the globally optimal hedging strategy.

If the stock returns are not IID the hedging formula (4.15) in general fails to yield the globally optimal hedging strategy, except for $t = T$. In this section we describe the final adjustment to the recursive least squares procedure, which is needed to handle the general case of non-IID returns. The final modification involves a change of measure from the original probability P to a new probability measure $P^\star$. In contrast to much of financial literature the new probability measure $P^\star$ is *not* a martingale measure, in the sense that taking expectations under this measure will *not* generate arbitrage-free prices.

The purpose of measure $P^\star$ is to internalize the stochastic changes in the *multi-period* Sharpe ratio which is related to process L in equation (4.20) (see section 6 for more details). We call L the *opportunity process*. If under P some states offer higher Sharpe ratio in the future than others, then the one-period realized hedging error in those states can be higher because better investment opportunities in the future will allow to make up for the higher error today. Whereas the least squares under P fail to incorporate the changing investment opportunities, $P^\star$ modifies the probability weights so that the least squares regression without intercept (4.15) taken under $P^\star$ yields the globally optimal hedging strategy. For this reason we call $P^\star$ the *opportunity-neutral measure*. Mildly extending the standard economic terminology we can talk of deterministic/predictable/stochastic opportunity set when the corresponding opportunity process is deterministic/predictable/stochastic.

Consider the problem

$$\min_{\vartheta_t} E_{t-1}(L_t(G_t^{v,\vartheta} - V_t^\star)^2) = \min_{\vartheta_t} E_{t-1}(L_t(G_{t-1}^{v,\vartheta} + \vartheta_t \Delta S_t - V_t^\star)^2). \tag{5.1}$$

Assume that L_t is $(0,1]$-valued and that $E_{t-1}(L_t(V_t^\star)^2) < \infty$. When L_t is stochastic as of $t-1$ equation (5.1) still represents a well-defined least squares regression because $E_{t-1}(L_t(\Delta S_t)^2) < \infty$. One can internalize the random weights L_t in a new probability measure $P^\star$ such that for any $\mathcal{F}_t$-measurable random variable Z we have

$$E_{t-1}^{P^\star}(Z) = E_{t-1}(L_t Z) / E_{t-1}(L_t). \tag{5.2}$$

From the Bayes' law we then see that the right definition of $P^\star$ is

$$dP^\star/dP = \prod_{t=1}^{T} L_t / E_{t-1}(L_t).$$

Now (5.1) reads $E_{t-1}(L_t) \min_{\vartheta_t} E_{t-1}^{P^\star}((G_{t-1}^{v,\vartheta} + \vartheta_t \Delta S_t - V_t^\star)^2)$. The solution of

$$\min_{\vartheta_t} E_{t-1}^{P^\star}((G_{t-1}^{v,\vartheta} + \vartheta_t \Delta S_t - V_t^\star)^2)$$

can be found from (4.2-4.6, 4.16, 4.19) if we replace P with $P^\star$. Specifically, on defining

$$\tilde{\lambda}_t^\star := \frac{E_{t-1}^{P^\star}(\Delta S_t)}{E_{t-1}^{P^\star}((\Delta S_t)^2)}, \tag{5.3}$$

$$\Delta \tilde{K}_t^\star := \frac{\left(E_{t-1}^{P^\star}(\Delta S_t)\right)^2}{E_{t-1}^{P^\star}((\Delta S_t)^2)}, \tag{5.4}$$

$$
\begin{aligned}
V_{t-1}^{\star} &:= E_{t-1}^{P^{\star}}\left(\frac{1-\tilde{\lambda}_t^{\star}\Delta S_t}{1-\Delta\tilde{K}_t^{\star}}V_t^{\star}\right) =: E_{t-1}^{Q^{\star}}\left(V_t^{\star}\right), && (5.5)\\
\frac{dQ^{\star}}{dP^{\star}} &:= \prod_{t=1}^{T}\frac{1-\tilde{\lambda}_t^{\star}\Delta S_t}{1-\Delta\tilde{K}_t^{\star}}, && (5.6)\\
\xi_t^{\star} &:= \frac{\operatorname{Cov}_{t-1}^{P^{\star}}\left(V_t^{\star},\Delta S_t\right)}{\operatorname{Var}_{t-1}^{P^{\star}}\left(\Delta S_t\right)} = \frac{E_{t-1}^{P^{\star}}\left(\left(V_t^{\star}-V_{t-1}^{\star}\right)\Delta S_t\right)}{E_{t-1}^{P^{\star}}((\Delta S_t)^2)}, && (5.7)\\
\psi_t^{\star} &:= E_{t-1}^{P^{\star}}(\left(V_t^{\star}-\xi_t^{\star}\Delta S_t-V_{t-1}^{\star}\right)^2). && (5.8)
\end{aligned}
$$

we conclude that the globally optimal strategy and the resulting hedging error satisfy

$$
\varphi_t(v) = \arg\min_{\vartheta_t} E_{t-1}^{P^{\star}}((G_{t-1}^{v,\varphi(v)}+\vartheta_t\Delta S_t-V_t^{\star})^2) = \xi_t^{\star}+\tilde{\lambda}_t^{\star}(V_{t-1}-G_{t-1}^{v,\varphi(v)}), \tag{5.9}
$$

$$
\begin{aligned}
E_{t-1}(L_t(G_t^{v,\varphi(v)}-V_t^{\star})^2) &= E_{t-1}(L_t)((1-\Delta\tilde{K}_t^{\star})(V_{t-1}^{\star}-G_{t-1}^{v,\varphi(v)})^2+\psi_t^{\star})\\
&= E_{t-1}(L_{t-1}(V_{t-1}^{\star}-G_{t-1}^{v,\varphi(v)})^2+L_t\psi_t^{\star}), && (5.10)
\end{aligned}
$$

where we have defined

$$
L_{t-1} := (1-\Delta\tilde{K}_t^{\star})E_{t-1}(L_t), \qquad L_T := 1. \tag{5.11}
$$

On applying (5.2)-(5.11) recursively with $V_T^{\star} := H$ we obtain

$$
\begin{aligned}
\min_{\vartheta} E(L_T(G_T^{v,\vartheta}-V_T^{\star})^2) &= E((G_T^{v,\varphi(v)}-V_T)^2)\\
&= L_0\left(v-V_0\right)^2+\sum_{t=1}^{T}E(L_t\psi_t^{\star}). && (5.12)
\end{aligned}
$$

A rigorous proof that all the quantities above are well-defined is given in section 8.2.

Example 5.1. *Consider the two-period trinomial model of Example 4.1, with modified objective probabilities. Suppose that in the up node at time* $t=1$ *the conditional distribution of one-period return* $R = (0.9\ 1.2\ 1.6)$ *is* $p_{2u} := (\frac{2}{3}\ \frac{1}{12}\ \frac{1}{4})$ *whereas in the middle and down node at time* 1 *it is* $p_{2m} = p_{2d} := (\frac{4}{5}\ \frac{1}{10}\ \frac{1}{10})$. *At time zero the conditional probability distribution of one-period return is assumed to be* $p_1 := (\frac{72}{91}\ \frac{9}{91}\ \frac{10}{91})$.

Since $L_2 = 1$ *we have* $p_2^{\star} = p_2$, *for all three conditional distributions of the one-period return at* $t=1$. *In the up node at time 1 the situation is the same as in*

Example 4.1, yielding

$$\begin{aligned}
(\tilde{\lambda}_2^\star|S_1 &= 160) = \frac{1}{160},\\
(\Delta\tilde{K}_2^\star|S_1 &= 160) = 0.1,\\
(L_1|S_1 &= 160) = 0.9,\\
(V_1^\star|S_1 &= 160) = 52,\\
(\xi_2^\star|S_1 &= 160) = 1,\\
(\psi_2^\star|S_1 &= 160) = 0.
\end{aligned}$$

In the middle and down nodes at time 1 *the stock price behaves like a martingale, i.e. we have* $E_1(\Delta S_2) = 0$*, implying* $\tilde{\lambda}_2^\star = 0$*,* $q^\star = p^\star$ *and*

$$\begin{aligned}
(L_1|S_1 = 120) &= (L_1|S_1 = 90) = 1,\\
(V_1^\star|S_1 = 120) &= E\left(V_2^\star|S_1 = 120\right) = 12,\\
(V_1^\star|S_1 = 90) &= E\left(V_2^\star|S_1 = 90\right) = 3\frac{3}{5},\\
(\xi_2^\star|S_1 = 120) &= \frac{E\left(\left(V_2^\star - V_1^\star\right)\left(S_2 - S_1\right)|S_1 = 120\right)}{E((S_2 - S_1)^2\,|S_1 = 120)} = 1,\\
(\xi_2^\star|S_1 = 90) &= \frac{E\left(\left(V_2^\star - V_1^\star\right)\left(S_2 - S_1\right)|S_1 = 90\right)}{E((S_2 - S_1)^2\,|S_1 = 90)} = \frac{1}{2},\\
(\psi_2^\star|S_1 = 120) &= E((V_2^\star - \xi_2^\star\Delta S_2 - V_1^\star)^2\,|S_1 = 120) = 0,\\
(\psi_2^\star|S_1 = 90) &= E((V_2^\star - \xi_2^\star\Delta S_2 - V_1^\star)^2\,|S_1 = 90) = 19\frac{11}{25}.
\end{aligned}$$

At $t = 0$ *the opportunity process one period ahead is stochastic,* $L_1 = (1\ 1\ 9/10)$*, and therefore the opportunity-neutral measure* $P^\star$ *will differ from the objective measure* P*. Specifically we have*

$$\frac{p_1^\star}{p_1} = \frac{L_1}{E(L_1)} = \begin{pmatrix} \frac{91}{90} & \frac{91}{90} & \frac{91}{100} \end{pmatrix},$$

implying $p_1^\star = \left(\frac{4}{5}\ \frac{1}{10}\ \frac{1}{10}\right)$*. Hence at time* $t = 0$ *we have* $E^{P^\star}(\Delta S_1) = 0, \tilde{\lambda}_1^\star = 0$ *and* $q^\star = p^\star$ *(but* $p^\star \neq p$*). This means*

$$\begin{aligned}
\Delta\tilde{K}_1^\star &= 0,\\
L_0 &= (1 - \Delta\tilde{K}_1^\star)E(L_1) = \frac{90}{91},\\
V_0 &= E^{P^\star}(V_1^\star) = 9\frac{7}{25},\\
\xi_1^\star &= \frac{E^{P^\star}\left(\left(V_1^\star - V_0^\star\right)\Delta S_1\right)}{E^{P^\star}((\Delta S_1)^2)} = \frac{16}{25},\\
\psi_1^\star &= E^{P^\star}((V_0^\star + \xi_1\Delta S_1 - V_1^\star)^2) = 12\frac{276}{625}.
\end{aligned}$$

By virtue of (5.12) the unconditional expected squared hedging error of the globally optimal strategy equals

$$E((G_T^{V_0^\star,\varphi(V_0^\star)} - V_T^\star)^2) = E\left(L_1\psi_1^\star + L_2\psi_2^\star\right) = 27\frac{7803}{11\,375}.$$

6. Opportunity process and the Sharpe ratio

In this section we will show that the opportunity process L is closely related to the Sharpe ratio of the globally optimal investment in the underlying asset.

Definition 6.1. *We call*

$$\rho := \sup\left\{\frac{E(\vartheta \bullet S_T)}{\sqrt{\operatorname{Var}(\vartheta \bullet S_T)}} : \vartheta \in \overline{\Theta}\right\}$$

the maximal unconditional Sharpe ratio, *where we set $\frac{0}{0} := 0$. Here $\overline{\Theta}$ is the set of admissible strategies with zero initial endowment from definition 2.2.*

Theorem 6.2. *The maximal unconditional Sharpe ratio is given by*

$$\rho^2 = L_0^{-1} - 1.$$

Consider the contingent claim $H = 1$ and the corresponding globally optimal hedging strategy $\varphi(0)$ with initial wealth 0,

$$\varphi_t(0) = \tilde{\lambda}_t^\star(1 - G_{t-1}^{0,\varphi(0)}). \tag{6.1}$$

$\varphi(0)$ regarded as an investment strategy is globally mean–variance efficient with unconditional moments

$$\begin{aligned} E(G_T^{0,\varphi(0)}) &= 1 - L_0, \\ \operatorname{Var}(G_T^{0,\varphi(0)}) &= L_0\left(1 - L_0\right). \end{aligned}$$

Proof. Easily,

$$\rho^2(X) := \frac{(E(X))^2}{\operatorname{Var}(X)} = \frac{1}{\inf_{\alpha\in\mathbb{R}}\{E\left((1-\alpha X)^2\right)\}} - 1 = \sup_{\alpha\in\mathbb{R}}\left\{\frac{1}{E\left((1-\alpha X)^2\right)} - 1\right\}.$$

For $X = \vartheta \bullet S_T$ we have

$$\begin{aligned} \rho^2 &= \sup_{\vartheta\in\overline{\Theta}}\{\rho^2(\vartheta \bullet S_T)\} = \sup_{\alpha\in\mathbb{R},\vartheta\in\overline{\Theta}}\left\{\frac{1}{E\left((1-\alpha(\vartheta \bullet S_T))^2\right)} - 1\right\} \\ &= \frac{1}{\inf_{\vartheta\in\overline{\Theta}}\{E\left((1-\vartheta \bullet S_T)^2\right)\}} - 1 = \frac{1}{L_0} - 1, \end{aligned}$$

where the last equality follows from (5.12) with contingent claim $H = 1$ and initial wealth $v = 0$. This also shows that $\varphi(0)$ in (6.1) is a globally mean–variance efficient investment strategy. An easy calculation (see 8.17) yields

$$E((1 - G_T^{0,\varphi(0)})G_T^{0,\varphi(0)}) = 0,$$

while (5.12) implies $E((1-G_T^{0,\varphi(0)})^2)=L_0$. From here

$$\begin{aligned} E(G_T^{0,\varphi(0)}) &= E((G_T^{0,\varphi(0)})^2)=1-L_0, \\ \operatorname{Var}(G_T^{0,\varphi(0)}) &= E((G_T^{0,\varphi(0)})^2)-(E(G_T^{0,\varphi(0)}))^2=L_0\,(1-L_0)\,. \end{aligned}$$

□

Remark 6.3. *Theorem 6.2 simplifies and generalizes characterization of globally mean–variance efficient portfolios in discrete time (cf. Li and Ng 2000 and Leippold et al. 2004). Continuous-time version of this result can be found in Černý and Kallsen (2007, Proposition 3.6) and Černý and Kallsen (2006). By contrast to global optimality, there is a parallel strand of literature where portfolio allocation and/or derivative pricing is based on local maximization of utility or of the Sharpe ratio (cf. Björk and Slinko 2006, Černý 2003, Christensen and Platen 2007, Cochrane and Saá-Requejo 2000, Kallsen 1999 and Kallsen 2002).*

7. The Gourieroux–Laurent–Pham change of numeraire method

The purpose of this section is to relate our sequential regressions to the numeraire method proposed by Gourieroux et al. (1998). We do not cover all technical details; these can be found in Arai (2005).

Suppose the wealth process $G^{0,\varphi(0)}$ defined in Theorem 6.2 is strictly less than 1. Then the variance-optimal measure $Q^\star$ is equivalent to P and its density process is given by the formula

$$E(dQ^\star/dP|\mathcal{F}_t)=L_t(1-G_t^{0,\varphi(0)})/L_0.$$

Gourieroux et al. proposed to take the process $1-G_t^{0,\varphi(0)}$ as a new numeraire, defining

$$\begin{aligned} \bar{S}_t &:= [1\ S_t]/(1-G_t^{0,\varphi(0)}), \\ d\bar{P}/dP &= (1-G_T^{0,\varphi(0)})^2/L_0. \end{aligned}$$

Exploiting the fact that $G^{0,\varphi(0)}$ is a $Q^\star$-martingale we find that the density process of the new measure $\bar{P}$ reads

$$E(d\bar{P}/dP|\mathcal{F}_t)=L_t(1-G_t^{0,\varphi(0)})^2/L_0.$$

Extending standard change of numeraire arguments Gourieroux et al. show the equivalence

$$\begin{aligned} \inf_{\vartheta} E((v+\vartheta\bullet S_T-H)^2) &= L_0\inf_{\bar{\vartheta}} E^{\bar{P}}((v+\bar{\vartheta}\bullet\bar{S}_T-\bar{H})^2), \\ \bar{H} &:= H/(1-G_T^{0,\varphi(0)}), \end{aligned}$$

for suitably chosen sets of admissible strategies. The equivalence is not as simple as it might seem because the dimension of $\bar{S}$ exceeds the dimension of S by one and special care has to be taken when interpreting $\bar{\vartheta}$ as a self-financing strategy.

S is a $Q^\star$-local martingale which implies $\bar{S}$ is a $\bar{P}$-local martingale, hence the solution of the mean–variance hedging problem under $\bar{P}$ becomes very simple, with

$\bar{L} = 1$, $\tilde{\bar{\lambda}} = 0$, $(\bar{P})^\star = \bar{P} = (\bar{Q})^\star$ and $\bar{V}_t = E^{\bar{P}}(\bar{H}|\mathcal{F}_t)$, corresponding to the solution of Föllmer and Sondermann (1986). Both the conditional and unconditional Sharpe ratios under $\bar{P}$ are zero, and myopic hedging with $\bar{S}$ under $\bar{P}$ is also globally optimal. In a model with IID returns this comes at the cost of more complicated dynamics of $\bar{S}$ and $\bar{V}$, notably $\bar{V}$ is now path-dependent. Also the relationship between the optimal strategy $\bar{\varphi}$ in the modified problem and the optimal strategy φ of the original problem is not straightforward.

By contrast, our solution does not require a change of numeraire and yields all components of the optimal strategy directly. Prices S are not $P^\star$-local martingales and in general $(P^\star)^\star \neq P^\star$ so one cannot claim that the opportunity set generated by S becomes deterministic (or even predictable) under $P^\star$. It is true, however, that myopic hedging under $P^\star$ leads to globally optimal hedging under P. Another important difference is that we obtain $\tilde{\lambda}^\star$ (Schweizer's adjustment process) as part of the recursive solution while in the numeraire method $\tilde{\lambda}^\star$ (and consequently $G_T^{0,\varphi(0)}$) are treated as inputs to be obtained elsewhere.

8. Technicalities

8.1. **Admissible strategies.** In this section we examine the notion of admissibility from definition 2.2 in the context of a general semimartingale model. Lemma A.2 in Černý and Kallsen (2007) shows that for every locally square-integrable semimartingale S there is an increasing sequence of stopping times $\{U_n\}_{n\in\mathbb{N}}$ converging to ∞ P-almost surely such that

$$\sup\{E\left(S_\tau^2\right) : \tau \leq U_n \text{ stopping time}\} < \infty$$

for any $n \in \mathbb{N}$.

The following two definitions taken from Černý and Kallsen (2007) can be applied equally in discrete or continuous time as long as the stock price process is a locally square-integrable semimartingale.

Definition 8.1. *A trading strategy ϑ is called* simple *if it is a linear combination of strategies $Y 1_{]\!]\tau_1,\tau_2]\!]}$ where $\tau_1 \leq \tau_2$ are stopping times dominated by U_n for some $n \in \mathbb{N}$ and Y is a bounded $\mathcal{F}_{\tau_1}$-measurable random variable. The set of terminal wealths attainable by simple self-financing strategies with initial endowment $v \in L^2(\Omega, \mathcal{F}_0, P)$ is denoted by*

$$K_2^s(v) := \{v + \vartheta \bullet S_T : \vartheta \textit{ simple }\}.$$

If the initial endowment is not fixed beforehand, we consider instead the set

$$K_2^s(\mathcal{F}_0) := \{v + \vartheta \bullet S_T : v \in L^2(\Omega, \mathcal{F}_0, P),\ \vartheta \textit{ simple }\}.$$

Since the hedging problems below concern the approximation of an arbitrary payoff H in $L^2(P)$ it makes perfect sense from an economic standpoint to view elements of the $L^2(P)$ closures

$$\begin{aligned} K_2(v) &:= \overline{K_2^s(v)}, \\ K_2(\mathcal{F}_0) &:= \overline{K_2^s(\mathcal{F}_0)}, \end{aligned}$$

as being attainable by a self-financing strategy. It is less immediately obvious that one can extend the definition of the closed subspace $K_2(v) \subseteq K_2(\mathcal{F}_0)$ to some initial endowments $v \notin L^2(\Omega, \mathcal{F}_0, P)$ as follows,

$$K_2(v) := \{H \in L^2(P) : \text{there exist } H^{(n)} \in K_2^s\left(v^{(n)}\right) \text{ with } H^{(n)} \xrightarrow{L^2(P)} H, v^{(n)} \xrightarrow{P} v\}.$$

Here $\xrightarrow{P}$ denotes convergence in probability. For the next definition we recall that if S is a semimartingale then $L(S)$ denotes the set of S-integrable predictable processes in the sense of Jacod and Shiryaev (2003), III.6.17.

Definition 8.2. *We call* $(v,\vartheta) \in L^0(\Omega, \mathcal{F}_0, P) \times L(S)$ *an* admissible endowment–strategy pair *if there exist some sequences* $(v^{(n)})_{n\in\mathbb{N}}$ *in* $L^2(\Omega, \mathcal{F}_0, P)$ *and* $(\vartheta^{(n)})_{n\in\mathbb{N}}$ *of simple strategies such that*

$$\begin{aligned} v^{(n)} + \vartheta^{(n)} \bullet S_t &\to v + \vartheta \bullet S_t \text{ in probability for any } t \in [0,T] \text{ and} \\ v^{(n)} + \vartheta^{(n)} \bullet S_T &\to v + \vartheta \bullet S_T \text{ in } L^2(P). \end{aligned}$$

We set

$$\begin{aligned} \overline{\Theta(v)} &:= \{\vartheta \in L(S) : (v,\vartheta) \text{ admissible}\}, \\ \overline{\Theta} &:= \overline{\Theta(0)}, \\ \overline{U} &:= \{v \in L^0(\Omega, \mathcal{F}_0, P) : \overline{\Theta(v)} \text{ not empty}\}, \\ \overline{U \times \Theta} &:= \{(v,\vartheta) \in L^0(\Omega, \mathcal{F}_0, P) \times L(S) : (v,\vartheta) \text{ admissible}\}. \end{aligned}$$

Clearly $\overline{U} \supseteq L^2(\Omega, \mathcal{F}_0, P)$. One easily verifies that $\overline{U \times \Theta} = \overline{U} \times \overline{\Theta} = \mathbb{R} \times \overline{\Theta}$ if the initial σ-field $\mathcal{F}_0$ is trivial. We recall the following result from Černý and Kallsen (2007).

Lemma 8.3. *1. For any* $v \in \overline{U}$ *the set* $K_2(v)$ *is closed in* $L^2(P)$ *and one has*

$$K_2(v) = \{v + \vartheta \bullet S_T : \vartheta \in \overline{\Theta(v)}\}.$$

For $v \in L^2(P)$ *we have in addition* $K_2(v) = \{v + \vartheta \bullet S_T : \vartheta \in \overline{\Theta}\}$.
2. $K_2(\mathcal{F}_0)$ *is closed in* $L^2(P)$ *and*

$$K_2(\mathcal{F}_0) = \{v + \vartheta \bullet S_T : (v,\vartheta) \text{ admissible}\}.$$

Definition 8.2 provides a general notion of admissibility for locally square-integrable semimartingale price processes. We will now demonstrate that the general definition is consistent with the simpler discrete-time definition 2.2 introduced at the beginning of the paper.

Proposition 8.4. *In a discrete-time model we have*

(1) $\overline{\Theta} = \{\vartheta$ *predictable*: $\vartheta \bullet S_T \in L^2(P)\}$
(2) $\overline{U \times \Theta} = \{(v,\vartheta) : v$ *is* $\mathcal{F}_0$*-measurable,* ϑ *is predictable,* $v + \vartheta \bullet S_T \in L^2(P)\}$

Proof. 1. The inclusion $\subseteq$ is obvious. To prove $\supseteq$ consider two arrays of natural numbers $\{K(t,m), n(t,m)\}_{t\in\mathcal{T},m\in\mathbb{N}}$ and for $i,j \in \mathcal{T}$, $m \in \mathbb{N}$ define random variables

$$\xi_{ij}^{(m)} := \begin{cases} \prod_{t=i}^{j} 1_{|\vartheta_t|<K(t,m)} 1_{t<U_{n(t,m)}} & \text{for } i \leq j \\ 1 & \text{for } i > j \end{cases},$$

where $|\vartheta_t|$ is the Euclidean norm of vector ϑ_t. We have

$$\xi_{j,t}^{(m)} \vartheta_t \Delta S_t \in L^2(P) \text{ for } j \leq t, \tag{8.1}$$

because $\xi_{j,t}^{(m)} \vartheta_t \Delta S_t$ is the gain from a simple strategy corresponding to $\tau_1 = \bigwedge_{i=j}^{t} U_{n(i,m)} \wedge (t-1)$, $\tau_2 = \bigwedge_{i=j}^{t} U_{n(i,m)} \wedge t$ and $Y = \vartheta_t \prod_{i=j}^{t} 1_{|\vartheta_i|<K(i,m)}$. Define

$$Y_j := \xi_{1,j-1}^{(m)} \sum_{t=j}^{T} \vartheta_t \Delta S_t \text{ for } j = 1, 2, \ldots, T.$$

A short calculation shows

$$Y_{j+1} = \xi_{j,j}^{(m)} Y_j - \xi_{1,j}^{(m)} \vartheta_j \Delta S_j \tag{8.2}$$

which together with (8.1) yields

$$Y_j \in L^2(P) \Rightarrow Y_{j+1} \in L^2(P) \text{ for } j = 1, 2, \ldots, T-1. \tag{8.3}$$

By assumption $Y_1 = \vartheta \bullet S_T \in L^2(P)$ and by virtue of (8.3) we have

$$Y_j \in L^2(P) \text{ for } j = 1, 2, \ldots, T. \tag{8.4}$$

Define $\vartheta_t^{(m)} := \xi_{1,t}^{(m)} \vartheta_t$. By construction $\vartheta^{(m)}$ is simple. We will now choose $\{K(t,m)\}_{t\in\mathcal{T},m\in\mathbb{N}}$, $\{n(t,m)\}_{t\in\mathcal{T},m\in\mathbb{N}}$ such that the strategy $\vartheta^{(m)}$ is an approximating sequence to ϑ as required by Definition 8.2. Equation (8.2) can be rearranged to yield

$$\vartheta_j \Delta S_j + Y_{j+1} - Y_j + (1 - \xi_{j,j}^{(m)}) Y_j = (1 - \xi_{1,j}^{(m)}) \vartheta_j \Delta S_j = \vartheta_j \Delta S_j - \vartheta_j^{(m)} \Delta S_j$$

and after summation from 1 to T we have

$$\vartheta \bullet S_T - \vartheta^{(m)} \bullet S_T = \sum_{j=1}^{T} (1 - \xi_{j,j}^{(m)}) Y_j. \tag{8.5}$$

Since $Y_1 \in L^2(P)$ we can now choose $K(1,m) > m$ and $n(1,m) > m$ such that $E((1-\xi_{1,1}^{(m)})^2 Y_1^2) < m^{-1}T^{-1}$. The values of $K(1,m)$ and $n(1,m)$ determine Y_2 and we then find $K(2,m) > m$ and $n(2,m) > m$ such that $E((1-\xi_{2,2}^{(m)})^2 Y_2^2) < m^{-1}T^{-1}$, etc. Such $K(i,m), n(i,m)$ exist by virtue of (8.4) and dominated convergence. Equation (8.5) implies

$$E((\vartheta \bullet S_T - \vartheta^{(m)} \bullet S_T)^2) < m^{-1},$$

and hence $\vartheta^{(m)} \bullet S_T \to \vartheta \bullet S_T$ in $L^2(P)$. Since $\lim_{m\to\infty} K(t,m) = \lim_{m\to\infty} n(t,m) = \infty$ for all t and $U_n \to \infty$ we have $\vartheta^{(m)} \bullet S_t \to \vartheta \bullet S_t$ a.s. and therefore in probability for all $t \in \mathcal{T}$.

2. The proof is analogous, but we start with $j = 0$ and define $\xi^{(m)}_{0,0} := 1_{|v|<K(0,m)}$. □

Remark 8.5. *1. Schweizer (1995) defines the following class of admissible strategies*

$$\Theta := \{\vartheta \text{ predictable} : \vartheta \bullet S_t \in L^2(P) \text{ for all } t \in \mathcal{T}\}.$$

He notes that the mean–variance hedging problem (1.1) may not have a solution in Θ*, and provides an example to that effect. His example is* not *arbitrage-free and therefore it does not fit the framework of the present paper.*

In Example 8.6 we use the main idea of Schweizer's example to construct an arbitrage-free *model with* $T = 3$*,* $\mathcal{F}_0$ *trivial,* $S \in L^2(P)$ *and* $\bar\vartheta$ *predictable such that* $H := \bar\vartheta \bullet S_3 \in L^2(P)$ *but* $\bar\vartheta \bullet S_2 \notin L^2(P)$*. By construction we have* $H \neq \vartheta \bullet S_3$ *for* $\vartheta \neq \bar\vartheta$*, but at the same time it follows from the proof of Proposition 8.4 that one can approximate* $\bar\vartheta \bullet S_3$ *with arbitrary precision in* $L^2(P)$ *by using simple strategies. Hence the expected squared hedging error can be made arbitrarily close to zero but never exactly zero within Schweizer's class of hedging strategies. This also shows that* H *does not have the standard F–S decomposition but it admits the extended F–S decomposition in the sense of Definition 4.2.*

2. Melnikov and Nechaev (1999) show that in discrete time mean–variance hedging can always be solved in the class of predictable strategies from Definition 2.2, under the somewhat stronger assumption $S \in L^2(P)$ *but notably without requiring the absence of arbitrage.*

3. We know from Černý and Kallsen (2007) that regardless of the setting (discrete time, continuous processes, general semimartingales) the mean–variance hedging problem has a solution in the class $\overline{\Theta}$ *from Definition 8.2 if there is an equivalent martingale measure with square-integrable density. In discrete time the existence of such a martingale measure follows from the absence of arbitrage via the Dalang–Morton–Willinger theorem. Proposition 8.4 shows that in discrete time* $\overline{\Theta}$ *coincides with the strategies of Melnikov and Nechaev (1999) if* $S \in L^2(P)$ *and that the Melnikov–Nechaev definition of admissibility can also be used when we only require local square-integrability of* S*.*

Example 8.6. *Here we modify Example 4 of Schweizer (1995) to come up with an arbitrage-free model with* $T = 3$*,* $S \in L^2(P)$ *and a strategy* $\bar\vartheta \in \overline{\Theta}$ *such that* $\bar\vartheta \bullet S_3 \in L^2(P)$ *but* $\bar\vartheta \bullet S_2 \notin L^2(P)$*.*

Let $\Omega = (0,1) \times \{-1,1\} \times \{-1,1\} \times \mathbb{R}$ *and let* $\mathcal{F}$ *be the Borel* σ*-algebra on* Ω*. By* $\omega = (u, y_1, y_2, z)$*,* $u \in (0,1)$*,* $y_1, y_2 \in \{-1,1\}$*,* $z \in \mathbb{R}$ *we denote the elements of* Ω*. Let* $U(\omega) = u$*,* $Y_i(\omega) = y_i$*,* $Z(\omega) = z$*,* $\mathcal{F}_0$ *trivial,* $\mathcal{F}_1 = \sigma(U, Y_1)$*,* $\mathcal{F}_2 = \sigma(U, Y_1, Y_2)$*,* $\mathcal{F}_3 = \mathcal{F}$*, and* P *be a measure on* Ω *such that* U, Y_1, Y_2 *are independent and uniformly distributed on their respective domains, while the conditional distribution of* Z *given* U, Y_1, Y_2 *is normal with mean* 0 *and variance* U^2*. Set* $S_0 := 0$*,* $\Delta S_1 := Y_1$*,* $\Delta S_2 := Y_2$*,* $\Delta S_3 := Z - Y_2$*. Clearly* S *is adapted and* $S \in L^2(P)$*. It is also arbitrage-free because*

ΔS_t take on both positive and negative values with non-zero probability, $t = 1, 2, 3$. The strategy $\bar{\vartheta}_1 := 0, \bar{\vartheta}_2 = \bar{\vartheta}_3 := U^{-1}$ is predictable and we have $H := \bar{\vartheta} \bullet S_3 = U^{-1}Z \in L^2(P)$, therefore $\bar{\vartheta} \in \overline{\Theta}$. At the same time $\bar{\vartheta} \bullet S_2 = U^{-1}Y_2 \notin L^2(P)$ which means $\bar{\vartheta} \notin \Theta$. Finally, by construction of the stock price process for any $\vartheta, \theta \in \overline{\Theta}$ such that $\vartheta \bullet S_3 = \theta \bullet S_3$ we have $\vartheta = \theta$ a.s. which implies that H cannot be hedged perfectly by trading strategies in Θ.

8.2. **Optimal hedging.** In this section we will treat the general case where S is a multidimensional process (cf. Bertsimas et al. 2001 and Černý 2004a). We do not assume that the conditional returns of individual assets are linearly independent. For any matrix A we denote by A^{-1} its Moore–Penrose inverse which is a particular matrix satisfying $AA^{-1}A = A$, cf. Albert (1972). Geometrically, $A^{-1}b$ is the shortest solution (in Euclidean norm) of the least squares problem $\min_x (Ax - b)^\top (Ax - b)$.

Theorem 8.7. *Under the assumptions of section 2 the process L given by*

$$\begin{aligned} L_T &= 1, \\ L_{t-1} &= E_{t-1}(L_t(1 - E_{t-1}(L_t\Delta S_t)^\top E_{t-1}((L_t\Delta S_t\Delta S_t^\top)^{-1}\Delta S_t))), \end{aligned}$$

is $(0, 1]$-valued and the opportunity-neutral measure $P^\star$,

$$\frac{dP^\star}{dP} := \prod_{t=1}^{T} \frac{L_t}{E_{t-1}(L_t)},$$

is well defined. The processes $\tilde{\lambda}^\star, V^\star$ and $\xi^\star$ given by

$$\begin{aligned} \tilde{\lambda}^\star_t &= E_{t-1}(L_t\Delta S_t)^\top E_{t-1}(L_t\Delta S_t\Delta S_t^\top)^{-1} \\ &= E^{P^\star}_{t-1}(\Delta S_t)^\top E^{P^\star}_{t-1}(\Delta S_t\Delta S_t^\top)^{-1}, && (8.6) \\ V^\star_{t-1} &= E^{P^\star}_{t-1}\left(\frac{1 - \tilde{\lambda}^\star_t\Delta S_t}{1 - \Delta\tilde{K}^\star_t}V^\star_t\right), \quad V^\star_T = H, && (8.7) \\ \Delta\tilde{K}^\star_t &= E^{P^\star}_{t-1}(\Delta S_t)^\top E^{P^\star}_{t-1}(\Delta S_t\Delta S_t^\top)^{-1}E^{P^\star}_{t-1}(\Delta S_t), && (8.8) \\ \xi^\star &= E^{P^\star}_{t-1}\left(\left(V^\star_t - V^\star_{t-1}\right)\Delta S_t\right)^\top E^{P^\star}_{t-1}(\Delta S_t\Delta S_t^\top)^{-1}. && (8.9) \end{aligned}$$

are well-defined. For a fixed admissible initial endowment $v \in \overline{U}$ the strategy $\varphi(v)$ given by

$$\varphi_t(v) = \xi^\star_t + \tilde{\lambda}^\star_t(V^\star_{t-1} - G^{v,\varphi(v)}_{t-1}), \tag{8.10}$$

is admissible and minimizes the expected squared hedging error among all admissible strategies with initial endowment v, while $(V^\star_0, \varphi(V^\star_0))$ is the optimal endowment–strategy pair if the hedging error is minimized over the initial endowment as well.

Proof. 1) By construction L_T is $(0, 1]$-valued. Suppose L_t is $(0, 1]$-valued. Since $E_{t-1}((\Delta S_t)^2) < \infty$ the least squares problem

$$L_{t-1} := \min\{E_{t-1}((\sqrt{L_t} - \lambda_t\sqrt{L_t}\Delta S_t)^2) : \lambda_t \text{ is } \mathcal{F}_{t-1}\text{-measurable}\} \tag{8.11}$$

has a (not necessarily unique) solution where the optimal value of λ_t can be chosen as $\tilde{\lambda}_t^\star = E_{t-1}\left(L_t \Delta S_t\right)^\top E_{t-1}(L_t \Delta S_t \Delta S_t^\top)^{-1}$. To see this define $A = E_{t-1}(L_t \Delta S_t \Delta S_t^\top)$, $\mu = E_{t-1}\left(L_t \Delta S_t\right)$ and denote by $A^{1/2}$ the unique symmetric matrix such that $\left(A^{1/2}\right)^2 = A$. We have $\lambda_t A \lambda_t^\top = 0 \Leftrightarrow E_{t-1}(L_t \left(\lambda_t \Delta S_t\right)^2) = 0 \Leftrightarrow \lambda_t \Delta S_t = 0 \Rightarrow \lambda_t \mu = 0$. This implies $\mu \in \operatorname{Span}(A)$ and consequently $AA^{-1}\mu = \mu$. Therefore we can write

$$\begin{aligned} E_{t-1}((\sqrt{L_t} - \lambda_t \sqrt{L_t} \Delta S_t)^2) &= E_{t-1}(L_t) - 2\lambda_t \mu + \lambda_t A \lambda_t^\top \\ &= E_{t-1}(L_t) - 2\lambda_t A A^{-1} \mu + \lambda_t A \lambda_t^\top \\ &= \left\| A^{1/2} \lambda_t^\top - A^{1/2} A^{-1} \mu \right\|^2 + E_{t-1}(L_t) - \mu^\top A^{-1} \mu, \end{aligned}$$

which proves the optimality of $\tilde{\lambda}_t^\star$ above.

2) Trivially, $0 \leq L_{t-1} \leq E_{t-1}\left(L_t\right) \leq 1$. By contradiction assume that $P(L_{t-1} = 0) > 0$. We have

$$1_{L_{t-1}=0} E_{t-1}((\sqrt{L_t} - \tilde{\lambda}_t^\star \sqrt{L_t} \Delta S_t)^2) = 0,$$

which in view of $L_t > 0$ is only possible if

$$1_{L_{t-1}=0} \left(1 - \tilde{\lambda}_t^\star \Delta S_t\right) = 0.$$

The latter contradicts the assumption of no arbitrage since $h := 1_{L_{t-1}=0} \tilde{\lambda}_t^\star$ is $\mathcal{F}_{t-1}$-measurable, $h \Delta S_t \geq 0$ but $h \Delta S_t \neq 0$. Hence L_{t-1} is $(0,1]$-valued and by induction this holds for all $t \in \mathcal{T}$.

3) By construction of $P^\star$ we have

$$\begin{aligned} E_{t-1}\left(L_t \Delta S_t\right) &= E_{t-1}\left(L_t\right) E_{t-1}^{P^\star}\left(\Delta S_t\right), \\ E_{t-1}(L_t \Delta S_t \Delta S_t^\top) &= E_{t-1}\left(L_t\right) E_{t-1}^{P^\star}(\Delta S_t \Delta S_t^\top), \end{aligned}$$

which shows the second equality in (8.6).

4) We have $E(L_T (V_T^\star)^2) < \infty$. Assume that $E(L_t (V_t^\star)^2) < \infty$. The least squares problem

$$\tilde{\psi}_{t-1}^\star := \min\{E_{t-1}((\alpha_t \sqrt{L_t} + \beta_t \sqrt{L_t} \Delta S_t - \sqrt{L_t} V_t^\star)^2) : \alpha_t, \beta_t \text{ are } \mathcal{F}_{t-1}\text{-measurable}\} \tag{8.12}$$

is well-defined and the optimal values can be chosen $\alpha_t = V_{t-1}^\star$ and $\beta_t = \xi_t^\star$. This is shown by the same method as in 1). The total sum of squares equals the explained sum of squares plus the residual sum of squares

$$E_{t-1}(L_t (V_t^\star)^2) = E_{t-1}((V_{t-1}^\star \sqrt{L_t} + \xi_t^\star \sqrt{L_t} \Delta S_t)^2) + \tilde{\psi}_{t-1}^\star. \tag{8.13}$$

We have seen in part 1) that $\sqrt{L_t}$ can be decomposed into P-orthogonal components

$$\sqrt{L_t} = \tilde{\lambda}_t^\star \sqrt{L_t} \Delta S_t + \sqrt{L_t} \left(1 - \tilde{\lambda}_t^\star \Delta S_t\right)$$

which in combination with (8.11, 8.13) yields

$$E_{t-1}(L_t (V_t^\star)^2) = L_{t-1} \left(V_{t-1}^\star\right)^2 + \left(\xi_t^\star + \tilde{\lambda}^\star V_{t-1}\right) E_{t-1}(L_t \Delta S_t \Delta S_t^\top) \left(\xi_t^\star + \tilde{\lambda}^\star V_{t-1}\right)^\top + \tilde{\psi}_{t-1}^\star. \tag{8.14}$$

The last two terms in (8.14) are non-negative which together with hypothesis implies

$$E(L_{t-1}\left(V_{t-1}^{\star}\right)^{2}) \leq E(L_{t}\left(V_{t}^{\star}\right)^{2}) < \infty.$$

By induction $\xi_t^\star$ and $V_{t-1}^\star$ are well-defined for all $t \in \mathcal{T}$.

5i) We now show that the strategy $(v, \varphi(v))$ defined in (8.10) is admissible for any $\mathcal{F}_0$-measurable v such that $E\left(L_0 v^2\right) < \infty$ and for any $H \in L^2(P)$. A straightforward calculation shows

$$\begin{aligned} (G_t^{v,\varphi(v)} - V_t^\star)\sqrt{L_t} &= (G_{t-1}^{v,\varphi(v)} - V_{t-1}^\star)\sqrt{L_t}\left(1 - \tilde{\lambda}_t^\star \Delta S_t\right) \\ &\quad + \sqrt{L_t}\left(V_{t-1}^\star + \xi_t^\star \Delta S_t - V_t^\star\right). \end{aligned} \tag{8.15}$$

On taking conditional expectations we find

$$\begin{aligned} E_{t-1}(L_t(G_t^{v,\varphi(v)} - V_t^\star)^2) &= L_{t-1}(G_{t-1}^{v,\varphi(v)} - V_{t-1}^\star)^2 + \tilde{\psi}_{t-1}^\star, \\ E((G_T^{v,\varphi(v)} - H)^2) &= E(L_T(G_T^{v,\varphi(v)} - V_T^\star)^2) \\ &= E(L_0\,(v - V_0^\star)^2) + \sum_{t=0}^{T-1} E(\tilde{\psi}_{t-1}^\star). \end{aligned} \tag{8.16}$$

It follows from (8.14) that $E(\sum_{t=1}^{T}\tilde{\psi}_t^\star) \leq E(L_T\left(V_T^\star\right)^2) = E\left(H^2\right)$ and by virtue of (8.16) $(v, \varphi(v))$ is admissible if and only if $E\left(L_0 v^2\right) < \infty$.

ii) Consider an admissible strategy $\vartheta \in \overline{\Theta}$. In view of the least squares regressions (8.11, 8.12) we have

$$\begin{aligned} 0 &= E\left(L_t\left(V_{t-1}^\star + \xi_t^\star \Delta S_t - V_t^\star\right)\Delta S_t\right) = E\left(L_t\left(V_{t-1}^\star + \xi_t^\star \Delta S_t - V_t^\star\right)\right) \\ &= E\left(L_t\left(1 - \tilde{\lambda}_t^\star \Delta S_t\right)\Delta S_t\right), \end{aligned}$$

which together with (8.15) yields

$$E_{t-1}(L_t(G_t^{v,\varphi(v)} - V_t^\star)G_t^{0,\vartheta}) = L_{t-1}(G_{t-1}^{v,\varphi(v)} - V_{t-1}^\star)G_{t-1}^{0,\vartheta}.$$

Since both $G_T^{v,\varphi(v)} - V_T^\star$ and $G_T^{0,\vartheta}$ are in $L^2(P)$ by Hölder's inequality $(G_T^{v,\varphi(v)} - V_T^\star)G_T^{0,\vartheta}$ is integrable and consequently

$$E((G_T^{v,\varphi(v)} - V_T^\star)G_T^{0,\vartheta}) = L_0(G_0^{v,\varphi(v)} - V_0^\star)G_0^{0,\vartheta} = 0. \tag{8.17}$$

iii) Consider an arbitrary admissible pair (v, ϑ). Using the arguments of 5i) with $H = v + \vartheta \bullet S_T$ we conclude $E\left(L_0 v^2\right) < \infty$, which means that $(v, \varphi(v))$ is also admissible and $\vartheta - \varphi(v) \in \overline{\Theta}$.

Now consider a general $H \in L^2\left(P\right)$ and the corresponding processes $V^\star, \varphi(v)$. We can write

$$\begin{aligned} E((G_T^{v,\vartheta} - V_T^\star)^2) &= E((G_T^{v,\varphi(v)} - V_T^\star + G_T^{0,\vartheta-\varphi(v)})^2) \\ &= E((G_T^{v,\varphi(v)} - V_T^\star)^2) + E((G_T^{0,\vartheta-\varphi(v)})^2), \end{aligned}$$

where the second equality follows from ii). Thus $(v, \varphi(v))$ has the smallest expected squared hedging error among all admissible strategies with fixed initial endowment v.

The optimality of $(V_0^\star, \varphi(V_0^\star))$ among all admissible endowment–strategy pairs then follows easily from (8.16). □

Corollary 8.8. *1. The set of admissible initial endowments is given explicitly by* $\overline{U} = \{v \in L(\mathcal{F}_0) : E\left(L_0 v^2\right) < \infty\}$.

2. $\overline{U} = L^2(\mathcal{F}_0)$ *if and only if* $\operatorname{ess\,inf} L_0 > 0$ *which holds if and only if the maximal Sharpe ratio defined in section 6 is bounded from above,* $\operatorname{ess\,sup}(\rho) < \infty$.

Proof. 1. $\subseteq$ This is shown in 5iii) above.

$\supseteq$ Take v such that $E\left(L_0 v^2\right) < \infty$ and fix a contingent claim $H = 0 \in L^2(P)$. By 5i) $(v, \varphi(v))$ is an admissible endowment–strategy pair.

2. Suppose $\operatorname{ess\,inf} L_0 = \varepsilon > 0$. Then $\varepsilon E(v^2) \leq E(L_0 v^2) \leq E(v^2)$ which implies $\overline{U} = L^2(\mathcal{F}_0)$. Conversely, if $\operatorname{ess\,inf} L_0 = 0$ define $p_n := P(L_0 \in [\frac{1}{n+1}, \frac{1}{n}))$. By passing to a subsequence $k_n \geq n$ we can restrict our attention to the values $p_{k_n} > 0$. Set

$$\begin{aligned} v(\omega) &= n^{-1}\sqrt{k_n/p_{k_n}} \text{ for } L_0(\omega) \in [1/(k_n+1), 1/k_n)\,, n \in \mathbb{N}, \\ v(\omega) &= 0 \text{ elsewhere.} \end{aligned}$$

Then $E(v^2) = \sum_{n\in\mathbb{N}} n^{-2} k_n \geq \sum_{n\in\mathbb{N}} n^{-1} = \infty$ while $E(L_0 v^2) \leq \sum_{n\in\mathbb{N}} n^{-2} < \infty$, implying $v \in \overline{U}$ but $v \notin L^2(\mathcal{F}_0)$. □

8.3. **The importance of $P^\star$.** We have argued in sections 5 and 6 that the opportunity-neutral measure $P^\star$ is an economically meaningful object playing an important role in optimal dynamic asset allocation. The following example shows that $P^\star$ is also technically indispensable, in the sense that the mean value process $V^\star$ always possesses first and second conditional moments under $P^\star$ (it is always $P^\star$-locally square integrable, cf. Černý and Kallsen 2007, Lemma 4.4) whereas it may fail to have both the first and second conditional moments under P. The same example shows that local risk minimization is in general ill-defined (V_1 is not square-integrable under P and V_0, ξ_0 fail to exist).

Example 8.9. *Let* $\Omega = (0,1) \times \{-1,1\} \times \{-1,1\}$ *and let* $\mathcal{F}$ *be the Borel* σ*-algebra on* Ω. *By* $\omega = (u, y_1, y_2)$, $u \in (0,1)$, $y_1, y_2 \in \{-1,1\}$, *we denote the elements of* Ω. *Let* $U(\omega) = u, Y_i(\omega) = y_i, \mathcal{F}_0$ *trivial,* $\mathcal{F}_1 = \sigma(U, Y_1), \mathcal{F}_2 = \mathcal{F}$, *and* P *be a measure on* Ω *such that* U *is a uniformly distributed random variable on* $[0,1]$, Y_1 *is independent of* U *and uniform on* $\{-1,1\}$ *and let the conditional distribution of* Y_2 *given* U, Y_1 *be* $P(Y_2 = 1|U, Y_1) = U^\gamma, P(Y_2 = -1|U, Y_1) = 1 - U^\gamma$ *for some* $\gamma > 0$. *Set* $S_0 = 0, \Delta S_1 = Y_1$, *and* $\Delta S_2 = Y_2$. *Consider measure* Q *under which* U, Y_1, Y_2 *are independent and uniformly distributed on their respective domains.* Q *is a martingale measure and* $Q \sim P$, *consequently the market is arbitrage-free.*

S *is bounded and trivially,* $S \in L^2(P)$. *Define the contingent claim* $H = U^{-\alpha} Y_2^+$ *for some* $\alpha \in \mathbb{R}$. *We have*

$$\begin{aligned} E\left(H^2|\mathcal{F}_1\right) &= U^{\gamma-2\alpha}, \\ E\left(H^2\right) &= E(U^{\gamma-2\alpha}) = \int_0^1 u^{\gamma-2\alpha} du, \end{aligned}$$

which means $E(H^2) < \infty \Leftrightarrow \gamma - 2\alpha > -1$*. To obtain* V_1 *we have*

$$\begin{aligned}
E(\Delta S_2|\mathcal{F}_1) &= 2U^\gamma - 1, \qquad E((\Delta S_2)^2|\mathcal{F}_1) = 1 \\
L_1 &= 1 - (2U^\gamma - 1)^2 = 4U^\gamma(1 - U^\gamma) \qquad (8.18) \\
V_1 &= V_1^\star = E\left(\left(1 - \frac{E(\Delta S_2|\mathcal{F}_1)\Delta S_2}{E((\Delta S_2)^2|\mathcal{F}_1)}\right)\frac{H}{L_1}\middle|\,\mathcal{F}_1\right) = \frac{1}{2}U^{-\alpha}. \qquad (8.19)
\end{aligned}$$

The result in (8.19) becomes obvious when one realizes that in the last period the market is complete with conditional risk-neutral probabilities of $\Delta S_2 = \pm 1$ *equal to* $1/2$*. Now*

$$E(V_1) = \int_0^1 \frac{1}{2}u^{-\alpha}du < \infty \Leftrightarrow \alpha < 1.$$

To construct the desired counterexample we therefore need $\alpha \geq 1$, $\gamma - 2\alpha > -1$ *and* $\gamma > 0$*. These conditions are met for example with* $\alpha = 1, \gamma > 1$*.*

Finally, we evaluate the conditional density of U *under* $P^\star$ *and compute* L_0*,*

$$\begin{aligned}
P^\star(U \in B) &= \frac{P(L_1 1_B(U))}{E(L_1)} =: \int_B p_1^\star(u)du, \\
p_1^\star(u) &= \frac{u^\gamma(1 - u^\gamma)}{\int_0^1 u^\gamma(1 - u^\gamma)du} = \frac{(\gamma + 1)(2\gamma + 1)}{\gamma}u^\gamma(1 - u^\gamma).
\end{aligned}$$

Note that V_1 *is square integrable under* $P^\star$ *whenever* $E(H^2) < \infty$ *in accordance with the general theory put forward in Černý and Kallsen (2007).*

To obtain L_0 *we evaluate* $E^{P^\star}(\Delta S_1) = 0$ *whereby (5.11) and (8.18) give*

$$L_0 = E(L_1) = \frac{4\gamma}{(\gamma + 1)(2\gamma + 1)},$$

which yields the maximal unconditional Sharpe ratio,

$$\rho = \sqrt{1/L_0 - 1} = \sqrt{\frac{2\gamma^2 - \gamma + 1}{4\gamma}} = \sqrt{\frac{2(\gamma - 1/4)^2 + 7/8}{4\gamma}}.$$

References

Albert, A. (1972). *Regression and the Moore-Penrose pseudoinverse.* New York: Academic Press. Mathematics in Science and Engineering, Vol 94.

Arai, T. (2005). An extension of mean-variance hedging to the discontinuous case. *Finance & Stochastics 9*, 129–139.

Bertsimas, D., L. Kogan, and A. W. Lo (2001). Hedging derivative securities and incomplete market: An ϵ-arbitrage approach. *Operations Research 49*(3), 372–397.

Björk, T. and I. Slinko (2006). Towards a general theory of good-deal bounds. *Review of Finance 10*, 221–260.

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

Černý, A. (2003). Generalized Sharpe Ratios and asset pricing in incomplete markets. *European Finance Review 7*(2), 191–233.

Černý, A. (2004a). Dynamic programming and mean-variance hedging in discrete time. *Applied Mathematical Finance 11*(1), 1–25.

Černý, A. (2004b). *Mathematical Techniques in Finance: Tools for Incomplete Markets*. Princeton: Princeton University Press.

Černý, A. and J. Kallsen (2006). Mean–variance hedging and optimal investment in Heston's model with correlation. SSRN working paper, `http://ssrn.com/abstract=909305`. To appear in *Mathematical Finance*.

Černý, A. and J. Kallsen (2007). On the structure of general mean–variance hedging strategies. *The Annals of Probability 35*(4), 1479–1531.

Christensen, M. and E. Platen (2007). Sharpe ratio maximization and expected utility when asset prices have jumps. *International Journal of Theoretical & Applied Finance 10*(8), 1339–1364.

Cochrane, J. H. and J. Saá-Requejo (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy 108*(1), 79–119.

Dalang, R., A. Morton, and W. Willinger (1990). Equivalent martingale measures and no-arbitrage in stochastic security market models. *Stochastics and Stochastics Reports 29*, 185–202.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*. Oxford University Press.

Delbaen, F. and W. Schachermayer (1998). A simple counterexample to several problems in the theory of asset pricing. *Mathematical Finance 8*(1), 1–11.

Föllmer, H. and M. Schweizer (1989). Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin 18*, 147–160.

Föllmer, H. and M. Schweizer (1991). Hedging of contingent claims under incomplete information. In *Applied stochastic analysis (London, 1989)*, Volume 5 of *Stochastics Monogr.*, pp. 389–414. New York: Gordon and Breach.

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

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

Hipp, C. (1993). Hedging general claims. In *Proceedings of the 3rd AFIR Colloquium, Rome*, Volume 2, pp. 603–613.

Jacod, J. and A. N. Shiryaev (2003). *Limit theorems for stochastic processes* (Second ed.), Volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Berlin: Springer-Verlag.

Kallsen, J. (1999). A utility maximization approach to hedging in incomplete markets. *Mathematical Methods of Operations Research 50*, 321–338.

Kallsen, J. (2002). Derivative pricing based on local utility maximization. *Finance & Stochastics 6*, 115–140.

Leippold, M., F. Trojani, and P. Vanini (2004). A geometric approach to multi-period mean variance optimization of assets and liabilities. *Journal of Economic Dynamics and Control 28*(6), 1079–1113.

Li, D. and W.-L. Ng (2000). Optimal dynamic portfolio selection: Dynamic multi-period formulation. *Mathematical Finance 10*(3), 387–406.

Melnikov, A. V. and M. L. Nechaev (1999). On the mean-variance hedging problem. *Theory of Probability and Its Applications 43*(4), 588–603.

Pham, H. (2000). On quadratic hedging in continuous time. *Mathematical Methods of Operations Research 51*(2), 315–339.

Schäl, M. (1994). On quadratic cost criteria for option hedging. *Mathematics of Operations Research 19*(1), 121–131.

Schweizer, M. (1991). Option hedging for semimartingales. *Stochastic Processes and their Applications 37*(2), 339–363.

Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research 20*, 1–32.

Schweizer, M. (2001). A guided tour through quadratic hedging approaches. In E. Jouini, J. Cvitanić, and M. Musiela (Eds.), *Option pricing, interest rates and risk management*, Handb. Math. Finance, pp. 538–574. Cambridge: Cambridge Univ. Press.

Cass Business School, City University London
Current address: 106 Bunhill Row, EC1Y 8TZ London, U.K.
E-mail address: cerny@martingales.info
URL: http://www.martingales.info

Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel
Current address: Christian-Albrechts-Platz 4, D-24098 Kiel, Germany
E-mail address: kallsen@math.uni-kiel.de
URL: http://www.numerik.uni-kiel.de/~jk