Hedging with Residual Risk: A BSDE Approach

Stefan Ankirchner and Peter Imkeller

Abstract. When managing energy or weather related risk often only imperfect hedging instruments are available. In the first part we illustrate problems arising with imperfect hedging by studying a toy model. We consider an airline's problem with covering income risk due to fluctuating kerosine prices by investing into futures written on heating oil with closely correlated price dynamics. In the second part we outline recent results on exponential utility based cross hedging concepts. They highlight in a generalization of the Black-Scholes delta hedge formula to incomplete markets. Its derivation is based on a purely stochastic approach of utility maximization. It interprets stochastic control problems in the BSDE language, and profits from the power of the stochastic calculus of variations.

Mathematics Subject Classification (2000). 91B28, 60H10, 60H07.

Keywords. Financial derivatives, hedging, minimal variance hedging, utility-based pricing, BSDE, sub-quadratic growth,differentiability, stochastic calculus of variations, Malliavin calculus.

Introduction

In recent years many financial instruments have been created which serve the purpose of transferring exogenous risk to capital markets in concepts of *securitization.* For instance in 1999 the Chicago Mercantile Exchange introduced weather futures contracts, the payoffs of which are based on average temperatures at specified locations. Another example are catastrophe futures based on an insurance loss index regulated by an independent agency.

The risk arising in hedging derivatives of this type, and equally in using them as hedging instruments, is impossible to perfectly replicate, since the underlying risk process carries independent uncertainty. To come close to a replication, in practice one often looks for a tradable asset that is well correlated to the non-tradable underlying of the derivative, and uses it to *cross hedge* the underlying risk. Since the correlation usually differs from one, a non-hedgeable *basis risk* remains.

In Section 1 of this paper, we will illustrate typical problems related to hedging the basis risk in a particular setting of cross hedging. We will consider the situation of an airline company facing the risk of increasing kerosine prices. It might cross hedge fluctuations in the kerosine price dynamics by holding heating oil futures the price evolution of which is closely correlated. Our analysis of the assessment of the problem the airline company faces starts with the intuitive approach of hedging the basis risk by minimizing the variance of the hedging error in a simple Gaussian setting. This approach, however, presents a counter-intuitive feature: though the correlation between the hedged asset and the hedging instrument may be very close to one, the percentage of the hedging error in units of the standard deviation of the uncertainty to be hedged is rather large. This calls for more efficient concepts of replicating the basis risk which in particular take into account its downside component.

In Section 2 we will give an overview of some recent work on utility based concepts of cross hedging. We consider models in which agents exposed to some exogenous risk generated by a non-homogeneous diffusion process buy or sell a financial derivative to set off a portion of it to a financial market with assets correlated to the risk index. We present *explicit* hedging strategies that optimize the expected exponential utility of an agent holding a portfolio of such derivatives. To this end we will establish some structure and smoothness properties of indifference prices such as the Markov property and differentiability with respect to the underlyings. Once these properties are established, we can explicitly describe the optimal hedging strategies in terms of the price gradient and correlation coefficients. This way we obtain a generalization of the classical *delta hedge* of the Black-Scholes model. The analytical tool for deriving the crucial smoothness properties of strategies and prices is provided by a BSDE based approach (see [8]), which can be seen as the probabilistic counterpart of the usually employed control theoretic methods whose more analytical touch finds its expression in the Hamilton-Jacobi-Bellman PDE (see for example [3, 4, 6, 7, 12, 13]). The BSDE approach culminates in a description of strategies and prices in terms of the solutions of tailor made BSDE with drivers of sub-quadratic growth, derived by applying the martingale optimality principle in a utility maximization or risk minimization context.

1. Hedging with residual risk

1.1. Imperfect hedging instruments

A hedging instrument is often unable to perfectly replicate the risk or uncertainty of the asset it is supposed to hedge. More precisely, the possible risky scenarios of its evolution cannot be mapped one-to-one to possible scenarios of hedging. In the context of hedging with futures on financial markets, the difference between the spot price of a risky asset and the price of the futures contract used to hedge it is called *basis*. More generally, we may consider the *basis* to be given by the difference between the *price of the asset to be hedged* and the *price of the hedging instrument*. That is why residual risk is frequently also referred to as *basis risk*.

A prominent example for financial derivatives that may entail residual risk are basket options. Basket options are written on stock market indices, for example the Dow Jones. In practice they are often hedged by trading some, but not all of its underlyings. Consequently they cannot be perfectly replicated, and there remains a basis risk.

Managing weather risk also often involves basis risk. Weather securities are highly, but in general not perfectly correlated with the risk the security holder bears. For example temperature derivatives may be used to hedge variations in the demand of heating oil. But the demand of heating oil may at least weakly depend on uncertainties not caused by weather and temperature fluctuations.

Hedging with futures provides the generic situation in which basis risk arises. In simple terms, a futures contract is an agreement to deliver (or to pay in cash the value of) a specified amount of a commodity, for example crude oil, on a future date at a price specified already today. To ensure their liquidity, futures are highly standardized, and as a consequence do not perfectly correlate with the risk the futures' holder bears. For example there may be a mismatch between the expiration date of the future and the date on which the futures' holder sells his commodity. Or the commodity underlying the future may not be exactly the commodity whose price has to be hedged.

One might be tempted to think that as the correlation between asset and hedging instrument increases, the significance of treating the related basis risk shrinks at the same pace. The example studied in the following subsection shows that this conjecture is surprisingly false, and that it is very important to take basis risk into account, even if this correlation is very high.

1.2. Case study: hedging jet fuel price fluctuations with heating oil futures

The revenues of airline companies strongly depend on the jet fuel spot price. Futures provide protection against price fluctuations. However, no futures on jet fuel are traded in Europe and the US. Heating oil and jet fuel prices are highly correlated (see Figure 1), and therefore in practice airlines buy heating oil futures to protect themselves against rising jet fuel spot prices. To display the role of high correlation in treating basis risk in a simple setting, let us assume that the daily price changes of jet fuel is given by a sequence of i.i.d. square integrable random variables $(\Delta J_i)_{i\geq 1}$. Similarly, assume the daily heating oil price changes $(\Delta H_i)_{i\geq 1}$ are i.i.d. and square integrable, and that ΔH_i is independent of ΔJ_k whenever $i \neq k$. Let $\sigma_J^2 = \mathrm{Var}(\Delta J_i)$ and $\sigma_H^2 = \mathrm{Var}(\Delta H_i)$. Figure 1 shows the daily spot price per Gallon, from January 2006 to December 2007, of No. 2 Heating Oil and Kerosene-Type Jet Fuel delivered at New York Harbor. The sample standard

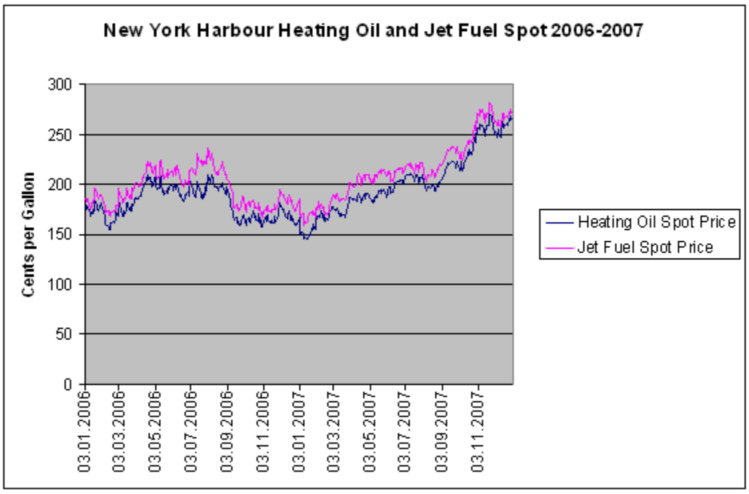

FIGURE 1. Daily spot prices.

deviation[1] of the price changes during this time period is given by

$$\widehat{\sigma}_J \approx 3,9986 \text{ and } \widehat{\sigma}_H \approx 3,8353.$$

Recall that the correlation between two random variables X and Y is defined by

$$\operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}},$$

and let $\rho = \operatorname{corr}(\Delta H_i, \Delta J_i)$. The empirical correlation between jet fuel and heating oil price changes, or more precisely the Pearson correlation coefficient[2], is given by

$$\widehat{\rho} \approx 0,896.$$

1.2.1. The minimum variance hedge ratio. The airline aims at hedging increasing fuel prices by buying heating oil futures. Suppose that it wants to hedge the price for N_J Gallons of jet fuel at a future date T. We assume that there exists a heating oil futures contract with matching delivery date T, and with a size of N_H Gallons. Let K be the price at time 0 of a heating oil futures contract. How many units of futures a shall the airline buy so that the variance of its fuel costs at time T are minimal?

[1]The *sample standard deviation* of a sample $x_1, \ldots, x_n$ of length $n \in \mathbb{N}$ is defined as $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (x_i - (\frac{1}{n}\sum_{j=1}^n x_j))^2}$. Notice that s^2 is an unbiased estimator of the variance.

[2]The *Pearson correlation coefficient*, also known as *sample correlation coefficient*, is defined by $\rho = \frac{n\sum_i x_i y_i - \sum_i x_i \sum_i y_i}{\sqrt{n\sum_i x_i^2 - (\sum_i x_i)^2}\sqrt{n\sum_i x_i^2 - (\sum_i x_i)^2}}$.

Let J_T and H_T denote the spot price at time T of jet fuel and heating oil, respectively. Notice that at time T the value of one futures contract is equal to $N_H \cdot H_T$. The airline's fuel costs amount to $(N_J J_T - a(N_H H_T - K))$, the variance of which is given by

$$\begin{aligned} & E\left[(N_J J_T - a(N_H H_T - K) - E[N_J J_T - a(N_H H_T - K)])^2\right] \\ = \; & T(N_J^2 \sigma_J^2 - 2a N_J N_H \rho \sigma_J \sigma_H + a^2 N_H^2 \sigma_H^2). \end{aligned}$$

The variance is minimal if the airline holds

$$a^* = \frac{N_J}{N_H} \cdot \rho \frac{\sigma_J}{\sigma_H}$$

units of the future. The first factor, $\frac{N_J}{N_H}$, adjusts the units of the futures to the quantity of jet fuel needed. The second factor,

$$h = \rho \frac{\sigma_J}{\sigma_H},$$

is referred to as *minimum variance hedge ratio* (see Hull [9, Chapter 4]), and determines the proportion of the jet fuel price risk that should be transferred to heating oil futures in order to minimize the variance of revenue fluctuations.

1.2.2. The hedging error. So far we have seen how many of the highly correlated heating oil futures an airline has to hold, in order to minimize the variance of its fuel expenses. Let us next discuss the hedging error or basis risk at time T. We will argue that although the correlation is 90%, the airline bears a *high* residual risk. To demonstrate this we assume in addition that the daily price changes of jet fuel and heating oil are *normally distributed*. Using the fact that two uncorrelated Gaussian random variables are independent, we can decompose the daily price changes of jet fuel into

$$\Delta J_i = \rho \frac{\sigma_J}{\sigma_H} \Delta H_i + \sqrt{1-\rho^2} N_i, \qquad i \geq 1,$$

where N_i is independent of ΔH_i, and normally distributed with variance $\mathrm{Var}(N_i) = \sigma_J^2$.

By the hedging error at time $T > 0$, when holding a futures, we mean the difference

$$\mathrm{error}(a) = N_J(J_T - J_0) - a N_H (H_T - H_0).$$

By holding $a^* = \frac{N_J}{N_H} \times \rho \frac{\sigma_J}{\sigma_H}$ futures, the hedging error at time $T > 0$, in Cent per Gallon, is given by

$$\mathrm{error} = \sum_{i=1}^{T} \sqrt{1-\rho^2} N_i.$$

Notice that the standard deviation of the error is given by

$$\sqrt{1-\rho^2}\sqrt{T}\sigma_J \approx 0.443\sqrt{T}\sigma_J.$$

The standard deviation of the jet fuel price at time T equals $\sqrt{T}\sigma_J$. This means that although the correlation between the prices of jet fuel and heating oil is almost 90%, only 56% of the standard deviation of the jet fuel price uncertainty can be hedged!

The conclusions we can draw from this case study are the following.

1. The hedge ratio provides a simple strategy to minimize the variance of price uncertainty. It is a *static* hedge, and depends only on the volatilities and the correlation of the processes.
2. *Even if the correlation is very high, there remains a considerably high hedging error!* If the correlation was as high as 98%, the standard deviation of the basis would still represent 19% of the total risk! The link between the correlation and the percentage contribution of the basis to the total risk is depicted in Figure 2.

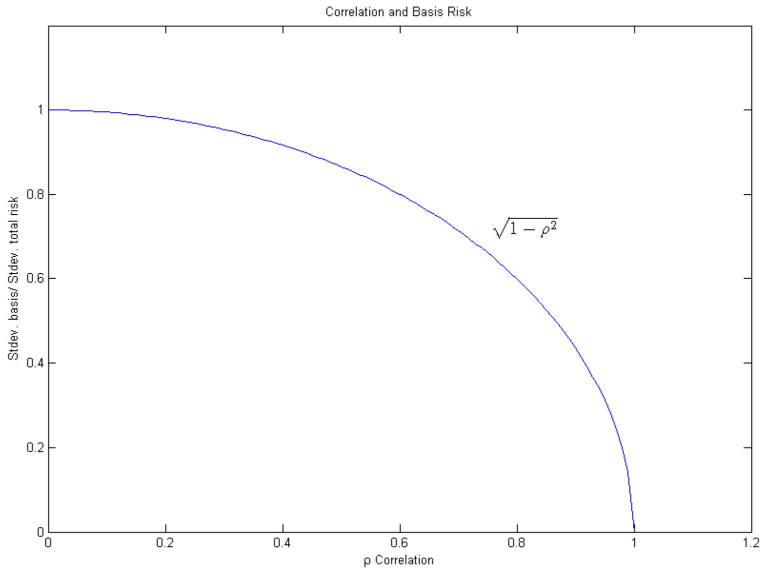

FIGURE 2. Basis risk in dependence of the correlation.

It clearly exhibits the following phenomenon. If the correlation is high, then a small change in the correlation leads to a large change in the percentage of basis risk relative to total risk. Conversely, if the correlation is low, a small change in the correlation leads to essentially no change in the percentage of basis risk relative to total risk.

2. A utility-based approach to hedging with basis risk

In this section, we shall sketch a utility based purely probabilistic approach of hedging the basis risk in a more sophisticated model for price processes of assets and hedging instruments. As an alternative to the intuitive and straightforward concept of minimizing the variance of the hedging cost discussed in Section 1, we shall minimize the expected loss of different hedging scenarios if revenues are measured with an exponential utility function. This way, we take into account the essential *downside risk* of the basis. Our approach provides optimal hedging strategies if the risk and the hedging instrument have non-linear payoffs. It allows to derive an explicit formula for the utility indifference price and the derivative hedge of a product designed to *cross hedge* the basis risk, generalizing the delta hedging formula in the solution of the Merton-Scholes problem to the setting of incomplete markets. The formula clarifies the role of correlation in hedging, and describes the reduction rate of risk by cross hedging. The method used to derive it translates the underlying optimization problem by *martingale optimality* into the language of *backward stochastic differential equations (BSDE)*. It profits from stochastic calculus of variations *(Malliavin's calculus)*, since the extension of the delta hedge formula is based on sensitivity of the BSDE providing the optimal hedges to system parameters such as initial states of a risk index process. In more formal terms, we shall investigate the following model.

2.1. The model

Let $d \in \mathbb{N}$ and let W be a d-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the P-completion of the filtration generated by W. Risk sources, for instance jet fuel price or temperature processes, will be described as diffusion processes with dynamics

$$dR_t = b(t, R_t)dt + \sigma(t, R_t)dW_t, \tag{2.1}$$

where $b : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ and $\sigma : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ are measurable functions. Throughout we assume that there exists a $C \in \mathbb{R}_+$ such that for all $t \in [0,T]$ and $x, x' \in \mathbb{R}^m$, denoting by $|\cdot|$ the norm in finite dimensional Euclidean spaces,

$$\text{(R1)} \qquad \begin{array}{rcl} |b(t,x) - b(t,x')| + |\sigma(t,x) - \sigma(t,x')| & \leq & C|x - x'|, \\ |b(t,x)| + |\sigma(t,x)| & \leq & C(1 + |x|). \end{array}$$

Suppose that an economic agent has expenses at time $T > 0$ of the form $F(R_T)$, where $F : \mathbb{R}^m \to \mathbb{R}$ is a bounded and measurable function. At time $t \in [0,T]$, the expected payoff of $F(R_T)$, conditioned on $R_t = r$, is given by $F(R_T^{t,r})$, where $R^{t,r}$ is the solution of the SDE

$$R_s^{t,r} = r + \int_t^s b(u, R_u^{t,r})du + \int_t^s \sigma(u, R_u^{t,r})dW_u, \quad s \in [t,T].$$

We assume that there exists a financial market on which k risky assets - such as heating oil futures or weather derivatives - are traded that may be correlated to the risk source. We further assume that there exists a non-risky asset, use it as numeraire and suppose that the prices of the risky assets in units of the numeraire evolve according to the SDE

$$dS_t^i = S_t^i(\alpha_i(t, R_t)dt + \beta_i(t, R_t)dW_t), \quad i = 1, \dots, k,$$

where $\alpha_i(t, r)$ is the ith component of a measurable and vector-valued map $\alpha : [0, T] \times \mathbb{R}^m \to \mathbb{R}^k$ and $\beta_i(t, r)$ is the ith row of a measurable and matrix-valued map $\beta : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{k\times d}$. Notice that W is the same $\mathbb{R}^d$-dimensional Brownian motion as the one driving the risk source (2.1), and hence the correlation between the risk and the tradable assets is determined by the matrices σ and β.

In order to exclude arbitrage opportunities in the financial market we assume $d \geq k$. For technical reasons we suppose that

(M1) α is bounded,
(M2) there exist constants $0 < \varepsilon < K$ such that $\varepsilon I_k \leq (\beta(t, r)\beta^*(t, r)) \leq K I_k$ for all $(t, r) \in [0, T] \times \mathbb{R}^m$,

where β^* denotes the transpose of β, and I_k is the k-dimensional unit matrix. If M and N are two square matrices of identical dimension, then we write $N < M$ if the difference $M - N$ is positive definite. (M2) implies that the symmetric matrix $\beta\beta^*$ is invertible. Moreover, the *Moore-Penrose pseudoinverse* of the matrix β is given by

$$\beta^+ = \beta^*(\beta\beta^*)^{-1} \ \in \mathbb{R}^{d\times k}.$$

Notice that β^+ is the right inverse of β, i. e.

$$\beta\beta^+ = I_k.$$

The *market price of risk* will be denoted by

$$\vartheta = \beta^+\alpha = \beta^*(\beta\beta^*)^{-1}\alpha.$$

The properties (M1) and (M2) imply that ϑ is uniformly bounded everywhere.

Suppose that our economic agent aims at reducing his risk exposure $F(R_T)$ by investing in the financial market. In order to determine an optimal hedge, we assume that the agent's preferences are described by the exponential utility function

$$U(x) = -e^{-\eta x}, \quad x \in \mathbb{R},$$

where $\eta > 0$ describes the risk aversion. By an *investment strategy*, or simply strategy, we mean any predictable process $\lambda = (\lambda^i)_{1\leq i\leq k}$ with values in $\mathbb{R}^k$ (row vectors) such that the integral process $\int_0^t \lambda_r^i \frac{dS_r^i}{S_r^i}$ is defined for all $i \in \{1, \dots, k\}$. We interpret λ^i as the value of the portfolio fraction invested into asset number i.

In what follows it will be convenient to embed the strategies into $\mathbb{R}^d$, the space of uncertainties. To this end let $C(t, r) = \{x\beta(t, r) : x \in \mathbb{R}^k\}, (t, r) \in [0, T] \times \mathbb{R}^m$. We denote by $p_t = \lambda_t\beta_t$ the image of any investment process λ with respect to β. For any image strategy $p = \lambda\beta$ we interpret

$$\int_0^t p_s(\vartheta_s ds + dW_s) = \sum_{i=1}^k \int_0^t \lambda_s^i \alpha_s^i ds + \sum_{i=1}^k \sum_{j=1}^d \int_0^t \lambda_s^i \beta_s^{ij} dW_s^j$$

as the increase of wealth up to time t. Moreover, the wealth at time t, conditioned on x at time s and $R_s = r$, $s \le t \le T$, is given by

$$X_t^{s,r,x,p} = x + \int_s^t p_u(\vartheta(u, R_u^{s,r})du + dW_u).$$

For $(t,r) \in [0,T] \times \mathbb{R}^m$ let $\mathcal{A}^{t,r}$ be the set of all predictable processes p with values in $\mathbb{R}^d$ such that $E \int_t^T |p_s|^2 ds < \infty$. The square integrability guarantees that there is no arbitrage (see Remark 2 in [8]). If $p \in \mathcal{A}^{t,r}$, then we say that p is *admissible* on $[t,T]$.

The value function is defined as

$$V^F(x) = \sup\{EU(X_T^{0,r,x,p} - F(R^{0,r})) : p \in \mathcal{A}^{0,r},\ p_s \in C(s, R_s^{0,r}) \text{ for all } s \in [0,T]\}.$$

Frequently we will need the conditional version of the value function given by

$$V^F(t,r,x) = \sup\{EU(x + \int_t^T p_s(\vartheta_s ds + dW_s) - F(R_T^{t,r})) : \\ p \in \mathcal{A}^{t,r},\ p_s \in C(s, R_s^{t,r}) \text{ for all } s \in [t,T]\}.$$

We recall briefly the Dynamic Programming or Bellman's Principle (for more details see f.ex. [5] and [11]). If one follows an optimal strategy up to a stopping time τ, the strategy will remain optimal, even by taking into account incoming new information. Mathematically, this may be expressed as follows. For all $(s,r) \in [0,T] \times \mathbb{R}^m$, $x \in \mathbb{R}^k$, and stopping times τ with values in $[s,T]$, we have

$$V^F(s,r,x) = \sup_p E\left[V^F\left(\tau, R_\tau^{s,r}, x + \int_s^\tau p_u(\vartheta_u du + dW_u)\right)\right]. \tag{2.2}$$

If V^F is a continuous function satisfying Bellman's principle (2.2), and if there exists an optimal strategy p^{opt} such that

$$V^F(0,r,x) = E\left[V^F\left(\tau, R_\tau^{0,r}, x + \int_0^\tau p_u^{\text{opt}}(\vartheta_u du + dW_u)\right)\right],$$

then $V^F(t, R_t^{0,r}, X_t^{0,r,x,p^{\text{opt}}})$ is a martingale. Moreover, if $V^F \in C^{1,2,2}$, then Ito's formula implies that V^F satisfies the associated HJB partial differential equation.

The standard approach of finding V^F and the optimal control p^{opt} is based on verification: Solve the HJB equation, and then show that the solution coincides with the value function V^F (*Verification Theorem*).

We don't work with the verification method here, but follow a purely probabilistic approach based on the *martingale optimality* of the process $V^F(t, R_t^{0,r}, X_t^{0,r,x,p^{\text{opt}}})$, $t \in [0,T]$. Notice that $V^F(\cdot, R^{0,r}, X^{0,r,x,p})$ is a supermartingale for any choice of p, and a martingale iff p is optimal. Moreover, the process satisfies the boundary condition $V^F(T, R_T^{0,r}, X_T^{0,r,x,p}) = U(X_T^{0,r,x,p} - F(R_T^{0,r}))$.

This motivates us to make the risky income $F(R_T^{0,r})$ dynamic, by finding a process $(Y_t)_{t\in[0,T]}$ that solves a BSDE with terminal condition $Y_T = F(R_T^{0,r})$, such that

- $(U(X_t^{0,r,x,p} - Y_t))_{0\le t\le T}$ is a supermartingale for all $p \in \mathcal{A}$,
- $(U(X_t^{0,r,x,p^{\text{opt}}} - Y_t))_{0\le t\le T}$ is a martingale for at least one $p^{\text{opt}} \in \mathcal{A}$.

2.2. Solving the control problem with BSDEs

The orthogonal projection of a vector $z \in \mathbb{R}^d$ onto the subspace $C = \{x\beta : x \in \mathbb{R}^k\}$ is given by

$$\Pi_C(z) = z\ \beta^*(\beta\beta^*)^{-1}\beta.$$

Notice that this can be deduced from the fact that $\Pi_C^2 = \Pi_C$. In terms of the pseudoinverse, the projection operator may be written as $\Pi_C(z) = z\ \beta^+\beta$.

Moreover, given an image strategy p with values in $\mathbb{R}^d$, the associated original strategy λ with values in $\mathbb{R}^k$ is given by

$$\lambda_t = p_t\beta^+(t,\cdot), \quad t \in [0,T]. \tag{2.3}$$

Indeed, we have $p\beta^+ = \lambda\beta\beta^+ = \lambda$.

The distance of a vector $z \in \mathbb{R}^d$ to the linear subspacet C will be defined as $\text{dist}(z, C) = \min\{|z - u| : u \in C\}$.

Let $f : [0,T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$ be the generator defined by

$$f(s,r,z) = \frac{1}{2}\eta \text{dist}^2(z + \frac{1}{\eta}\vartheta^*(s,r), C(s,r)) - z\vartheta(s,r) - \frac{1}{2\eta}|\vartheta(s,r)|^2. \tag{2.4}$$

Notice that f is a generator with sub-quadratic growth in z, for which there exists a well established theory (see Kobylanski [10]). Let us recall some notation needed to formulate its results. For $p \ge 1$ and $n \in \mathbb{N}$ we denote by $\mathcal{H}^p(\mathbb{R}^n)$ the set of all $\mathbb{R}^n$-valued predictable processes ζ such that $E\left(\int_0^1 |\zeta_t|^2 dt\right)^{\frac{p}{2}} < \infty$, and by $\mathcal{S}^p$ the set of all $\mathbb{R}$-valued predictable processes δ satisfying $E\left(\sup_{s\in[0,1]} |\delta_s|^p\right) < \infty$. By $\mathcal{S}^\infty$ we denote the set of all essentially bounded $\mathbb{R}$-valued predictable processes.

Recall that we assumed the payoff function F and the market price of risk ϑ to be bounded. According to one of the central results of the theory of BSDE with generators of sub-quadratic growth, there exists a unique solution $(Y,Z) \in \mathcal{S}^\infty(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE

$$Y_t = F(R_T^{0,r}) - \int_t^T Z_s dW_s + \int_t^T f(s, R_s^{0,r}, Z_s) ds. \tag{2.5}$$

Lemma 2.1. *For every locally square integrable and $(\mathcal{F}_t)$-predictable p, $U(X^{0,r,x,p} - Y)$ is a local supermartingale. Moreover, if for $(t,r) \in [0,T] \times \mathbb{R}^m$ we take $p_t = \Pi_{C(t,R_t^{0,r})}(Z_t + \frac{1}{\eta}\vartheta^*(t, R_t^{0,r}))$, then $U(X^{0,r,x,p} - Y)$ is a local martingale.*

Proof. For all $(s,r) \in [0,T] \times \mathbb{R}^m$, $p \in \mathbb{R}^k$ and $z \in \mathbb{R}^d$ let

$$h(s,r,z,p) = -p\vartheta_s + \frac{1}{2}\eta|p - z|^2,$$

and notice that

$$\min_{p\in C(s,r)} h(s,r,z,p) = f(s,r,z), \tag{2.6}$$

where the maximum is attained at $p = \Pi_{C(s,r)}(z + \frac{1}{\eta}\vartheta^*(s,r))$.

Now let p be a locally square integrable and $(\mathcal{F}_t)$-predictable process. To simplify notation we use the abbreviation $X^p = X^{0,r,x,p}$. An application of Ito's formula to $U(X^p - Y)$ yields for $t \in [0,T]$

$$\begin{aligned} U(X_t^p - Y_t) &= U(x - Y_0) + \int_0^t U'(X_s^p - Y_s)(p_s - Z_s)dW_s \\ &\quad + \int_0^t U'(X_s^p - Y_{s-})(p_s\vartheta_s + f(s, Z_s))ds \\ &\quad + \frac{1}{2}\int_0^t U''(X_s^p - Y_s)(|p_s|^2 - 2p_s Z_s^* + |Z_s|^2)ds. \end{aligned}$$

Moreover, we may write

$$U(X_t^p - Y_t) = U(x - Y_0) + \text{ local martingale} \tag{2.7}$$

$$\quad + \int_0^t U'(X_s^p - Y_s)(f(s, R_s^{0,r}, Z_s) - h(s, p_s, Z_s))ds \tag{2.8}$$

Equation (2.6) implies that the bounded variation process in (2.8) is decreasing and hence that $U(X^p - Y)$ is a local supermartingale. By choosing $p_t = \Pi_{C(t,R_t^{0,r})}(Z_t + \frac{1}{\eta}\vartheta^*(t, R_t^{0,r})), (t,r) \in [0,T] \times \mathbb{R}^m$ the integrand in (2.8) vanishes, and therefore in this case $U(X^p - Y)$ is a local martingale. □

With the help of Lemma 2.1 we can express the maximal expected utility $V^F(x)$ and the optimal investment strategy in terms of the solution of (2.5).

Theorem 2.2. *The value function satisfies*

$$V^F(x) = U(x - Y_0),$$

and there exists an optimal image strategy p*, given by*

$$p_t = \Pi_{C(t,R_t^{0,r})}(Z_t + \frac{1}{\eta}\vartheta^*(t, R_t^{0,r})), \quad t \in [0,T].$$

From (2.3) we immediately obtain the following expression for the optimal investment strategy.

Corollary 2.3. *The optimal strategy* π *is given by*

$$\pi_t = Z_t\beta^+ \ (t, R_t^{0,r}) + \frac{1}{\eta}\alpha^*(\beta\beta^*)^{-1} \ (t, R_t^{0,r}), \quad t \in [0,T].$$

We remark that Theorem 2.2 can be generalized to the situation where the constraint sets C are arbitrary closed sets (see [8]).

Proof of Theorem 2.2. For $t \in [0,T]$ let $p_t = \Pi_{C(t,R_t^{0,r})}(Z_t + \frac{1}{\eta}\vartheta^*(t, R_t^{0,r}))$. According to the preceding lemma there exists a sequence of stopping time τ_n converging to T, a.s. such that for all $n \geq 1$, the stopped process $U(G^{\widehat{\pi}}_{\cdot \wedge \tau_n} - Y_{\cdot \wedge \tau_n})$ is a martingale. Now observe that

$$U(X^p - Y) \quad = \quad e^{\eta(Y_0 - x)} \mathcal{E}\left(-\eta \int_0^\cdot \left(p_s - (Z_s + \frac{\vartheta_s}{\eta}) + \frac{\vartheta_s}{\eta} \right) dW_s \right).$$

The definition of p implies that $|p| \leq |Z| + \frac{1}{\eta}\|\vartheta\|_\infty$, and hence for every stopping time τ we have $\int_\tau^T |p|_s^2 ds \leq \int_\tau^T |Z_s|_s^2 ds + T\frac{1}{\eta^2}\|\vartheta\|_\infty^2$. This means that $(p \cdot W)$ is a BMO martingale (for further details see [8]). This further yields that $\{U(X^p_\rho - Y_\rho) : \rho$ stopping time with values in $[0,T]\}$ is uniformly integrable, and hence $p \in \mathcal{A}$. Moreover, $\lim_n EU(X^p_{\cdot \wedge \tau_n} - Y_{\cdot \wedge \tau_n}) = EU(X^p - Y)$, from which we deduce $EU(X^p_T - Y_T) = EU(x - Y_0)$.

Note that for all $\tilde{p} \in \mathcal{A}$ we have

$$EU(G^{\tilde{p}}_T - Y_T) \leq EU(G^{\tilde{p}}_0 - Y_0) = EU(x - Y_0),$$

which shows that p is indeed the optimal image strategy. Finally, it follows that $V^F(0, r, x) = EU(x - Y_0)$. □

2.3. Indifference price and optimal hedge

The optimal strategy π can be decomposed into the sum of a pure investment strategy and a pure hedging component. In order to describe the pure hedging component, we shall consider the utility maximization problem with and without the additional obligation $F(R^{0,r}_T)$, compute the optimal strategies in both cases, and then take their difference. So let $(\widehat{Y}^{t,r}, \widehat{Z}^{t,r}) \in \mathcal{S}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ be the solution of the BSDE with generator f, defined as in (2.4), but terminal condition equal to 0,

$$\widehat{Y}^{t,r}_s = -\int_s^T \widehat{Z}^{t,r}_u dW_u + \int_s^T f(u, R^{t,r}_u, \widehat{Z}^{t,r}_u) du, \quad s \in [t, T].$$

From Theorem 2.2 we obtain that

$$V^0(t, x, r) = -e^{-\eta(x - \widehat{Y}^{t,r}_t)}, \quad (t, r) \in [0, T] \times \mathbb{R}^m, x \in \mathbb{R}^k,$$

and the optimal strategy $\widehat{\pi}$ on $[t, T]$ satisfies

$$\widehat{\pi}_s \beta(s, R^{t,r}_s) = \Pi_{C(s, R^{t,r}_s)}[\widehat{Z}^{t,r}_s + \frac{1}{\eta}\vartheta(s, R^{t,r}_s)], \quad s \in [t, T].$$

The presence of the derivative $F(R_T)$ leads to a change in the optimal strategy from $\widehat{\pi}$ to π. More precisely, let $(Y^{t,r}, Z^{t,r})$ be unique solution of the BSDE

$$Y^{t,r}_s = F(R^{t,r}_T) - \int_s^T Z^{t,r}_u dW_u + \int_s^T f(u, R^{t,r}_u, Z^{t,r}_u) du, \quad s \in [t, T].$$

Theorem 2.2 implies

$$V^F(t, x, r) = -e^{-\eta(x - Y^{t,r}_t)},$$

and the optimal strategy π on $[t, T]$ satisfies

$$\pi_s \beta(s, R^{t,r}_s) = \Pi_{C(s, R^{t,r}_s)}[Z^{t,r}_s + \frac{1}{\eta}\vartheta(s, R^{t,r}_s)], \quad s \in [t, T].$$

The Markov property of our risk process R guarantees that the optimal strategies depend only on time and the actual value of R.

Lemma 2.4. *There exist measurable deterministic functions ν and $\widehat{\nu}$, defined on $[0,T] \times \mathbb{R}^m$ and taking values in $\mathbb{R}^d$, such that for $(t,r) \in [0,T] \times \mathbb{R}^m$, the optimal strategies, conditioned on $R_t = r$, are given by $\pi_s^{t,r} = \nu(s, R_s^{t,r})$ and $\widehat{\pi}_s^{t,r} = \widehat{\nu}(s, R_s^{t,r})$ for all $s \in [t,T]$.*

Proof. See Theorem 5.13 in [2]. □

Next we define for all $(t,r) \in [0,T] \times \mathbb{R}^m$

$$\Delta(t,r) = \nu(t,r) - \widehat{\nu}(t,r).$$

Then the optimal investment π satisfies

$$\pi(t,r) = \widehat{\pi}(t,r) + \Delta(t,r).$$

$\widehat{\pi}$ represents a pure investment part, and Δ is the part of the strategy that compensates the random obligation $F(R_T^{t,r})$. We therefore call Δ *optimal hedge.*

Since $\Pi_{C(s,R_s^{t,r})}$ is a linear operator, the optimal hedge satisfies

$$\Delta(s, R_s^{t,r}) = \Pi_{C(s,R_s^{t,r})}[Z_s^{t,r} - \widehat{Z}_s^{t,r}] = \left(Z_s^{t,r} - \widehat{Z}_s^{t,r}\right)(\beta^*(\beta\beta^*)^{-1})(s, R_s^{t,r}),$$

which will be further simplified in the subsequent section.

It turns out that the optimal hedge Δ is closely related to the indifference price of the obligation $F(R_T)$. As usual, we mean by *indifference price* the amout of money $p \in \mathbb{R}$ such that the economic agent is indifferent between having $F(R_T)$ in his portfolio or receiving the riskless payment p.

The difference between $\widehat{\pi}$ and π measures the diversifying impact of $F(R_T)$, also called *diversification pressure*. We will see that we can express the diversification pressure in terms of a price sensitivity multiplied with the *hedge ratio* we encountered already in Section 1. To this end define for all $(t,r) \in [0,T] \times \mathbb{R}^m$,

$$p(t,r) = Y_t^{t,r} - \widehat{Y}_t^{t,r}.$$

It turns out that $p(t,r)$ is the indifference price of $F(R_T^{t,r})$.

Theorem 2.5. *For $(t,r) \in [0,T] \times \mathbb{R}^m$ the quantity $p(t,r)$ represents the indifference price of $F(R_T^{t,r})$, i.e.*

$$V^F(t,x,r) = V^0(t, x - p(t,r), r).$$

Proof. Let $x \in \mathbb{R}^k$, $(t,r) \in [0,T] \times \mathbb{R}^m$ be given. Recall that $V^F(x,t,r) = -e^{-\eta(x - Y_t^{t,r})}$ and $V^0(x,t,r) = -e^{-\eta(x - \widehat{Y}_t^{t,r})}$. Setting $V^F(t,x,r) = V^0(t, x - p(t,r), r)$, immediately gives the result. □

2.4. Delta hedging

If we impose stronger smoothness conditions on the coefficients of the index process R and the function F, then we can show that the price function p is differentiable in r, and we can obtain an explicit representation of the optimal hedge in terms of the price gradient. To this end we need to introduce the following class of functions.

Definition 2.6. Let $n, p \geq 1$. We denote by $\mathbf{B^{n\times p}}$ the set of all functions $h : [0,T] \times \mathbb{R}^m \to \mathbb{R}^{n\times p}$, $(t,x) \mapsto h(t,x)$, differentiable in x, for which there exists a constant $C > 0$ such that $\sup_{(t,x)\in[0,T]\times\mathbb{R}^m} \sum_{i=1}^m \left|\frac{\partial h(t,x)}{\partial x_i}\right| \leq C$, for all $t \in [0,T]$ we have $\sup_{x\in\mathbb{R}^m} \frac{|h(t,x)|}{1+|x|} \leq C$, and $x \mapsto \frac{\partial h(t,x)}{\partial x}$ is Lipschitz continuous with Lipschitz constant C.

We will assume that the coefficients of the index diffusion satisfy in addition to (R1) the following two conditions

(R2) $\sigma \in \mathbf{B^{m\times d}}$, $b \in \mathbf{B^{m\times 1}}$,

(R3) F is a bounded and twice differentiable function such that

$$\nabla F \cdot \sigma \in \mathbf{B}^{1\times d} \text{ and } \sum_{i=1}^m b_i(t,r)\frac{\partial}{\partial r_i}F(r) + \frac{1}{2}\sum_{i,j=1}^m [\sigma\sigma^*]_{ij}(t,r)\frac{\partial^2}{\partial r_i \partial r_j}F(r) \in \mathbf{B}^{1\times 1}.$$

Theorem 2.7. *Suppose that (R1), (R2) and (R3) are satisfied. Besides, suppose that the volatility matrix β and the drift density α are bounded, Lipschitz continuous in r, differentiable in r and that for all $1 \leq i \leq k$, $1 \leq j \leq d$ the derivatives $\nabla_r\beta_{ij}$ and $\nabla_r\alpha_i$ are also Lipschitz continuous in r. Then the optimal hedge satisfies, for all $(t,r) \in [0,T]\times\mathbb{R}^m$,*

$$\Delta(t,r) = \nabla_r p\ \sigma\beta^+\ (t,r).$$

Proof. Under conditions (R1)-(R3) we can show that the solution processes (Y,Z) resp. $(\hat{Y},\hat{Z})$ are differentiable with respect to the initial states of the index process, and that Z resp. $\hat{Z}$ is the Malliavin trace of Y resp. $\hat{Y}$. This smoothness transfers to p via its representations by means of the BSDE solutions. The identification of the control processes Z resp. $\hat{Z}$ by the Malliavin traces of Y resp. $\hat{Y}$ then directly relates Δ with ∇p. For details see [1] and [2]. □

The matrix $\sigma\beta^+(t,r)$ can be interpreted as hedge ratio. To illustrate this, let $k = m = 1$, $d = 2$, $\sigma = \begin{pmatrix} a & 0 \end{pmatrix}$, $\beta = \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix}$. Then the risk process is driven by the martingale $M = \int_0^\cdot a(s,r)dW_s^1$, and the financial asset by $N = \int_0^\cdot(\gamma_1(t,r)dW_t^1 + \gamma_2(t,r)dW_t^2)$. The instantaneous correlation between the driving martingales M and N at time t, conditioned on the risk process to be r, is given by

$$\rho(t,r) = \frac{dE(M_tN_t - M_0N_0)}{\sqrt{dE(\langle M,M\rangle_t)}\sqrt{dE(\langle N,N\rangle_t)}} = \frac{\gamma_1}{\sqrt{\gamma_1^2+\gamma_2^2}}(t,r)$$

The volatility of the risk source is $\text{vola}_R = a$, and the one of the financial asset is $\text{vola}_S = \sqrt{\gamma_1^2 + \gamma_2^2}$. Now observe that

$$\sigma\beta^*(\beta\beta^*)^{-1}(t,r) = \rho \, \frac{\text{vola}_R}{\text{vola}_S} \, (t,r),$$

which, in accordance with Section 1, we call again *hedge ratio*. In dimension 1 we may thus reformulate Theorem 2.7 as follows.

Theorem 2.8. *Let $k = m = 1$, $d = 2$. Then the optimal hedge is equal to the hedge ratio h multiplied with the sensitivity of the indifference price with respect to the risk source, i.e.*

$$\Delta = \frac{\partial p}{\partial r} \cdot h.$$

References

[1] S. Ankirchner, P. Imkeller, and G. Dos Reis, *Classical and variational differentiability of BSDEs with quadratic growth*, Electron. J. Probab., **12 (53)** (2007), 1418–1453 (electronic).

[2] S. Ankirchner, P. Imkeller, and G. Dos Reis, *Pricing and hedging of derivatives based on non-tradable underlyings*, to appear in Mathematical Finance, 2008.

[3] S. Ankirchner, P. Imkeller, and A. Popier, *Optimal cross hedging of insurance derivatives*, Stochastic Analysis and Applications, **26 (4)** (2008), 679–709.

[4] M. H. A. Davis, *Optimal hedging with basis risk*, In: *From Stochastic Calculus to Mathematical Finance*, Springer, Berlin, (2006), 169–187.

[5] W. Fleming and M Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer Verlag, **1993**.

[6] V. Henderson, *Valuation of claims on nontraded assets using utility maximization*, Math. Finance, **12 (4)** (2002), 351–373.

[7] V. Henderson and D. Hobson, *Real options with constant relative risk aversion*, J. Econom. Dynam. Control, **27 (2)** (2002), 329–355.

[8] Y. Hu, P. Imkeller, and M. Müller, *Utility maximization in incomplete markets*, Ann. Appl. Probab., **15 (3)** (2005), 1691–1712.

[9] J. C. Hull, *Options, Futures, and Other Derivatives*, Prentice Hall, 5th edition, **2003**.

[10] M. Kobylanski, *Backward stochastic differential equations and partial differential equations with quadratic growth*, Ann. Probab., **28 (2)** (2000), 558–602.

[11] N. Krylov, *Controlled Diffusion Processes*, Springer Verlag, **1980**.

[12] M. Monoyios, *Performance of utility-based strategies for hedging basis risk*, Quant. Finance, **4 (3)** (2004), 245–255.

[13] M. Musiela and T. Zariphopoulou, *An example of indifference prices under exponential preferences*, Finance Stoch., **8 (2)** (2004), 229–239.

Stefan Ankirchner
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany