

# Simplified mean-variance portfolio optimisation

Claudio Fontana<sup>1</sup>      Martin Schweizer<sup>2</sup>

June 4, 2012

## Abstract

We propose a simplified approach to mean-variance portfolio problems by changing their parametrisation from trading strategies to final positions. This allows us to treat, under a very mild no-arbitrage-type assumption, a whole range of quadratic optimisation problems by simple mathematical tools in a unified and model-independent way. We provide explicit formulas for optimal positions and values, connections between the solutions to the different problems, two-fund separation results, and explicit expressions for indifference values.

**Key words:** mean-variance, portfolio choice, hedging, indifference valuation, Markowitz problem, two-fund separation, no approximate profits, minimum variance, Sharpe ratio

**JEL classification codes:** G11, C61, D81

This version: June 4, 2012

*(to appear in Mathematics and Financial Economics)*

---

<sup>1</sup>Politecnico di Milano, Department of Mathematics, p.zza L. da Vinci 32, I-20133 Milano (Italy), and University of Padova, Department of Pure and Applied Mathematics, via Trieste 63, I-35121 Padova (Italy), [claudio.fontana@polimi.it](mailto:claudio.fontana@polimi.it)

<sup>2</sup>ETH Zurich, Department of Mathematics, Rämistrasse 101, CH-8092 Zurich (Switzerland), and Swiss Finance Institute, Walchestrasse 9, CH-8006 Zurich (Switzerland), [martin.schweizer@math.ethz.ch](mailto:martin.schweizer@math.ethz.ch)

## 0 Introduction

Mean-variance portfolio optimisation is one of the classical problems in financial economics. It has recently gained a lot of popularity in mathematical finance and has been studied by diverse and highly sophisticated tools from stochastic calculus. To which extent is that really needed?

To make the question more precise, consider one standard formulation of the classical Markowitz problem (there are other versions and we discuss them all in the paper): Given a financial market, find a portfolio with maximal return (mean) given a constraint on its risk (variance). The familiar mathematical description is to search for a (self-financing) strategy  $\vartheta$  whose resulting gains from trade  $G_T(\vartheta)$  maximise  $E[G_T(\vartheta)]$  over all allowed  $\vartheta \in \Theta$  subject to  $\text{Var}[G_T(\vartheta)] \leq \sigma^2$  for some constant  $\sigma^2 > 0$ . The control variable is the strategy  $\vartheta$ , and in particular, “portfolio” is interpreted as “trading strategy”. In a one-period model with returns given by an  $\mathbb{R}^d$ -valued random variable  $\Delta S$ , a strategy is simply a constant vector  $\vartheta \in \mathbb{R}^d$ , and trading gains are the scalar product  $\vartheta^\top \Delta S$ . In a continuous-time model with (discounted) asset prices described by an  $\mathbb{R}^d$ -valued semimartingale  $(S_t)_{0 \leq t \leq T}$ , a strategy is an  $\mathbb{R}^d$ -valued predictable  $S$ -integrable stochastic process (satisfying some technical conditions), and  $G_T(\vartheta)$  is given by the real-valued stochastic integral  $\int_0^T \vartheta_u dS_u$ .

In those terms, what do we want to know? It is important to realise that this depends on our goals. To achieve the optimum for the Markowitz problem by actually trading in the market, we need to know the optimal *strategy*  $\vartheta^*$ . The same holds if we consider a quadratic hedging problem and want to implement the optimal hedge. For those purposes, stochastic calculus techniques are indispensable if we work in continuous time. But perhaps we are only interested (e.g. for theoretical analysis) in the (structure of the) optimal final *position*  $g^* = G_T(\vartheta^*)$ , or even only in the *value* (here, the maximal expected gain  $E[g^*]$ ) of the optimisation problem. The latter situation occurs for instance if we use a mean-variance criterion to determine an indifference valuation rule.

The message of our paper is that finding optimal *positions and values* for mean-variance problems is very simple and in particular does not need stochastic calculus nor any specific modelling structure. The key point is that we need not look at  $S$  and  $\vartheta$  separately — all that matters for our purposes is  $G_T(\vartheta)$ . As this depends linearly on  $\vartheta$ , the set  $\mathcal{G}$  of all gains  $g$  from trade in a frictionless financial market is simply a linear space. (Of course, frictions or transaction costs will complicate this; but then we already leave the classical setting.) Moreover, that space  $\mathcal{G}$  should be a subset of  $L^2(P)$  since our problem formulation involves mean and variance. In other words, we no longer look at trading *strategies* as control variables, but only at the resulting final *positions* — and in particular, “portfolio” is now interpreted as “position”. It turns out that this change of parametrisation from strategies in  $\Theta$  to positions in  $\mathcal{G}$  makes everything very simple and tractable and leads to a unified perspective on mean-variance portfolio optimisation.

Of course, this idea is not new. As we discuss in more detail in Section 5, it can be traced back at least as far as Kreps [20]. In the context of describing the mean-variance

frontier, switching from portfolio space to asset space is explicitly recommended in Chapter 5 of Cochrane [7]. Nevertheless, we offer more than just an abstract rewriting or rederivation of existing results. We make a deliberate effort to reduce assumptions to a minimum and to keep the setting as well as the arguments as simple and transparent as possible. We systematically use Hilbert space arguments at the level of positions to solve and relate a whole range of four mean-variance optimisation problems. We explicitly work out the connections between the four solutions, and we do all this carefully in an undiscounted model-independent frictionless framework under a very mild no-arbitrage-type condition. Our results include explicit formulas for optimal positions and values, two-fund separation and beta representation results, a CAPM formula, and explicit indifference valuation rules.

The paper is structured as follows. Section 1 presents the setup and the formulation of our four mean-variance optimisation problems. Section 2 contains the mathematics — it solves the four problems explicitly and provides a number of connections between their solutions. We even do this more generally than discussed above, by replacing  $g$  with  $g - Y$  for some exogenous extra financial position  $Y$ . Section 3 starts on the applications to financial economics; it determines the mean-variance frontier, derives two different but related two-fund separation results, and presents a CAPM formula. Section 4 introduces mean-variance indifference valuation. Because we can compute the *values* of our optimisation problems explicitly, we can also explicitly obtain, for a suitably chosen  $Y$ , the compensation (financial amount)  $h$  at which an agent is indifferent, under a mean-variance criterion and at optimal investment, between either selling a contingent claim  $H$  for an amount of  $h$  or not selling  $H$  and not getting extra money. Finally, Section 5 contains a detailed discussion of related work in the literature.

## 1 Setup and problem

This section describes the framework and introduces the mean-variance portfolio optimisation problems we are interested in. For a probability space  $(\Omega, \mathcal{F}, P)$ , call  $L^2 := L^2(\Omega, \mathcal{F}, P)$  the space of all (equivalence classes of) real-valued square-integrable random variables with scalar product  $(X, Y) = E[XY]$  and norm  $\|X\|_{L^2} = (E[X^2])^{1/2}$ . Let  $\mathcal{G} \neq \emptyset$  be a subset of  $L^2$ , denote by  $\mathcal{G}^\perp := \{X \in L^2 \mid (X, Y) = 0 \text{ for all } Y \in \mathcal{G}\}$  its orthogonal complement in  $L^2$ , and write  $\overline{\mathcal{G}}$  for its closure in  $L^2$ . Finally, let  $B$  be a real-valued random variable in  $L^2$  with  $B > 0$   $P$ -a.s.

The financial interpretation is as follows. Think of a time horizon  $T \in (0, \infty)$  and let  $t = 0$  be the initial time. Then  $\mathcal{G}$  represents the set of all undiscounted cumulated gains from trade (evaluated at time  $T$ ) generated by suitable self-financing trading strategies starting at  $t = 0$  from zero initial capital. The element  $B$  represents the strictly positive value (at the final time  $T$ ) of a *numeraire* asset having initial value 1 and can, but need not, be interpreted as the final value of a *savings account*. We avoid calling  $B$  “riskless”; investing one unit of money in this asset only guarantees that we end up at  $T$  with a strictly positive amount  $B$ , which is however

random and can be strictly less than the initial investment of 1. The set

$$\mathcal{A} := \{cB + g \mid c \in \mathbb{R}, g \in \mathcal{G}\} = \mathbb{R}B + \mathcal{G}$$

then consists of all *attainable* undiscounted final wealths, i.e. those square-integrable payoffs which can be replicated in the abstract financial market  $(B, \mathcal{G})$  by following a self-financing strategy starting from some initial capital  $c$ . We emphasise that our focus is on final results or positions at time  $T$ ; the notion of a trading strategy need not be and is not specified. Note also that we do not assume that  $\mathcal{G}$  is closed in  $L^2$ . Square-integrability is imposed to ensure existence of means and variances, which is necessary when dealing with mean-variance problems. Finally, the Hilbert space structure of  $L^2$  allows an easy and efficient derivation of general solutions to several mean-variance problems, as will be seen in Section 2.

**Remark 1.1. 1)** It is worth emphasising that apart from linearity of  $\mathcal{G}$  and the obvious requirement of square-integrability, the present setup for an abstract financial market does not rely on any underlying modelling structure. As a consequence, all the results we are going to present are model-independent, and in particular hold for both discrete-time and continuous-time frictionless models. We refer the reader to Examples 1–3 in Schweizer [35] for an illustration of how typical financial models can be embedded into the present abstract setting.

2) In the terminology used in Chapter 4 of Cochrane [7],  $\overline{\mathcal{G}}$  might be called *payoff space*.  $\diamond$

Let us now introduce a basic **standing assumption** for the rest of the paper.

**Assumption I.** The two following conditions hold:

- (a)  $\mathcal{G}$  is a linear subspace of  $L^2$ .
- (b) There are *no approximate riskless profits in  $L^2$* , meaning that  $\overline{\mathcal{G}}$  does not contain 1.

Intuitively, part (a) of Assumption I amounts to considering a frictionless financial market without constraints, transaction costs or other nonlinearities in trading. The same condition appears in Chapter 4 of Cochrane [7] on p.62 as assumption (A1). The condition  $1 \notin \overline{\mathcal{G}}$  in part (b) represents an abstract and very mild no-arbitrage-type condition. It can be equivalently formulated as  $\mathbb{R} \cap \overline{\mathcal{G}} = \{0\}$ , and this amounts to excluding the clearly undesirable situation where an agent is able to reach, or approximate in the  $L^2$ -sense, a deterministic riskless final wealth from zero initial capital. As will be shown in the next section, (b) is necessary and sufficient for the solvability of the quadratic problems we are now going to introduce.

We shall be mainly concerned with four mean-variance portfolio optimisation problems, denoted as Problems (A)–(D) and formulated in the following abstract terms. Let  $Y \in L^2$  represent the final undiscounted value of a generic financial position/liability,  $\alpha \in (0, \infty)$  a given risk-aversion coefficient,  $\mu \in \mathbb{R}$  a target minimal expected value and  $\sigma^2 \in (0, \infty)$  a target maximal variance. Then we consider

**Problem (A')**  $E[g - Y] - \alpha \text{Var}[g - Y] = \max!$  over all  $g \in \mathcal{G}$ .

**Problem (B')**  $\text{Var}[g - Y] = \min!$  over all  $g \in \mathcal{G}$  such that  $E[g - Y] \geq \mu$ .

**Problem (C')**  $E[g - Y] = \max!$  over all  $g \in \mathcal{G}$  such that  $\text{Var}[g - Y] \leq \sigma^2$ .

**Problem (D')**  $\|Y - g\|_{L^2} = \min!$  over all  $g \in \mathcal{G}$ .

We shall argue below that each Problem (X') has the same optimal value as the corresponding Problem (X) where we optimise over  $\overline{\mathcal{G}}$  instead of  $\mathcal{G}$ .

The financial interpretations of Problems (A')–(D') are obvious. (A') is the portfolio optimisation problem faced by an agent with mean-variance preferences and risk-aversion coefficient  $\alpha$ . (B') and (C') are the classical Markowitz portfolio selection problems, slightly extended by including the random liability  $Y$ . More specifically, in (B'), the agent is interested in minimising the variance of her/his final net position, given a minimal target level  $\mu$  for its expected value. Symmetrically, in (C'), the agent wants to maximise the expected value of her/his final net position, given a maximal target level  $\sigma^2$  for its variance. Finally, (D') consists of finding the optimal quadratic hedge for  $Y$ . We illustrate at the end of Section 3 how different investment situations can be represented via suitable choices of  $Y$ . Note that all these problems, in the language of Cochrane [7], are formulated directly over the payoff space, bypassing the introduction of and need for a strategy space.

**Remark 1.2.** Part (b) of Assumption I excludes the case  $1 \in \overline{\mathcal{G}}$ , but not the case  $1 \in \mathcal{G}^\perp$ . However, the latter situation is neither mathematically interesting nor realistic from an economic point of view. In fact,  $1 \in \mathcal{G}^\perp$  means that  $E[g] = (g, 1) = 0$  for all  $g \in \mathcal{G}$ . But then there is nothing to optimise in (C'), and the constraint in (B') is trivially always or never satisfied, depending on whether  $E[Y] \leq -\mu$  or  $E[Y] > -\mu$ . Finally, (A') reduces to the simpler problem of minimising the variance. In financial terms, the case  $1 \in \mathcal{G}^\perp$  corresponds to the situation where all undiscounted cumulated gains have zero expectation under the original (real-world) probability measure  $P$ . Loosely speaking, this means that the basic instruments available for trade are martingales under  $P$ . In this case, there is no proper notion of a trade-off between risk (variance) and return (expected value), and so we cannot meaningfully consider mean-variance portfolio optimisation problems.  $\diamond$

Due to Remark 1.2, there is no loss of generality in introducing the following additional **standing assumption** for the sequel.

**Assumption II.**  $1 \notin \mathcal{G}^\perp$ , or equivalently  $\{g \in \mathcal{G} \mid E[g] \neq 0\} \neq \emptyset$ .

## 2 Mathematical basics

This section contains the mathematical ingredients for solving Problems (A)–(D). The Hilbert space structure of our framework makes the results both general and easy to obtain. We postpone to later sections all economic considerations and financial applications.

Recall that the orthogonal complement  $\mathcal{G}^\perp$  is a closed linear subspace of  $L^2$ , and denote by  $\pi$  the orthogonal projection in  $L^2$  on  $\mathcal{G}^\perp$ . Since  $\mathcal{G}$  is a linear subspace of  $L^2$  by part (a) of Assumption I, we have  $(\mathcal{G}^\perp)^\perp = \overline{\mathcal{G}}$ . This yields the direct sum decomposition  $L^2 = \overline{\mathcal{G}} \oplus \mathcal{G}^\perp$ , meaning that any  $Y \in L^2$  can be uniquely written as

$$Y = g^Y + N^Y = g^Y + \pi(Y) \quad \text{with } g^Y \in \overline{\mathcal{G}} \text{ and } N^Y = \pi(Y) \in \mathcal{G}^\perp. \quad (2.1)$$

(We use both notations  $\pi(Y)$  and  $N^Y$  to facilitate later comparisons to the literature.) Using this basic orthogonal decomposition, we can already tackle Problem (D'). Note first that

$$\inf_{g \in \mathcal{G}} \|Y - g\|_{L^2} = \inf_{g \in \overline{\mathcal{G}}} \|Y - g\|_{L^2}. \quad (2.2)$$

In fact, “ $\geq$ ” is clear from  $\mathcal{G} \subseteq \overline{\mathcal{G}}$ , and conversely, any  $\bar{g} \in \overline{\mathcal{G}}$  admits a sequence  $(g_n)_{n \in \mathbb{N}} \subseteq \mathcal{G}$  converging in  $L^2$  to  $\bar{g}$ . So  $\inf_{g \in \mathcal{G}} \|Y - g\|_{L^2} \leq \|Y - g_n\|_{L^2} \rightarrow \|Y - \bar{g}\|_{L^2}$  as  $n \rightarrow \infty$ , and as  $\bar{g} \in \overline{\mathcal{G}}$  is arbitrary, we also get “ $\leq$ ” in (2.2). In other words, the optimal value in (D') does not change if we replace  $\mathcal{G}$  by its closure  $\overline{\mathcal{G}}$  in  $L^2$ . The projection theorem then gives

$$\inf_{g \in \mathcal{G}} \|Y - g\|_{L^2} = \inf_{g \in \overline{\mathcal{G}}} \|Y - g\|_{L^2} = \min_{g \in \overline{\mathcal{G}}} \|Y - g\|_{L^2} = \|Y - g^Y\|_{L^2} = \|N^Y\|_{L^2}. \quad (2.3)$$

Optimising over the closed subspace  $\overline{\mathcal{G}}$  ensures existence and uniqueness for the solution to Problem (D), where  $\overline{\mathcal{G}}$  replaces  $\mathcal{G}$  in (D'), and the solution is the projection in  $L^2$  of  $Y$  on  $\overline{\mathcal{G}}$ ,

$$\operatorname{argmin}_{g \in \overline{\mathcal{G}}} \|Y - g\|_{L^2} = g^Y. \quad (2.4)$$

**Remark 2.1.** Also for Problems (A')–(C'), the optimal values do not depend on whether we optimise over  $\mathcal{G}$  or  $\overline{\mathcal{G}}$ . This is easily checked by the same arguments as for (2.2), using that  $g_n \rightarrow \bar{g}$  in  $L^2$  implies  $E[g_n - Y] \rightarrow E[\bar{g} - Y]$  and  $\operatorname{Var}[g_n - Y] \rightarrow \operatorname{Var}[\bar{g} - Y]$ , for any  $Y \in L^2$ .  $\diamond$

In view of Remark 2.1, we henceforth consider Problems (A)–(D) instead of (A')–(D'), where the optimisation now goes over the closed linear subspace  $\overline{\mathcal{G}}$  instead of  $\mathcal{G}$ . As a preliminary to deriving the solutions to (A)–(C), we introduce the following variance-minimisation problem.

**Problem (MV)**  $\quad \operatorname{Var}[Y - g] = \min! \quad \text{over all } g \in \overline{\mathcal{G}}.$

The solution to (MV) is given in the following result and denoted by  $g_{\text{mv}}^Y$ , where the subscript “mv” stands for “minimum variance”. It is obtained via the solution to (D) derived in (2.4).

**Proposition 2.2.** *For  $Y \in L^2$ , Problem (MV) admits in  $\overline{\mathcal{G}}$  the unique solution*

$$g_{\text{mv}}^Y := \operatorname{argmin}_{g \in \overline{\mathcal{G}}} \operatorname{Var}[Y - g] = g^Y - a_Y^* (1 - \pi(1)), \quad \text{where } a_Y^* := \frac{E[N^Y]}{E[\pi(1)]}. \quad (2.5)$$

*Proof.* The key idea for reducing (MV) to (D) is the simple fact that

$$\text{Var}[Y - g] = \min_{a \in \mathbb{R}} E[(Y - g - a)^2].$$

Hence we can write

$$\min_{g \in \overline{\mathcal{G}}} \text{Var}[Y - g] = \min_{g \in \overline{\mathcal{G}}} \min_{a \in \mathbb{R}} E[(Y - g - a)^2] = \min_{a \in \mathbb{R}} \min_{g \in \overline{\mathcal{G}}} E[(Y - a - g)^2].$$

But the inner minimisation over  $\overline{\mathcal{G}}$  is just Problem (D) for  $Y - a$  and is solved by  $g^{Y-a}$ . By linearity of the projection and (2.1), we have  $g^{Y-a} = g^Y - ag^1 = g^Y - a(1 - \pi(1))$  and so

$$\min_{g \in \overline{\mathcal{G}}} \text{Var}[Y - g] = \min_{a \in \mathbb{R}} E[(Y - a - g^{Y-a})^2] = \min_{a \in \mathbb{R}} E[(N^Y - a\pi(1))^2]. \quad (2.6)$$

Now observe that because  $1 - \pi(1)$  is in  $\overline{\mathcal{G}}$ , we have  $(N^Y, 1 - \pi(1)) = 0$  and  $(\pi(1), 1 - \pi(1)) = 0$ . This gives  $E[N^Y \pi(1)] = E[N^Y]$  and  $E[\pi(1)] = E[(\pi(1))^2] = \|\pi(1)\|_{L^2}^2 > 0$  since  $1 \notin \overline{\mathcal{G}}$  by part (b) of Assumption I. Squaring out and completing the square therefore yields

$$E[(N^Y - a\pi(1))^2] = E[\pi(1)] \left( a - \frac{E[N^Y]}{E[\pi(1)]} \right)^2 - \frac{(E[N^Y])^2}{E[\pi(1)]} + E[(N^Y)^2]. \quad (2.7)$$

So the optimal  $a \in \mathbb{R}$  is uniquely given by

$$a_Y^* := \operatorname{argmin}_{a \in \mathbb{R}} E[(N^Y - a\pi(1))^2] = \frac{E[N^Y]}{E[\pi(1)]},$$

and we obtain

$$g_{\text{mv}}^Y := \operatorname{argmin}_{g \in \overline{\mathcal{G}}} \text{Var}[Y - g] = g^{Y-a_Y^*} = g^Y - a_Y^*(1 - \pi(1)).$$

The uniqueness of the solution  $g_{\text{mv}}^Y \in \overline{\mathcal{G}}$  follows from the projection theorem via the uniqueness of  $g^{Y-a} \in \overline{\mathcal{G}}$  for all  $a \in \mathbb{R}$ .  $\square$

Let us now introduce the notation  $R_{\text{mv}}^Y := g_{\text{mv}}^Y - Y$ , where “ $R$ ”, not to be confused with return, stands for the final “result” of a financial position. Then, for any  $g \in \overline{\mathcal{G}}$ , we can write

$$g - Y = R_{\text{mv}}^Y + g - g_{\text{mv}}^Y$$

and hence

$$E[g - Y] = E[R_{\text{mv}}^Y] + E[g - g_{\text{mv}}^Y]. \quad (2.8)$$

Furthermore, due to the optimality of  $g_{\text{mv}}^Y \in \overline{\mathcal{G}}$  and the linearity of  $\overline{\mathcal{G}}$ , the first order condition for (MV) gives for the element  $R_{\text{mv}}^Y$  the fundamental zero-covariance property

$$\text{Cov}(R_{\text{mv}}^Y, g) = 0 \quad \text{for all } g \in \overline{\mathcal{G}}. \quad (2.9)$$

Since  $g - g_{\text{mv}}^Y \in \overline{\mathcal{G}}$  for any  $g \in \overline{\mathcal{G}}$ , this implies that we have

$$\text{Var}[g - Y] = \text{Var}[R_{\text{mv}}^Y] + \text{Var}[g - g_{\text{mv}}^Y]. \quad (2.10)$$

Equations (2.8) and (2.10) show that in Problems (A)–(C), we can isolate the part coming from the minimum variance element  $R_{\text{mv}}^Y$ . Furthermore, since  $g_{\text{mv}}^Y$  is in  $\overline{\mathcal{G}}$  and  $\overline{\mathcal{G}}$  is a linear space, the mapping  $g \mapsto g' := g - g_{\text{mv}}^Y$  is a bijection of  $\overline{\mathcal{G}}$  to itself. These observations suggest that we can reduce the general versions of our mean-variance problems to the particular case  $Y \equiv 0$ . We exploit this in the proofs of the three following propositions.

**Remark 2.3.** For future use in later sections, we compute the mean and variance of the optimal position  $R_{\text{mv}}^Y = g_{\text{mv}}^Y - Y = -N^Y - a_Y^*(1 - \pi(1))$ . From the expression for  $a_Y^*$  in (2.5), we get

$$E[R_{\text{mv}}^Y] = -\frac{E[N^Y]}{E[\pi(1)]}, \quad (2.11)$$

and using (2.6) and (2.7) yields

$$\text{Var}[R_{\text{mv}}^Y] = E[(N^Y)^2] - \frac{(E[N^Y])^2}{E[\pi(1)]}. \quad (2.12)$$

◇

We start with the solution to Problem (A), denoted by  $g_{\text{opt},A}^Y(\gamma)$ , where  $\gamma := 1/\alpha$  is the risk-tolerance corresponding to the risk-aversion  $\alpha$ . Since  $\alpha \in (0, \infty)$ , we have also  $\gamma \in (0, \infty)$ , but it will be useful later to include in Proposition 2.4 the case  $\gamma = 0$  (which intuitively corresponds to infinite risk-aversion). Note that after replacing  $\alpha$  by  $\gamma = 1/\alpha$ , Problem (A) does make sense for  $\gamma = 0$ .

**Proposition 2.4.** For  $Y \in L^2$  and  $\gamma \in [0, \infty)$ , Problem (A) has a unique solution  $g_{\text{opt},A}^Y(\gamma) \in \overline{\mathcal{G}}$ . It is explicitly given by

$$g_{\text{opt},A}^Y(\gamma) = \underset{g \in \overline{\mathcal{G}}}{\text{argmin}}\{\text{Var}[g - Y] - \gamma E[g - Y]\} = g_{\text{mv}}^Y + g_{\text{opt},A}^0(\gamma), \quad (2.13)$$

where  $g_{\text{opt},A}^0(\gamma) \in \overline{\mathcal{G}}$  is the solution to Problem (A) for  $Y \equiv 0$ , explicitly given by

$$g_{\text{opt},A}^0(\gamma) = \underset{g \in \overline{\mathcal{G}}}{\text{argmin}}\{\text{Var}[g] - \gamma E[g]\} = \frac{\gamma}{2} \frac{1}{E[\pi(1)]} (1 - \pi(1)). \quad (2.14)$$

*Proof.* Notice first that with  $\gamma = 1/\alpha$ , Problem (A) can be equivalently formulated as

$$\text{Var}[g - Y] - \gamma E[g - Y] = \min! \quad \text{over all } g \in \overline{\mathcal{G}}.$$

Moreover, equations (2.8) and (2.10) allow us to write, for any  $g \in \overline{\mathcal{G}}$ ,

$$\text{Var}[g - Y] - \gamma E[g - Y] = \text{Var}[R_{\text{mv}}^Y] - \gamma E[R_{\text{mv}}^Y] + \text{Var}[g - g_{\text{mv}}^Y] - \gamma E[g - g_{\text{mv}}^Y].$$

Since  $\overline{\mathcal{G}}$  is linear and contains  $g_{\text{mv}}^Y$ , we can thus reduce (A) to the basic problem

$$\text{Var}[g] - \gamma E[g] = \min! \quad \text{over all } g \in \overline{\mathcal{G}}. \quad (2.15)$$



More precisely, if  $g_{\text{opt},A}^0(\gamma) \in \overline{\mathcal{G}}$  denotes the solution to (2.15), then the solution  $g_{\text{opt},A}^Y(\gamma) \in \overline{\mathcal{G}}$  to (A) in its original formulation is given by (2.13). Hence it only remains to solve (2.15). Following the same idea as in the proof of Proposition 2.2, we write

$$\text{Var}[g] - \gamma E[g] = \min_{a \in \mathbb{R}} E[(g - a)^2] - \gamma E[g] = \min_{a \in \mathbb{R}} \left( E \left[ \left( g - \left( a + \frac{\gamma}{2} \right) \right)^2 \right] - \frac{\gamma^2}{4} - a\gamma \right). \quad (2.16)$$

But for  $Y \equiv a + \frac{\gamma}{2}$ , the solution of (D) is by (2.4) and linearity of the projection

$$g^{a+\frac{\gamma}{2}} = \left( a + \frac{\gamma}{2} \right) g^1 = \left( a + \frac{\gamma}{2} \right) (1 - \pi(1)). \quad (2.17)$$

Combining this with (2.16) and completing the square gives

$$\begin{aligned} \min_{g \in \overline{\mathcal{G}}} \{ \text{Var}[g] - \gamma E[g] \} &= \min_{a \in \mathbb{R}} \left\{ \min_{g \in \overline{\mathcal{G}}} E \left[ \left( g - \left( a + \frac{\gamma}{2} \right) \right)^2 \right] - \frac{\gamma^2}{4} - a\gamma \right\} \\ &= \min_{a \in \mathbb{R}} \left\{ E \left[ \left( \left( a + \frac{\gamma}{2} \right) \pi(1) \right)^2 \right] - \frac{\gamma^2}{4} - a\gamma \right\} \\ &= \min_{a \in \mathbb{R}} E[\pi(1)] \left( a - \frac{\gamma}{2} \frac{1 - E[\pi(1)]}{E[\pi(1)]} \right)^2 - \frac{\gamma^2}{4} \frac{E[1 - \pi(1)]}{E[\pi(1)]}. \end{aligned} \quad (2.18)$$

Note that as in the proof of Proposition 2.2, part (b) of Assumption I that  $1 \notin \overline{\mathcal{G}}$  gives  $E[\pi(1)] > 0$ . The last expression in (2.18) is clearly minimised over  $a$  by the unique value

$$a_\gamma^* := \frac{\gamma}{2} \frac{E[1 - \pi(1)]}{E[\pi(1)]},$$

and together with (2.17), this yields

$$g_{\text{opt},A}^0(\gamma) = \operatorname{argmin}_{g \in \overline{\mathcal{G}}} \{ \text{Var}[g] - \gamma E[g] \} = g^{a_\gamma^* + \frac{\gamma}{2}} = \frac{\gamma}{2} \frac{1}{E[\pi(1)]} (1 - \pi(1)).$$

The uniqueness of the solution again follows from the projection theorem via the uniqueness of  $g^{a+\frac{\gamma}{2}} \in \overline{\mathcal{G}}$  for all  $a \in \mathbb{R}$  and  $\gamma \in [0, \infty)$ .  $\square$

**Remark 2.5. 1)** The proofs of Propositions 2.2 and 2.4 both rely on the elementary identity  $\text{Var}[X] = \min_{a \in \mathbb{R}} E[(X - a)^2]$  for  $X \in L^2$ . This allows us to reduce variance-minimisation problems to particular cases of Problem (D).

**2)** The above trick of expressing the variance as an optimal value for a minimisation problem over  $\mathbb{R}$  is also at the root of the appearance of the quantity  $1 - \pi(1)$ ; in fact, this is simply the projection in  $L^2$  of the constant  $1 \in \mathbb{R}$  on  $\overline{\mathcal{G}}$ .

**3)** The solution to (MV) can be recovered from the solution to (A), and this also illustrates why it is useful to pass from risk-aversion  $\alpha$  to risk-tolerance  $\gamma = 1/\alpha$ . If we let  $\gamma = 0$ , Proposition 2.4 yields

$$g_{\text{opt},A}^Y(0) = \operatorname{argmin}_{g \in \overline{\mathcal{G}}} \text{Var}[g - Y] = g_{\text{mv}}^Y.$$

This simple relation is in line with intuition, because  $\gamma = 0$  corresponds to infinite risk-aversion ( $\alpha = \infty$ ), which means in (A) that one is only interested in minimising the risk (variance).

4) As seen in the proof of Proposition 2.4, the condition  $1 \notin \overline{\mathcal{G}}$  of no approximate riskless profits in  $L^2$  is *sufficient* for the existence of a (unique) solution to Problem (A). But it is also *necessary*. In fact, suppose to the contrary that  $\tilde{g} \in \overline{\mathcal{G}}$  solves (A), but  $1 \in \overline{\mathcal{G}}$ . Then  $g' := \tilde{g} + 1 \in \overline{\mathcal{G}}$  satisfies  $\text{Var}[g' - Y] = \text{Var}[\tilde{g} - Y]$  and  $E[g' - Y] = 1 + E[\tilde{g} - Y] > E[\tilde{g} - Y]$ , contradicting the optimality of  $\tilde{g}$ .

5) One can prove the uniqueness of the solution to (A) directly by using only its optimality. But the argument above via the projection theorem leads to a more compact proof.  $\diamond$

The results obtained so far do not rely on Assumption II that  $1 \notin \mathcal{G}^\perp$ . It is easy to see from Proposition 2.4 that if  $1 \in \mathcal{G}^\perp$ , the solutions to Problems (MV) and (A) coincide, since  $1 \in \mathcal{G}^\perp$  implies  $\pi(1) \equiv 1$ . But for tackling Problems (B) and (C), we shall exploit Assumption II. The basic idea is well known; it is folklore that the solutions to the Markowitz problems (B) and (C) are obtained by choosing for the risk-aversion  $\alpha$  in (A) a particular value, in such a way that it matches up with the respective constraint in (B) or (C). In more detail, this goes as follows. In analogy to  $R_{\text{mv}}^Y = g_{\text{mv}}^Y - Y$ , we first introduce the notation

$$R_{\text{opt},A}^Y(\gamma) := g_{\text{opt},A}^Y(\gamma) - Y = R_{\text{mv}}^Y + g_{\text{opt},A}^0(\gamma) = R_{\text{mv}}^Y + \frac{\gamma}{2} \frac{1}{E[\pi(1)]} (1 - \pi(1)).$$

Using

$$\text{Var}[1 - \pi(1)] = \text{Var}[\pi(1)] = E[(\pi(1))^2] - (E[\pi(1)])^2 = E[\pi(1)](1 - E[\pi(1)]) \quad (2.19)$$

and recalling from (2.9) the zero-covariance property of  $R_{\text{mv}}^Y$ , we then compute

$$E[R_{\text{opt},A}^Y(\gamma)] = E[R_{\text{mv}}^Y] + \frac{\gamma}{2} \frac{E[1 - \pi(1)]}{E[\pi(1)]}, \quad (2.20)$$

$$\text{Var}[R_{\text{opt},A}^Y(\gamma)] = \text{Var}[R_{\text{mv}}^Y] + \frac{\gamma^2}{4} \frac{E[1 - \pi(1)]}{E[\pi(1)]}. \quad (2.21)$$

So for  $1 \in \mathcal{G}^\perp$ , we obtain  $E[R_{\text{opt},A}^Y(\gamma)] = E[R_{\text{mv}}^Y]$  and  $\text{Var}[R_{\text{opt},A}^Y(\gamma)] = \text{Var}[R_{\text{mv}}^Y]$  for all  $\gamma \in [0, \infty)$ . But if Assumption II holds, then we have  $E[1 - \pi(1)] = \|1 - \pi(1)\|_{L^2}^2 > 0$ , and therefore the functions  $\gamma \mapsto E[R_{\text{opt},A}^Y(\gamma)]$  from  $[0, \infty)$  to  $[E[R_{\text{mv}}^Y], \infty)$  and  $\gamma \mapsto \text{Var}[R_{\text{opt},A}^Y(\gamma)]$  from  $[0, \infty)$  to  $[\text{Var}[R_{\text{mv}}^Y], \infty)$  are both surjective. This implies that for any  $\mu \in [E[R_{\text{mv}}^Y], \infty)$ , there exists  $\gamma_\mu \in [0, \infty)$  such that  $E[R_{\text{opt},A}^Y(\gamma_\mu)] = \mu$ , and analogously, any  $\sigma^2 \in [\text{Var}[R_{\text{mv}}^Y], \infty)$  admits some  $\gamma_{\sigma^2} \in [0, \infty)$  such that  $\text{Var}[R_{\text{opt},A}^Y(\gamma_{\sigma^2})] = \sigma^2$ . Under the (standing) Assumptions I and II, this simple observation allows us to derive the solutions to Problems (B) and (C) from the solution to Problem (A), as shown in the next two results.

**Proposition 2.6.** *Let  $Y \in L^2$  and  $\mu \in \mathbb{R}$ . If  $\mu > E[R_{\text{mv}}^Y]$ , then Problem (B) admits a unique solution  $g_{\text{opt},B}^Y(\mu) \in \overline{\mathcal{G}}$ . It is explicitly given by*

$$g_{\text{opt},B}^Y(\mu) = g_{\text{mv}}^Y + g_{\text{opt},B}^0(\mu - E[R_{\text{mv}}^Y]), \quad (2.22)$$

where  $g_{\text{opt},B}^0(m)$  is the solution to Problem (B) for  $Y \equiv 0$  and constraint  $m$ , explicitly given by

$$g_{\text{opt},B}^0(m) = \frac{m}{E[1 - \pi(1)]} (1 - \pi(1)). \quad (2.23)$$

If  $\mu \leq E[R_{\text{mv}}^Y]$ , then Problem (B) has  $g_{\text{mv}}^Y$  as unique solution.

*Proof.* As in the proof of Proposition 2.4, (B) can be reduced to the basic version

$$\text{Var}[g] = \min! \quad \text{over all } g \in \overline{\mathcal{G}} \text{ such that } E[g] \geq m, \quad (2.24)$$

where  $m$  in (2.24) stands for  $\mu - E[R_{\text{mv}}^Y]$ . More precisely, if  $g_{\text{opt},B}^0(m) \in \overline{\mathcal{G}}$  denotes the solution to (2.24), then the solution  $g_{\text{opt},B}^Y(\mu) \in \overline{\mathcal{G}}$  to (B) in its original formulation is given by (2.22), due to (2.10) and (2.8).

If  $m \leq 0$ , then (2.24) is trivially solved by  $g \equiv 0$ , which proves the last assertion. For any  $m > 0$ , there is some  $\gamma_m \in (0, \infty)$  with  $m = E[R_{\text{opt},A}^0(\gamma_m)] = E[g_{\text{opt},A}^0(\gamma_m)]$ ; in fact, (2.20) gives due to  $R_{\text{mv}}^0 = 0$  that

$$\gamma_m = 2m \frac{E[\pi(1)]}{E[1 - \pi(1)]}. \quad (2.25)$$

We claim that  $g_{\text{opt},B}^0(m) = g_{\text{opt},A}^0(\gamma_m)$ , i.e. that  $g_{\text{opt},A}^0(\gamma_m)$  solves (2.24). To see this, take  $g' \in \overline{\mathcal{G}}$  with  $E[g'] \geq m = E[g_{\text{opt},A}^0(\gamma_m)]$ . Because  $g_{\text{opt},A}^0(\gamma_m)$  solves (A) for  $\gamma_m$  and  $Y \equiv 0$ , we then get

$$m - \frac{\text{Var}[g']}{\gamma_m} \leq E[g'] - \frac{\text{Var}[g']}{\gamma_m} \leq E[g_{\text{opt},A}^0(\gamma_m)] - \frac{\text{Var}[g_{\text{opt},A}^0(\gamma_m)]}{\gamma_m} = m - \frac{\text{Var}[g_{\text{opt},A}^0(\gamma_m)]}{\gamma_m}.$$

Since  $\gamma_m > 0$ , this implies  $\text{Var}[g'] \geq \text{Var}[g_{\text{opt},A}^0(\gamma_m)]$  which shows that  $g_{\text{opt},A}^0(\gamma_m)$  solves (2.24). The uniqueness of the solution to (B) then follows from the uniqueness of the solution to (A). The explicit expression (2.23) is obtained by inserting (2.25) into (2.14).  $\square$

The solution for (C) is derived next; the proof is symmetric to that of Proposition 2.6.

**Proposition 2.7.** *Let  $Y \in L^2$  and  $\sigma^2 \in [0, \infty)$ . If  $\sigma^2 \geq \text{Var}[R_{\text{mv}}^Y]$ , then Problem (C) admits a unique solution  $g_{\text{opt},C}^Y(\sigma^2) \in \overline{\mathcal{G}}$ . It is explicitly given by*

$$g_{\text{opt},C}^Y(\sigma^2) = g_{\text{mv}}^Y + g_{\text{opt},C}^0(\sigma^2 - \text{Var}[R_{\text{mv}}^Y]), \quad (2.26)$$

where  $g_{\text{opt},C}^0(v)$  is the solution to Problem (C) for  $Y \equiv 0$  and constraint  $v$ , explicitly given by

$$g_{\text{opt},C}^0(v) = \sqrt{\frac{v}{\text{Var}[1 - \pi(1)]}} (1 - \pi(1)). \quad (2.27)$$

If  $\sigma^2 < \text{Var}[R_{\text{mv}}^Y]$ , Problem (C) cannot be solved.

*Proof.* As in the proof of Proposition 2.4, we use (2.8) and (2.10). In view of (2.10), the last assertion and the case  $\sigma^2 = \text{Var}[R_{\text{mv}}^Y]$  are clear; so we focus on the case where  $\sigma^2 > \text{Var}[R_{\text{mv}}^Y]$ . Then (C) can be reduced to the basic version

$$E[g] = \max! \quad \text{over all } g \in \overline{\mathcal{G}} \text{ such that } \text{Var}[g] \leq v, \quad (2.28)$$

where  $v$  stands for  $\sigma^2 - \text{Var}[R_{\text{mv}}^Y]$ . More precisely, if  $g_{\text{opt},C}^0(v) \in \overline{\mathcal{G}}$  denotes the solution to (2.28), then the solution  $g_{\text{opt},C}^Y(\sigma^2) \in \overline{\mathcal{G}}$  to (C) in its original formulation is given by (2.26).

To solve (2.28), note that (2.19) and (2.21) with  $Y \equiv 0$ , hence  $R_{\text{mv}}^Y = 0$ , give for  $v > 0$  that

$$\gamma_v = 2\sqrt{v} \frac{E[\pi(1)]}{\sqrt{\text{Var}[1 - \pi(1)]}} \in (0, \infty) \quad (2.29)$$

satisfies  $v = \text{Var}[R_{\text{opt},A}^0(\gamma_v)] = \text{Var}[g_{\text{opt},A}^0(\gamma_v)]$ . We claim that  $g_{\text{opt},C}^0(v) = g_{\text{opt},A}^0(\gamma_v)$ , i.e. that  $g_{\text{opt},A}^0(\gamma_v)$  solves (2.28). Indeed, for any  $g' \in \overline{\mathcal{G}}$  with  $\text{Var}[g'] \leq v = \text{Var}[g_{\text{opt},A}^0(\gamma_v)]$ , we obtain from the fact that  $g_{\text{opt},A}^0(\gamma_v)$  solves Problem (A) for  $\gamma_v$  and  $Y \equiv 0$  that

$$E[g'] - \frac{v}{\gamma_v} \leq E[g'] - \frac{\text{Var}[g']}{\gamma_v} \leq E[g_{\text{opt},A}^0(\gamma_v)] - \frac{\text{Var}[g_{\text{opt},A}^0(\gamma_v)]}{\gamma_v} = E[g_{\text{opt},A}^0(\gamma_v)] - \frac{v}{\gamma_v}.$$

This yields  $E[g'] \leq E[g_{\text{opt},A}^0(\gamma_v)]$ , showing that  $g_{\text{opt},A}^0(\gamma_v)$  solves (2.28). Uniqueness follows again from the uniqueness of the solution to (A), and the explicit expression (2.27) is obtained by inserting (2.29) into (2.14).  $\square$

**Remark 2.8.** 1) Note that the solutions to Problems (B) and (C) both satisfy their constraints with equalities, at least in the genuinely interesting cases where  $\mu \geq E[R_{\text{mv}}^Y]$  and  $\sigma^2 \geq \text{Var}[R_{\text{mv}}^Y]$ . As a consequence, (B) and (C) could equivalently be formulated with equality constraints. Alternatively, this could be seen by checking directly that an element  $g \in \overline{\mathcal{G}}$  satisfying the constraints with strict inequality cannot be optimal.

2) Propositions 2.4, 2.6 and 2.7 show that the solutions to Problems (A)–(C) have a very similar and simple structure — they all are linear combinations of the minimum variance element  $g_{\text{mv}}^Y$  and  $1 - \pi(1)$ . If one knows a priori the key role played by the element  $1 - \pi(1)$ , the solutions to Problems (A)–(C) can be quickly derived as follows. Notice first that  $\overline{\mathcal{G}} = \overline{\mathcal{G}} + g_{\text{mv}}^Y$  since  $\overline{\mathcal{G}}$  is a linear space and  $g_{\text{mv}}^Y \in \overline{\mathcal{G}}$ . Furthermore, the space  $\overline{\mathcal{G}}$  can be represented as

$$\overline{\mathcal{G}} = \mathbb{R}(1 - \pi(1)) \oplus \mathcal{N}, \quad \text{where } \mathcal{N} := \{g \in \overline{\mathcal{G}} \mid E[g] = 0\}. \quad (2.30)$$

Indeed, this direct sum decomposition is obtained by noting that  $(\text{span}\{1 - \pi(1)\})^\perp \cap \overline{\mathcal{G}} = \mathcal{N}$ , because  $E[g] = E[g(1 - \pi(1))]$  for  $g \in \overline{\mathcal{G}}$ . So we can write  $\overline{\mathcal{G}} = g_{\text{mv}}^Y + \overline{\mathcal{G}} = g_{\text{mv}}^Y + \mathbb{R}(1 - \pi(1)) + \mathcal{N}$ , and hence all  $g \in \overline{\mathcal{G}}$  admit the decomposition

$$g = g_{\text{mv}}^Y + w(1 - \pi(1)) + n \quad \text{for some } w \in \mathbb{R} \text{ and } n \in \mathcal{N}. \quad (2.31)$$

Because  $\text{Cov}(R_{\text{mv}}^Y, g) = 0$  for all  $g \in \overline{\mathcal{G}}$  by (2.9) and  $\text{Cov}(1 - \pi(1), n) = E[(1 - \pi(1))n] = 0$  for all  $n \in \mathcal{N}$ , we obtain for  $R_{\text{mv}}^Y = g_{\text{mv}}^Y - Y$  that

$$\begin{aligned} E[g - Y] &= E[R_{\text{mv}}^Y] + wE[1 - \pi(1)], \\ \text{Var}[g - Y] &= \text{Var}[R_{\text{mv}}^Y] + w^2 \text{Var}[1 - \pi(1)] + \text{Var}[n]. \end{aligned}$$

But then it is an easy exercise to check that optimising over  $w \in \mathbb{R}$  and  $n \in \mathcal{N}$  directly yields the solutions to Problems (A)–(C) as given in Propositions 2.4, 2.6 and 2.7.

The above reasoning does not yet explain how the special element  $1 - \pi(1)$  comes up. For that, note that  $1 \notin \overline{\mathcal{G}}$  by Assumption I and  $1 \notin \mathcal{G}^\perp$  by Assumption II. So  $1 - \pi(1)$  is simply the projection of 1 on  $\overline{\mathcal{G}}$ , and (2.30) is the orthogonal decomposition of  $\overline{\mathcal{G}}$  into the span of this element and its orthogonal complement. A similar comment appears in Remark 2.5.

**3)** A decomposition like (2.31) also appears in the theorem on p.85 in Section 5.3 of Cochrane [7], where it is used to describe the mean-variance frontier. We go a little further here since we use (2.31) to relate the solutions to (A)–(C).  $\diamond$

### 3 Applications to financial economics

We now discuss financial implications of the results obtained in Section 2. In particular, we derive some properties of the solutions to Problems (A)–(D) which are abstract versions of classical results from mean-variance portfolio selection. Consider a fixed element  $Y \in L^2$ . In order to focus on the more interesting cases, we assume throughout this section that the parameters  $\mu \in \mathbb{R}$  and  $\sigma^2 \in (0, \infty)$  appearing in (B) and (C) are such that

$$\mu > E[R_{\text{mv}}^Y] \quad \text{and} \quad \sigma^2 > \text{Var}[R_{\text{mv}}^Y].$$

We first make a crucial observation. As can be seen from Propositions 2.2, 2.4, 2.6 and 2.7, the solutions to (A)–(D) all have the same fundamental structure

$$g_{\text{opt},i}^Y = g_{\text{mv}}^Y + c_{\text{opt},i}^Y(1 - \pi(1)) \quad \text{for } i \in \{A, B, C, D\} \quad (3.1)$$

for suitable constants  $c_{\text{opt},i}^Y \in \mathbb{R}$  and where  $g_{\text{opt},D}^Y := g^Y$ . This can be seen as an abstract generalisation of the classical *two-fund separation theorem*, in the sense that the solutions to (A)–(D) can all be decomposed into the sum of the “minimum variance payoff”  $g_{\text{mv}}^Y$  and an additional term proportional to  $1 - \pi(1)$ . The latter represents the best  $L^2$ -approximation in  $\overline{\mathcal{G}}$  of the constant payoff 1, and we see that only the *amount* invested there depends on the problem under consideration (and on the specific values of the parameters  $\alpha$ ,  $\mu$  and  $\sigma^2$ ). Alternatively,  $1 - \pi(1)$  can be characterised as the unique element of  $\overline{\mathcal{G}}$  in the Riesz representation of the continuous linear functional  $E[\cdot]$  on  $\overline{\mathcal{G}}$ ; in fact,  $E[g] = E[g(1 - \pi(1) + \pi(1))] = (g, 1 - \pi(1))$  for all  $g \in \overline{\mathcal{G}}$ . These observations are close to Section 6.5 in Cochrane [7], but go further in that we actually study and relate different optimisation problems.

Using the notation  $R_{\text{opt},i}^Y := g_{\text{opt},i}^Y - Y$  and omitting the dependence on  $\alpha$ ,  $\mu$  and  $\sigma^2$  gives

$$E[R_{\text{opt},i}^Y] = E[R_{\text{mv}}^Y] + c_{\text{opt},i}^Y E[1 - \pi(1)], \quad (3.2)$$

$$\text{Var}[R_{\text{opt},i}^Y] = \text{Var}[R_{\text{mv}}^Y] + (c_{\text{opt},i}^Y)^2 \text{Var}[1 - \pi(1)], \quad (3.3)$$

where (3.3) follows from the zero-covariance property of  $R_{\text{mv}}^Y$  in (2.9). Recall that Assumption II implies  $E[1 - \pi(1)] = \|1 - \pi(1)\|_{L^2}^2 > 0$ . Thus we can solve (3.2) for  $c_{\text{opt},i}^Y = \frac{E[R_{\text{opt},i}^Y] - E[R_{\text{mv}}^Y]}{E[1 - \pi(1)]}$  and insert this expression into (3.3) to get, for  $i \in \{A, B, C, D\}$ ,

$$\text{Var}[R_{\text{opt},i}^Y] = \text{Var}[R_{\text{mv}}^Y] + (E[R_{\text{opt},i}^Y] - E[R_{\text{mv}}^Y])^2 \frac{E[\pi(1)]}{E[1 - \pi(1)]}. \quad (3.4)$$

Similarly, using (2.19) to solve for  $c_{\text{opt},i}^Y$  in (3.3) and plugging that into (3.2), we obtain

$$E[R_{\text{opt},i}^Y] = E[R_{\text{mv}}^Y] + \sqrt{\text{Var}[R_{\text{opt},i}^Y] - \text{Var}[R_{\text{mv}}^Y]} \sqrt{\frac{E[1 - \pi(1)]}{E[\pi(1)]}}. \quad (3.5)$$

The equivalent equations (3.4) and (3.5) represent abstract versions of the classical *mean-variance frontier*, which provides a simple relationship between the mean (“return”) and the variance (“risk”) of any element  $R_{\text{opt},i}^Y$  which is an optimal outcome according to a mean-variance criterion. In particular, they show a linear relationship between the “excess return”, with respect to  $R_{\text{mv}}^Y$ , of a mean-variance optimal element  $R_{\text{opt},i}^Y$  and the square root of its “excess risk”. This is similar to Chapter 5 of Cochrane [7], but a bit more explicit.

The coefficients  $c_{\text{opt},i}^Y$  appearing in (3.1) also admit an interesting characterisation as “*beta factors*”. To see this, notice first that the zero-covariance property of  $R_{\text{mv}}^Y$  in (2.9) yields

$$\text{Cov}(R_{\text{opt},i}^Y, 1 - \pi(1)) = \text{Cov}(R_{\text{mv}}^Y + c_{\text{opt},i}^Y(1 - \pi(1)), 1 - \pi(1)) = c_{\text{opt},i}^Y \text{Var}[1 - \pi(1)].$$

Because  $\text{Var}[1 - \pi(1)] > 0$  due to Assumption II, we thus obtain

$$c_{\text{opt},i}^Y = \frac{\text{Cov}(R_{\text{opt},i}^Y, 1 - \pi(1))}{\text{Var}[1 - \pi(1)]}.$$

We have therefore proved for  $i \in \{A, B, C, D\}$  the relation

$$E[g_{\text{opt},i}^Y] - E[g_{\text{mv}}^Y] = E[R_{\text{opt},i}^Y] - E[R_{\text{mv}}^Y] = \frac{\text{Cov}(R_{\text{opt},i}^Y, 1 - \pi(1))}{\text{Var}[1 - \pi(1)]} E[1 - \pi(1)]. \quad (3.6)$$

Together with (3.1), this can be regarded as a *beta representation* as in Chapter 6 of Cochrane [7], with  $1 - \pi(1)$  playing the role of a “market portfolio” or reference asset. The relation (3.6) says that the excess expected value (with respect to  $g_{\text{mv}}^Y$ ) of the solution to any of the Problems (A)–(D) is proportional to the expected value of the “market portfolio”  $1 - \pi(1)$ , with a proportionality factor having the typical structure “ $\beta = \text{Cov} / \text{Var}$ ”.

**Remark 3.1.** Because  $\pi(1)$  is in  $\mathcal{G}^\perp$ , we have for every  $g \in \overline{\mathcal{G}} = (\mathcal{G}^\perp)^\perp$  that

$$E[g] = E[g(1 - \pi(1))] = \text{Cov}(g, 1 - \pi(1)) + E[g]E[1 - \pi(1)].$$

Using (2.19) and solving for  $E[g]$  thus yields

$$E[g] = \frac{\text{Cov}(g, 1 - \pi(1))}{\text{Var}[1 - \pi(1)]} E[1 - \pi(1)] \quad \text{for all } g \in \overline{\mathcal{G}}. \quad (3.7)$$

A similar relation can be found in Proposition 1.29 of Fontana [11], which in turn generalises a result due to Courtault et al. [8]. As above,  $1 - \pi(1)$  plays the role of an abstract market portfolio. However, while (3.7) holds for any  $g \in \overline{\mathcal{G}}$ , the analogous relation (3.6) is more subtle since it only holds for the *optimal* elements  $g_{\text{opt},i}^Y$ . This is due to the presence of the extra position (random endowment)  $Y$ .  $\diamond$

The zero-covariance property of the minimum variance element  $R_{\text{mv}}^Y$  also implies another interesting relation. For any  $R^Y := g - Y$  with  $g \in \overline{\mathcal{G}}$  and all  $i \in \{A, B, C, D\}$ , we have

$$\begin{aligned} \text{Cov}(R_{\text{opt},i}^Y, R^Y) &= \text{Cov}\left(R_{\text{mv}}^Y + c_{\text{opt},i}^Y(1 - \pi(1)), R_{\text{mv}}^Y - g_{\text{mv}}^Y + g\right) \\ &= \text{Var}[R_{\text{mv}}^Y] + c_{\text{opt},i}^Y \text{Cov}(1 - \pi(1), g - g_{\text{mv}}^Y) \\ &= \text{Var}[R_{\text{mv}}^Y] + c_{\text{opt},i}^Y E[\pi(1)] E[g - g_{\text{mv}}^Y] \\ &= \text{Var}[R_{\text{mv}}^Y] + c_{\text{opt},i}^Y E[\pi(1)] (E[R^Y] - E[R_{\text{mv}}^Y]). \end{aligned} \quad (3.8)$$

For  $i = B$  (with constraint  $\mu$ ), where  $c_{\text{opt},B}^Y = \frac{\mu - E[R_{\text{mv}}^Y]}{E[1 - \pi(1)]}$  due to (2.22) and (2.23), this gives

$$\text{Cov}(R_{\text{opt},B}^Y, R^Y) = \text{Var}[R_{\text{mv}}^Y] + (\mu - E[R_{\text{mv}}^Y]) (E[R^Y] - E[R_{\text{mv}}^Y]) \frac{E[\pi(1)]}{E[1 - \pi(1)]}. \quad (3.9)$$

Now take any  $\hat{g} \in \overline{\mathcal{G}}$  such that  $\hat{R}^Y := \hat{g} - Y$  and  $R_{\text{opt},B}^Y$  are uncorrelated. Then (3.9) yields

$$E[\hat{R}^Y] = E[R_{\text{mv}}^Y] - \frac{\text{Var}[R_{\text{mv}}^Y]}{\mu - E[R_{\text{mv}}^Y]} \frac{E[1 - \pi(1)]}{E[\pi(1)]}. \quad (3.10)$$

Solving (3.8) for  $E[R^Y]$ , plugging in (3.10), using (3.4) for  $i = B$  and again (3.10) gives

$$\begin{aligned} E[R^Y] &= E[R_{\text{mv}}^Y] + \frac{\text{Cov}(R_{\text{opt},B}^Y, R^Y) - \text{Var}[R_{\text{mv}}^Y] E[1 - \pi(1)]}{\mu - E[R_{\text{mv}}^Y]} \frac{E[\pi(1)]}{E[\pi(1)]} \\ &= E[\hat{R}^Y] + \frac{\text{Cov}(R_{\text{opt},B}^Y, R^Y) E[1 - \pi(1)]}{\mu - E[R_{\text{mv}}^Y]} \frac{E[\pi(1)]}{E[\pi(1)]} \\ &= E[\hat{R}^Y] + \frac{\text{Cov}(R_{\text{opt},B}^Y, R^Y) \text{Var}[R_{\text{mv}}^Y] + (\mu - E[R_{\text{mv}}^Y])^2 \frac{E[\pi(1)]}{E[1 - \pi(1)]} E[1 - \pi(1)]}{\text{Var}[R_{\text{opt},B}^Y] (\mu - E[R_{\text{mv}}^Y]) E[\pi(1)]} \\ &= E[\hat{R}^Y] + \frac{\text{Cov}(R_{\text{opt},B}^Y, R^Y)}{\text{Var}[R_{\text{opt},B}^Y]} (\mu - E[\hat{R}^Y]) \end{aligned} \quad (3.11)$$

for any  $g \in \overline{\mathcal{G}}$ . For any  $\mu \in \mathbb{R}$ , the expected value of an arbitrary  $R^Y := g - Y$  can thus be written as a generalised convex combination of  $\mu = E[R_{\text{opt},B}^Y] = E[g_{\text{opt},B}^Y - Y]$  and  $E[\hat{R}^Y]$ , where  $g_{\text{opt},B}^Y$  is the solution to (B) with constraint  $\mu$ , and  $\hat{R}^Y = \hat{g} - Y$  is an element having zero correlation with  $R_{\text{opt},B}^Y$ . We emphasise that this holds for *any*  $R^Y := g - Y$  with  $g \in \overline{\mathcal{G}}$ . In particular,  $R^Y$  need not be optimal according to any of our mean-variance criteria.

To obtain a CAPM-type formula, we can argue as follows. Consider an economy of  $I$  individuals who all have a random endowment  $Y$  (the same across all individuals) and who choose their portfolios according to the optimality criterion in Problem (B), with parameters  $\mu_i > E[R_{\text{mv}}^Y]$  for  $i = 1, \dots, I$ . By Proposition 2.6, the optimal choice for the  $i$ -th individual is

$$g_{\text{opt},B}^{i,Y} = g_{\text{mv}}^Y + \frac{\mu_i - E[R_{\text{mv}}^Y]}{E[1 - \pi(1)]}(1 - \pi(1)).$$

Let  $\bar{\mu} := \frac{1}{I} \sum_{i=1}^I \mu_i$  and let  $\bar{g}_{\text{opt},B}^Y := \frac{1}{I} \sum_{i=1}^I g_{\text{opt},B}^{i,Y}$  be the ‘‘average position’’ of the economy. Since  $\mu_i > E[R_{\text{mv}}^Y]$  for all  $i$ , we also have  $\bar{\mu} > E[R_{\text{mv}}^Y]$ . Observe also that

$$\bar{g}_{\text{opt},B}^Y = \frac{1}{I} \sum_{i=1}^I \left( g_{\text{mv}}^Y + \frac{\mu_i - E[R_{\text{mv}}^Y]}{E[1 - \pi(1)]}(1 - \pi(1)) \right) = g_{\text{mv}}^Y + \frac{\bar{\mu} - E[R_{\text{mv}}^Y]}{E[1 - \pi(1)]}(1 - \pi(1)).$$

Due to Proposition 2.6, we can conclude that  $\bar{g}_{\text{opt},B}^Y$  solves Problem (B) with constraint  $\bar{\mu}$ . This means that if all agents in the economy choose their positions according to (B), the ‘‘average position’’ of the economy is mean-variance efficient in the sense that it solves (B) with ‘‘average constraint’’  $\bar{\mu}$ . Using now (3.11), we obtain for *any*  $g \in \overline{\mathcal{G}}$  that

$$E[g - Y] = E[R^Y] = E[\hat{R}^Y] + \frac{\text{Cov}(\bar{g}_{\text{opt},B}^Y - Y, g - Y)}{\text{Var}[\bar{g}_{\text{opt},B}^Y - Y]}(\bar{\mu} - E[\hat{R}^Y]).$$

This can be regarded as an abstract version of the *zero-beta CAPM formula*; see Section 5.1 of Barucci [1]. Note that here, we interpret as a *market portfolio* the average position  $\bar{g}_{\text{opt},B}^Y$  of the economy.

We next consider a related mean-variance portfolio optimisation problem, namely

$$\mathbf{Problem (SR)} \quad \frac{E[g - Y]}{\sqrt{\text{Var}[g - Y]}} = \max! \quad \text{over all } g \in \overline{\mathcal{G}} \text{ such that } \text{Var}[g - Y] > 0.$$

The quantity to be maximised in (SR) is an abstract counterpart of the classical *Sharpe ratio*, a typical measure for the trade-off between risk and return. The solution to (SR) is as follows.

**Proposition 3.2.** *Let  $Y \in L^2$  and suppose that  $E[R_{\text{mv}}^Y] > 0$  and  $\text{Var}[R_{\text{mv}}^Y] > 0$ . Then Problem (SR) admits a unique solution  $g_{\text{sr}}^Y \in \overline{\mathcal{G}}$ , explicitly given by*

$$g_{\text{sr}}^Y = g_{\text{mv}}^Y + \frac{\text{Var}[R_{\text{mv}}^Y]}{E[R_{\text{mv}}^Y]} \frac{1}{E[\pi(1)]}(1 - \pi(1)). \quad (3.12)$$



*Proof.* Since  $\text{Var}[R_{\text{mv}}^Y] > 0$ , the same reasoning via (2.8) and (2.10) as in the proofs of Propositions 2.4, 2.6 and 2.7 allows us to reduce (SR) to the basic version

$$\frac{E[R_{\text{mv}}^Y] + E[g]}{\sqrt{\text{Var}[R_{\text{mv}}^Y] + \text{Var}[g]}} = \max! \quad \text{over all } g \in \bar{\mathcal{G}}. \quad (3.13)$$

If we denote by  $g^* \in \bar{\mathcal{G}}$  the solution to (3.13), then the solution to the original problem is  $g_{\text{sr}}^Y = g_{\text{mv}}^Y + g^*$ , where  $g_{\text{mv}}^Y$  is the solution to (MV). To solve (3.13), we proceed in two steps. We first fix  $\mu \in (0, \infty)$  and want to minimise  $\text{Var}[g]$  over all  $g \in \bar{\mathcal{G}}$  satisfying the extra constraint  $E[g] = \mu$ . Due to Proposition 2.6, this problem is uniquely solved by

$$g_{\text{opt},B}^0(\mu) = \frac{\mu}{E[1 - \pi(1)]}(1 - \pi(1)),$$

and so we get with the help of (2.19) that

$$\max_{g \in \bar{\mathcal{G}}: E[g]=\mu} \frac{E[R_{\text{mv}}^Y] + E[g]}{\sqrt{\text{Var}[R_{\text{mv}}^Y] + \text{Var}[g]}} = \frac{E[R_{\text{mv}}^Y] + \mu}{\sqrt{\text{Var}[R_{\text{mv}}^Y] + \text{Var}[g_{\text{opt},B}^0(\mu)]}} = \frac{E[R_{\text{mv}}^Y] + \mu}{\sqrt{\text{Var}[R_{\text{mv}}^Y] + \mu^2 \frac{E[\pi(1)]}{E[1-\pi(1)]}}.$$

Since  $E[R_{\text{mv}}^Y] > 0$ , it can be readily checked that the last expression is maximised over  $\mu$  by

$$\mu_* = \frac{\text{Var}[R_{\text{mv}}^Y] E[1 - \pi(1)]}{E[R_{\text{mv}}^Y] E[\pi(1)]} =: c_{\text{sr}}^Y E[1 - \pi(1)], \quad (3.14)$$

and so (3.13) is uniquely solved by  $g^* = g_{\text{opt},B}^0(\mu_*)$ . Problem (SR) is therefore uniquely solved by  $g_{\text{sr}}^Y = g_{\text{mv}}^Y + g_{\text{opt},B}^0(\mu_*)$ .  $\square$

**Remark 3.3.** It can be checked that if  $E[R_{\text{mv}}^Y] < 0$ , the element  $g_{\text{sr}}^Y \in \bar{\mathcal{G}}$  given in Proposition 3.2 can be characterised as the unique *minimiser* of the ratio  $E[g - Y]/\sqrt{\text{Var}[g - Y]}$ .  $\diamond$

Combining (3.1) (or Propositions 2.4, 2.6 and 2.7) with Proposition 3.2 yields an alternative formulation of a *two-fund separation* result. In fact, writing (3.1) and (3.12) via (3.14) as

$$\begin{aligned} g_{\text{opt},i}^Y &= g_{\text{mv}}^Y + c_{\text{opt},i}^Y(1 - \pi(1)), \\ g_{\text{sr}}^Y &= g_{\text{mv}}^Y + c_{\text{sr}}^Y(1 - \pi(1)) \end{aligned}$$

allows us to solve for  $1 - \pi(1)$  and obtain, for  $i \in \{A, B, C, D\}$ ,

$$g_{\text{opt},i}^Y = g_{\text{mv}}^Y + \frac{c_{\text{opt},i}^Y}{c_{\text{sr}}^Y}(g_{\text{sr}}^Y - g_{\text{mv}}^Y) = \frac{c_{\text{opt},i}^Y}{c_{\text{sr}}^Y}g_{\text{sr}}^Y + \left(1 - \frac{c_{\text{opt},i}^Y}{c_{\text{sr}}^Y}\right)g_{\text{mv}}^Y.$$

So the solutions to Problems (A)–(D) can all be written as generalised convex combinations of  $g_{\text{mv}}^Y$  and  $g_{\text{sr}}^Y$ , the solutions of *minimising the variance* and of *maximising the Sharpe ratio* for  $g - Y$ , respectively. This extends the well-known fact that the mean-variance frontier can be spanned by any two elements that lie on it; see Section 5.4 of Cochrane [7].

In preparation for the next section, we now specialise the abstract results from Section 2 to a more concrete financial situation. We replace the abstract random variable  $Y \in L^2$  by

$$Y = -cB + (H - hB) - H_0 \quad \text{with } c, h \in \mathbb{R} \text{ and } H, H_0 \in L^2. \quad (3.15)$$

This describes the net financial balance (outflows minus incomes) at the final time  $T$  faced by an agent who, at the starting time  $t = 0$ , is endowed with initial capital  $c$  and sells for a compensation of  $h$  the contingent claim  $H$ , to be paid at  $T$ . In addition, the agent has a position  $H_0$  (evaluated at  $T$ ), which can be interpreted as an existing book of options or as a random endowment. We can then study what happens if the agent trades in the market by choosing an optimal  $g \in \overline{\mathcal{G}}$  according to one of the mean-variance rules formalised as Problems (A)–(D). Of course, this includes “pure investment” problems without trading a contingent claim by simply letting  $H \equiv 0$  and  $h = 0$ .

For later use in solving mean-variance indifference valuation problems, we now give explicit formulas for the optimal values of Problems (A)–(D) for the specific  $Y$  given in (3.15). Recall that  $R_x^Y = g_x^Y - Y$  and note that (3.15) yields  $N^Y = N^H - N^{H_0} - (c + h)\pi(B)$ . First, (2.11) and (2.12) in Remark 2.3 give for the minimum variance result  $R_{\text{mv}}^Y$  the mean and variance as

$$\begin{aligned} \mu_{\text{mv}}(c, H, h, H_0) &:= E[R_{\text{mv}}^{-cB+(H-hB)-H_0}] \\ &= (c + h) \frac{E[\pi(B)]}{E[\pi(1)]} - \frac{E[N^H] - E[N^{H_0}]}{E[\pi(1)]}, \end{aligned} \quad (3.16)$$

$$\begin{aligned} \sigma_{\text{mv}}^2(c, H, h, H_0) &:= \text{Var}[R_{\text{mv}}^{-cB+(H-hB)-H_0}] \\ &= E[((c + h)\pi(B) - N^H + N^{H_0})^2] \\ &\quad - \frac{((c + h)E[\pi(B)] - E[N^H] + E[N^{H_0}])^2}{E[\pi(1)]}. \end{aligned} \quad (3.17)$$

Next, the optimal value for (A) with risk-aversion  $\alpha \in (0, \infty)$  is given by

$$\begin{aligned} v_*(c, H, h, H_0; \alpha) &:= E[R_{\text{opt},A}^{-cB+(H-hB)-H_0}(1/\alpha)] - \alpha \text{Var}[R_{\text{opt},A}^{-cB+(H-hB)-H_0}(1/\alpha)] \\ &= \mu_{\text{mv}}(c, H, h, H_0) - \alpha \sigma_{\text{mv}}^2(c, H, h, H_0) + \frac{1}{4\alpha} \frac{E[1 - \pi(1)]}{E[\pi(1)]} \end{aligned} \quad (3.18)$$

from (2.20) and (2.21). The Markowitz problem (B) of minimising the variance given a constraint  $\mu \in \mathbb{R}$  on the mean leads via (3.3), (2.19), (2.22) and (2.23) to the optimal variance

$$\begin{aligned} \sigma_*^2(c, H, h, H_0; \mu) &:= \text{Var}[R_{\text{opt},B}^{-cB+(H-hB)-H_0}(\mu)] \\ &= \sigma_{\text{mv}}^2(c, H, h, H_0) + \left( (\mu - \mu_{\text{mv}}(c, H, h, H_0))^+ \right)^2 \frac{E[\pi(1)]}{E[1 - \pi(1)]}. \end{aligned} \quad (3.19)$$

Finally, the optimal mean in Problem (C), given a constraint  $\sigma^2 \in (0, \infty)$  on the variance with  $\sigma^2 \geq \sigma_{\text{mv}}^2(c, H, h, H_0)$ , is due to (3.2), (2.26), (2.27) and (2.19) given by

$$\begin{aligned} \mu_*(c, H, h, H_0; \sigma^2) &:= E[R_{\text{opt},C}^{-cB+(H-hB)-H_0}(\sigma^2)] \\ &= \mu_{\text{mv}}(c, H, h, H_0) + \sqrt{\sigma^2 - \sigma_{\text{mv}}^2(c, H, h, H_0)} \sqrt{\frac{E[1 - \pi(1)]}{E[\pi(1)]}}. \end{aligned} \quad (3.20)$$

**Remark 3.4.** The mean-variance hedging problem for an initial capital  $c \in \mathbb{R}$  and a contingent claim  $H \in L^2$  is usually written as

$$\|H - cB - g\|_{L^2} = \min! \quad \text{over all } g \in \overline{\mathcal{G}}$$

(see e.g. Schweizer [33, 36]). In our notation, this is (D) for  $Y := H - cB$ . The corresponding minimal value is by (2.3) given by

$$\min_{g \in \overline{\mathcal{G}}} \|H - cB - g\|_{L^2} = \|N^H - c\pi(B)\|_{L^2}. \quad (3.21)$$

Instead of fixing  $c$ , we could optimise with respect to the initial capital as well and consider

$$\|H - cB - g\|_{L^2} = \min! \quad \text{over all } (c, g) \in \mathbb{R} \times \overline{\mathcal{G}}.$$

If  $B \notin \overline{\mathcal{G}}$  so that  $E[B\pi(B)] \neq 0$ , the optimal initial capital  $c_*(H) \in \mathbb{R}$  is by (3.21) given by

$$c_*(H) := \operatorname{argmin}_{c \in \mathbb{R}} \{\|N^H - c\pi(B)\|_{L^2}\} = \frac{E[N^H\pi(B)]}{E[B\pi(B)]} = \frac{E[H\pi(B)]}{E[B\pi(B)]} = \left( \frac{d\tilde{P}}{dP}, \frac{H}{B} \right), \quad (3.22)$$

where  $\tilde{P}$  denotes the so-called *variance-optimal signed  $(\mathcal{G}, B)$ -martingale measure*; see Lemma 6 in Schweizer [36] where  $\tilde{P}$  is given explicitly for the (undiscounted) case of  $B \neq 1$ . The value  $c_*(H)$  is also called the  *$L^2$ -approximation value* of the payoff  $H$ .  $\diamond$

## 4 Mean-variance indifference valuations

In the last section, we have introduced a financial position of the form  $Y = -cB + (H - hB) - H_0$ , where  $H \in L^2$  represents a contingent claim sold by our agent for a compensation amount  $h$ . However,  $h$  has been considered as exogenously given. In the present section, we study how a value for  $h$  can be determined endogenously. As an application of the mean-variance theory developed so far, we analyse several *mean-variance indifference valuation rules*, i.e. we determine the amount  $h$  at which an agent is indifferent, in terms of optimal value according to a mean-variance criterion, between the two following alternatives:

1. Sell the contingent claim  $H$ , receive the amount  $h$  and optimise over  $g \in \overline{\mathcal{G}}$ , for the chosen criterion, the value of the final net position  $g - Y = (c + h)B + g - H + H_0$ .
2. Ignore the contingent claim  $H$  and just optimise the value of  $cB + g + H_0$  over  $g \in \overline{\mathcal{G}}$ .

**Remark 4.1.** 1) We emphasise that the outcome of the approach in this section is a subjective *value* for  $H$  (to a potential seller). This should not be confused with a price since we are not assuming that  $H$  is available for trade in a liquid market. At most, one could view the value as an OTC quote.

2) It is well known that preferences based on quadratic criteria exhibit a sort of risk-loving behaviour due to their non-monotonicity. As a consequence, we might obtain for some positive

payoffs  $H$  a negative value for  $h$  — our agent might be willing to pay (!) for entering into a risky position. This could be avoided by adding appropriate constraints in our valuation approach, as discussed below in part 4) of Remark 4.12; but unfortunately, that would destroy all the analytical tractability we otherwise get from a quadratic criterion. We also explain below (after Corollary 4.11) how aspects of diversification might lead to negative valuation outcomes.  $\diamond$

To derive explicit results for the above approach, we need some preliminaries. We recall from Section 1 the set  $\mathcal{A} := \mathbb{R}B + \mathcal{G}$  and denote its  $L^2$ -closure by  $\overline{\mathcal{A}}$ . Intuitively,  $\mathcal{A}$  contains all undiscounted final wealths generated by a trading strategy for some  $g \in \mathcal{G}$  starting from some initial capital  $c \in \mathbb{R}$ . So  $\overline{\mathcal{A}}$  consists of those undiscounted payoffs which can be approximately attained in the financial market  $(B, \mathcal{G})$ , in the sense that they are  $L^2$ -limits of a sequence of attainable final wealths. Then we introduce

**Assumption III.** There exist  $\bar{g} \in \overline{\mathcal{G}}$  and a constant  $\delta \neq 0$  such that  $\delta B + \bar{g} = 1$   $P$ -a.s.

With the above interpretation of  $\overline{\mathcal{A}}$ , Assumption III is equivalent to saying that a riskless *zero-coupon bond* can be approximately attained in the abstract financial market  $(B, \mathcal{G})$  (from an initial investment of  $\delta$ ).

**Remark 4.2. 1)** An easy extension (taking into account both the cases  $B \notin \overline{\mathcal{G}}$  and  $B \in \overline{\mathcal{G}}$ ) of the arguments used in Lemma 2 of Schweizer [36] allows to show that  $\overline{\mathcal{A}} = \mathbb{R}B + \overline{\mathcal{G}}$ . Hence Assumption III can be equivalently formulated as  $1 \in \overline{\mathcal{A}}$  (or, equivalently,  $\mathbb{R} \cap \overline{\mathcal{A}} \neq \{0\}$ ).

**2)** Due to the linearity of  $\overline{\mathcal{G}}$ , it is easy to check that Assumption III is equivalent to the condition  $\mathbb{R}B + \overline{\mathcal{G}} = \mathbb{R} + \overline{\mathcal{G}}$ .  $\diamond$

**In this section, we always suppose that Assumption III is satisfied, with  $\delta > 0$**  (in fact, the case  $\delta < 0$  can be seen as a pathological *arbitrage* situation). This is motivated on the one hand by the fact that it makes the theory particularly simple and elegant, as we shall see below. On the other hand, it is also reasonable to expect that such an assumption will be satisfied in many financial markets. One could still solve mean-variance indifference valuation problems without Assumption III, but this would lead to more involved formulae without a clear economic interpretation. Hence we omit the details.

It is interesting to note that Assumption III is related to the notion of *no approximate profits in  $L^2$* , formally defined as the condition  $B \notin \overline{\mathcal{G}}$ ; see Schweizer [35, 36].

**Lemma 4.3.** *If Assumption III holds, the conditions of “no approximate riskless profits in  $L^2$ ” and “no approximate profits in  $L^2$ ” are equivalent, i.e. we have  $1 \notin \overline{\mathcal{G}}$  if and only if  $B \notin \overline{\mathcal{G}}$ .*

*Proof.* This follows directly from the linearity of  $\overline{\mathcal{G}}$ , since  $1 = \delta B + \bar{g}$  can be rewritten as  $B = (1 - \bar{g})/\delta$ .  $\square$

Lemma 4.3 implies that as soon as Assumption III is satisfied, we can equivalently work with any of the two no-arbitrage-type conditions  $1 \notin \overline{\mathcal{G}}$  and  $B \notin \overline{\mathcal{G}}$ . Moreover, the condition  $B \notin \overline{\mathcal{G}}$  can be shown to be equivalent to an abstract version of the classical *law of one price*; see Courtault et al. [8] and Fontana [11], Section 1.4.1. Finally, we can use Assumption III to obtain a more detailed version of the orthogonal decomposition (2.1), as follows.

**Lemma 4.4.** *Under Assumption III, the terms  $g^Y \in \overline{\mathcal{G}}$  and  $N^Y \in \mathcal{G}^\perp$  in the decomposition (2.1) of  $Y \in L^2$  can be uniquely represented as*

$$g^Y = \tilde{g}^Y + c^Y(B - \pi(B)) \quad \text{and} \quad N^Y = c^Y \pi(B) + L^Y, \quad (4.1)$$

where  $c^Y = \frac{E[Y\pi(B)]}{E[B\pi(B)]}$ , the element  $\tilde{g}^Y \in \overline{\mathcal{G}}$  is the orthogonal projection in  $L^2$  of  $Y - c^Y B$  on  $\overline{\mathcal{G}}$ , and  $L^Y \in \overline{\mathcal{A}}^\perp$  is given by  $L^Y = Y - c^Y B - \tilde{g}^Y$ . Furthermore, we have  $E[L^Y] = 0$ .

*Proof.* Because  $L^2 = \overline{\mathcal{A}} \oplus \overline{\mathcal{A}}^\perp$ , any  $Y \in L^2$  can be uniquely decomposed as

$$Y = a^Y + L^Y, \quad \text{where } a^Y \in \overline{\mathcal{A}} \text{ and } L^Y \in \overline{\mathcal{A}}^\perp.$$

Moreover,  $a^Y \in \overline{\mathcal{A}} = \mathbb{R}B + \overline{\mathcal{G}}$  gives  $a^Y = c^Y B + \tilde{g}^Y$  with  $c^Y \in \mathbb{R}$  and  $\tilde{g}^Y \in \overline{\mathcal{G}}$  and therefore

$$Y = c^Y(B - \pi(B)) + \tilde{g}^Y + L^Y + c^Y \pi(B). \quad (4.2)$$

Note that  $c^Y(B - \pi(B)) + \tilde{g}^Y \in \overline{\mathcal{G}}$  and  $L^Y \in \mathcal{G}^\perp$ , since  $L^Y \in \overline{\mathcal{A}}^\perp$  and  $\mathcal{G} \subseteq \overline{\mathcal{A}}$ . The assertion (4.1) thus follows from the uniqueness of the decomposition (2.1), and we have  $E[L^Y] = (L^Y, 1) = 0$  since  $L^Y \in \overline{\mathcal{A}}^\perp$  and  $1 \in \overline{\mathcal{A}}$ . Finally, because  $B = (1 - \bar{g})/\delta$  is in  $\overline{\mathcal{A}}$ ,  $L^Y \in \overline{\mathcal{A}}^\perp$  implies that  $(L^Y, B) = 0$ . Since also  $L^Y \in \mathcal{G}^\perp$ , we get  $E[L^Y \pi(B)] = (L^Y, B) - (L^Y, B - \pi(B)) = 0$  and therefore  $E[Y\pi(B)] = c^Y E[B\pi(B)]$  due to (4.2). Because  $B \notin \overline{\mathcal{G}}$  by Lemma 4.3, we have  $E[B\pi(B)] > 0$ , and solving for  $c^Y$  thus completes the proof.  $\square$

**Remark 4.5.** If we think of  $Y := H$  as a contingent claim, the term  $c^H$  in Lemma 4.4 represents in financial terms the ‘‘replication price’’ of the attainable part  $a^H \in \overline{\mathcal{A}}$  of  $H$ . (The term ‘‘price’’, for  $c^H$ , is justified here since  $a^H$  can be replicated at that cost.) Moreover,  $c^H$  also coincides with the quantity  $c_*(H)$  in (3.22) because

$$c^H = \frac{E[H\pi(B)]}{E[B\pi(B)]} = \frac{E[N^H \pi(B)]}{E[B\pi(B)]} = c_*(H).$$

Thus the constant  $c^H$  can also be interpreted as the  $L^2$ -approximation value of  $H$ .  $\diamond$

Using Lemma 4.4, we can obtain more explicit expressions for the optimal values of our mean-variance problems. Since Assumption III gives  $\pi(B) = \pi(1)/\delta$  and we have by (2.19) that  $E[1 - \pi(1)] = \text{Var}[\pi(1)]/E[\pi(1)]$ , we can rewrite (3.16)–(3.20) by simple computations as

$$\mu_{\text{mv}}(c, H, h, H_0) = \frac{c + h - c^H + c^{H_0}}{\delta}, \quad (4.3)$$

$$\sigma_{\text{mv}}^2(c, H, h, H_0) = \text{Var}[L^H - L^{H_0}] = \text{Var}[L^H] + \text{Var}[L^{H_0}] - 2 \text{Cov}(L^H, L^{H_0}), \quad (4.4)$$

$$v_*(c, H, h, H_0; \alpha) = \frac{c + h - c^H + c^{H_0}}{\delta} - \alpha \text{Var}[L^H - L^{H_0}] + \frac{1}{4\alpha} \frac{\text{Var}[\pi(B)]}{(E[\pi(B)])^2}, \quad (4.5)$$

$$\sigma_*^2(c, H, h, H_0; \mu) = \left( \left( \mu - \frac{c + h - c^H + c^{H_0}}{\delta} \right)^+ \right)^2 \frac{(E[\pi(B)])^2}{\text{Var}[\pi(B)]} + \text{Var}[L^H - L^{H_0}], \quad (4.6)$$

$$\mu_*(c, H, h, H_0; \sigma^2) = \frac{c + h - c^H + c^{H_0}}{\delta} + \sqrt{\sigma^2 - \text{Var}[L^H - L^{H_0}]} \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]}. \quad (4.7)$$

For the “pure investment case”  $H \equiv 0$  and  $h = 0$ , this simplifies to

$$\mu_{\text{mv}}(c, 0, 0, H_0) = \frac{c + c^{H_0}}{\delta}, \quad (4.8)$$

$$\sigma_{\text{mv}}^2(c, 0, 0, H_0) = \text{Var}[L^{H_0}], \quad (4.9)$$

$$v_*(c, 0, 0, H_0; \alpha) = \frac{c + c^{H_0}}{\delta} - \alpha \text{Var}[L^{H_0}] + \frac{1}{4\alpha} \frac{\text{Var}[\pi(B)]}{(E[\pi(B)])^2}, \quad (4.10)$$

$$\sigma_*^2(c, 0, 0, H_0; \mu) = \left( \left( \mu - \frac{c + c^{H_0}}{\delta} \right)^+ \right)^2 \frac{(E[\pi(B)])^2}{\text{Var}[\pi(B)]} + \text{Var}[L^{H_0}], \quad (4.11)$$

$$\mu_*(c, 0, 0, H_0; \sigma^2) = \frac{c + c^{H_0}}{\delta} + \sqrt{\sigma^2 - \text{Var}[L^{H_0}]} \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]}. \quad (4.12)$$

We now formally introduce the mean-variance valuation rules we analyse in this section.

**Definition 4.6.** Let  $c \in \mathbb{R}$  and  $H, H_0 \in L^2$ . For a given risk-aversion coefficient  $\alpha \in (0, \infty)$ , the (A)-indifference value of  $H$  is defined by

$$h_A(H; c, H_0, \alpha) := \inf\{h \in \mathbb{R} \mid v_*(c, H, h, H_0; \alpha) \geq v_*(c, 0, 0, H_0; \alpha)\}. \quad (4.13)$$

For  $\mu \in \mathbb{R}$ , the (B)-indifference value of  $H$  is defined by

$$h_B(H; c, H_0, \mu) := \inf\{h \in \mathbb{R} \mid \sigma_*^2(c, H, h, H_0; \mu) \leq \sigma_*^2(c, 0, 0, H_0; \mu)\}. \quad (4.14)$$

For  $\sigma^2 \in (0, \infty)$ , the (C)-indifference value of  $H$  is defined by

$$h_C(H; c, H_0, \sigma^2) := \inf\{h \in \mathbb{R} \mid \mu_*(c, H, h, H_0; \sigma^2) \geq \mu_*(c, 0, 0, H_0; \sigma^2)\}. \quad (4.15)$$

We use here the notation introduced in (3.18)–(3.20) and the convention  $\inf \emptyset = +\infty$ .

**Remark 4.7. 1)** We repeat that the mean-variance indifference *values* introduced above should not be regarded as market *prices* for  $H$ ; they are outcomes of subjective valuation mechanisms.

**2)** As can be seen from (4.5), the function  $h \mapsto v_*(c, H, h, H_0; \alpha)$  is continuous (even affine) and strictly increasing, since  $\delta > 0$ . Consequently,  $h_A(H; c, H_0, \alpha)$  satisfies the relation

$$v_*(c, H, h_A(H; c, H_0, \alpha), H_0; \alpha) = v_*(c, 0, 0, H_0; \alpha). \quad (4.16)$$

This means that the (A)-indifference value could also be defined by the implicit requirement that it makes the agent indifferent, in terms of maximal values for Problem (A), between the two alternatives of selling or not selling  $H$ , as explained at the beginning of this section. An analogous result holds true for the (B)- and (C)-indifference values, at least in the more interesting cases where the functions  $h \mapsto \sigma_*^2(c, H, h, H_0; \mu)$  and  $h \mapsto \mu_*(c, H, h, H_0; \sigma^2)$  are continuous and strictly monotonic. See the proofs of Propositions 4.9 and 4.10 for more details.

**3)** We have defined all our indifference values from the point of view of a *seller* of the contingent claim  $H$ . One can also consider the *buyer* versions by simply replacing  $H$  and  $h$

with  $-H$  and  $-h$ , respectively, and “inf” with “sup” in the definitions. In the case of the (A)-indifference value, we have for instance

$$h_A^{\text{buyer}}(H; c, H_0, \alpha) := \sup\{h \in \mathbb{R} \mid v_*(c, -H, -h, H_0; \alpha) \geq v_*(c, 0, 0, H_0; \alpha)\}.$$

It is easy to check that one has between the seller and buyer versions the intuitive relation

$$h_i^{\text{seller}}(H) := h_i(H) = -h_i^{\text{buyer}}(-H) \quad \text{for } i \in \{A, B, C\}.$$

4) Let us briefly consider the case where  $1 \notin \overline{\mathcal{G}}$ , but  $B \in \overline{\mathcal{G}}$ . In particular, due to Lemma 4.3, Assumption III cannot hold. Since  $B \in \overline{\mathcal{G}}$  implies that  $\pi(B) \equiv 0$ , (3.16)–(3.20) show that  $\mu_{\text{mv}}$  and  $\sigma_{\text{mv}}^2$  and hence also the optimal values of Problems (A)–(C) do not depend on  $h$ . In this case, the indifference valuation problems formulated above are not well-posed and we always have  $h_i(H) \in \{-\infty, +\infty\}$  for any  $H \in L^2$  and  $i \in \{A, B, C\}$ .  $\diamond$

We are now ready to solve the mean-variance indifference valuation problems explicitly. To focus on the financially meaningful cases, we always impose Assumption II that  $1 \notin \mathcal{G}^\perp$ . With all the work done so far, the proofs of the next three results are very simple; we just use the explicit expressions for the optimal values of Problems (A)–(C) given in (4.3)–(4.12).

**Proposition 4.8.** *Let  $c \in \mathbb{R}$  and  $H_0 \in L^2$ . For any risk-aversion coefficient  $\alpha \in (0, \infty)$  and any  $H \in L^2$ , the (A)-indifference value is explicitly given by*

$$h_A(H; c, H_0, \alpha) = c^H + \delta\alpha(\text{Var}[L^H] - 2\text{Cov}(L^H, L^{H_0})),$$

where  $c^H$ ,  $L^H$  and  $L^{H_0}$  are from Lemma 4.4.

*Proof.* Use (4.13) and compare (4.5) and (4.10).  $\square$

**Proposition 4.9.** *Let  $c \in \mathbb{R}$  and  $H_0 \in L^2$ . For  $\mu \in \mathbb{R}$  and  $H \in L^2$ , the (B)-indifference value is explicitly given by*

$$h_B(H; c, H_0, \mu) = \begin{cases} +\infty & \text{if } \text{Var}[L^H - L^{H_0}] > \sigma_*^2(c, 0, 0, H_0; \mu), \\ h_*(c, H, H_0; \mu) & \text{if } \text{Var}[L^H - L^{H_0}] \leq \sigma_*^2(c, 0, 0, H_0; \mu), \end{cases}$$

where

$$h_*(c, H, H_0; \mu) := c^H - (c + c^{H_0}) + \delta\mu - \delta\sqrt{\sigma_*^2(c, 0, 0, H_0; \mu) - \text{Var}[L^H - L^{H_0}]} \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]}.$$

*Proof.* Comparing (4.6) and (4.11) shows that we have  $\sigma_*^2(c, H, h, H_0; \mu) > \sigma_*^2(c, 0, 0, H_0; \mu)$  for all  $h \in \mathbb{R}$  if  $\text{Var}[L^H - L^{H_0}] > \sigma_*^2(c, 0, 0, H_0; \mu)$ ; so (4.14) then gives  $h_B(H; c, H_0, \mu) = +\infty$ . On the other hand, if  $\text{Var}[L^H - L^{H_0}] \leq \sigma_*^2(c, 0, 0, H_0; \mu)$ , then  $h_*(c, H, H_0; \mu)$  above is well defined and due to (4.6) and (4.11) satisfies  $\sigma_*^2(c, H, h_*(c, H, H_0; \mu), H_0; \mu) = \sigma_*^2(c, 0, 0, H_0; \mu)$ . This implies  $h_B(H; c, H_0, \mu) = h_*(c, H, H_0; \mu)$ .  $\square$

**Proposition 4.10.** *Let  $c \in \mathbb{R}$  and  $H_0 \in L^2$ . For  $\sigma^2 \geq \text{Var}[L^{H_0}]$  and  $H \in L^2$ , the (C)-indifference value is explicitly given by*

$$h_C(H; c, H_0, \sigma^2) = \begin{cases} +\infty & \text{if } \text{Var}[L^H - L^{H_0}] > \sigma^2, \\ h_*(c, H, H_0; \sigma^2) & \text{if } \text{Var}[L^H - L^{H_0}] \leq \sigma^2, \end{cases}$$

where

$$h_*(c, H, H_0; \sigma^2) := c^H - \delta \left( \sqrt{\sigma^2 - \text{Var}[L^H - L^{H_0}]} - \sqrt{\sigma^2 - \text{Var}[L^{H_0}]} \right) \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]}.$$

*Proof.* Proposition 2.7 and (4.4) show that if  $\sigma^2 < \sigma_{\text{mv}}^2(c, H, h, H_0) = \text{Var}[L^H - L^{H_0}]$ , Problem (C) for  $Y = -cB + (H - hB) - H_0$  cannot be solved and hence  $h_C(H; c, H_0, \sigma^2) = +\infty$  by (4.15). On the other hand, if  $\sigma^2 \geq \text{Var}[L^H - L^{H_0}]$ , then  $h_*(c, H, H_0; \sigma^2)$  above is well defined and satisfies  $\mu_*(c, H, h_*(c, H, H_0; \sigma^2), H_0; \sigma^2) = \mu_*(c, 0, 0, H_0; \sigma^2)$  due to (4.7) and (4.12). This implies  $h_C(H; c, H_0, \sigma^2) = h_*(c, H, H_0; \sigma^2)$ .  $\square$

The next result shows that in the nontrivial cases, all mean-variance indifference values share the same fundamental structure. For ease of notation, we omit most arguments of the  $h_i$ .

**Corollary 4.11.** *Let  $c \in \mathbb{R}$  and  $H_0 \in L^2$ . For any  $\alpha \in (0, \infty)$ ,  $\mu > \frac{c+cH_0}{\delta}$  and  $\sigma^2 > \text{Var}[L^{H_0}]$  and any  $H \in L^2$  such that  $\text{Var}[L^H - L^{H_0}] \leq \sigma_*^2(c, 0, 0, H_0; \mu)$  and  $\text{Var}[L^H - L^{H_0}] \leq \sigma^2$ , we have for some  $\alpha_i \in (0, \infty)$  that*

$$h_i(H) = c^H + \delta \alpha_i (\text{Var}[L^H] - 2 \text{Cov}(L^H, L^{H_0})) =: c^H + \varrho_i(H) \quad \text{for } i \in \{A, B, C\}. \quad (4.17)$$

*Note, however, that  $\alpha_i$  can depend on  $H$  via  $L^H$  in the cases  $i \in \{B, C\}$ .*

*Proof.* For  $i = A$ , this is immediate from Proposition 4.8 with  $\alpha_A := \alpha$ . For  $i = B$  and  $i = C$ , one simply checks by direct computation that (4.17) holds, respectively, with

$$\alpha_B := \left( \sqrt{\sigma_*^2(c, 0, 0, H_0; \mu) - \text{Var}[L^{H_0}]} + \sqrt{\sigma_*^2(c, 0, 0, H_0; \mu) - \text{Var}[L^H - L^{H_0}]} \right)^{-1} \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]},$$

$$\alpha_C := \left( \sqrt{\sigma^2 - \text{Var}[L^{H_0}]} + \sqrt{\sigma^2 - \text{Var}[L^H - L^{H_0}]} \right)^{-1} \frac{\sqrt{\text{Var}[\pi(B)]}}{E[\pi(B)]}.$$

$\square$

The representation in Corollary 4.11 has an interesting financial interpretation. Indeed, (4.17) shows that all mean-variance indifference values can be written as the sum of  $c^H$  and an additional risk premium  $\varrho_i(H)$ . By Remark 4.5, the term  $c^H$  is the replication price for the attainable part of the contingent claim  $H$ , or the  $L^2$ -approximation value of  $H$ . The risk premium depends on  $H$  only via  $L^H$ , which is the unhedgeable part of  $H$ , and it also takes into account the covariance between  $L^H$  and the unhedgeable part  $L^{H_0}$  of the existing position  $H_0$ . The indifference value  $h_i(H)$  itself is always increasing with respect to the difference



$\text{Var}[L^H] - 2\text{Cov}(L^H, L^{H_0})$ . So an agent might be willing to *pay* for selling a payoff  $H$  if in his view, its unhedgeable part  $L^H$  has a *diversification* or *insurance* effect on his overall position.

In view of part 3) in Remark 4.7, Proposition 4.8 also yields an explicit expression for the *bid-ask spread*  $s_A$  between the seller and buyer versions of the (A)-indifference value; we have

$$s_A(H) := h_A^{\text{seller}}(H) - h_A^{\text{buyer}}(H) = h_A(H) + h_A(-H) = 2\delta\alpha_A \text{Var}[L^H] = 2\delta\alpha \text{Var}[L^H].$$

It is interesting to observe that this bid-ask spread depends only on the risk associated to the unhedgeable part  $L^H$  of the contingent claim  $H$ ; the existing position  $H_0$  plays no role.

**Remark 4.12. 1)** For  $i \in \{B, C\}$ , the bid-ask spread  $s_i(H)$  has a more complicated form because then  $\alpha_i(H) \neq \alpha_i(-H)$ , as seen in the proof of Corollary 4.11. We do not write out the formulas, but we mention that we do obtain  $\alpha_i(H) = \alpha_i(-H)$ , and hence  $s_i(H) = 2\delta\alpha_i(H) \text{Var}[L^H]$ , if  $\text{Cov}(L^H, L^{H_0}) = 0$ . Note that  $\text{Var}[L^H]$  also represents the remaining risk in the quadratic hedging problem for  $H$ , because Lemma 4.4, Remark 4.5 and Remark 3.4 yield

$$\text{Var}[L^H] = \|L^H\|_{L^2}^2 = \|N^H - c^H\pi(B)\|_{L^2}^2 = \|N^H - c_*(H)\pi(B)\|_{L^2}^2 = \min_{(c,g) \in \mathbb{R} \times \bar{\mathcal{G}}} \|H - cB - g\|_{L^2}^2.$$

**2)** It is worth pointing out that the indifference values satisfy the following very intuitive *iterativity property*: For any  $H_1, H_2 \in L^2$  and  $i \in \{A, B, C\}$ , we have

$$h_i(H_1 + H_2; c, H_0) = h_i(H_1; c, H_0) + h_i(H_2; c + h_i(H_1; c, H_0), H_0 - H_1), \quad (4.18)$$

at least in the nontrivial cases. This can be shown by the same arguments as in Section 5.3 of Schweizer [36]. The reason why it holds is the description of  $h_i$  via an implicit equation as in part 2) of Remark 4.7; see (4.16) for the example case  $i = A$ . In financial terms, (4.18) says that the value for selling the sum claim  $H_1 + H_2$  equals the sum of the value for first selling the claim  $H_1$  plus the value for then selling the claim  $H_2$ , if we adjust before the second sale both initial capital and initial position to take into proper account the effect of the first sale.

**3)** Consider the case where  $H \in \bar{\mathcal{A}}$  so that  $H = c^H B + \tilde{g}^H$  for some  $c^H \in \mathbb{R}$  and  $\tilde{g}^H \in \bar{\mathcal{G}}$ . Intuitively, such a contingent claim  $H$  is (approximately) attainable with initial capital  $c^H$ . Under the assumptions of Corollary 4.11, all mean-variance indifference values then coincide with the replication price  $c^H$  since  $L^H \equiv 0$ . This result is of course expected — the value of an attainable payoff does not depend on preferences, but is determined by arbitrage arguments alone.

**4)** Suppose  $c \geq 0$  (and not only  $c \in \mathbb{R}$ ), so that the initial capital is nonnegative. Let us also restrict the definitions of all indifference values to the interval  $[-c, +\infty)$  and denote by  $h_i^c(H)$  the resulting version of  $h_i(H)$ , for  $i \in \{A, B, C\}$ . Intuitively, this amounts to excluding the undesirable situation where an agent is allowed to start with  $c + h_i(H) < 0$ , i.e. in a debt position. It is then easy to verify that we have the natural relation  $h_i^c(H) = \max(-c, h_i(H))$ .  $\diamond$

## 5 Connections to the literature

As already mentioned in the introduction, mean-variance portfolio optimisation problems have always represented a classical topic in financial economics. In the traditional and simplest formulation, beginning with the seminal work of Markowitz [26], one considers a single-period model with a random vector in  $\mathbb{R}^d$  representing the returns on a finite number of assets. One then derives the mean-variance optimal strategy, represented by a deterministic vector in  $\mathbb{R}^d$ , and the equations describing the mean-variance frontier. For standard textbook accounts, we refer the reader to Chapter 4 of Ingersoll [17], the book by Markowitz [27], Chapter 3 of Huang/Litzenberger [16], Chapter 6 of Luenberger [24] or Chapter 3 of Barucci [1]. The survey by Steinbach [39] contains a more detailed treatment and an extensive bibliography. Another standard account is in Section 5.2 of Cochrane [7], and we come back to this later in this section.

In the last two decades, quadratic portfolio optimisation problems have also drawn the attention of researchers in the mathematical finance community. Typically, one considers more or less general continuous-time semimartingale models and uses the powerful tools of stochastic calculus to characterise the optimal *strategy*, which is represented by a predictable process (satisfying suitable technical conditions). We do not attempt here a detailed survey of the extensive relevant literature, but only refer to Schweizer [38]. We just mention that one large body of literature on mean-variance hedging is based on projection techniques and martingale methods; see for instance the survey papers by Pham [30] and Schweizer [37]. In addition, stochastic control techniques and backward stochastic differential equations have been used to solve Markowitz problems in continuous-time models; see for instance Zhou [43] for an overview of the Itô process case, or Czichowsky/Schweizer [10] for some recent results in a general semimartingale framework. In the context of discrete-time multiperiod models, Markowitz problems and mean-variance optimal strategies have been studied in Li/Ng [23] via recursive techniques, and in Leippold et al. [21] by a geometric approach. Duality methods have also been employed, to obtain characterisations of mean-variance optimal strategies in terms of optimal (signed) martingale measures; see for instance Leitner [22], Hou/Karatzas [15], Xia/Yan [42] and Czichowsky/Schweizer [9]. All the authors from the last group work in general semimartingale settings which do not assume specific modelling structures. In that respect, and also in some of its techniques, this latter strand of literature is rather close to our abstract approach.

While the stochastic calculus approach has made possible some significant advances, it has also sometimes obscured by technicalities some of the key ideas. The present paper aims to redress the balance by going back to a simple and model-independent setup with easy mathematics where the main ideas become better visible. To put this into a broader perspective, let us consider the problem of maximising expected utility from terminal wealth by dynamic trading in a financial market. We make no attempt to survey or even mention the extensive literature on this topic; our purpose is better served if we focus on just three key contributions.

The most important insight from our perspective is due to Pliska [31] who in turn gives credit to Kreps [20]. The idea in Pliska [31] is that optimal portfolio problems can be separated into

two steps: (1) maximise the objective function over the set of all outcomes that are attainable in the given financial market, and use convex duality tools to characterise the optimal final position; (2) describe in more detail the set of attainable outcomes. Pliska [31] considers a classical (concave and strictly increasing) utility function and explains how to tackle step (1).

The next step forward was taken by Karatzas et al. [18] who worked out step (2) from above in an Itô process setting with a Brownian filtration. They introduced the concept of fictitious completions and showed that the solution to the original portfolio problem (even under constraints) could be obtained by working in the least favourable fictitious complete market. It must be emphasised that their approach crucially needs the underlying Brownian structure, via a martingale representation result.

The third and final contribution we want to highlight is due to Kramkov/Schachermayer [19]. They tackle the optimal portfolio problem in a general semimartingale market where they prove existence of a solution and characterise its structure via the first order conditions for optimality. They use functional analysis and convex duality techniques, and an important point is that they focus on the indirect utility (or value function) of the portfolio optimisation problem. In particular, strategies become of secondary importance.

The main point of the present paper is that for mean-variance optimisation problems without constraints, the above overall approach becomes very simple and leads to very explicit results. The attainable positions naturally form a linear subspace of  $L^2$ , and it is enough to use simple Hilbert space arguments to work out the value functions and the optimal final positions for a whole range of quadratic problems. In that sense, our paper is also a tribute to the insights in the earlier economics and finance literature.

More precisely, the setup of our paper lies on a middle ground between the classical approach in  $\mathbb{R}^d$  and the more sophisticated one via semimartingales and stochastic calculus. On the one hand, we essentially have a one-period model. On the other hand, we avoid any specific description of the underlying (frictionless) financial market because we do not model assets, returns and strategies, but only work with the abstract (linear) space  $\mathcal{A} := \mathbb{R}B + \mathcal{G} \subseteq L^2$  of attainable final wealths. Put differently, we *parametrise* our variables not via strategies, but directly via the resulting final *positions*. The key advantage of this approach is that it allows to describe in a simple way the general structure of all mean-variance optimal positions, together with their fundamental economic properties. Moreover, our results are by construction model-independent and hence hold for any (frictionless) semimartingale financial model. But of course, there is a price to pay: We can describe the optimal wealth *positions* and their general properties, but we cannot give the corresponding *trading strategies* — there *are* no strategies in our setup because these depend on the financial market model. We provide some more comments on this aspect in Remark 5.3.

Of course, there are also precursors of our work in the economics and finance literature. Perhaps the best reference is Section 5.3 in Cochrane [7] which makes the key observation that arguments and derivations become easier if one switches from “portfolio space” to “asset space” (or “return space”, in Cochrane [7]). This can be traced back to earlier work by

Chamberlain/Rothschild [6] and Hansen/Richard [12] who introduced Hilbert space techniques in the context of mean-variance problems to study the mean-variance frontier; see Chapters 5 and 6 of Cochrane [7] for a textbook account. A related approach can be found in Luenberger [25], under the standard assumption that the market is generated by a finite number of assets. The specific  $L^2$ -setup adopted in this paper has been introduced in Schweizer [33, 34] and used in several later works; see for instance Schweizer [35, 36], Møller [29] and Sun/Wang [40, 41].

In comparison with the last group of papers, our results here provide two innovations. We systematically tackle and solve a whole range of quadratic optimisation problems in a unified way, including connections between the different problems and their solutions. Moreover, we systematically deal with *undiscounted* quantities. This contrasts with the standard mathematical finance literature where one typically (“without loss of generality”) works from the beginning with already discounted quantities. If we interpret  $B$  as the final value of a savings account, discounting corresponds to letting  $B \equiv 1$ . As a consequence, several well-known results for the discounted case (see for instance Møller [29], Theorem 4.3) can be recovered by specialising our general results to the case  $B \equiv 1$ . However, using undiscounted terms seems to us more natural from a financial economics point of view.

Earlier work on abstract financial markets with stochastic interest rates by Schweizer [36] and Sun/Wang [40, 41] has interpreted  $B$  as the final value of a savings account and then considered mean-variance problems in terms of  $B$ -discounted quantities, under the no-arbitrage-type condition  $B \notin \overline{\mathcal{G}}$  of no approximate profits in  $L^2$ . For related work, compare also Section 3.5 in Rheinländer [32] and Chapter 1 in Fontana [11]. Because we do not discount and give no specific interpretation to  $B$ , we impose instead the no-arbitrage-type condition  $1 \notin \overline{\mathcal{G}}$  of no approximate *riskless* profits in  $L^2$ . As we have seen, only the latter is necessary for solving our general mean-variance problems. Of course, the distinction only matters if  $B$  is random.

**Remark 5.1.** The issue of discounting is actually more subtle than the bland phrase “without loss of generality” suggests. Several papers introduce  $B$ -discounted quantities and then solve mean-variance portfolio optimisation problems with respect to the measure  $P^B$  defined by  $dP^B/dP := B^2/E[B^2]$ , instead of the original  $P$ . For mean-variance hedging, this is fine because  $\|g/B\|_{L^2(P^B)}$ , the second moment of  $B$ -discounted gains with respect to the measure  $P^B$ , corresponds (up to a normalising factor) to  $\|g\|_{L^2(P)}$ . But this does not hold for the mean, since

$$E^B \left[ \frac{g}{B} \right] := E \left[ \frac{dP^B}{dP} \frac{g}{B} \right] = E \left[ \frac{B^2}{E[B^2]} \frac{g}{B} \right] = \frac{1}{E[B^2]} E[Bg].$$

Because  $g$  models undiscounted gains, it seems not clear if the last quantity has a meaningful economic interpretation under the original measure  $P$ , nor why an agent with mean-variance preferences should be interested in it. In that sense, the approach first suggested in Schweizer [34] and later followed by Sun/Wang [40, 41], among others, is mathematically elegant but seems economically flawed. Our current approach does not suffer from this inconsistency.  $\diamond$

Despite the last remark, let us briefly return to the discounted case  $B \equiv 1$ . As can be seen from Schweizer [33, 37], Pham [30] and Møller [29], mean-variance optimisation problems are

via duality closely linked to the so-called *variance-optimal (signed) martingale measure*  $\tilde{P}$ . In our abstract terms, this is defined (for  $B \equiv 1$ ) by  $d\tilde{P}/dP := \tilde{D}$ , where  $\tilde{D} \in \mathcal{G}^\perp$  denotes the element which minimises  $\|D\|_{L^2}$  over all  $D \in \mathcal{G}^\perp$  such that  $E[D] = 1$ . The following result is known; but the proof we give here, and especially the insight behind it, seems to us much more elegant than previous ones (e.g. in Schweizer [33]).

**Corollary 5.2.** *Let  $B \equiv 1$ . If Assumption I holds, the variance-optimal (signed) martingale measure  $\tilde{P}$  can be uniquely characterised by*

$$\frac{d\tilde{P}}{dP} = \frac{\pi(1)}{E[\pi(1)]}.$$

*Proof.* Equivalently to the definition,  $\tilde{D}$  minimises  $\text{Var}[D]$  over all  $D \in \mathcal{G}^\perp$  such that  $E[D] = 1$ . But this is simply a particular case of Problem (B), with  $Y \equiv 0$  and with  $\bar{\mathcal{G}}$  exchanged for  $\mathcal{G}^\perp$ . In Proposition 2.6, we thus have to replace  $\pi$  by  $\text{Id} - \pi$ , hence  $1 - \pi(1)$  by  $\pi(1)$ , and so the result follows directly from (2.23) with  $m = 1$ .  $\square$

We conclude this section with a brief literature review for Section 4. Utility-based indifference valuation rules were introduced in the mathematical finance literature by Hodges/Neuberger [14] and then studied in a variety of settings; see for instance Henderson/Hobson [13] and Becherer [4] for recent overviews. However, explicit results are available only in a handful of cases; this mainly includes exponential utility as in Becherer [2, 3] and mean-variance preferences as in the present paper. More specifically, the indifference valuation rules analysed in Section 4 are closely related to the *utility indifference prices* under mean-variance preferences used in Mercurio [28], Møller [29], Schweizer [36], Sun/Wang [40] and Section 1.3 of Fontana [11]. By letting  $B \equiv 1$  throughout Section 4, we easily obtain mean-variance indifference values with respect to discounted quantities, recovering the case studied in Mercurio [28] and Møller [29]. In particular, if  $B \equiv 1$ , Assumption III is automatically satisfied with  $\delta = 1$  and  $\bar{g} \equiv 0$ . Definition 4.6 is inspired by the notion of *mean-variance price* introduced by Bielecki et al. [5] in the context of credit risk modelling, and our Proposition 4.9 can be regarded as a generalised and abstract counterpart to their Proposition 18.

**Remark 5.3.** A comparison between Section 4 and the results in Part II of Bielecki et al. [5] makes advantages and disadvantages of our approach rather transparent. In contrast to Bielecki et al. [5], we can obtain mean-variance indifference *valuation formulas* without ever specifying a precise credit risk model, and this also allows us to bypass, as in Section 4.2 of Fontana [11], most of the sometimes technical and complicated arguments in Bielecki et al. [5]. On the other hand, those latter arguments not only produce the valuation formulas, but also, for the specific chosen model, the dynamic *trading strategies* needed for hedging. Our approach inherently cannot deliver that.

In more general terms, our approach is easy and fully general and remains so as long as we are content with *values* or *positions*. If we want the corresponding *strategies*, however, we have

to choose a model — and this is then where all the stochastic calculus techniques must come in, and where extra work is required.  $\diamond$

**Acknowledgements.** Claudio Fontana gratefully acknowledges financial support from the “Nicola Bruti Liberati” scholarship for studies in Quantitative Finance. For Martin Schweizer, financial support by the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management) is gratefully acknowledged. The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

## References

- 1 BARUCCI, E. (2003), *Financial Markets Theory. Equilibrium, Efficiency and Information*, Springer, London.
- 2 BECHERER, D. (2003), Rational hedging and valuation of integrated risks under constant absolute risk aversion, *Insurance: Mathematics and Economics*, 33: 1–28.
- 3 BECHERER, D. (2006), Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging, *Annals of Applied Probability*, 16: 2027–2054.
- 4 BECHERER, D. (2010), Utility indifference valuation, in: Cont, R. (ed.), *Encyclopedia of Quantitative Finance*, 1854–1869, Wiley, Chichester.
- 5 BIELECKI, T.R., JEANBLANC, M. & RUTKOWSKI, M. (2004), Hedging of defaultable claims, in: Carmona, R. A. et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2003, Lecture Notes in Mathematics 1847*, 1–132, Springer, Berlin Heidelberg.
- 6 CHAMBERLAIN, G. & ROTHSCILD, M. (1983), Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica*, 51: 1281–1304.
- 7 COCHRANE, J.H. (2005), *Asset Pricing*, revised edition, Princeton University Press, Princeton.
- 8 COURTAULT, J.M., DELBAEN, F., KABANOV, Y. & STRICKER, C. (2004), On the law of one price, *Finance and Stochastics*, 8: 525–530.
- 9 CZICHOWSKY, C. & SCHWEIZER, M. (2012), Convex duality in mean-variance hedging under convex trading constraints, *NCCR FINRISK working paper No. 667, ETH Zürich*, [http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP667\\_D1.pdf](http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP667_D1.pdf), to appear in the *Applied Probability journals*
- 10 CZICHOWSKY, C. & SCHWEIZER, M. (2012), Cone-constrained continuous-time Markowitz problems, *NCCR FINRISK working paper No. 683, ETH Zürich*,

- [http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP683\\_D1.pdf](http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP683_D1.pdf), to appear in *Annals of Applied Probability*
- 11 FONTANA, C. (2010), Mean-Variance Problems with Applications to Credit Risk Models, *MAS Finance thesis, ETH Zurich and University of Zurich*,  
[http://www.math.unipd.it/~fontana/Files/MASF\\_CFontana.pdf](http://www.math.unipd.it/~fontana/Files/MASF_CFontana.pdf).
  - 12 HANSEN, L.P. & RICHARD, S.F. (1987), The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica*, 55: 587–613.
  - 13 HENDERSON, V. & HOBSON, D. (2009), Utility indifference pricing: An overview, in: Carmona, R. (ed.), *Indifference Pricing: Theory and Applications*, 44–74, Princeton University Press, Princeton.
  - 14 HODGES, S.D. & NEUBERGER, A. (1989), Optimal replication of contingent claims under transaction costs, *Review of Futures Markets*, 8: 222–239.
  - 15 HOU, C. & KARATZAS, I. (2004), Least-squares approximation of random variables by stochastic integrals, in: Kunita, H. et al. (eds.), *Stochastic Analysis and Related Topics in Kyoto*, Advanced Studies in Pure Mathematics, vol. 41: 141–166, Mathematical Society of Japan, Tokyo.
  - 16 HUANG, C. & LITZENBERGER, R.H. (1988), *Foundations for Financial Economics*, North-Holland, New York.
  - 17 INGERSOLL, J.E. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa.
  - 18 KARATZAS, I., LEHOCZKY, J.P., SHREVE, S.E. & XU, G.-L. (1991), Martingale and duality methods for utility maximization in an incomplete market, *SIAM Journal on Control and Optimization*, 29: 702–730.
  - 19 KRAMKOV, D. & SCHACHERMAYER, W. (1999), The asymptotic elasticity of utility functions and optimal investment in incomplete markets, *Annals of Applied Probability*, 9: 904–950.
  - 20 KREPS, D.M. (1979), Three essays on capital markets, *Research Paper No. 499, Graduate School of Business, Stanford University*,  
[http://gsbapps.stanford.edu/researchpapers/detail1.asp?Document\\_ID=2200](http://gsbapps.stanford.edu/researchpapers/detail1.asp?Document_ID=2200).
  - 21 LEIPPOLD, M., TROJANI, F. & VANINI, P. (2004), A geometric approach to multiperiod mean variance optimization of assets and liabilities, *Journal of Economic Dynamics and Control*, 28: 1079–1113.
  - 22 LEITNER, D. (2000), Mean-variance efficiency and intertemporal price for risk, *Discussion Paper CoFE 35, University of Konstanz*.

- 23** LI, D. & NG, W.L. (2000), Optimal dynamic portfolio selection: Multiperiod mean-variance formulation, *Mathematical Finance*, 10: 387–406.
- 24** LUENBERGER, D.G. (1998), *Investment Science*, Oxford University Press, New York.
- 25** LUENBERGER, D.G. (2001), Projection pricing, *Journal of Optimization Theory and Applications*, 109: 1–25.
- 26** MARKOWITZ, H. (1952), Portfolio selection, *Journal of Finance*, 7: 77–91.
- 27** MARKOWITZ, H. (1987), *Mean-variance Analysis in Portfolio Choice and Capital Markets*, Basil Blackwell, Oxford New York.
- 28** MERCURIO, F. (2001), Claim pricing and hedging under market incompleteness and “mean-variance” preferences, *European Journal of Operations Research*, 133: 635–652.
- 29** MØLLER, T. (2001), On transformations of actuarial valuation principles, *Insurance: Mathematics and Economics*, 28: 281–303.
- 30** PHAM, H. (2000), On quadratic hedging in continuous time, *Mathematical Methods of Operations Research*, 51: 315–339.
- 31** PLISKA, S.R. (1986), A stochastic calculus model of continuous trading: Optimal portfolios, *Mathematics of Operations Research*, 11: 371–382.
- 32** RHEINLÄNDER, T. (1999), Optimal Martingale Measures and their Applications in Mathematical Finance, *PhD thesis, Technische Universität Berlin*.
- 33** SCHWEIZER, M. (1996), Approximation pricing and the variance-optimal martingale measure, *Annals of Probability*, 24: 206–236.
- 34** SCHWEIZER, M. (1997), From actuarial to financial valuation principles, *Proceedings of the 7th AFIR Colloquium and the 28th ASTIN Colloquium, Cairns (1997), Joint Day Volume*, 261–282.
- 35** SCHWEIZER, M. (1999), Risky options simplified, *International Journal of Theoretical and Applied Finance*, 2: 59–82.
- 36** SCHWEIZER, M. (2001a), From actuarial to financial valuation principles, *Insurance: Mathematics and Economics*, 28: 31–47.
- 37** SCHWEIZER, M. (2001b), A guided tour through quadratic hedging approaches, in: Jouini, E. et al. (eds.), *Option Pricing, Interest Rates and Risk Management*, 538–574, Cambridge University Press, Cambridge.
- 38** SCHWEIZER, M. (2010), Mean-variance hedging, in: Cont, R. (ed.), *Encyclopedia of Quantitative Finance*, 1177–1181, Wiley, Chichester.



- 39** STEINBACH, M.C. (2001), Markowitz revisited: Mean-variance models in financial portfolio analysis, *SIAM Review*, 43: 31–85.
- 40** SUN, W.G. & WANG, F.W. (2005), Claim hedging in an incomplete market, *Journal of Systems Science and Complexity*, 18: 193–201.
- 41** SUN, W.G. & WANG, F.W. (2006), The mean-variance investment problem in a constrained financial market, *Journal of Mathematical Economics*, 42: 885–895.
- 42** XIA, J. & YAN, J.A. (2006), Markowitz’s portfolio optimization in an incomplete market, *Mathematical Finance*, 16: 203–216.
- 43** ZHOU, X.Y. (2003), Markowitz’s world in continuous time, and beyond, in: Yao, D.D. et al. (eds.), *Stochastic Modeling and Optimization*, 279–310, Springer, New York.