Finance Stoch (2011) 15: 27–55
DOI 10.1007/s00780-009-0117-9

Co-monotonicity of optimal investments and the design of structured financial products

Marc Oliver Rieger

Received: 18 June 2008 / Accepted: 9 January 2009 / Published online: 7 January 2010

Abstract We prove that, under very weak conditions, optimal financial products on complete markets are co-monotone with the reversed state price density. Optimality is meant in the sense of the maximization of an arbitrary preference model, e.g., expected utility theory or prospect theory. The proof is based on a result from transport theory. We apply the general result to specific situations, in particular the case of a market described by the Capital Asset Pricing Model or the Black–Scholes model, where we derive a generalization of the two-fund-separation theorem and give an extension to APT factor models and structured products with several underlyings. We use our results to derive a new approach to optimization in wealth management, based on a direct optimization of the return distribution of the portfolio. In particular, we show that optimal products can (essentially) be written as monotonic functions of the market return. We provide existence and nonexistence results for optimal products in this framework. Finally we apply our results to the study of bonus certificates, show that they are not optimal, and construct a cheaper product yielding the same return distribution.

Keywords Co-monotonicity · Structured products · Portfolio optimization · No-arbitrage condition · Decision theory

Mathematics Subject Classification (2000) 91B28 · 49J45 · 49M25 · 91B06

JEL Classification G11 · C61

The author thanks Thorsten Hens for his steady and very valuable contributions to this project and Bernard Dumas, Alexander Schied, Martin Schweizer, and Mei Wang for their helpful comments.

M.O. Rieger (✉)
ISB, University of Zurich, Zurich, Switzerland
e-mail: rieger@isb.uzh.ch

1 Introduction

Structured financial products have gained a large popularity in many countries in the last years. In 2007, a volume of 100 billion USD has been issued in the USA, and in Germany even 200 billion Euro have been reached. Yearly growth rates (at least before the financial crisis) have been in the order of 30%, and structured products are quickly becoming a standard form of investment for private investors. Their success sometimes challenges traditional financial models but can often be explained by behavioral theories that take effects like loss aversion into account.

On the other hand, many banks have understood today that selling separate financial products is not the best way to achieve an overall optimal portfolio for the client, since it does not consider correlations between different products in the client's portfolio. Therefore an integrated wealth management that optimizes the overall portfolio of a client or offers a tailor-made collection of assets and structured products is more and more frequently offered. A systematic approach to this optimization that takes into account potential behavioral biases of the client is therefore of high importance.

In this article we lay the theoretical foundation of such a systematic approach for wealth management and study properties that an optimized portfolio should have. We consider this in the framework of structured products, where we assume that the total wealth of the client is invested into this structured product and there is hence no background risk to be hedged. We impose only the mildest possible conditions on the preferences of the client. In particular we do not assume that the client will want to optimize its investment according to mean–variance theory or to the rational framework of expected utility theory, but we allow explicitly for other decision models, e.g., behavioral models like prospect theory. We also allow for "benchmarking," i.e., for variable reference points that are set, for instance, by some index (like a stock market index), as well as for decision models based on the total return.

One main result of our work is that even in this general setting certain properties of an optimal financial product are always present (Theorem 2.12). In particular, we show that optimal products "follow the market," i.e., they are co-monotone with the market portfolio (in the case of a CAPM market or Black–Scholes pricing) or with the reversed state price density (in the general case). This result has immediate consequences for the design of financial products in the context of wealth management, and we use it to develop a new method for finding optimal investments. Another main result of our work is an existence theorem for such optimal investments (Theorem 3.1).

Let us have a closer look at the model we are studying. We consider complete and efficient financial markets in which all market participants have homogenous beliefs and act according to a maximization of their utility. The main focus of this article lies on the question what properties a financial product on such a market has to satisfy if it is optimal in the sense that it maximizes a given utility of an investor. Before we make this question more precise, we first review some properties of such markets (compare, e.g., [8] for details and [2] for generalizations). First, we need a general pricing formula. One obtains such a formula based on the assumption of no-arbitrage. A precise derivation can be found, e.g., in the book by Le Roy and Werner [21, Chap. 20]. We give a short heuristic below.

Let the return distribution of an asset be given by the probability measure p on $\mathbb{R}$, let the state price density be π, and let their mean and variance be given by $\mathbb{E}(p)$,

$\mathbb{E}(\pi)$ and $\mathrm{var}(p)$, $\mathrm{var}(\pi)$, respectively. (In this article we interpret probability measures on the space of possible returns often as random variables by identifying the sample space with the space of possible returns.) Let R be the return of the risk-free asset (i.e., the interest rate or risk-free rate). Then we can derive from a suitable no-arbitrage condition that all financial products that are available for a fixed price (for simplicity, we set this price to one) and that can be described by a joint probability measure T on $\mathbb{R}^m \times \mathbb{R}^n$ such that $p = \int_{\mathbb{R}} dT(\cdot, y)$ and $\pi = \int_{\mathbb{R}} dT(x, \cdot)$ satisfy the constraint

$$\mathbb{E}(p) - R = -\beta_{p\pi}\big(\mathbb{E}(\pi) - R\big), \qquad \text{where } \beta_{p\pi} = \frac{\operatorname{cov} T}{\operatorname{var}(\pi)}. \tag{1.1}$$

Heuristically, this formula can be understood in terms of risk factors. Then π is the risk of an investment in a given situation. The returns of the asset p above the risk-free rate are then approximated by a linear regression with π, and $\beta_{p\pi}$ is the slope of this linear approximation, i.e., the regression coefficient. Alternatively, one can think of $\pi(x)$ as the price for a fixed payoff in state x. If x corresponds to the market return at maturity, then payoffs in states with low market return will be more expensive, and thus we expect π to be decreasing in x. For a rigorous derivation of (1.1), see [21, Chap. 20].

We make henceforth the general assumption that the state prices are nonnegative (compare [8]), which corresponds to assuming that the preferences of the market participants are weakly monotonic. An optimal product is defined as a product that maximizes a given utility subject to condition (1.1). The utility could here be given according to expected utility theory, prospect theory, or a different model, depending on the application one has in mind.

We define the *reversed state price density* $\tilde{\pi}$ by $\tilde{\pi}(\mathbb{E}(\pi) + x) := \pi(\mathbb{E}(\pi) - x)$. It has been observed in the literature that in certain cases an optimal asset is co-monotone to $\tilde{\pi}$. (We give a precise definition of co-monotonicity in Sect. 2.1.) This is similar to a classical result in the context of Pareto efficiency, see, e.g., [20] and [21, Sect. 15.5], and has been generalized by a very simple and neat result of Dybvig [9] to optimal portfolio design in the case of finite state spaces with equal probability for each state and an expected utility maximizer. He states that then “any cheapest way to achieve a lottery assigns the outcomes of the lottery to the states in reverse order of the state price density”, in other words that an optimal portfolio and the reversed state price density are co-monotone. He also mentions that “the analysis still works when probabilities are unequal if we assume that agents are risk averse” [9, p. 389]. Furthermore, he suggests a generalization to a nonexpected utility setting of Machina preferences [22] and to continuous state spaces with nonatomic state prices. (A proof for the latter extension in the case of a strictly concave utility can be found in [11, Chap. 3].)

Dybvig presents interesting applications of co-monotonicity on dynamic portfolio strategies in a follow-up paper [10], but no additional results on co-monotonicity are derived there. A first generalization of co-monotonicity results to markets with frictions has been given by Jouini and Kallal [14] for finite state spaces and by Jouini and Porte [17] for arbitrary state spaces, in both cases under the assumption of a concave utility function. The concept of co-monotonicity has also been applied to insurance risks by Dhaene et al. [7] and by Carlier and Dana [4, 5].

In this article we generalize the observation of Dybvig [9] in a mathematically rigorous way into four directions:

1. We remove the restriction of equal probabilities for the states (without adding additional assumptions on the preferences).
2. We give a rigorous generalization to arbitrary state spaces, without assuming concavity of the utility function or nonatomic measures. In fact, we study general probability measures on the infinite state space of returns (i.e., the real numbers).
3. Instead of considering only expected utility theory, we allow for arbitrary decision models for investors without background risk, compare Definition 2.11.[1]
4. We allow for "benchmarking," i.e., the investor's utility may depend on a variable reference point, e.g., the market index (see Sect. 2.5).

Our results are based on general mathematical methods for transport plans (compare Ambrosio [1]), which we summarize in Sect. 2.1 of this article. We need to rely on this result since rearrangement techniques (going back to Hardy, Littlewood, and Pólya [12] and later applied in [5] and [3]) generally require nonatomic marginals. We apply these mathematical results in Sect. 2.2 to derive general conditions under which the outcome distribution of an optimal financial product is co-monotone with the reversed state price density. It turns out that this property is much more universal than had been anticipated. We discuss special cases in Sect. 2.3–2.5, in particular the Capital Asset Pricing Model (CAPM), and prices according to the Black–Scholes model. We extend a variant of the two-fund-separation theorem to the case of general preferences.

The co-monotonicity result opens the path for a new approach to the design of optimal financial products, which is based not on the optimization of asset allocations, but instead of a direct optimization of the underlying return distribution. This approach is explained in Sect. 3.1, and new existence results for optimal financial products are derived. Since this optimization idea extends the set of admissible investments substantially, the optimization cannot rely on classical methods but needs to invoke results from nonlinear analysis. We briefly sketch some numerical methods for the computation of such optimal products in Sect. 3.2. There are, however, limitations to this approach, which provide interesting insights into the shortfalls of pricing formulas based on the no-arbitrage condition. These limitations will be discussed in Sect. 3.3.

We conclude this article by a practical application, namely the study of bonus certificates (Sect. 4). We demonstrate that these products are not optimal and can be improved by a monotonizing procedure that can be performed explicitly.

2 Co-monotonicity

In this section we present some of our main results. We start out from a mathematical analysis of joint probability measures in Sect. 2.1. The mathematically less inclined

[1] In particular, there is no need to assume a strictly concave utility maximization, contrary to the traditional intuition to co-monotonicity that one might get from the equivalence between state price density and marginal rate of substitution [20].

reader might skip this section and continue with Sect. 2.2 at first reading, relying on the intuitive idea that joint probability measures that minimize (or maximize) certain quantities are co-monotone, which means roughly that the probability measures that are connected via the joint probability measure "follow" each other: a larger outcome of one of them always corresponds to a larger outcome of the other and vice versa.

In Sect. 2.2 we apply these results to the study of optimal financial products and prove that such optimal products are co-monotone with the reversed state price density under general conditions. In Sects. 2.3–2.5, we study special cases of this general statement which are of particular interest.

2.1 Co-monotonicity of joint probability distributions

The main mathematical tool that we apply in this article is the so-called "transport theory." This theory originally dealt with optimizing transports of soil, e.g., in construction or mining, and goes back to the 18th century when the French mathematician Monge [24] introduced the first version of this problem. Major progress on this problem has been achieved in the 1940s with the seminal work by the Russian economist and mathematician Kantorovich [19]. We state his formulation for the one-dimensional case that is of particular interest for our purpose.

Definition 2.1 (Transport problem) Let $\mu \in \mathcal{P}(\mathbb{R}^m), \nu \in \mathcal{P}(\mathbb{R}^n)$ (i.e., probability measures on $\mathbb{R}^m$ and $\mathbb{R}^n$), and let $c\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be a lower semi-continuous function (the *cost function*). Then the *transport problem* consists of finding a joint probability measure $T \in \mathcal{P}(\mathbb{R}^m \times \mathbb{R}^n)$ which minimizes

$$C(T) := \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} c(x, y)\, dT(x, y) \tag{2.1}$$

and such that the marginals of T are given by μ and ν, i.e., $\mathrm{pr}_1 T = \mu, \mathrm{pr}_2 T = \nu$, where pr_1 is the projection on the first m coordinates, and pr_2 the projection on the last n coordinates, i.e.,

$$\mathrm{pr}_1 T := \int_{\mathbb{R}^m} dT(\cdot, y), \qquad \mathrm{pr}_2 T := \int_{\mathbb{R}^n} dT(x, \cdot).$$

At the optimum, the mass μ is transported to ν for least cost according to the transport plan T. It is well known that the above transport problem admits a solution, see [1, Sect. 2].

Before we continue, we have to define the following useful notation.

Definition 2.2 (Push forward) Let $n \in \mathbb{N}$, and let $f: \mathbb{R}^n \to \mathbb{R}^n$ be a measurable function, and μ a measure on $\mathbb{R}^n$. Then we define the *push forward* $f_{\#}\mu$ by

$$(f_{\#}\mu)(B) := \mu\big(f^{-1}(B)\big) \quad \text{for all Borel sets } B \subset \mathbb{R}^n.$$

In other words, $f\#\mu = \mu \circ f^{-1}$ is simply the image measure of μ under f.

The central property we use in this article is co-monotonicity. We present first a one-dimensional formulation, before we generalize to c-monotonicity.

Definition 2.3 (Co-monotonicity of joint probability measures) Let $T \in \mathcal{P}(\mathbb{R}, \mathbb{R})$ be a joint probability measure with marginals $\mu, \nu \in \mathcal{P}(\mathbb{R})$. Then T is called *co-monotone* if for all $(x_1, y_1), (x_2, y_2) \in \operatorname{supp} T$ with $x_2 > x_1$, we have $y_2 \geq y_1$. (We also sometimes say that μ and ν are co-monotone.)

It is easy to see that this property can be expressed equivalently as a monotonicity property as follows:

Remark 2.4 T is co-monotone if and only if it satisfies the following condition: For all Borel sets $A, B \subset \mathbb{R}$ with $\mu(A) > 0$, $\mu(B) > 0$, and $\inf B > \sup A$, we define

$$A' := \operatorname{supp} \operatorname{pr}_2(T|_{A\times\mathbb{R}}), \qquad B' := \operatorname{supp} \operatorname{pr}_2(T|_{B\times\mathbb{R}}).$$

Then $\inf B' \geq \sup A'$.

Co-monotonicity hence means than that one marginal "follows" the other: the larger the outcome of one, the larger the outcome of the other (and vice versa). The definition of co-monotonicity is a natural extension of the usual notion of monotonicity in the following sense.

Proposition 2.5 *Let T be a joint probability measure with marginals μ and ν. If T can be interpreted as a map that maps the measure μ pointwise to ν, then T is co-monotone if this map is monotone as a function on $\mathbb{R}$. More precisely, if there exists a map $\tilde{T}\colon \operatorname{supp} \mu \to \mathbb{R}$ such that $T = (\mathrm{Id} \times \tilde{T})_{\#}\mu$, then T is co-monotone iff $\tilde{T}$ is μ-a.e. monotone.*[2]

Proof It is straightforward to prove that $\operatorname{supp} T$ satisfies the condition of Definition 2.3 if and only if $\tilde{T}$ is monotone. □

In the higher-dimensional case we introduce the following:

Definition 2.6 (c-monotonicity) Let $T \in \mathcal{P}(\mathbb{R}^m \times \mathbb{R}^n)$ be a joint probability measure from $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $c\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be a lower semi-continuous function. Then T is called *c-monotone* if for all $I \geq 1$, all $(x_i, y_i) \in \operatorname{supp} T$, and all permutations σ of the set $\{1, 2, \ldots, I\}$, we have

$$\sum_{i=1}^{I} c(x_i, y_i) \leq \sum_{i=1}^{I} c(x_i, y_{\sigma(i)}).$$

There are extensions of this concept to time-continuous settings [15, 16], but we do not pursue this issue in the current article, where we only study two-period models.

The central mathematical result that we apply is that solutions of transport problems under quite general conditions exist and are c-monotone [1, Theorems 2.1–2.2]. We summarize these results as follows:

[2] We remark that it is necessary to allow $\tilde{T}$ to be nonmonotone on a set N with $\mu(N) = 0$, since $\tilde{T}$ can be defined arbitrarily on such sets.

Theorem 2.7 *Let $c\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be a lower semi-continuous function. Then the transport problem (2.1) with marginals $\mu \in \mathcal{P}(\mathbb{R}^m)$ and $\nu \in \mathcal{P}(\mathbb{R}^n)$ admits a solution $T \in \mathcal{P}(\mathbb{R}^m \times \mathbb{R}^n)$ which is c-monotone.*

From these fundamental results it is easy to derive the following corollary for the case $m = n = 1$ which we use in the next section.[3]

Corollary 2.8 *There exists a co-monotone minimizer $T \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ of the transport problem with cost function $c(x, y) = -(x - \mathbb{E}(\mu))(y - \mathbb{E}(\nu))$.*

Proof From Theorem 2.7 we know that there exists an optimal c-monotone T. If we take two arbitrary points (x_1, y_1) and (x_2, y_2) in $\operatorname{supp} T$ with $x_2 \geq x_1$, then we have the inequality

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1).$$

Using $c(x, y) = -(x - \mathbb{E}(\mu))(y - \mathbb{E}(\nu))$ and simplifying the resulting inequality, we arrive at $(x_1 - x_2)(y_1 - y_2) \geq 0$. Since $x_2 \geq x_1$, it follows that $y_2 \geq y_1$. Thus T is co-monotone. □

We should like to mention that a direct proof of this result by constructing the co-monotone joint probability measure for a discrete problem and passing to the limit is possible but lengthy [25]. "Obvious" methods, e.g., the use of rearrangement techniques or Hardy–Littlewood inequalities, unfortunately need further assumptions, e.g., absolutely continuous marginals.

We conclude this section with a useful approximation lemma, which essentially states that co-monotone distributions can be approximated by functions.

Lemma 2.9 *Let T be a co-monotone joint probability measure with marginals μ and ν. Then there is a sequence of co-monotone joint probability measures*

$$T_n = (\mathrm{Id} \times \psi_n)_\# \mu,$$

where $\psi_n\colon \operatorname{supp} \mu \to \mathbb{R}$ are such that $T_n \overset{\star}{\rightharpoonup} T$.

The proof uses a standard approximation procedure and can be found, e.g., in [25].

2.2 Optimal investments: the general case

We have now all mathematical tools at hand to study co-monotonicity of financial products. We first define the prototypical optimization problem we want to study.

Definition 2.10 (Optimal financial products) Let $T \in \mathcal{P}(\mathbb{R}, \mathbb{R})$ be a joint probability measure with the marginals $\mathrm{pr}_2 T = p$ and $\mathrm{pr}_1 T = \pi$, where p is the return distribution of a financial product, and π the state price density. Let $U\colon \mathcal{P}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ be a

[3] We come back to higher-dimensional problems in Sect. 2.6 to discuss factor models.

function that assigns to every joint probability distribution a utility. Then we call T *optimal* if it maximizes $U: \mathcal{P}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ under the no-arbitrage condition (1.1).

In the following we restrict the class of admissible utility functions. The main underlying assumptions are a "positive attitude" regarding additional returns and that there is no background risk involved in the investment decision. These two assumptions are made rigorous in the following definition.

Definition 2.11 (Admissible utility functions) We call $U: \mathcal{P}(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ *admissible* if (i) and one of (ii) or (ii′) hold, where the conditions are:

(i) $U(T) \le U(T(\cdot, \cdot - c))$ for all $c > 0$.

(ii) There are a nondecreasing function $h: \mathbb{R} \to \mathbb{R}$ and a functional $\tilde{U}: \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ such that $U(T) = \tilde{U}(p_h)$, where

$$p_h(y) := \mathrm{pr}_2 T\big(x, y + h(x)\big) = \int T\big(x, y + h(x)\big)\, dy.^{4}$$

(ii′) U depends only on the return distribution p, i.e., there is a function $\tilde{U}: \mathbb{R} \to \mathbb{R}$ such that $U(T) = \tilde{U}(\mathrm{pr}_2 T)$.

Condition (i) can be summarized as "more money is better": If a financial product is described by T, then an alternative product that is only different in that it yields an *additional* sure return of $c > 0$ is always at least weakly preferred. Condition (ii′) says simply that the investor's utility only depends on the return distribution. This is the usual case and occurs, e.g., when his utility is state-independent. Condition (ii) seems at first glance quite restrictive; however, it is in fact a generalization of (ii′), which we can see if we set $h(y) := 0$.

Before we discuss these conditions in more detail, we present the main result of this section.

Theorem 2.12 (Co-monotonicity of general optimal financial products) *Let $T \in \mathcal{P}(\mathbb{R}, \mathbb{R})$ be a joint probability measure describing a financial product, where the marginals $p := \mathrm{pr}_2 T$ and $\pi := \mathrm{pr}_1 T$ of T are the return distribution and the state price density. Let T be optimal with respect to an admissible utility function. Then T is co-monotone with the reversed state price density $\tilde{\pi}$, i.e., $\hat{T}$ defined by $\hat{T}(x, y) := T(-x, y)$ is co-monotone.*

It is important to notice here that we have made no assumptions on the precise form of T and its regularity. T could (a priori!) for instance be an absolutely continuous measure, or a finite weighted sum of Dirac measures, or generally *any* probability measure. Our result, however, states that if T is optimal, it has to be of a very specific type, namely co-monotone.

[4]This is a small abuse of notation but helps to make the following more readable. More precisely, we write $T(x, y + z)$ for the measure that assigns the value $T(\{(x, y) \in \mathbb{R}^2 : (x, y - z) \in A\})$ to the set $A \subset \mathbb{R}^2$.

We remind the reader that we only made assumptions (i) and (ii) on the investor's preferences. Both include and even generalize the classical case corresponding to (i) and (ii′). We shall see in Sect. 2.5 why this generalization is useful. For the moment, let us see why we cannot generalize even more, in other words, why condition (ii) is in fact needed.

Remark 2.13 If condition (ii) of Definition 2.11 is violated, the utility U could be chosen such that joint probability measures which fail to satisfy co-monotonicity have particularly large utility; as a trivial example, we simply define $U(T) = 1$ for all T which are not co-monotone with $\tilde{\pi}$ and $U(T) = 0$ otherwise. This would obviously satisfy condition (i), but an optimal product for U could not be co-monotone with $\tilde{\pi}$.

We shall now prove Theorem 2.12. Afterwards we discuss important special cases of this result. The main idea of the proof is to apply Theorem 2.7 to prove that the covariance of T is maximized when T is co-monotone. For given p, we can then monotonize T in a way which leaves the utility unchanged but at the same time decreases the price of the product according to the no-arbitrage condition (1.1). The price reduction can then be used to improve the product by adding a sure return. Condition (i) implies that this new product has a larger utility. We conclude that an optimal T is co-monotone. To make this idea work by only using condition (ii) and not adding any assumptions to the state-price density $\tilde{\pi}$ will be the main task of the proof where we need the mathematical results of Sect. 2.1.

Proof of Theorem 2.12 Let $p \in \mathcal{P}$, and let $T \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ be a joint probability measure with marginals p and $\tilde{\pi}$ that maximizes the covariance, i.e.,

$$\operatorname{cov} T = \max\bigl\{\operatorname{cov} T \,|\, T \in \mathcal{P}(\mathbb{R} \times \mathbb{R}),\ \mathrm{pr}_2 T = \mathbb{E}(p),\ \mathrm{pr}_1 T = \mathbb{E}(\tilde{\pi})\bigr\}. \tag{2.2}$$

We prove that T is co-monotone. First, we reformulate the problem (2.2) as a transport problem, i.e., we want to find $T \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ minimizing

$$C(T) := \int_{\mathbb{R}} \int_{\mathbb{R}} c(x, y)\, dT(x, y)$$

such that $\mathrm{pr}_2 T = p$, $\mathrm{pr}_1 T = \tilde{\pi}$, where

$$c(x, y) := -\bigl(x - \mathbb{E}(p)\bigr)\bigl(y - \mathbb{E}(\tilde{\pi})\bigr). \tag{2.3}$$

Due to Corollary 2.8, we can assume that the minimizing T, i.e., the T with the largest covariance, is co-monotone. In other words, we have found a joint probability measure that maximizes the covariance, given its marginals, and is co-monotone.

Suppose now that T is a joint probability measure that maximizes the utility but is not co-monotone. Denoting $T^h(x, y) := T(x, y + h(x))$, we compute

$$\begin{aligned}\operatorname{cov} T &= \int_{\mathbb{R}} \int_{\mathbb{R}} \bigl(x - \mathbb{E}(\tilde{\pi})\bigr)\bigl(y - \mathbb{E}(p)\bigr)\, dT(x, y)\\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \bigl(x - \mathbb{E}(\tilde{\pi})\bigr)\bigl(y + h(x) - \mathbb{E}(p)\bigr)\, dT^h(x, y)\end{aligned}$$

$$= \operatorname{cov} T^h + \int_{\mathbb{R}} \int_{\mathbb{R}} \big(x - \mathbb{E}(\tilde{\pi})\big) h(x)\, dT^h(x, y)$$
$$= \operatorname{cov} T^h + \int_{\mathbb{R}} \big(x - \mathbb{E}(\tilde{\pi})\big) h(x)\, d\tilde{\pi}(x).$$

We use $p^h = \mathrm{pr}_2 T^h$ and $\mathrm{pr}_1 T = \mathrm{pr}_1 T^h = \tilde{\pi}$. We can maximize $\operatorname{cov} T$ without changing the marginals of T^h by maximizing $\operatorname{cov} T^h$, since only the first term in the above equation depends on T^h, whereas the second term only depends on T^h via $\tilde{\pi}$ which is fixed. Applying our above derivation, we see that $\operatorname{cov} T^h$ can be maximized by monotonizing T^h. We call the resulting co-monotone joint probability measure $\tilde{T}^h$ and denote $\tilde{T}^h(x, y - h(x))$ by $\tilde{T}$. Since $U(T)$ depends only on $\mathrm{pr}_2 T^h$, the utility is unchanged, i.e., $U(T) = U(\tilde{T})$. The covariance, however, has increased, i.e., $\operatorname{cov} \tilde{T} > \operatorname{cov} T$, since otherwise T^h would have been already co-monotone; but then T would have been co-monotone as well, since h is by assumption a monotone function. We define

$$d := R + \frac{\mathbb{E}(\tilde{\pi}) - R}{\operatorname{var}(\tilde{\pi})} \operatorname{cov} \tilde{T} - \mathbb{E}(p).$$

Since p satisfies the no-arbitrage condition (1.1), we have

$$R + \frac{\mathbb{E}(\tilde{\pi}) - R}{\operatorname{var}(\tilde{\pi})} \operatorname{cov} T - \mathbb{E}(p) = 0,$$

and therefore $d > 0$. Now define a new product $S \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ by

$$S(x, y) := \tilde{T}(x, y - d).$$

Then S satisfies the no-arbitrage condition (1.1), and its utility is by assumption (i) larger than the utility of $\tilde{T}$. Hence $U(S) > U(\tilde{T}) = U(T)$, and therefore T cannot be an optimal financial product. Thus every maximizer has to be co-monotone. □

In the following sections we apply Theorem 2.12 to several problems, but first we state a natural corollary that has some importance for applications [13].

Corollary 2.14 (Payoff function) *Let the likelihood ratio ℓ be a nonincreasing function of the market return. Then under the assumptions of Theorem 2.12, every optimal structured product can be written a.e. as a function of the market return (a "payoff function"). If the market return m is a nonatomic measure, then there even exists an optimal structured product that can be written everywhere as a (not necessarily continuous) function.*

Proof Let $S \subset \mathbb{R}$ be the set on which T cannot be represented by a function, i.e., $S := \{x \in \mathbb{R}; \#\operatorname{supp} T|_{\{x,\mathbb{R}\}>1}\}$. The set S has measure zero on $\mathbb{R}$, so if m is nonatomic, then T has measure zero on $S \times \mathbb{R}$. We can therefore set $T = 0$ on $S \times \mathbb{R}$ without changing the marginals. The modified T can then be represented by a function of the market return. □

2.3 The special cases of CAPM and Black–Scholes

A simplifying assumption on financial markets often made in applications is to equal the reversed state price density with the market return. This is based on the Capital Asset Pricing Model (CAPM) which follows from the mean–variance approach introduced by Markowitz [23]. Its fundamental assumption is that every investor in the market selects his portfolio according to mean–variance preferences, i.e., considers only mean and variance of the assets. In such a market every investor would hold only assets of a market portfolio and the risk-free asset, as the two-fund-separation theorem shows. The natural question is how to invest in such a market in order to maximize a utility function that is not necessarily of mean–variance type but, for instance, follows expected utility theory. Obviously, the two-fund-separation theorem will not hold in this case, but can we find some other general results describing optimal investments?

Based on the results of the previous sections, we first state the following variant of Theorem 2.12.

Proposition 2.15 (Co-monotonicity in CAPM markets) *Every financial product on a CAPM market* (*i.e.*, *a financial market where asset prices follow the CAPM*) *which is optimal for an arbitrary admissible utility has a return distribution that is co-monotone with the market return.*

Proof Note that the no-arbitrage condition (1.1) in the case of a CAPM market becomes

$$\mathbb{E}(p) - R = \beta_{pm}\big(\mathbb{E}(m) - R\big), \tag{2.4}$$

where $\beta_{pm} = \operatorname{cov}(p, m)/\operatorname{var}(m)$, and m is the market return. The market return takes therefore the role of the reversed state price density in Theorem 2.12. □

We can now extend the two-fund-separation to the case of arbitrary admissible decision models. Note that the two-fund-separation theorem implies that every product that is optimal in the mean–variance framework has a return (adjusted by a constant depending on the risk-free rate) which depends linearly on the market return. A product which is optimal for an arbitrary admissible decision model does not necessarily satisfy this, but its return depends monotonically on the market return. In other words, we have the following result.

Theorem 2.16 (Generalized "two-fund-separation") *Consider a CAPM market. An optimal investment for an investor with admissible utility gives a return which is co-monotone with the market return. If the joint probability measure can be described by a function in the sense of Proposition 2.5, this function is monotonic. If the investor is a mean–variance maximizer, this function is affine and* $R_p(x) = (1-\lambda)R + \lambda x$ *with* $\lambda \in \mathbb{R}$.

Proof The first part of this result is a reformulation of Proposition 2.15. The second part simply follows from two-fund-separation: Every optimal product has a return of the form $\lambda x + (1-\lambda)R$, where x is the market return, R the risk-free rate, and $\lambda \in \mathbb{R}$ the number of shares invested in the market portfolio. □

An immediate consequence of this result is that an optimal financial product should never speculate on falling prices, since this would violate co-monotonicity.[5]

Finally, this result also holds in the Black–Scholes framework, since the state price density is still a decreasing function of the market return.

Proposition 2.17 *The generalized two-fund-separation also holds if we use Black–Scholes pricing.*

2.4 Performance based on the outcome

In this section we study an important special case of the results of Sect. 2.2, namely when utility is only based on the outcome. We call an investor with such a utility a "private investor," since prototypical private investors would fall into this category.

We start with a precise definition what we understand by a "private investor."

Definition 2.18 (Private investor) A *private investor* is described by a utility functional U satisfying the following conditions:

1. The utility functional depends only on the return distribution of the investment, i.e., $U = U(p)\colon \mathcal{P}(\mathbb{R}) \to \mathbb{R}$. ("Only the result matters.")
2. If we shift the return distribution to the right, the utility increases, i.e., if the return distribution is given by $p \in \mathcal{P}(\mathbb{R})$ and $p_c := p(\cdot - c)$ for $c > 0$, then $U(p) < U(p_c)$. ("The more, the better.")

It is easy to see that a private investor's utility function is admissible; we just have to set $h := 0$ in Definition 2.11. Therefore, we can apply Theorem 2.12 and obtain the following result.

Proposition 2.19 *An optimal financial product for a private investor is co-monotone with the reversed state price density.*

This implies, e.g., that for all private investors in a CAPM market, the results of the previous section, in particular Proposition 2.15, apply, i.e., an optimal investment should "follow the market."

At this point it seems worthwhile to discuss how these results relate to the usual portfolio optimization strategy that tries to identify investments which are uncorrelated with the market return to improve the overall performance of the portfolio. The key differences are:

– Our first assumption implies the absence of background risk. This means that we consider the *overall investment* of a person, rather than an additional position that he might or might not add to his portfolio. Our results do not say anything about

[5] Recall that we have assumed homogeneous beliefs and the absence of background risk.

the structure of such additional positions. It might even be useful for the investor to take into his portfolio an additional position which is anti-co-monotone with the market (e.g., by going short in an asset) in order to hedge a certain risk induced by a different part of his portfolio.

– Another misunderstanding may arise from the word "market." This means of course the entirety of all possible investments, not only stocks. In particular, an investment which is uncorrelated with the stock market is usually still correlated with the "market" in this general sense.
– We have assumed that beliefs are homogenous and (in order to derive Proposition 2.15) that the market can be described by the CAPM or Black–Scholes model. In reality we might profit from anomalies of the market that are not described by the models of classical finance. For such situations, Proposition 2.15 is not applicable.

Although these limitations set a caveat on applications of our results, co-monotonicity with the reversed state prices (or the market return if we can describe the market by the CAPM) should still hold in practice if we do not aim to exploit market anomalies and if we consider our investment portfolio as a whole.

In the next section we shall see that this is even the case if we do not think in absolute returns but instead if we measure returns with respect to a benchmark index.

2.5 Performance based on a benchmark

Let us now consider the somehow opposite case of a private investor, namely an investor whose utility only depends on the return of his investment *relative to the state price*. In the case of a CAPM market, this corresponds to an investor who is only interested in the excess return of his investment compared to the market return. (Since this is the practically relevant case, we assume in this section a CAPM market.) We call such an investor a "fund manager," imagining an agent who is paid depending on the performance of his fund *with respect to the market return*. More precisely, we introduce for the case of a CAPM market

Definition 2.20 (Fund manager) A *fund manager* is described by a utility U satisfying the following conditions:

1. The utility depends only on the difference between the return of investment and market, i.e., $U = U(p_m)\colon \mathcal{P}(\mathbb{R}) \to \mathbb{R}$, where $p_m := \mathrm{pr}_2 T(x, y - x)$. In other words, the market index is used as benchmark.
2. If we shift the return distribution to the right, the utility increases, i.e., if the return distribution is given by $p \in \mathcal{P}(\mathbb{R})$ and we set $p_c := p(\cdot - c)$ for $c > 0$, then $U(p - m) < U(p_c - m)$.

We can now easily see that this is just another special case of Theorem 2.12; in fact, a fund manager's utility is admissible so that we can choose h as identity in Definition 2.11. Therefore, we obtain the following:

Proposition 2.21 *In a CAPM market, an optimal portfolio for a fund manager has a return distribution that is co-monotone with the market return.*

There is another interesting consequence of the assumptions on the preferences of a fund manager.

Proposition 2.22 (Benchmarking leads to risky products) *An optimal product for a fund manager in a CAPM market is at least as risky as the market portfolio, i.e., the difference between the return distribution and the market return is a nondecreasing function of the return.*

Proof This follows immediately from the co-monotonicity of the optimal T^h in the proof of Theorem 2.12. □

In the case of an investor with expected utility preferences with respect to the market return, i.e., a utility of the form

$$U = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u(x - y)\, dT(x, y), \tag{2.5}$$

we can strengthen the last proposition.

Proposition 2.23 *A product that maximizes the utility function U in (2.5) with $u \in C^1$, $u'(0) > 0$ in a CAPM market must be riskier than the market portfolio, i.e., the difference between the return of the product and the market return is a nondecreasing and nonconstant function of the market return.*

Proof Due to the co-monotonicity of any optimal joint probability distribution and Lemma 2.9, it is again sufficient to consider the case where $\operatorname{supp} T$ is of the form $\{(x, R_p(x))\}$. By Proposition 2.22 we already know that $R_p(x) - x$ must be nondecreasing. We want to show that $R_p(x) - x$ is nonconstant. The pricing constraint implies that $R_p(x) - x$ can only be constant if $R_p(x) = x$, i.e., if the optimal product is the market portfolio itself. Let us suppose that $R_p(x) = x$ is optimal. Our goal is to construct an improved product that satisfies the pricing constraint and has a higher utility for the fund manager.

For simplicity, we assume $u(0) = 0$. Moreover, we assume that m is absolutely continuous. (Otherwise we could approximate m by a sequence of absolutely continuous measures.) Let $\varepsilon \geq 0$ and define

$$R_{p_\varepsilon}(x) := \begin{cases} x - \varepsilon + \delta, & x < m_0, \\ x + \varepsilon + \delta, & x \geq m_0, \end{cases}$$

where m_0 is defined by $\int_{-\infty}^{m_0} dm(x) = \int_{m_0}^{+\infty} dm(x) = \frac{1}{2}$, and $\delta > 0$ is given by the no-arbitrage condition and will be computed below. (δ is in some sense the "risk premium" that we get for taking the additional risk expressed by ε.) Let p_ε be the joint probability measure induced by R_{p_ε}. Then its mean value is

$$\begin{aligned}\mathbb{E}(p_\varepsilon) &= \int_{-\infty}^{+\infty} R_{p_\varepsilon}(x)\, dm(x) \\ &= \int_{-\infty}^{+\infty} (x + \delta)\, dm(x) + \varepsilon \int_{m_0}^{+\infty} dm(x) - \varepsilon \int_{-\infty}^{m_0} dm(x) \; = \mathbb{E}(m) + \delta.\end{aligned}$$

Therefore the covariance of the co-monotone joint probability measure with marginals m and p_ε can be computed as

$$\begin{aligned}\operatorname{cov}(p_\varepsilon, m) &= \int_{-\infty}^{m_0} \big(x - \mathbb{E}(m)\big)\big(x - \varepsilon + \delta - \mathbb{E}(p_\varepsilon)\big)\, dm(x) \\ &\quad + \int_{-\infty}^{m_0} \big(x - \mathbb{E}(m)\big)\big(x + \varepsilon + \delta - \mathbb{E}(p_\varepsilon)\big)\, dm(x) \\ &= \operatorname{var}(m) + \int_{-\infty}^{m_0} \big(x - \mathbb{E}(m)\big)\big(-\varepsilon + \delta - \mathbb{E}(p_\varepsilon) + \mathbb{E}(m)\big)\, dm(x) \\ &\quad + \int_{-\infty}^{m_0} \big(x - \mathbb{E}(m)\big)\big(\varepsilon + \delta - \mathbb{E}(p_\varepsilon) + \mathbb{E}(m)\big)\, dm(x).\end{aligned}$$

We insert this and the formula for $\mathbb{E}(p_\varepsilon)$ into the no-arbitrage condition and obtain

$$\begin{aligned}\mathbb{E}(m) + \delta - R &= \big(\mathbb{E}(m) - R\big)\bigg(1 + \frac{1}{\operatorname{var}(m)} \int_{-\infty}^{m_0} -\big(x - \mathbb{E}(m)\big)\varepsilon\, dm(x) \\ &\quad + \frac{1}{\operatorname{var}(m)} \int_{-\infty}^{m_0} \big(x - \mathbb{E}(m)\big)\varepsilon\, dm(x)\bigg) \\ &= \frac{\mathbb{E}(m) - R}{\operatorname{var}(m)} \bigg(\operatorname{var}(m) + \int_{-\infty}^{m_0} x\, dm(x) - \int_{m_0}^{+\infty} x\, dm(x)\bigg)\varepsilon.\end{aligned}$$

We can resolve this to obtain a formula for δ, namely

$$\delta(\varepsilon) = \frac{\mathbb{E}(m) - R}{\operatorname{var}(m)} \bigg(\int_{-\infty}^{m_0} x\, dm(x) - \int_{m_0}^{+\infty} x\, dm(x)\bigg)\varepsilon.$$

We see from this that $\delta(0) = 0$ and $\delta'(0) > 0$. We use a Taylor expansion to compute the utility difference of p and p_ε:

$$\begin{aligned}U(p_\varepsilon) - U(p) &= \frac{1}{2}u(\delta - \varepsilon) + \frac{1}{2}u(\delta + \varepsilon) \\ &= u'(0)\delta(\varepsilon) + O\Big(\big(\delta(\varepsilon) - \varepsilon\big)^2, \big(\delta(\varepsilon) + \varepsilon\big)^2\Big) \\ &= u'(0)\big(\delta'(0)\varepsilon\big) + O\big(\varepsilon^2\big).\end{aligned}$$

Therefore, for $\varepsilon > 0$ sufficiently small, this difference is positive, i.e., $U(p_\varepsilon) > U(p)$, which shows that p defined by $R_p(x) = x$ cannot be optimal. □

We want to stress that a crucial condition for the derivation of this result was the differentiability of u at zero. If we replace the utility function u by a value function that has a kink at zero ("loss aversion") as suggested, e.g., by Kahneman and Tversky [18], our result does not apply, and it is conceivable that an investor would indeed stick exactly to the market portfolio.

2.6 Asset baskets and factor pricing

So far we have studied purely one-dimensional problems. Many structured products, however, give payments according to the payoff of several underlyings. Can we describe such products within our framework?

It is useful to put this question into a more general setup and to connect it with the arbitrage pricing theory (APT) introduced by Ross [27]. This model extends the CAPM in that it takes more than one potential risk factor into account. Factors can be, e.g., the stock market price development and also changes in the interest rate, the GDP, etc. The price of an asset p is then approximated by a sum of the linear contributions of all these factors, i.e.,

$$p = \mathbb{E}(p) + \sum_{f=1}^{F} \beta_{p,r_f} r_f,$$

where F is the number of factors, r_f is the return of the factor f, and β_{p,r_f} is the sensitivity of p to the factor f. Using a linear regression to estimate the β-terms, one obtains

$$\beta_{p,r_f} = \frac{\operatorname{cov}(p, r_f)}{\operatorname{var}(r_f)},$$

and the CAPM can be found as a special case for $F = 1$ where r_1 is the market return. See [21, Chap. 20] and [6, Chap. J] for further details and a derivation of the APT.

If we have a structured product whose payment depends on more than one underlying, we arrive at the same formula, where r_f denotes the returns of the F underlyings instead of the factor returns. Can we derive a co-monotonicity result in this framework? In fact we can obtain the following result.

Proposition 2.24 (Co-monotonicity with several factors or underlyings) *If asset prices are given by a factor model or depend on several underlyings, an optimal structured product is co-monotone with the* weighted sum *of the factors. More precisely, if the factor returns are given by r_f, then the optimal product is co-monotone with $\sum_{f=1}^{F} d_f r_f$, where we write $d_f := (\mathbb{E}(r_f) - R)/\operatorname{var}(r_f)$.*

Proof The key idea is to evoke the multidimensional Theorem 2.7 with $m = F \geq 1$, $n = 1$, and

$$c(x, y) := \sum_{f=1}^{F} \frac{\mathbb{E}(r_f) - R}{\operatorname{var}(r_f)} \big(y - \mathbb{E}(p)\big)\big(x - \mathbb{E}(r_f)\big),$$

where $x = (x_1, \dots, x_F) \in \mathbb{R}^F$. Here we use that

$$\sum_{f=1}^{F} \frac{\mathbb{E}(r_f) - R}{\operatorname{var}(r_f)} \int_{\mathbb{R}^F} \int_{\mathbb{R}} \big(y - \mathbb{E}(p)\big)\big(x - \mathbb{E}(r_f)\big)\, dT_f(x, y)$$

$$= \int_{\mathbb{R}^F} \int_{\mathbb{R}} \sum_{f=1}^{F} \frac{\mathbb{E}(r_f) - R}{\operatorname{var}(r_f)} \big(y - \mathbb{E}(p)\big)\big(x - \mathbb{E}(r_f)\big)\, dT(x, y)$$

with T_f denoting the projection of T on the (x_f, y)-plane. By Theorem 2.7 an optimal joint probability measure for the corresponding transport problem is c-monotone. Using the same computation as in the one-dimensional case, we see that it is also $\tilde{c}$-monotone with

$$\tilde{c}(x, y) := \sum_{f=1}^{F} d_f x y.$$

It follows that for (x^a, y^a) and $(x^b, y^b) \in \operatorname{supp} T$, we have

$$\tilde{c}(x^a, y^a) + \tilde{c}(x^b, y^b) \le \tilde{c}(x^a, y^b) + \tilde{c}(x^b, y^a).$$

Inserting the definition of $\tilde{c}$, we obtain

$$\left(\sum_{f=1}^{F} d_f \left(x_f^a - x_f^b\right)\right)\left(y^a - y^b\right) \ge 0,$$

which implies the co-monotonicity of p with respect to $\sum_{f=1}^{F} d_f r_f$. □

There is of course no co-monotonicity with the single factors, since the factors are not perfectly correlated.

A limitation of the above result is that the factors d_f are difficult to estimate. Nevertheless, there are many payoff profiles that can be immediately discarded as nonoptimal, regardless of the precise value of the weighting factors d_f. Examples for this include all worst-off structures where (at least in some scenarios) the payoff depends entirely on the underlying with the lowest return, since this would give a weight of zero to all other underlyings. Similarly, separate capital protection for each underlying (i.e., a payoff of the form $\sum_f \max(1, g r_f)$ where $g < 1$ is the participation rate in gains) cannot be optimal.

3 Designing optimal financial products

Applying the results of the previous section, we can introduce a new method for portfolio optimization. The co-monotonicity makes our task of finding optimal investments a lot easier, since co-monotone joint probability measures with given marginals are unique. Therefore, it is sufficient to optimize the return distribution p, assuming that the joint probability measure T is the unique co-monotone joint probability measure with marginals p and $\tilde{\pi}$. We are therefore only left with maximizing utility over all probability measures that satisfy the no-arbitrage condition. We outline in Sect. 3.2 how this could be done numerically, but before that, in Sect. 3.1, we study existence and some properties of solutions for the resulting optimization problem. In Sect. 3.3 we finally consider some cases where existence fails.

3.1 Existence of optimal financial products

In this section, we mainly deal for simplicity with rational investors in the sense of [28], i.e., we assume that the utility U can be expressed by

$$U(T) = \int_0^\infty u(x)\, d(\mathrm{pr}_2 T)(x) = \int_0^\infty u(x)\, dp(x),$$

where x is the final wealth of the investor who invests in a product with the return p, and u is a von Neumann–Morgenstern utility function. We assume that u is continuous and increasing.

From now on, we implicitly assume that, for given p and π, T is the joint probability measure with marginals p and $\tilde{\pi}$ which is co-monotone with $\tilde{\pi}$. Thus we can define the *maximum covariance* between p and $\tilde{\pi}$ as

$$\mathrm{mcov}(p, \tilde{\pi}) := \mathrm{cov}\, T.$$

Our optimization problem can now be stated as finding $p \in \mathcal{P}$ that maximizes

$$U(p) := \int_0^\infty u(x)\, dp(x)$$

subject to

$$\mathbb{E}(p) - R = \frac{\mathrm{mcov}(p, \tilde{\pi})}{\mathrm{var}(\tilde{\pi})} \big(\mathbb{E}(\tilde{\pi}) - R\big). \tag{3.1}$$

We formulate the following existence result.

Theorem 3.1 (Existence of optimal financial products) *Let the preferences of the investor be given by expected utility theory with utility function u. Assume that u is continuous, increasing, and of sublinear growth, i.e., that $u(x)/x \to 0$ as $x \to \infty$. Assume furthermore that the state price density $\tilde{\pi}$ vanishes outside an interval $[0, M]$ and that $\mathrm{var}(\tilde{\pi}) > M(\mathbb{E}(\tilde{\pi}) - R)$. Then there exists an optimal financial product, i.e., the above problem admits a maximizer $p \in \mathcal{P}$. Moreover, $\mathbb{E}(p) < \infty$.*

We shall see in Sect. 3.3 that some of the conditions made in this theorem are not purely technical but indeed necessary.

Proof of Theorem 3.1 The proof consists of the following steps:

1. There is a constant $C > 0$ such that all $p \in \mathcal{P}$ that satisfy (3.1) have a finite expected value $\mathbb{E}(p) \leq C$.
2. For every sequence (p_n) satisfying (3.1), there exists a $p \in \mathcal{P}$ such that a subsequence $(p_{n'})$ of (p_n) converges weakly-$\star$ to p, i.e., $p_{n'} \overset{\star}{\rightharpoonup} p$.
3. The supremum of $U(p)$ over all $p \in \mathcal{P}$ that satisfy (3.1) is finite.
4. The maximization problem admits a maximizer.

Step 1: First, we use the approximation result in Lemma 2.9 and assume therefore without loss of generality that all joint probability measures can be expressed by functions. We denote the function corresponding to a return p by R_p. We observe the useful identity

$$\begin{aligned}\operatorname{mcov}(p,\tilde{\pi}) &= \int_0^\infty \big(x-\mathbb{E}(\tilde{\pi})\big)\big(R_p(x)-\mathbb{E}(p)\big)\,d\pi(x)\\ &= \int_0^\infty xR_p(x)\,d\pi(x)-\mathbb{E}(\tilde{\pi}).\end{aligned}$$

To simplify notation, we write $\sigma^2 := \operatorname{var}(\tilde{\pi})$. Let $p \in \mathcal{P}$ and assume that p satisfies the no-arbitrage condition (3.1). To obtain an estimate on $\mathbb{E}(p)$ using the estimate on the support of $\tilde{\pi}$, we write

$$\begin{aligned}\mathbb{E}(p) &= R + \frac{\operatorname{mcov}(p,\tilde{\pi})}{\sigma^2}\big(\mathbb{E}(\tilde{\pi})-R\big)\\ &= R + \frac{\int_0^\infty xR_p(x)\,d\pi(x)-\mathbb{E}(\tilde{\pi})}{\sigma^2}\big(\mathbb{E}(\tilde{\pi})-R\big)\\ &\le R + M\mathbb{E}(p)-\mathbb{E}(\tilde{\pi})\sigma^2\big(\mathbb{E}(\tilde{\pi})-R\big).\end{aligned}$$

Resolving this while using the assumption $\sigma^2 = \operatorname{var}(\pi) > M(\mathbb{E}(\tilde{\pi})-R)$, we get

$$\mathbb{E}(p) \le \frac{R-\frac{\mathbb{E}(\tilde{\pi})}{\sigma^2}(\mathbb{E}(\tilde{\pi})-R)}{1-\frac{M}{\sigma^2}} < \infty, \tag{3.2}$$

thus arriving at the desired uniform bound for p.

Step 2: Let (p_n) be a sequence of probability measures satisfying (3.1). We want to prove that we can select a subsequence $(p_{n'})$ which is converging weakly-$\star$ to a probability measure p. By Prokhorov's theorem, it is sufficient to prove that (p_n) is tight, i.e., that for all $\eta > 0$, there is a compact subset K_η of $\mathbb{R}_+$ such that $p_n(K_\eta) > 1-\eta$. Suppose that (p_n) is not tight. Then for all $L > 0$, there exists an $n_0(L) \in \mathbb{N}$ such that $p_{n_0(L)}((L,+\infty)) > \eta_0$. Our strategy is now to estimate $\operatorname{mcov}(p_n,\tilde{\pi})$ from below to show that under this assumption, it must diverge. This then implies via the no-arbitrage condition (3.1) that also $\mathbb{E}(p_n)$ diverges, in contradiction to the uniform bound we have derived in Step 1.

So define M_1,M_2 such that $(x-\mathbb{E}(\tilde{\pi})(R_{p_n}(x)-\mathbb{E}(p_n)) =: h(x)$ is positive on $(0,M_1)$ and (M_2,M) and negative on (M_1,M_2). This is possible since $x-\mathbb{E}(\tilde{\pi})$ and $R_{p_n}(x)-\mathbb{E}(p_n)$ are nondecreasing functions with sign changes in $\mathbb{E}(\tilde{\pi})$ and $r := R_{p_n}^{-1}(\mathbb{E}(p_n))$, respectively. Then we write

$$\operatorname{mcov}(p_n,\tilde{\pi}) = \int_0^M \big(x-\mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x)-\mathbb{E}(p_n)\big)\,d\pi(x)$$

$$
\begin{aligned}
&= \int_0^{M_1} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&\quad + \int_{M_1}^{M_2} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&\quad + \int_{M_2}^{M} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&=: I_0 + I_1 + I_2.
\end{aligned}
$$

We have to distinguish two cases, depending on whether or not $r \geq \mathbb{E}(\tilde{\pi})$.

Case A: $r \geq \mathbb{E}(\tilde{\pi})$, i.e., $M_1 = \mathbb{E}(\tilde{\pi})$, $M_2 = r$. In this case, we have

$$
\begin{aligned}
I_0 + I_1 \geq& \int_0^{M_1} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}\big(\mathbb{E}(\tilde{\pi})\big) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&+ \int_{M_1}^{M_2} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}\big(\mathbb{E}(\tilde{\pi})\big) - \mathbb{E}(p_n)\big)\, d\pi(x) \geq 0.
\end{aligned}
$$

Now we use the assumption that (p_n) is not tight. It implies that, for all $L > R_{p_n}(r)$ and $n \geq n_0(L)$,

$$
\begin{aligned}
I_2 &= \int_{M_2}^{M} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&\geq \int_{R_{p_n}^{-1}(L)}^{M} \big(R_{p_n}(L) - \mathbb{E}(\tilde{\pi})\big)\big(L - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&\geq \int_{R_{p_n}^{-1}(L)}^{M} \big(R_{p_{n_0(L)}}^{-1}(L) - \mathbb{E}(\tilde{\pi})\big)\big(L - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&= p_{n_0(L)}\big((L, +\infty)\big)\big(R_{p_{n_0(L)}}^{-1}(L) - \mathbb{E}(\tilde{\pi})\big)\big(L - \mathbb{E}(p_n)\big) \\
&> \text{const.}\big(L - \mathbb{E}(p_n)\big) \to +\infty \quad \text{as } L \to \infty.
\end{aligned}
$$

Taking both estimates together, we have proved in this case that $\text{mcov}(p_n, m) \to +\infty$ as $n \to \infty$.

Case B: $r < \mathbb{E}(\tilde{\pi})$, i.e., $M_1 = r$, $M_2 = \mathbb{E}(\tilde{\pi})$. We decompose the maximum covariance analogously to case A and estimate

$$
\begin{aligned}
I_0 + I_1 =& \int_0^{r} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&+ \int_{r}^{\mathbb{E}(\tilde{\pi})} \big(x - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x)
\end{aligned}
$$

$$
\begin{aligned}
&\geq \int_0^r \big(r - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&\qquad + \int_r^{\mathbb{E}(\tilde{\pi})} \big(r - \mathbb{E}(\tilde{\pi})\big)\big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\, d\pi(x) \\
&= \big(r - \mathbb{E}(\tilde{\pi})\big)\bigg(\int_0^{\mathbb{E}(\tilde{\pi})} R_{p_n}(x)\, d\pi(x) - \mathbb{E}(p_n)\int_0^{\mathbb{E}(\tilde{\pi})} d\pi(x)\bigg),
\end{aligned}
$$

which is positive since R_{p_n} is nondecreasing, $\mathbb{E}(p_n) = \int_0^M R_{p_n}(x)\, d\pi(x)$, and hence $\int_0^{\mathbb{E}(\tilde{\pi})} R_{p_n}(x)\, d\pi(x) \leq \mathbb{E}(p_n)\int_0^{\mathbb{E}(\tilde{\pi})} d\pi(x)$. For I_2, we can now use essentially the same estimate as in Step A, which proves that also in this case $\operatorname{mcov}(p_n, \tilde{\pi}) \to \infty$ as $n \to \infty$.

From the no-arbitrage condition (3.1) we immediately see that because $\operatorname{mcov}(p_n, \tilde{\pi}) \to \infty$, also $\mathbb{E}(p_n) \to \infty$ as $n \to \infty$, which contradicts the uniform bound of $\mathbb{E}(p_n)$ in (3.2). Therefore (p_n) is tight, and we can apply Prokhorov's theorem to obtain the existence of a weak-$\star$ limit $p \in \mathcal{P}$ for a subsequence of (p_n).

Step 3: We denote the concave envelope of the utility function u by u^c. For every $p \in \mathcal{P}$ which satisfies (3.1) and therefore also (3.2), we estimate with the help of Jensen's inequality that

$$
U(p) \leq \int_0^\infty u^c(x)\, dp(x) \leq u^c\big(\mathbb{E}(p)\big) \leq u^c\left(\frac{R - \frac{\mathbb{E}(\tilde{\pi})}{\sigma^2}(\mathbb{E}(\tilde{\pi}) - R)}{1 - \frac{M}{\sigma^2}}\right) < +\infty.
$$

Due to this uniform bound, we can find a maximizing sequence (p_n) of probability measures satisfying (3.1), and using the results of Step 2, we can extract a subsequence that converges weakly-$\star$ to a limit $p \in \mathcal{P}$. It remains to prove that this limit is indeed a solution of our maximization problem. In the remaining part of the proof, we write, for simplicity, (p_n) for the subsequence $(p_{n'})$ of (p_n).

Step 4: Here we use the sublinear growth of u and estimate, for any $L > 0$, that

$$
\begin{aligned}
\left|\int_0^\infty u(x)\, dp - \int_0^\infty u(x)\, dp_n\right| &\leq \left|\int_0^L u(x)\, dp - \int_0^L u(x)\, dp_n\right| \\
&\quad + \left|\int_L^\infty \frac{u(x)}{x} x\, dp - \int_L^\infty \frac{u(x)}{x} x\, dp_n\right|.
\end{aligned}
$$

While the first term converges to zero as $n \to \infty$ since $p_n \overset{\star}{\rightharpoonup} p$, the second can be estimated by

$$
\begin{aligned}
\left|\int_L^\infty \frac{u(x)}{x} x\, dp - \int_L^\infty \frac{u(x)}{x} x\, dp_n\right| &\leq \frac{u(L)}{L}\int_L^\infty x\, d(p - p_n) \\
&\leq \frac{u(L)}{L}\big|\mathbb{E}(p) - \mathbb{E}(p_n)\big|.
\end{aligned}
$$

Using again estimate (3.2), the last term is bounded in n. If we consider the limit $L \to \infty$, this expression becomes arbitrarily small; therefore, using an appropriate sequence of $L = L(n)$, we have proved that $U(p_n) \to U(p)$.

It only remains to prove that p satisfies the no-arbitrage condition (3.1). We first show that

$$\mathbb{E}(p) \le R + \frac{\mathrm{mcov}(\tilde{\pi}, m)}{\sigma^2}\big(\mathbb{E}(\tilde{\pi}) - R\big). \tag{3.3}$$

To this end, we use the no-arbitrage condition for p_n and (3.2) and estimate, for any sufficiently small $\varepsilon > 0$, that

$$\begin{aligned}
\mathbb{E}(p) &= \mathbb{E}(p) - \mathbb{E}(p_n) + \mathbb{E}(p_n) \\
&= \int_0^M \big(R_p(x) - R_{p_n}(x)\big)\, d\pi + R \\
&\quad + \frac{\mathbb{E}(\tilde{\pi}) - R}{\sigma^2}\bigg(\int_0^{M-\varepsilon} \big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi \\
&\quad + \int_{M-\varepsilon}^M \big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi\bigg) \\
&\le \int_0^{M-\varepsilon} \big(R_p(x) - R_{p_n}(x)\big)\, d\pi + R \\
&\quad + \frac{\mathbb{E}(\tilde{\pi}) - R}{\sigma^2} \int_0^{M-\varepsilon} \big(R_{p_n}(x) - \mathbb{E}(p_n)\big)\big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi \\
&\quad + \int_{M-\varepsilon}^M \big(R_p(x) - R_{p_n}(x)\big)\, d\pi + \int_{M-\varepsilon}^M \big(R_{p_n}(x) - \mathbb{E}(p_n) x\big)\, d\pi.
\end{aligned}$$

Since $\mathbb{E}(p_n)$ is uniformly bounded, it converges, at least for a subsequence, to some constant K. We can therefore pass to the limit as $n \to \infty$ and obtain

$$\begin{aligned}
\mathbb{E}(p) \le \int_{M-\varepsilon}^M R_p(x)\, d\pi + R + \frac{\mathbb{E}(\tilde{\pi}) - R}{\sigma^2}\bigg(\int_0^{M-\varepsilon} \big(R_p(x)\big(x - \mathbb{E}(\tilde{\pi})\big)\big)\, d\pi \\
- K \int_0^{M-\varepsilon} \big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi\bigg) + K \int_{M-\varepsilon}^M d\pi.
\end{aligned}$$

We apply the identity

$$\int_{M-\varepsilon}^M \big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi = -\int_0^{M-\varepsilon} \big(x - \mathbb{E}(\tilde{\pi})\big)\, d\pi$$

to derive

$$\mathbb{E}(p) \leq \int_{M-\varepsilon}^{M} R_p(x)\,d\pi + R + \frac{\mathbb{E}(\tilde{\pi}) - R}{\sigma^2} \int_0^{M-\varepsilon} \big(R_p(x)\big(x - \mathbb{E}(\tilde{\pi})\big)\big)\,d\pi$$
$$+ K \frac{\mathbb{E}(\tilde{\pi}) - R}{\sigma^2} \int_{M-\varepsilon}^{M} \big(x - \mathbb{E}(\tilde{\pi})\big)\,d\pi + K \int_{M-\varepsilon}^{M} d\pi.$$

This inequality holds for all sufficiently small $\varepsilon > 0$. We can therefore pass to the limit and obtain (3.3).

Now let us suppose that this inequality were *strict*. Then we could improve p by adding a certain outcome for sure while at the same time satisfying the no-arbitrage condition (3.1) exactly. (We have seen in the proof of Theorem 2.12 how to do this.) This improved product p' would by assumption have a larger utility than p, i.e., $U(p') > U(p)$; but $U(p_n) \to U(p)$ as $n \to \infty$ and (p_n) was defined as a maximizing sequence for U. Therefore inequality (3.3) must in fact be an equality, and (3.1) holds for p.

Thus p is indeed a solution of our maximization problem, and we have proved the existence result. □

It is also possible to characterize some features of solutions. In particular, we have the following result.

Proposition 3.2 *If $u'(x) \to 0$ as $x \to \infty$ and $\tilde{\pi}$ satisfies the assumptions from Theorem 3.1, then an optimal financial product has a finite maximum return, i.e.,* supp p *is bounded.*

Proof Assuming that there exists an optimal $p \in \mathcal{P}$ with unbounded support and which can be represented by a function R_p, we define, for $\varepsilon > 0$,

$$R_{p'}(x) := \begin{cases} R_p(x) + \delta & \text{for } x \in (0, M - \varepsilon), \\ R_p(x) - \varepsilon & \text{for } x \in (M - \varepsilon, M), \end{cases}$$

where δ is chosen such that p' satisfies the no-arbitrage condition (3.1). We define $\kappa := \int_{M-\varepsilon}^{M} d\pi(x)$. A lengthy but straightforward computation shows then that $\delta = O(\varepsilon\kappa)$ and that

$$U(p') - U(p) \geq u'\big(\mathbb{E}(\tilde{\pi})\big)(1 - \kappa)\delta - u'\big(R_p(M - \varepsilon)\big)\kappa\varepsilon + O\big(\varepsilon^2\kappa\big).$$

This is positive if $\varepsilon > 0$ is chosen small enough, since in this case $R_p(M - \varepsilon) \to +\infty$ and therefore $u'(R_p(M - \varepsilon)) \to 0$. Thus p cannot be optimal. □

What kind of problems may arise for the existence result if one replaces expected utility theory by prospect theory? One main difference is the probability weighting that will overweight extreme small probability events. This might compensate for the diminishing marginal utility of large outcomes. It is, e.g., possible to construct a $p \in \mathcal{P}$ with finite expected value but infinite PT-utility. This phenomenon is essentially a new variant of the St. Petersburg paradox; compare [26] for details.

3.2 Numerical approximation

In order to compute optimal financial products numerically, the existence proof of the previous section can give some rough guidance: The main idea is to optimize in p rather than in T and to compute T for every given p and $\tilde{\pi}$ as the unique co-monotone joint probability measure with marginals p and $\tilde{\pi}$. This approach is much more efficient than an optimization in T, since the number of necessary variables for an approximation is much smaller, as we shall see in a moment.

We formulate this method for the case of finitely many states, since in a numerical approximation T would be replaced by a matrix, and p and $\tilde{\pi}$ would be approximated by vectors. This corresponds mathematically to approximating p and $\tilde{\pi}$ by sums of weighted Dirac measures. Let $x_1, \dots, x_N$ be the set of possible outcomes, where $x_1 < x_2 < \cdots < x_N$. We want to find the optimal vector $(p_1, \dots, p_N)$ of probabilities for these outcomes, where $p_i \geq 0$, such that

(i) The total probability is one: $p_1 + \cdots + p_N = 1$.
(ii) The probability measure $p = \sum_{i=1}^N p_i \delta_{x_i}$ maximizes (among all probability measures of this form) a given utility $U(p)$ subject to the constraint implied by (3.2).

It is now clear why our approach is more efficient than a direct optimization of T: If we approximate p and $\tilde{\pi}$ each by N weighted Dirac measures, then T is an $N \times N$-matrix. A direct optimization of T would therefore be an optimization in N^2 rather than in N variables.

However, we also have to compute the co-monotone T (or at least its covariance), given its marginals p and $\tilde{\pi}$, in an efficient way to make this idea work. Such an algorithm could be obtained from the construction of Theorem 2.7 by starting from an arbitrary joint probability measure with given marginals. It is, however, possible to compute the covariance of the co-monotone joint probability measure directly and at the same time more efficiently by applying a simple algorithm used by [25] in the context of transport plans. For that, set $i = j = 1$, $L = \tilde{\pi}_1$, and $C = 0$. Then, as long as $i \leq n$ or $j \leq N$, do the following:

– If $L > p_j$, then set $L = L - p_j$, $C = C + p_j(x_i - \mathbb{E}(p))(x_j - \mathbb{E}(\tilde{\pi}))$.
– If $L \leq p_j$, then set $L = 0$, $C = C + L(x_i - \mathbb{E}(p))(x_j - \mathbb{E}(\tilde{\pi}))$.
– If $L = 0$, then set $i = i + 1$, $L = \tilde{\pi}_i$, otherwise set $j = j + 1$.

This algorithm terminates since $\sum_{i=1}^N \tilde{\pi}_i = 1 = \sum_{j=1}^N p_j$. The variable C returns the maximum covariance of p and $\tilde{\pi}$, i.e., $C = \text{mcov}(p, \tilde{\pi})$.

Using this algorithm, the constraint (3.1) can be realized without explicitly knowing the joint probability measure T. The resulting finite constrained maximization problem can be solved with standard algorithms for nonconcave maximization. This rough sketch of ideas should be sufficient to demonstrate the possibility of solving this optimization problem also in a practical application, but naturally there are still interesting open questions, e.g., regarding the quality of convergence of this approximation.

3.3 Potential nonexistence

In this section we come briefly back to the existence theorem of Sect. 3.1 and sketch an example that demonstrates how one could get nonexistence in certain situations.

The source of potential problems is most easily seen when relaxing the constraint of positive final wealth (i.e., optimizing in $\mathcal{P}(\mathbb{R})$ rather than in $\mathcal{P}(\mathbb{R}_+)$). For simplicity, we set

$$\tilde{\pi} := \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_2, \qquad R = 0.$$

In this case, we can construct a sequence of probability measures

$$p_n := \frac{1}{2}\delta_{-n+c(n)} + \frac{1}{2}\delta_{n+c(n)},$$

where $c(n)$ is chosen such that p_n satisfies the no-arbitrage condition (3.1). Based on this condition and using that $\mathbb{E}(p_n) = c(n)$, $\mathbb{E}(\tilde{\pi}) = 1/2$, $\mathrm{var}(\tilde{\pi}) = 3/2$, and that $R = 0$, we compute

$$\begin{aligned} c(n) &= \frac{\mathrm{mcov}(p_n, \tilde{\pi})}{\mathrm{var}(\tilde{\pi})}\mathbb{E}(\tilde{\pi}) \\ &= \frac{1}{3}\Big(\big(-n + c(n) - c(n)\big)(-3/2) + \big(n + c(n) - c(n)\big)(3/2)\Big) \\ &= n. \end{aligned}$$

Therefore, if u is unbounded, $U(p_n) \to +\infty$ as $n \to \infty$, and we have obviously nonexistence. But even if u is bounded, but strictly increasing, $U(p_n)$ is strictly increasing as $n \to \infty$, but (p_n) converges weakly-$\star$ to $p := \frac{1}{2}\delta_0$; thus $U(p) < \lim_{n\to\infty} U(p_n)$. Hence also in this case, an existence proof is not possible.

We see that the fundamental problem is that $\mathrm{mcov}(p_n, \tilde{\pi})$ and $\mathbb{E}(p_n)$ simultaneously tend to infinity. This problematic phenomenon that can, as we have just seen, lead to nonexistence can only be excluded under additional conditions like those we introduced in our existence result. It is an interesting question for future work to what extent these conditions can be relaxed.

4 Applications to bonus certificates

In this section we want to have a look at a specific type of structured products with co-monotone payoffs, *bonus certificates*. In their simplest form, they are an investment in a certain underlying that guarantees capital protection as long as the price of the underlying does not fall below a certain threshold, the "barrier level." Once the underlying is below the threshold, the capital protection is gone and is also not recovered by future increases above the barrier level. The payoff diagram of such a product at maturity is schematically illustrated in Fig. 1. Such bonus certificates are nowadays routinely sold by banks (in particular in Europe) to private clients who want to invest on the stock market but, at the same time, want to minimize their exposure to potential losses. Such private investors typically do not have background risk to be hedged by this product, nor could one convincingly argue that they have an information advantage that would allow them to profit from speculation. Thus our

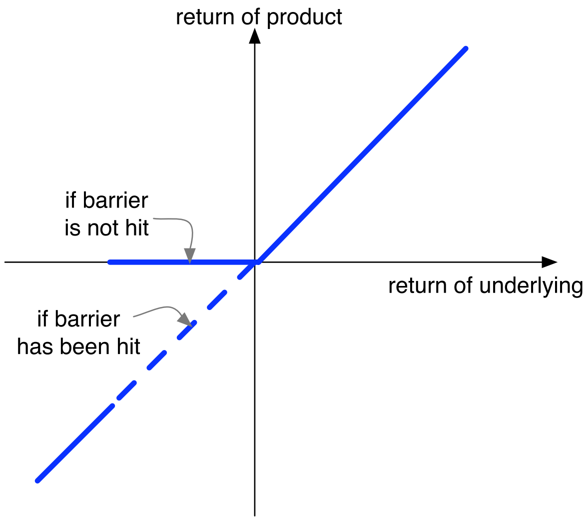

Fig. 1 Payoff diagram of a bonus certificate

two assumptions are most likely satisfied: The clients have no background risk, and they have homogeneous beliefs.

The joint probability distribution of these products, however, is obviously not co-monotone, since its support contains the whole diagonal and also parts of the x-axis of the above diagram. Therefore, independently of the stochastic process of the underlying, the product is not optimal, neither regarding the zero reference point nor the reference point of the underlying (compare Propositions 2.19 and 2.21 or Corollary 2.8). In fact, we can compute analytically the "optimized" variant of this product that has the same return distribution but lower hedging costs.

Proposition 4.1 (Monotonized bonus certificate) *Let $T \in \mathcal{P}(\mathbb{R}^2)$ be the joint probability measure of the market return $m \in \mathcal{P}$ and the return $p \in \mathcal{P}$ of the bonus certificate at maturity. Assume for simplicity that m is absolutely continuous. Let B be the barrier level. Let R_p be defined by*

$$R_p(x) := \begin{cases} 0 & \textit{for } x \in [x_0, 0], \\ P^{-1}(M(x)) & \textit{for } x \in [B, x_0), \\ x & \textit{otherwise}, \end{cases}$$

where we set $x_0 := \sup\{x \in \mathbb{R} \,|\, m([x_0, 0]) \geq T([B, 0) \times \{0\})\}$, $P(x) := p((-\infty, x])$, and $M(x) := m((-\infty, x])$. Then the alternative product $T_{\mathrm{opt}} := (\mathrm{Id} \times R_p)_{\#}m$ that yields the payoff $R_p(x)$ when the market return is x generates the same return p, i.e., $\mathrm{pr}_2(T_{\mathrm{opt}}) = p$, but has a lower price.

Proof We need to verify that $\mathrm{pr}_2(T_{\mathrm{opt}}) = p$. It is clear that

$$\big(\mathrm{pr}_2(T_{\mathrm{opt}})\big)\big(\{0\}\big) = m\big([x_0, 0]\big) = T\big([B, 0] \times \{0\}\big) = p\big(\{0\}\big).$$

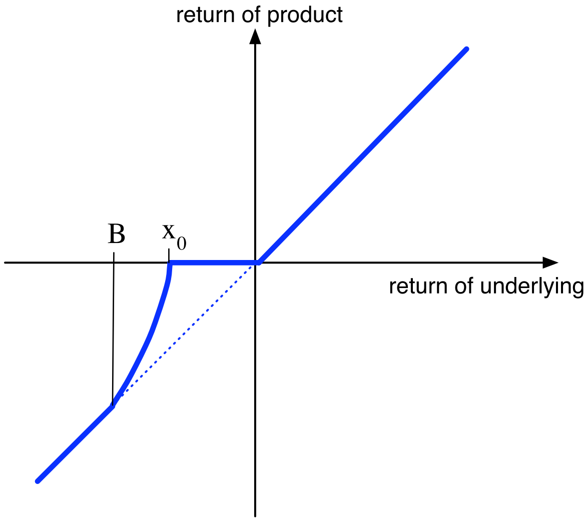

Fig. 2 Payoff diagram of the optimized product, yielding the same return distribution for a lower price

Hence, we only need to check the condition for $y \in (B, 0)$. There, we have

$$\big(\mathrm{pr}_2(T_{\mathrm{opt}})\big)(y) = \frac{m(R_p^{-1}(y))}{R_p'(R_p^{-1}(y))}.$$

We compute

$$R_p'(x) = \frac{d}{dx}\big(P^{-1}\big(M(x)\big)\big) = \big(P^{-1}\big)'\big(M(x)\big)m(x),$$

and thus

$$\big(\mathrm{pr}_2(T_{\mathrm{opt}})\big)(y) = \frac{m(R_p^{-1}(y))}{(P^{-1})'(M(R_p^{-1}(y)))\; m(R_p^{-1}(y))}.$$

Now since $(P^{-1})' = 1/P'(P^{-1})$ and $P^{-1}(M(x)) = R_p(x)$, this simplifies to

$$\big(\mathrm{pr}_2(T_{\mathrm{opt}})\big)(y) = P'\big(P^{-1}\big(M\big(R_p^{-1}(y)\big)\big)\big) = p(y).$$

Since T_{opt} is co-monotone by construction (R_p is nondecreasing) and T is not co-monotone (as we have seen), but $U(T_{\mathrm{opt}}) = U(p) = U(T)$, we can apply Proposition 2.19 to see that T cannot be optimal. □

The optimized product (see Fig. 2) yields precisely the same return distribution for the customer but will be cheaper to hedge for the bank. (There is no arbitrage opportunity here, since the optimized product will yield a lower return in *some* states!) In fact, this analysis shows that the success of products involving barriers cannot be explained without considering behavioral biases as has been pointed out in [13].

5 Conclusions

We have seen that every optimal financial product is co-monotone with the reversed state price density. This holds regardless of the preferences under consideration, as long as the investor only considers absolute returns of his investment *or* if he measures performance relative to an index which is a nondecreasing function of the reversed state price density.

In the special case of a market that can (at least to some extent) be described by the Capital Asset Pricing Model or the Black–Scholes model, this implies that optimal portfolios for an investor who is only interested in the *absolute returns* of his portfolio "follow the market," i.e., their return is the better, the better the return of the market portfolio. This monotonicity also holds for an investor who is only interested in *relative performance* with respect to the market return. Again, in both instances the underlying decision model can be chosen arbitrarily (e.g., expected utility theory, mean–variance or prospect theory). Moreover, we have shown that for an investor whose utility solely depends on the *relative return* of his investment, only financial products that are at least as risky as the market portfolio can be optimal. In the case of a smooth, concave von Neumann–Morgenstern utility investor, an optimal investment even has to be riskier than the market portfolio.

The fundamental assumptions of these results were homogeneous beliefs and no background risk. Moreover we assumed complete and arbitrage-free markets. The assumptions on the investor's preferences were, however, very weak. In this way, we have extended previous work by Dybvig [9] in several directions: allowing for arbitrary state spaces (as Föllmer and Schied and also Jouini and Porte did already for (strictly) concave utility functions in [11, 17]), general probability measures (including nonabsolutely continuous ones), general preferences, and even preferences that are depending on the relative performance with respect to a benchmark (e.g., the state price density or the market return). This shows in particular that optimal products can be described (essentially) by a simple increasing payoff function which makes it easier to study optimal structured products, see [13].

An extension of this co-monotonicity result to several underlyings (or factor models, like APT) is possible. The optimal product is then co-monotone with a weighted sum of the underlyings (or factors).

We studied a new method for the construction of optimal financial products, based on the idea of finding the optimal return distribution among all probability measures satisfying the no-arbitrage condition. This approach makes it necessary to study existence of optimal financial products. We proved an existence result using ideas from the calculus of variations and outlined a numerical algorithm for obtaining optimal financial products based on the investor's preferences. Some remarks on situations where existence fails underline the role of some of the assumptions in the existence theorem.

Another application was in the context of bonus certificates. Under the simplifying assumption of a CAPM market or the Black–Scholes model, we could show that bonus certificates cannot be optimal and can be optimized analytically to reduce hedging costs.

References

1. Ambrosio, L.: Lecture notes on optimaltransport problems. In: Colli, P., Rodrigues, C. (eds.) CIME Summer School in Madeira. Lecture Notes in Mathematics, vol. 1813, pp. 1–52. Springer, Berlin (2003)
2. Bizid, A., Jouini, E., Koehl, P.-F.: Pricing of non-redundant derivatives in a complete market. Rev. Deriv. Res. **2**, 287–314 (1998)
3. Cambanis, S., Simons, G., Strout, W.: Inequalities for $Ek(X,Y)$ when the marginals are fixed. Z. Wahrscheinlichkeitstheor. Verw. Geb. **36**, 285–294 (1976)
4. Carlier, G., Dana, R.: Core of convex distortions of a probability. J. Econ. Theory **113**, 199–222 (2003)
5. Carlier, G., Dana, R.-A.: Rearrangement inequalities in non-convex insurance models. J. Math. Econ. **41**, 483–503 (2005)
6. Copeland, T.E., Weston, J.F.: Financial Theory and Corporate Policy. Addison-Wesley, Reading (1979)
7. Dhaene, J., Denuit, M., Goovaerts, M., Kaas, R., Vyncke, D.: The concept of comonotonicity in actuarial science and finance: theory. Insur. Math. Econ. **31**, 3–33 (2002)
8. Duffie, D.: Dynamic Asset Pricing Theory. Princeton University Press, Princeton (2001)
9. Dybvig, P.H.: Distributional analysis of portfolio choice. J. Bus. **61**, 369–393 (1988)
10. Dybvig, P.H.: Inefficient dynamic portfolio strategies or how to throw away a million dollars in the stock market. Rev. Financ. Stud. **1**, 67–88 (1988)
11. Föllmer, H., Schied, A.: Stochastic Finance, 2nd edn. De Gruyter, Berlin (2004)
12. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge University Press, Cambridge (1952)
13. Hens, T., Rieger, M.O.: The dark side of the moon—Structured products from the customer's perspective (February 10, 2009). Available at SSRN: http://ssrn.com/abstract=1342360 (2009)
14. Jouini, E., Kallal, H.: Efficient trading strategies in the presence of market frictions. Rev. Financ. Stud. **14**, 343–369 (2001)
15. Jouini, E., Napp, C.: Comonotonic processes. Insur. Math. Econ. **32**, 255–265 (2003)
16. Jouini, E., Napp, C.: Conditional comonotonicity. Bus. Econ. **27**(2), 153–166 (2004)
17. Jouini, E., Porte, V.: Efficient trading strategies (October 26, 2005). Available at SSRN: http://ssrn.com/abstract=1000208 (2005)
18. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica **47**, 263–291 (1979)
19. Kantorovich, L.V.: On the transfer of masses. Dokl. Akad. Nauk. SSSR **37**, 227–229 (1942)
20. Landsberger, M., Meilijson, I.: Co-monotone allocations, Bickel–Lehmann dispersion and the Arrow–Pratt measure of risk aversion. Ann. Oper. Res. **52**(2), 97–106 (1994)
21. LeRoy, S.F., Werner, J.: Principles of Financial Economics. Cambridge University Press, Cambridge (2001)
22. Machina, M.J.: "Expected utility" analysis without the independence axiom. Econometrica **50**, 277–323 (1982)
23. Markowitz, H.M.: Portfolio selection. J. Finance **7**, 77–91 (1952)
24. Monge, G.: Géométrie descriptive. Éditions Jacques Gabay, Sceaux (1989) (reprint of the 1799 original)
25. Rieger, M.O.: Monotonicity of transport plans and applications. CVGMT preprint, Pisa, http://cvgmt.sns.it/cgi/get.cgi/papers/rie06a/ (2006)
26. Rieger, M.O., Wang, M.: Cumulative prospect theory and the St. Petersburg paradox. Econ. Theory **28**, 665–679 (2006)
27. Ross, S.: The arbitrage theory of capital asset pricing. J. Econ. Theory **13**, 341–360 (1976)
28. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)