Reward Schemes

David Lagziel (Tel Aviv University)

---

16 May 2016, 11:15-12:15, bld. 72, room 465

Abstract: An investor has some funds invested through investment firms. By the end of the year, she uses the net profits for her personal needs, and reallocates the funds among these firms according to the firms' performance. While the investor tries to maximize her total expected earnings (given some subjective utility function), each investment firm tries to maximize the overall amount of funds it will be allocated to manage. A reward scheme is a rule that determines how funds should be allocated among the investment firms based on their performance. A reward scheme is optimal if it induces the (self-interested) firms to act in accordance with the interests of the investor. We show that an optimal reward scheme exists under quite general conditions.

Reward Schemes*

David Lagziel† and Ehud Lehrer‡

March 13, 2016

Abstract:

An investor has some funds invested through investment firms. By the end of the year, she uses the net profits for her personal needs, and reallocates the funds among these firms according to the firms' performance. While the investor tries to maximize her total expected earnings (given some subjective utility function), each investment firm tries to maximize the overall amount of funds it will be allocated to manage. A *reward scheme* is a rule that determines how funds should be allocated among the investment firms based on their performance. A reward scheme is *optimal* if it induces the (self-interested) firms to act in accordance with the interests of the investor. We show that an optimal reward scheme exists under quite general conditions.

Journal of Economic Literature classification numbers: C72, D47, G11, G14, G23, G24.

Keywords: Reward Schemes; Optimal Reward Schemes; Investment firms; Investment game; Market design; Portfolio management.

---

*The authors wish to thank Eddie Dekel, Ehud Kalai, Rani Spiegler, Yaron Azrieli, and Jörgen Weibull for their valuable comments. Lehrer acknowledges the support of the Israel Science Foundation, Grant #963/15. Lagziel acknowledges the support of the Israel Science Foundation, Grant #538/11.

†Tel Aviv University, Tel Aviv 69978, Israel. e-mail: davidlag@post.tau.ac.il.

‡Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France. e-mail: lehrer@post.tau.ac.il.

1 Introduction

There is a whole range of papers[1] offer a variety of theories and models in an attempt to explain the 2007–2009 financial crisis. This paper highlights and focuses on one key aspect of that sort: the gap between the objectives of investors on one hand, and the incentives of investment funds and brokers on the other.

Consider an investor who has some funds already invested through investment firms. She wants to reallocate her funds among the firms according to their performance. While the goal of the investor is to maximize her total expected earnings, each investment firm tries to maximize the overall amount of funds bestowed in its hands to manage. The rule by which the investor reallocates her funds determines the environment in which the investment firms operate: it determines their incentives and ultimately their modus operandi. This rule is referred to as a *reward scheme*. Reward schemes are supposed to guarantee that financial managers will make every effort necessary to produce the optimal possible investments for their investors.

Even though these goals are rather clear, their effects are relatively vague. When the outcomes are stochastic and the schemes are based on past performance, the agents[2] (in order to serve their own interests) might take unnecessary risks, from the investors' perspectives. Therefore, the formulation of rewards schemes must guarantee that the agents will act according to the best interests of the investors.

In the current paper we address this issue through a simple set-up. We assume that a decision maker (DM) invests her funds through several investment firms. By the end of the year, she uses the net profits for her personal needs, and reallocates the funds according to a *reward scheme* that depends on the firms' performance.

This general set-up may also accommodate for other scenarios. One example is a situation where the DM wishes the invest additional funds based on the earnings the firms produced in the past year (see Section 2). Yet another example that applies to out model is a scenario where the DM is a manager who wishes to expand divisions in her corporation based on their annual return, and so on.

The core of the problem we address lies in the discrepancy between the motivations

[1] See, e.g., Fligstein and Goldstein (2010); Hansen (2009); and Simpson (2011). In addition, Akerlof and Romer (1993) gives some theoretical insight into the subject of opportunism in finicial institutions. For a general survey on the wrong incentives that led to the financial crisis of 2007–2009, see Fligstein and Roehrkasse (2013).

[2] We sometimes refer to the firms as agents, and to the investor as decision maker.

of the economic entities involved. While the firms wish to maximize the total expected funds they manage, the DM wishes to maximize the expected return of her investment.[3] Typically, these two motivations do not agree. By and large, the competition between firms pushes them to take riskier actions. To make things even worse, the DM cannot fully monitor the precise actions of the firms. She can typically observe only the quarterly, or annual, earnings reports. As a result, the investment firms may abuse this situation to increase their own expected payoffs at the expense of the DM's profit. Our objective in this paper is twofold: to establish a formal model for the analysis of this problem and to introduce constructive methods that will incentivize the investment firms to act in accordance with the goals of the DM.

The share of the funds a specific firm gets to manage depends not only on its own past performance, but on other firms' performance as well. The reward scheme introduced by the DM will actually induce a competition between the firms, or an *investment game*, as we call it. A reward scheme is said to be *optimal* if in every equilibrium (of the investment game) all firms act according to the best interest of the DM.

We first prove that for every market, i.e., for every set of possible actions of the investment firms, the DM can find an optimal reward scheme. This means that by properly designing the reward scheme, the DM can have the firms act in any equilibrium so as to maximize the DM's profits. The proof we provide is constructive and holds for a general number of firms and actions. More specifically, we present an optimal reward scheme that is linear in the sum of differences of the firms' earnings.

When the market changes frequently and the set of actions changes considerably, the DM might not be able to design, in advance, an optimal scheme. The question arises whether there is a *universal* optimal reward scheme that could cater to any set of possible actions. It turns out that here things are less optimistic. We show that one cannot devise a reward scheme which remains optimal for every set of actions. In other words, in order to be able to design an optimal reward scheme, the DM should either know what investment options are available for the firms (for investment), or at least the bound of their yields.

This paper differs from the related literature in several respects. In most previous

[3] We first consider a risk-neutral DM, i.e., an investor who wishes only to maximize her expected profits. Subsequently, we generalize our results to *any* expected utility maximizer DM. In other words, we extend our results to include risk aversion, risk loving, and so on.

works (see Subsection 1.1 ahead), the DM faces agents of various types and various abilities. While the DM cannot distinguish between agents of different types, their types do affect the DM's utility. For instance, different workers vary in their productivity rates, thus affecting their employers' profit. A leading question in the literature is whether the DM can design a reward scheme that rewards skilled agents, and screens out unskilled ones.

In our setting, in contrast, all agents are potentially of the same type. They are all experts, all exposed to the same data and, most importantly, all have the same set of possible actions.[4] This assumption distinguishes our work from most previous financial studies. In our set-up, using agents of the same type requires almost no assumptions over the information structure and over the probability space. However, the effects of private information and different trading abilities, that are the main focus in most previous studies, remain vague.

Nevertheless, the idea that all the information is available to the agents is consistent with studies showing that returns are, mostly in the short term and to some extent also in the long run, predictable.[5] For this reason, we assume that investment firms have value-enhancing trading abilities. From a clear theoretical point of view, this is not unprecedented, as can be seen in Treynor and Black (1973); Admati and Pfleiderer (1997); and Elton and Gruber (2004).

We show that the outcome, in a scenario of homogeneous agents, could still be unfavourable to the DM. Such a scenario is illustrated in the motivating example (Section 2). In this example, an investor has her funds managed by some investment firms. In order to induce a profit-maximizing competition between the firms, she considers reallocating the funds according to a winner-takes-all reward scheme. It turns out that in the unique equilibrium of the investment game induced, all firms invest in the portfolio that serves worst the interests of the DM.

In our context, even in the presence of skilled and unskilled fund managers, and even if the investor succeeds in screening out the unskilled ones, the skilled fund managers can still fail to deliver significant returns. This may happen due to the fact that their incentives are ill-matched with the objectives of the investor. In this paper we focus

[4]In contrast, e.g., to the set-up of Sharpe (1981); Berk and Green (2004); and Dybvig et al. (2010), among many others.

[5]See, e.g.,Dai and Singleton (2002); Lewellen (2004); Torous et al. (2004); Cochrane and Piazzesi (2005); Campbell and Yogo (2006); Ang and Bekaert (2007); and Lettau and Van Nieuwerburgh (2008).

on the issue of ensuring an optimal outcome for the DM, independently of the agents' skill level. In a broad sense, our main goal resembles the optimal auction design of Myerson (1981), whose goal is to provide an optimal mechanism that serves best the goals of the seller (and in our case, maximizes the DM's revenue).

Though we mainly use the terms 'investor' and 'investment firms', our model is not limited to this environment alone. The model we propose applies to many environments, and is especially relevant to the relations between shareholders and managers of financial institutions. These relations were, and still are, a central issue in the economic world. A key aspect in these relations is the disparity between the managers' self-interests and those of the company, and thus of the shareholders. A clear evidence for this tension could be found in the words of the 13$^{\text{th}}$ Federal Reserve chairman, Alan Greenspan: "I made a mistake in presuming that the self-interests of organizations, specifically banks and others, were such as that they were best capable of protecting their own shareholders and their equity in the firms".[6]

1.1 Related literature

Taking a broad perspective of the subject, our work lies between the economic literature and the financial one. On one hand, we are using basic game theory and mechanism design approach to tackle the problem of an optimal incentive scheme in a general principle-agent problem. On the other hand, we apply our results mainly to the problem of delegated portfolio management in financial markets. Both fields are well-studied, and therefore we will not be able to relate specifically to all previous works (See Stracca (2006) for a comprehensive survey of the field).

Numerous studies were conducted on the importance of performance-based payoffs and reputation in non-deterministic markets. These studies were conducted for a good reason: the need to sustain a high-level of reputation is the one objective nearly all agents have in common in almost any market, and specifically so in non-deterministic, performance-based markets.

This need is sometimes translated to a well-known phenomenon called "herding", where agents tend to mimic other agents, though they can perform better independently. Mimicking other competitors enables the agents to level their returns relative to the market, and therefore preserve their reputation. One can even track back this

[6]New York Times, "Greenspan 'shocked' that free markets are flawed", October 23, 2008.

concept to the words of Keynes (1936): "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally." (Book 4, Chapter 12, page 158).

This phenomenon is a key element in our work, and is best exemplified by the motivating example in Section 2. In our set-up, the agents can always choose an optimal portfolio for the investor. Yet they chose to balance their gains and losses relative to the other agents by investing (like their competitors) in a sub-optimal manner.

In this context, a reward scheme is a sophisticated ranking mechanism by which firms are ranked according to their past performance. Nonetheless, an important distinction should be made. Our result show that the reputation should not be limited to an ordinal ranking. In the optimal reward scheme we propose, the cardinality is crucial. Specifically, our optimal reward scheme states that every firm starts out with the same basic share. Any deviation from this basic share is proportional to the firm's excess return relative to the market's average.

One of the first papers to deal with the problem of herding is Scharfstein and Stein (1990). It shows that under certain circumstances agents simply mimic the behaviour of others, ignoring substantive information they posses. Specifically, when the market assesses the capabilities of an agent according to his performance, as well as his conventionality, agents may mimic each other, thus leading to a sub-optimal outcome.

Dasgupta and Prat (2006) continues this line of work. It studies the effect of fund-managers' career concerns on their performance. Given a specific model, Dasgupta and Prat show that without career concerns, only fund managers with special abilities will trade. However, they also prove that career concerns may lead to an equilibrium where unskilled fund managers have to trade in order to stand out, which may involve taking unnecessary risks. This work is generalized to a dynamic model in Dasgupta and Prat (2008). However, the latter focuses on the effect of career concerns on market prices and trading volume. In both cases the investors are risk-neutral return maximizers.

In general, the manipulability abilities of unskilled agents were proven to exist in more than a few papers, such as Lehrer (2001); Sandroni et al. (2003); Sandroni (2003); Shmaya (2008); and Olszewski and Sandroni (2008).

Foster and Young (2010) proves that it is almost impossible to generate a reward scheme in which skilled agents are rewarded while unskilled ones are eliminated from the market. This result is based on the assumption that the agents' strategies and tactics are not observable. Recently, He et al. (2015) showed that this result could be inverted when a liquidation boundary is set along with a requirement from the agents to deposit their own money to offset potential losses.

In the financial literature, Sharpe (1981) is one of the first studies that deals with the issue of using multiple asset managers, and the need to provide good incentives (or otherwise coordinate between different managers). This work was later followed by Barry and Starks (1984) and more recently by Van Binsbergen et al. (2008). In general, these papers focus on the impact of decentralized investment management when using multiple investment firms. These papers require much more structure over the available assets in the entire market, and for each investment firm, separately.

Our non-existence theorem regarding a universal optimal reward scheme also has some parallels in the literature. Holmstrom and Milgrom (1991) studies the principal-agent problem in a set-up completely different from ours, yet it reaches a similar conclusion according to which, sometimes the only optimal reward scheme is constant. It proves that in order to increase their payoff, agents would divert their effort to where it is easier to measure their performance, and derives implications with respect to job design. For example, teachers that receive bonuses according to their students' test scores, might neglect other important aspects.

Assuming that some actions are known to the risk-neutral investor, Carroll (2015) examines contracts where the investor evaluates the performance of the agent through a worst-case criterion consistent with her own knowledge. Under these conditions, it shows that linear contracts are optimal. Nevertheless, one should not confuse this linearity and our proposed linear reward scheme. In Carroll (2015) the conclusion is that a fixed share of the return is optimal, whereas we suggest a contract where the linearity is taken with comparison to other agents.

Although we mainly consider a market with competing firms, the problem we discuss relates to studies where economic incentives, in accordance with social norms, affect the production within firms. Huck et al. (2012) uses a simple model of team production to show that contract designing can increase or decrease the total effort exerted by workers. They study team-pay contracts- i.e., contracts that depend on the total effort of the workers, when employees are either selfish, or team spirited- and relative-performance contracts that depend on individual efforts. In their setup, they show that a competition between workers can decrease the overall production of the firm, while team incentives can increase the production of each worker. They also prove that higher bonuses can reduce efforts by reducing the social pressure of other workers on the individual worker.[7] The model of Huck et al. (2012) and similar ones relate to the dynamic model that we present later, in Section 5.

1.2 Outline of the paper

The paper is organized as follows. Section 2 presents a simple 2-firm reward-scheme problem that illustrates the drawbacks of results-based incentives in competitive non-deterministic markets. In Section 3 we present the model along with the main assumptions. Section 4 includes the main results, divided into three parts: in Subsection 4.1 we show how to formulate an optimal reward scheme for any specific market and a risk-neutral DM. These results are extended in Subsection 4.2 to an expected-utility maximizer DM. In Subsection 4.3 we prove that, unless independent of the outcomes, every reward scheme might fail as the market evolves. Concluding remarks and additional comments are presented in Section 5.

2 A motivating example: a 2-firm reward-scheme problem

An investor wishes to invest some funds through one of two investment firms: Firm 1 and Firm 2. The DM has already some funds invested through these firms and she wishes to allocate additional funds according to some predetermined rule that depends on the firms' yearly earnings.

The goal of the DM is to maximize the expected earnings of current and future investments. However, as she is not aware of the possible bonds in the market, she chooses to allocate the entire available amount to the firm which presents the highest earnings by the end of the year (i.e., a winner-takes-all reward scheme). In case both firms present the same earnings, the funds are equally divided between the two firms.

Suppose that the firms can invest either in Bond X_1, which yields 5% per year with probability (w.p.) 1, or in Bond X_2, which yields 5.1% per year w.p. 0.6 and 0% per year w.p. 0.4%. The goal of the firms is to maximize their expected earnings and to

[7] This line of investigation could be traced back to the work of Kandel and Lazear (1992), on how profit sharing and peer pressure affect the actions of workers in a company. In addition, one should acknowledge the contributions of Holmstrom and Milgrom (1990); Bacharach (1999); and Fischer and Huddart (2008), to this field.

maximize the overall amount of funds they manage, according to the utility functions given below.

In this example we prove that, although investing in Bond X_2 is substantially worse than investing in Bond X_1 (in terms of both expected return and risk), the unique equilibrium in the induced game is when both firms invest in X_2.

Formally, let $A = \{X_1, X_2\}$ be the set of (pure) actions available to the firms. The distributions of X_1 and X_2 (when considered as random variables) is the following:

$$X_1 = 1.05 \text{ per year w.p. } 1, \qquad X_2 = \begin{cases} 1.051, & \text{per year w.p. } \frac{3}{5}, \\ 1.0, & \text{per year w.p. } \frac{2}{5}. \end{cases}$$

These distributions are common knowledge between the firms. The firms can also mix between X_1 and X_2. That is, Firm i may decide to invest, say, a portion $\alpha_i \in [0, 1]$ of the money it manages in X_1 and $1 - \alpha_i$ in X_2. To such a strategy we refer as a diversified strategy.

The utility functions of the investment firms depend on a parameter $\lambda \in [0, 1]$. Let $\sigma_i = \alpha_i X_1 + (1 - \alpha_i)X_2$ be the strategy of Firm i. The utility of Firm 1, U_1 is defined as follows:[8]

$$\begin{aligned} U_1(\sigma_1, \sigma_2) &= \lambda \mathbf{E}(\sigma_1) + (1-\lambda)\mathbf{E}\left(\mathbf{1}_{\{\sigma_1 > \sigma_2\}} + \frac{\mathbf{1}_{\{\sigma_1 = \sigma_2\}}}{2}\right) = \\ &= \lambda \mathbf{E}(\alpha_1 X_1 + (1-\alpha_1)X_2) + \\ &+ (1-\lambda)\left(\Pr\left((\alpha_1 - \alpha_2)[X_1 - X_2] > 0\right) + \frac{\Pr\left((\alpha_1 - \alpha_2)[X_1 - X_2] = 0\right)}{2}\right). \end{aligned} \tag{1}$$

In words, the utility function of Firm 1 is a weighted average (λ vs. $1 - \lambda$) of its earnings (e.g., $\mathbf{E}(\alpha_1 X_1 + (1 - \alpha_1)X_2)$) and the probability that the additional funds are allocated to Firm 1. Firms 2's utility function is defined in a similar fashion.

The following lemma shows that while the firms maximize their profits, the result is unfavorable to the DM.

Lemma 1. *For every* $0 \leqslant \lambda < \frac{1}{1.194} \approx 0.83$*, the unique equilibrium is when both firms choose to invest only in* Bond X_2.

While the expected earnings per year of Bond X_2 is 3.006%, that of Bond X_1 is 5%. It turns out that the reward scheme (i.e., the winner-takes-all mechanism) is adversarial

[8] In what follows $\mathbf{E}$ stands for the expectation.

to the interests of the DM: the unique equilibrium is (X_2, X_2), which from the DM's perspective is the worst possible result. Moreover, even from the firms' perspective the equilibrium (X_2, X_2) is Pareto-dominated by any other profile (σ_1, σ_2), such that $\sigma_1 = \sigma_2$.

Proof. Fix $\lambda \in \left[0, \frac{1}{1.194}\right)$. Assume that Firm 1 employs $\sigma_1 = \alpha_1 X_1 + (1-\alpha_1)X_2$ and Firm 2 employs $\sigma_2 = \alpha_2 X_1 + (1-\alpha_2)X_2$, where $0 \leqslant \alpha_1, \alpha_2 \leqslant 1$. The first term $\lambda\mathbf{E}(\sigma_1)$ of Firm 1's utility equals $\lambda(1.0306 + 0.0194\alpha_1)$, which is linearly increasing in α_1 when $\lambda > 0$. The second term equals

$$(1-\lambda)\left(\Pr(\sigma_1 > \sigma_2) + \frac{1}{2}\Pr(\sigma_1 = \sigma_2)\right) = (1-\lambda)\cdot\begin{cases} 3/5, & \text{if } \alpha_1 < \alpha_2, \\ 1/2, & \text{if } \alpha_1 = \alpha_2, \\ 2/5, & \text{if } \alpha_1 > \alpha_2. \end{cases}$$

When $\lambda = 0$, the strategy $\sigma_1 = X_2$ (i.e., $\alpha_1 = 0$) is a dominant strategy and the result holds. We may thus assume that $\lambda \in \left(0, \frac{1}{1.194}\right)$. By the linearity in α_i, a profile of strategies in which $\alpha_i < \alpha_{-i}$ cannot be an equilibrium, since Firm i has a profitable deviation to $\frac{\alpha_i + \alpha_{-i}}{2}$. This deviation increases the first term of Firm i's utility without affecting the second term.

In addition, if $\alpha_1 = \alpha_2 > 0$, then any firm can make an infinitesimal deviation to $\alpha_i - \epsilon \in (0, \alpha_i)$ and gain $\frac{3(1-\lambda)}{5}$ instead of $\frac{1-\lambda}{2}$ with an infinitesimal loss in $\lambda\mathbf{E}(\sigma_i)$. Therefore, we only need to consider the profile where $\alpha_1 = \alpha_2 = 0$, that is, (X_2, X_2). A direct computation shows that no profitable deviation exists and the result holds. We point out that the linearity of $\lambda\mathbf{E}(\sigma_i)$ in α_i implies that we only need to verify that deviating to X_1 is not profitable.

It is important to note that the utility functions in this example are not linear. If we would let the firms invest their entire allocation either in X_1 or in X_2, without allowing diversified investments, we would get a game where each firm has only two possible actions at its disposal. In such case the payoff matrix would be the following:

	X_1	X_2
X_1	$0.5 + 0.55\lambda$, $0.5 + 0.55\lambda$	$0.4 + 0.65\lambda$, $0.6 + 0.4306\lambda$
X_2	$0.6 + 0.4306\lambda$, $0.4 + 0.65\lambda$	$0.5 + 0.5306\lambda$, $0.5 + 0.5306\lambda$

Table 1: The two pure-action game.

Table 1 presents a 2-player game with the relevant expected values. It is easy to see that X_2 is dominating X_1 for every $\lambda < \frac{1}{1.194}$. Note, however, that this is not sufficient for showing that X_2 is a dominant strategy. The reason is that the utility function of Firm 1 (see Eq. (1)), as well as that of Firm 2, are not linear. That is, $U_1(\alpha_1 X_1 + (1-\alpha_1)X_2, \sigma_2)$ is typically not equal to $\alpha_1 U_1(X_1, \sigma_2) + (1-\alpha_1)U_1(X_2, \sigma_2)$. This implies that although X_2 is dominant over X_1, the action X_2 is not necessarily dominant over all other mixtures between the two. The lack of linearity (let alone the lack of continuity) might also impede the existence of an equilibrium . In our example, one can verify that an equilibrium does not exist[9] when $\frac{1}{1.194} < \lambda < 1$. ■

This example illustrates the problem in case there are two firms and two pure actions, and a specific winner-takes-all reward scheme. The model presented in the following section concerns a more general case. In order to maintain simplicity we assume that firms care only about the volume of their allocations (related to the second summand of the RHS of Eq. (1)) and not about the actual performance (the LHS of Eq. (1)) of the fund they manage. That is, $\lambda = 0$.

3 The model

There are k investment firms in the market. Let $A = \{X_1, \ldots, X_n\}$ be a set of random variables with a finite expectation. This is the set of possible investment options available to every firm. The yield of the i-th investment[10] is represented by the random variable X_i. The elements composing A will be referred to later as assets or *pure strategy*.

A *diversified strategy* σ_i of player i is a mixture of random variables in A. Formally,[11] $\sigma_i = \sum_{j=1}^{n} \sigma_i^j X_j$, where $\sigma_i^j \geqslant 0$ and $\sum_{j=1}^{n} \sigma_i^j = 1$. The set of diversified strategies is denoted by Q. For instance, a firm taking a pure action X_j invests all its managed funds in the j-th asset, where as in case it chooses to use the diversified strategy

[9]A simple computation shows that (X_2, X_2) is no longer an equilibrium, because a deviation to X_1 is profitable. All other profiles are not equilibria by the same reasoning given in the proof of Lemma 1.

[10]Any investment in a financial asset, such as a bond, a stock, or an option, as well as any other sort of investment, such as in real estate or in a commodity.

[11]We sometimes denote a diversified strategy σ_i as a distribution $(\sigma_i^1, \ldots, \sigma_i^n)$ over the set of pure actions A. Nevertheless, the formal definition states that σ_i is the new random variable $\sum_{j=1}^{n} \sigma_i^j X_j$, which is a convex combination of pure strategies given the weights $(\sigma_i^1, \ldots, \sigma_i^n)$.

$\sigma_i = \sum_{j=1}^{n} \sigma_i^j X_j$, it invests a proportion σ_i^j of its managed funds in the j-th asset.[12]

Initially, the investor, or the decision maker (DM), has some funds invested through the k firms. For simplicity and without loss of generality, we assume that the funds are equally divided between the firms and each firm gets a normalized initial amount of 1.

The DM is willing to reallocate her funds among the investment firms based on their performance. For this purpose we introduce the notion of a reward scheme. Let r_i be the measurement of Firm i's performance. That is, r_i denotes the realization of player i's diversified strategy σ_i. Note that the DM is not familiar with the assets included in A. She does not know their distributions, nor their expected payoffs, and can only observe the performances $(r_1, \ldots, r_k)$ of the firms at the end of a single time period.

Definition 1. *A* reward scheme *is a function* $f : \mathbb{R}^k \to [0,1]^k$ *such that for every* $r \in \mathbb{R}^k$*,*

$$\sum_{i=1}^{k} f_i(r) = 1. \tag{2}$$

In words, given a vector $(r_1, \ldots, r_k)$ of the firms' performances, a proportion $f_i(r_1, \ldots, r_k)$ of the available funds is to be allocated to Firm i.

The DM publicly commits to a reward scheme f. This, in turn, defines a k-player game, called an *investment game* and denoted G_f, as follows. Firm i (referred to also as Player i) chooses a strategy $\sigma_i \in Q$. Player i's payoff depends not only on its own strategy, but on all other players' strategies as well. When $\sigma = (\sigma_1, \ldots, \sigma_k) \in Q^k$ is the profile of strategies used by the players, the expected payoff of Player i is

$$\mathbf{E}\left[f_i\left(\sigma\right)\right].$$

In words, the payoff of Firm i is the expected proportion of the funds it is going to manage. This game is symmetric in all respects: all the players are homogeneous in their utility function and have the same set of strategies.

Definition 2. *A profile of strategies* $\sigma \in Q^k$ *is a* Nash equilibrium *in the investment game* G_f *if*

$$\boldsymbol{E}\left[f_i\left(\sigma_i, \sigma_{-i}\right)\right] \geqslant \boldsymbol{E}\left[f_i\left(\sigma_i', \sigma_{-i}\right)\right],$$

[12]In game theory, a diversified strategy is commonly perceived as a pure strategy.

for every Player i and for every strategy $\sigma_i' \in Q$.

In the situation under consideration, the DM is actually a mechanism designer. She announces a reward scheme and thereby defines an investment game. The investment firms are the players in this game. They wish to maximize their expected payoffs. The goal of the DM, on the other hand, is to design a game G_f in a way that in any Nash equilibrium $\sigma = (\sigma_1, ..., \sigma_k)$, the strategy of Firm i maximizes the expected return of the DM:

$$\mathbf{E}\left[\sigma_i\right] = \max_{X_j \in A} \mathbf{E}\left[X_j\right].$$

Remark 1. *In Subsection 4.2, we introduce a utility function $U : \mathbb{R} \to \mathbb{R}$ for the DM. Given a utility function U, the goal of the DM is to maximize $\boldsymbol{E}\left[U\left(\sum_{i=1}^k \sigma_i\right)\right]$. In the current set-up, we assume that the DM is risk-neutral, e.g., $U(x) = x$ for all $x \in \mathbb{R}$. Hence,*

$$\boldsymbol{E}\left[U\left(\sum_{i=1}^k \sigma_i\right)\right] = \boldsymbol{E}\left[\sum_{i=1}^k \sigma_i\right].$$

By linearity, the RHS term $\boldsymbol{E}\left[\sum_{i=1}^k \sigma_i\right]$ is maximized if and only if $\boldsymbol{E}\left[\sigma_i\right] = \max_{X_j \in A} \boldsymbol{E}\left[X_j\right]$ for every Firm i.

Remark 2. *In general, the DM consumes the net profit and then redistributes the funds according to a reward scheme f. The actual goal of the DM is, therefore, to maximize the expected net profit and not the expected value $\sum_{i=1}^k \sigma_i$. Since the two are equal, up to a constant, we use the latter for the maximization problem. Alternatively, one could change the set of assets A such that $\sum_{i=1}^k \sigma_i$ represents the net profit of the DM.*

3.1 Optimal Reward Schemes

Incentivizing investment firms through a reward scheme bears some similarities with the principal-agent problem. The DM can be thought of as a principal who is interested in motivating her agents (e.g., the investments firms) to produce optimal expected earnings. The DM cannot monitor the investment strategies used by the firms; she can observe only the investment results (performance). Based on those results alone, she wants to create incentives for the firms so that the latter will serve best her interests. The example given in Section 2 shows that intuitive methods, such as winner-takes-all, might be counterproductive: they might generate a sub-optimal result for the DM. In the following section we show how to properly design a reward scheme and induce productive incentives.

Let $O \subseteq A$ be the set of optimal assets:

$$O = \{X_j \in A;\ \mathbf{E}\left[X_j\right] \geqslant \mathbf{E}\left[X_\ell\right] \text{ for every } X_\ell \in A\}. \tag{3}$$

Thus, there exists an $\epsilon > 0$ such that for every $X_j \in O$ and $X_\ell \notin O$,

$$\mathbf{E}[X_j] > \mathbf{E}[X_\ell] + \epsilon. \tag{4}$$

Definition 3. *A reward scheme f is* optimal*, if*
(i) *an equilibrium exists in G_f; and*
(ii) *in every equilibrium $\sigma = (\sigma_1, ..., \sigma_k)$ and for every i,*

$$\boldsymbol{E}\left[\sigma_i\right] = \max_{X_j \in A} \boldsymbol{E}\left[X_j\right]. \tag{5}$$

In words, f is optimal if in every equilibrium σ in G_f the diversified actions of all the players produce the maximal expected value; these diversified actions are weighted averages of pure actions from O.

Remark 3. *Due to Eq.* (2)*, for every reward scheme f and every profile of strategies σ it holds that,*

$$\boldsymbol{E}\left[\sum_{i=1}^{k} f_i(\sigma)\right] = 1.$$

Therefore, the investment game G_f is a fixed-sum game in the sense that the sum of the expected payoffs of all the players is 1*. For every reward scheme f and for every game G_f we can subtract* $1/k$ *from each player's utility function and obtain a symmetric zero-sum k-player game.*

3.2 Interpreting the reward scheme

Though the concept of incentive schemes in the problem of delegated portfolio management is well-studied, the idea of a reward scheme that sustains the conditions given in Definition 1, is relatively novel. These conditions go to the very core of our model, and therefore need to be explained thoroughly.

Most previous works considered mainly incentive plans that are simple monetary rewards. These plans do not necessarily sustain the budget constraint given by Equation (2). However, we think that our reward-scheme model captures an important property of the portfolio management market.

As most investment firms get a certain percentage of the portfolio they manage, the ability of the DM to negotiate their incentives is limited. An optimal reward scheme keeps the firms in accordance with the best interests of the DM, by taking advantage of her elementary ability to redistribute the funds.

One could also give a different interpretation for a reward scheme. Consider a scenario where the DM has a fixed sum of money to be used for bonus purposes. The reward scheme, in this case, is a results-based bonus plan. This exemplifies another important aspect of the reward scheme - it is bounded. A reward scheme is applicable for any fixed sum, and the DM need not commit to unbounded rewards.

In addition, the implementation of a reward scheme is trivial. The DM needs but to observe the actual performances $(r_1, \ldots, r_k)$ of the firms at the end of a single time period. Furthermore, we do not require any exogenous benchmark portfolio, as do most previous works. This simplicity allows a complete layman (in terms of investment) to employ such an incentives-coordinating mechanism.

4 Main results

In this section we prove the central results of this paper. The first shows that for every finite set of random variables A, there exists an optimal reward scheme f. The second result states that for every non-trivial (i.e., non-constant) reward scheme f, there exists a set A such that f is not optimal.

The combination of these results is significant. On the positive side, in any non-deterministic market the DM can design rules that ensure that her interests are accomplished by the players. On the negative side however, when the market is dynamic- meaning that the market is constantly changing in terms of possible actions and assets' yields are constantly growing- and when the DM cannot keep track of these changes, then no single reward scheme can produce optimal results (in any possible market). That is, any non-trivial reward scheme can lead to suboptimal results. More formally, the only reward scheme that induces also equilibria in which players act according to the DM's preferences, is a reward scheme which, paradoxically, is independent of the players' actions.

4.1 A fixed set of actions

The first theorem we prove is constructive. For every set of actions A, we specify an optimal reward scheme f. The optimal reward scheme f that we propose is linear in terms of the differences between the players' earnings.

4.1.1 Bounded random variables

We start with the case of bounded random variables. Assume that there exists an $M \in \mathbb{R}$ such that $\Pr(|X_i| \leqslant M) = 1$ for every asset $X_i \in A$. When such an M exists, we say that A is *uniformly bounded.*

Define the *Linear Reward Scheme* f as,

$$f_i(r) = \frac{1}{k} + \begin{cases} \frac{\sum_{j \neq i}(r_i - r_j)}{2k(k-1)M}, & \text{if } \forall i, |r_i| \leqslant M, \\ 0, & \text{if } \exists i \text{ s.t. } M < |r_i|. \end{cases}$$

One can verify that f is well-defined, since for every $r \in \mathbb{R}^k$, the equality $\sum_i f_i(r) = 1$ holds and $f(r) \in [0,1]^k$.

The Linear Reward Scheme f can be rewritten as

$$f_i(r) = \frac{1}{k} + \frac{1}{2Mk}\left[r_i - \frac{1}{k-1}\sum_{j \neq i} r_j\right] \mathbf{1}_{\{\forall i,\ |r_i| \leqslant M\}}.$$

This presentation provides an important economic insight on the optimal reward scheme f: this reward scheme distributes to all players, before the results are considered, the same basic share $1/k$. When the results are considered, every player gains or loses relatively to the basic share, a portion that depends on the difference between his result and the average result of the other players.

In other words, the performance of every player is assessed relatively to the other players' average performance, and not to some exogenous benchmark portfolio (which is the case in most previous studies). This is one crucial aspect of the Linear Reward Scheme. It generates a competition that, in return, generates a specific benchmark for each player, which is the average performance of his competitors. In reality, one can implement the Linear Reward Scheme just by employing an estimated large M (e.g., as the maximal result obtained in previous years).

Theorem 1. *For every uniformly-bounded set of assets A, the* Linear Reward Scheme *is optimal.*

Remark 4. *For the sake of simplicity, and without loss of generality, we assume in the proof of* Theorem 1 *that $O = \{X_1\}$. That is,*

$$\boldsymbol{E}[X_1] > \boldsymbol{E}[X_i], \tag{6}$$

for every $2 \leqslant i \leqslant n$. This implies that a reward scheme f is optimal if and only if $(X_1, \ldots, X_1)$ is a unique equilibrium in G_f.

Proof. We prove that for every Player i, for every profile of diversified actions $(\sigma_1, \sigma_2, \ldots, \sigma_k) \in Q^k$ of players $1, \ldots, k$ respectively, and for every strategy $\sigma_i \neq X_1$ of Player i, the inequality

$$\mathbf{E}\left[f_i\left(\sigma_1, \ldots, \sigma_{i-1}, X_1, \sigma_{i+1}, \ldots, \sigma_k\right)\right] > \mathbf{E}\left[f_i\left(\sigma_1, \ldots, \sigma_k\right)\right],$$

holds.

Without loss of generality, assume that $i = 1$. Therefore

$$\begin{aligned}
\mathbf{E}\left[f_1\left(\sigma_1, \sigma_2, \ldots, \sigma_k\right)\right] &= \mathbf{E}\left[\frac{\sum_{j=1}^{k}(\sigma_1 - \sigma_j)}{2k(k-1)M} + \frac{1}{k}\right] \\
&= \mathbf{E}\left[\frac{(k-1)\sigma_1 - \sum_{j=2}^{k}\sigma_j}{2k(k-1)M} + \frac{1}{k}\right] \\
&< \mathbf{E}\left[\frac{(k-1)X_1 - \sum_{j=2}^{k}\sigma_j}{2k(k-1)M} + \frac{1}{k}\right] \\
&= \mathbf{E}\left[\frac{\sum_{j=2}^{k}\left(X_1 - \sigma_j\right)}{2k(k-1)M} + \frac{1}{k}\right] \\
&= \mathbf{E}\left[f_1\left(X_1, \sigma_2, \ldots, \sigma_k\right)\right],
\end{aligned}$$

where the first and the last equalities follow from the definition of f, and the inequality follows from the fact that $\sigma_1 \neq X_1$ and $\mathbf{E}[\sigma_1] = \mathbf{E}\left[\sum_{i=1}^{n} \sigma_1^j X_j\right] < \mathbf{E}[X_1]$. ■

Remark 5. *The uniqueness of the equilibrium $(X_1, \ldots, X_1)$ is not surprising. It follows directly from the linearity of $f_i(r)$ in r_i. This linearity implies that X_1 is a dominant strategy for every player. Therefore, the proof of Theorem 1 actually shows that the Linear Reward Scheme f induces a* Dominance-Solvable *investment game, and that the unique Nash equilibrium is in fact a dominant-strategy equilibrium.*

4.1.2 The self-induced benchmark

The result of Theorem (1) is consistent with the well-known fact that, under risk-neutrality and no effort costs, paying a share of the output induces the first-best choice. In this sense, the optimality of any linear reward scheme might seem obvious. However, our Linear Reward Scheme has another important and novel property we wish to discuss: a *self-induced benchmark*.

Admati and Pfleiderer (1997) presents the problem of using exogenous benchmarks in various scenarios. The main problem that arises is that the benchmark is not calibrated to the effort, and sometimes the risk, that the players incur. The same problem is met in Holmstrom and Milgrom (1987) and Sappington (1991). In fact, in most frameworks it is well-established that the linear scheme is only a second-best solution, as it leads to an underinvestment in effort and information, when needed.

The scheme we propose relates to these issues by generating a competition between the firms. The Linear Reward Scheme defines a competition in which every player is measured in comparison to the other players, who are also evaluated in the same way. Therefore, the effort and risk invested by one player are taken with respect to those invested bu the other players.

In case most firms exert some effort to increase their share, a marginal player is motivated, due to the competition-based benchmark, to act in the same way. The competition also solves the problem of risk-sharing, as the strategy of every player is correlated with its benchmark, via the equilibrium of the investment game.

We believe that this approach can solve the second-best problem presented in Admati and Pfleiderer (1997), where is it proven that the best use of information is achieved when the benchmark is optimal. Though we leave the model of costly efforts to future research, the problem of optimality under risk and uncertainty is discussed in Section 4.2.

4.1.3 The motivating example — revisited

Before we generalize the Linear Reward Scheme f to unbounded random variables, we show how it affects the example presented in Section 2.

Recall that there are two players ($k = 2$), and two pure actions X_1 and X_2. For the sake of simplicity, fix $M = 2$ and note that $X_i < M$ for every $i = 1, 2$. The utility function of Player 1, given the profile of actions (σ_1, σ_2), is

$$\begin{aligned} U_1(\sigma_1, \sigma_2) &= \lambda \mathbf{E}(\sigma_1) + (1-\lambda)\mathbf{E}(f_1(\sigma_1, \sigma_2)) \\ &= \lambda \mathbf{E}(\sigma_1) + (1-\lambda)\left[\frac{1}{2} + \frac{\mathbf{E}(\sigma_1 - \sigma_2)}{8}\right] \\ &= \frac{(1-\lambda)(4 - \mathbf{E}(\sigma_2)) + (1+7\lambda)\mathbf{E}(\sigma_1)}{8}. \end{aligned}$$

This utility increases with $\mathbf{E}(\sigma_1)$, rendering X_1 a dominant strategy, and (X_1, X_1) the unique Nash equilibrium.

4.1.4 Unbounded random variables

Theorem 1 proves that the linear reward scheme is optimal when A is uniformly bounded. However, the same reward scheme cannot be applied to unbounded random variables. Theorem 2 presents a modified linear reward scheme for a general set of assets, namely for the case where A is not uniformly bounded, and states that it is optimal.

In the unbounded case the assumption that X_1 is a unique optimal action (see Ineq. (6)) limits generality. We therefore analyse the general case: O may contain several optimal pure actions rather than a unique optimal pure action.

Let $Q_1 = \{q \in Q : q^i = 0, \forall X_i \in O\}$ be the set of diversified actions in Q where all the sub-optimal pure actions are taken with probability 0, and let $Q_2 = \{q \in Q : q^i = 0, \forall X_i \notin O\}$ be the set of diversified actions in Q where all the optimal pure actions in O are assigned probability 0. The following lemma enables us to define f. Recall ϵ from Ineq. (4).

Lemma 2. *There exists an $M > 0$ such that for every $(X, q, \alpha) \in Q_1 \times Q_2 \times [0,1]$ and every $m \geqslant M$,*

$$\boldsymbol{E}\left[|X - q|\mathbf{1}_{\{|(1-\alpha)X + \alpha q| > m\}}\right] < \frac{\epsilon}{2}.$$

The proof is given in the Appendix.

We use the M given in Lemma 2 to define a new reward scheme f, similar to the Linear Reward Scheme used in Theorem 1. First, define the real-valued function $\phi : \mathbb{R} \to \mathbb{R}$ as

$$\phi(x) = \begin{cases} -M, & \text{if } x < -M, \\ x, & \text{if } -M \leqslant x \leqslant M, \\ M, & \text{if } x > M. \end{cases}$$

For every player i and every vector $r = (r_1, \ldots, r_k) \in \mathbb{R}^k$, define the *Truncated Reward Scheme* f such that

$$\begin{aligned} f_i(r) &= \frac{1}{k} + \frac{\sum_{j=1}^{k} [\phi(r_i) - \phi(r_j)]}{2k(k-1)M} \\ &= \frac{1}{k} + \frac{1}{2Mk} \left[\phi(r_i) - \frac{1}{k-1} \sum_{j \neq i} \phi(r_j) \right]. \end{aligned}$$

One can verify that f is well-defined, since

$$\begin{aligned} 0 \leqslant \frac{1}{k} + \frac{1}{2Mk} \left[\phi(r_i) - \frac{1}{k-1} \sum_{j \neq i} \phi(r_j) \right] \leqslant \frac{2}{k} \quad &\Leftrightarrow \quad -\frac{1}{k} \leqslant \frac{1}{2Mk} \left[\phi(r_i) - \frac{1}{k-1} \sum_{j \neq i} \phi(r_j) \right] \leqslant \frac{1}{k} \\ &\Leftrightarrow \quad -2M \leqslant \phi(r_i) - \frac{1}{k-1} \sum_{j \neq i} \phi(r_j) \leqslant 2M \\ &\Leftrightarrow \quad \left| \phi(r_i) - \frac{1}{k-1} \sum_{j \neq i} \phi(r_j) \right| \leqslant 2M, \end{aligned}$$

and the last inequality holds for every $r_i, r_j \in \mathbb{R}$.

The next theorem concludes subsection 4.1. It states that an optimal reward scheme exists also when the set A is not uniformly bounded.

Theorem 2. *For every set of pure actions A, the* Truncated Reward Scheme *f is optimal.*

The proof is given in the Appendix.

In this subsection we presented the simple case of optimal reward schemes, in which the players are concerned only with the share they obtain from the managed funds. Using the terminology of the example in Section 2, we actually assumed that $\lambda = 0$ for every player. It is important to emphasize, though, that our results could be extended beyond this restriction. One can take $\lambda > 0$ and show that the previously-defined reward schemes are still considered optimal. The intuition is clear: once λ is greater than 0, every player has an additional incentive to choose an optimal asset rather than a sub-optimal one. Therefore, the reward schemes that were previously considered optimal remain so also when players are concerned with their actual performance and not only with their share.

4.1.5 Uniqueness

Theorems 1 and 2 prove that our proposed linear reward schemes are optimal. These theorems suggest that a sufficient condition for an optimal reward scheme is linearity. The next question we wish to address is whether this is also a necessary condition. The following theorem proves that linearity is indeed crucial for optimality. Specifically, if one wants a reward scheme f to generate an investment game G_f with a dominant-strategy optimal equilibrium (i.e., an equilibrium that sustains Eq. (5)), then $f_i(r)$ must be linear in r_i.

Theorem 3. *Fix $M > 0$ and assume that for every subset A of strictly-bounded actions (i.e., $\Pr(|X_i| > M) = 1$ for every action X_i), the investment game G_f has a dominant-strategy optimal equilibrium. Then, $f_i(r)$ is linear in r_i for every r_{-i}.*

The proof is given in the Appendix.

4.2 An expected utility maximizer DM

Though Theorems 1 and 2 apply for a wide range of assets, the risk-neutrality set-up is quite restrictive in the context of financial markets. Clearly, it is very difficult to make sense of financial markets as they are by assuming investors are risk-neutral. For this reason, we now extend the previous theorems for cases where the DM is an expected utility maximizer with some non-linear utility function.

Assume that the DM has a general utility function $U : \mathbb{R} \to \mathbb{R}$. Her main goal, as mentioned in Remark 1 and explained in Definition 4, is to maximize $\mathbf{E}\left[U\left(\sum_{i=1}^{k} \sigma_i\right)\right]$ where $\sigma = (\sigma_1, \dots, \sigma_k)$ is a Nash equilibrium in the induced investment game G_f.

Definition 4. *A reward scheme f is* U-optimal*, if*
(i) *an equilibrium exists in G_f; and*
(ii) *in every equilibrium $\sigma = (\sigma_1, ..., \sigma_k)$,*

$$\boldsymbol{E}\left[U\left(\sum_{i=1}^{k}\sigma_i\right)\right] = \max_{q \in Q} \boldsymbol{E}\left[U(kq)\right]. \tag{7}$$

In words, f is U-optimal if in every equilibrium σ in G_f, the combined portfolio $\sum_{i=1}^{k} \sigma_i$ produce the maximal expected utility for the DM.[13]

[13]The factor k in the RHS of Eq. (7) follows from the initial amount of 1 given to each firm.

Note that we do not restrict our attention solely to risk-averse or to risk-loving decision making. Nevertheless, we do need to draw two assumptions over the utility function U.

The first assumption is that the utility function is bounded by some large constant M. Specifically, a utility function U is *bounded* if there exists an $M > 0$ such that $|U(x)| \leqslant M$ for every $x \in \mathbb{R}$.

The second assumption relates to the number of optimal diversified actions w.r.t. U. Formally, a utility function U is *uniquely maximized* if there exists a diversified action $q^* \in Q$ such that

$$E[U(kq^*)] > E[U(kq)], \quad \forall q \in Q \backslash \{q^*\}.$$

In words, we say that U is *uniquely maximized* in case the maximal expected utility in the RHS of Eq. (7) is obtained by a unique diversified action.

In the following theorem, we define the *General Reward Scheme*, and along with these two assumptions, we extend the previous results of Theorems 1 and 2.

Theorem 4. *For every set of actions A such that U is bounded and uniquely maximized, the following* General Reward Scheme *f is U-optimal:*

$$f_i(r) = \frac{1}{k} + \frac{1}{2Mk}\left[U(kr_i) - \frac{1}{k-1}\sum_{j \neq i} U(kr_j)\right]\mathbf{1}_{\{\forall i,\ |U(kr_i)| \leqslant M\}}.$$

The proof is given in the Appendix, and follows the same steps as the proof of Theorem 1.

The economic interpretation of the General Reward Scheme is therefore clear. The DM uses his own utility function in order to produce an incentive scheme that is linear with respect to U. The performance of each player is first assessed via the DM's utility function, and later compared to the utility-based performance of the other players.

Remark 6. *The requirement that every equilibrium is optimal in Definitions* (3) *and* (4) *is quite restrictive, especially in the context of delegated portfolio management. Another possibility is to define optimality by the existence of an optimal equilibrium, in the sense of Equations* (5) *and* (7)*. This alternation is given in Definition* (5) *that follows. Note that under this weaker assumption the uniquely-maximized condition in Theorem* (4) *is redundant.*

Remark 7. *Under the weaker definition, given in the previous remark, and without the uniquely-maximized condition, the result of Theorem 4 still holds, even when the players are not risk-neutral, but using a constant absolute risk aversion (CARA) utility function. Specifically, since we do not restrict ourselves to normal distributions, assume that player i tries to maximize* $\boldsymbol{E}[X]-\frac{\gamma}{2}\mathrm{Var}[X]$ *when* X *is player i's payoff and assume that the DM has a bounded utility function* U*. In this case, the optimal equilibrium given in the proof of Theorem (4) still holds. The proof is straightforward since a deviation of one player can only increase his variance,*[14] *without increasing the expected value. This extends previously-known results based only on a risk-averse DM with a CARA utility function and a CARA utility maximizing players (under the assumption of normal distributions).*

4.3 A universal reward scheme

Theorems 1, 2, and 4 show that for a given set of actions, one can design an optimal reward scheme. The next question that we address is whether or not there exists a reward scheme that is optimal[15] for every set of actions A.

Definition 5. *A reward scheme* f *is said to be* universal *if for every set of actions* $A = \{X_1, \ldots, X_n\}$*, the induced investment game* G_f *has an equilibrium* $(\sigma_1, ..., \sigma_k)$ *where* $\boldsymbol{E}[\sigma_i] = \max_{X_j \in A} \boldsymbol{E}[X_j]$ *for every* $i = 1, \ldots, k$.

In words, f is a universal reward scheme if for every finite set of actions there exists an equilibrium that sustains the optimality condition given in Eq. (5). When comparing a universal reward scheme and an optimal reward scheme, one should notice two differences. A reward scheme is optimal if *every* equilibrium is optimal, whereas a reward scheme is universal if an optimal equilibrium *exists*. Secondly, a reward scheme f is universal if for every set of actions A the induced investment game G_f has an optimal equilibrium. An optimal reward scheme, on the other hand, relates only to a specific set of actions.

The following theorem states that a non-constant universal reward scheme does not exist.

[14] The variance is minimal since the strategy of every player and his benchmark are correlated, as explained in Subsection 4.1.2

[15] In this subsection we consider a risk-neutral DM.

Theorem 5. *In the case of two players, there is no non-constant universal reward scheme.*

The proof is given in the Appendix.

A generalization of Theorem 5 to any number of players $k \geqslant 3$ is not trivial though. For example, take any non-constant reward scheme $f : \mathbb{R}^k \to \mathbb{R}^k$ such that $f_i(r) = 1/k$ for every $r \in \mathbb{R}^k$ that has at least two identical coordinates. In this case, for every action X_j, the profile of strategies $(X_j, \ldots, X_j)$ is an equilibrium, because a unilateral deviation would still leave at least two identical coordinates of at least two other players, in which case the share would be determined as $1/k$. In other words, any deviation of a single player will not bear any influence on the payoffs. On the other hand, requiring that all equilibria satisfy Eq. (5) will not hold when f is constant, because in this case all profiles are equilibria. Therefore, when the number of players is three or more, we must introduce a stronger requirement .

Definition 6. *A reward scheme f is said to be* strongly universal *if for every set $A = \{X_1, \ldots, X_n\}$, every optimal profile of strategies (i.e., satisfying Eq. (5)) in G_f constitutes necessarily an equilibrium.*

Note that if a Nash equilibrium exists, then any strongly universal reward scheme is a universal reward scheme.

Theorem 6. *If f is a strongly universal reward scheme, then every profile of strategies is an equilibrium.*

The proof is given in the Appendix.

5 Concluding remarks

5.1 Varying portfolio preferences

The results presented in this paper have two complementary aspects: practical and theoretical. On the practical level, we provide a specific description of a reward scheme that guarantees that agents are motivated to act according to the DM's interests. On the theoretical side, we show that an always-optimal reward scheme simply does not exist.

This work focuses on an investor who naturally tries to maximize her expected utility. Although this assumption is common in the literature, one can still follow the line of Holmstrom and Milgrom (1991) and assume that the DM simultaneously incorporates into her utility function different aspects of the managed profile, such as expected return and some measurement of riskiness. One can assume in general that the investor has a preference relation over the set of possible portfolios, and thus try to find a reward scheme that induces an optimal equilibrium (with respect to this preference relation). We leave this problem for future research.

5.2 A dynamic model

The model we consider in this paper in static. One can naturally extend this model to a dynamic one, where the firms take into account the future volume of the funds they manage. At any stage the DM redistributes available funds (including her yearly earnings and her already-allocated funds) according to the firms' previous performance, ans following a reward scheme she has conceived. In return, each firm receives a fixed percentage of the entire volume it manages. As a result, each firm wants to excel at the present period in order to receive more funds to manage in the future. In other words, the firms are primarily concerned with getting as greater a portion of the available funds as possible.

In more formal terms, the stage-payoff of a firm resembles the utility function discussed in Section 3 — it takes into account only the volume of the funds allocated for it to manage. In regard to the example in Section 2, the stage-payoff of a firm in a dynamic setting may resemble that in Eq. (1) with $\lambda = 0$. This implies that the actual performance of a firm is important only as long as it affects the volume of the funds allocated for that firm to manage. However, a more general model could take into consideration a situation where the firms balance between making the cake bigger and at the same time, getting a larger slice of it in the future. To a certain extent, this latter model resembles the one in Huck et al. (2012), where workers consider both the common prosperity and their own personal good.

5.3 Non-homogeneous firms

In this paper we analysed a model in which the assets available (i.e., set A) are common to all players. This is quite a natural assumption. The model, however, can accommodate for a more general scenario, where actions sets are firm specific. The results above can be easily stated in these terms as well.

What about the case where firms may have private information or different levels of expertise? A dynamic Bayesian model with asymmetric reward schemes is left for a future investigation.

5.4 Multiple decision makers

Chevalier and Ellison (1997) study how the inflow of investments and the actions of investment firms relate to one another. In general, it is clear that the behaviour of additional investors can distort incentives generated by a single DM. These effects are significant and have been demonstrated in many empirical studies. We leave the problem of multiple investors and the implications of these issues on regulatory decisions for future research.

References

Admati, Anat R and Paul Pfleiderer, "Does it all add up? Benchmarks and the compensation of active portfolio managers," *The Journal of Business*, 1997, *70* (3), 323–350.

Akerlof, George A. and Paul M. Romer, "Looting: the economic underworld of bankruptcy for profit," *Brookings Papers on Economic Activity*, 1993, *24* (2), 1–74.

Ang, Andrew and Geert Bekaert, "Stock return predictability: Is it there?," *Review of Financial studies*, 2007, *20* (3), 651–707.

Bacharach, Michael, "Interactive team reasoning: a contribution to the theory of co-operation," *Research in Economics*, 1999, *53* (2), 117–147.

Barry, Christopher B. and Laura T. Starks, "Investment Management and Risk Sharing with Multiple Managers," *Journal of Finance*, 1984, *39* (2), 477–91.

Berk, Jonathan B. and Richard C. Green, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 2004, *112* (6), 1269–1295.

Campbell, John Y. and Motohiro Yogo, "Efficient tests of stock return predictability," *Journal of financial economics*, 2006, *81* (1), 27–60.

Carroll, Gabriel, "Robustness and linear contracts," *American Economic Review*, 2015, *105* (2), 536–563.

Chevalier, Judith and Glenn Ellison, "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 1997, *105* (6), 1167–1200.

Cochrane, John H and Monika Piazzesi, "Bond Risk Premia," *American Economic Review*, mar 2005, *95* (1), 138–160.

Dai, Qiang and Kenneth J. Singleton, "Expectation puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics*, mar 2002, *63* (3), 415–441.

Dasgupta, Amil and Andrea Prat, "Financial equilibrium with career concerns," *Theoretical Economics*, 2006, *1* (1), 67–93.

_ **and** _ , "Information aggregation in financial markets with career concerns," *Journal of Economic Theory*, 2008, *143* (1), 83–113.

Dybvig, Philip H., Heber K. Farnsworth, and Jennifer N. Carpenter, "Portfolio Performance and Agency," *Review of Financial Studies*, jan 2010, *23* (1), 1–23.

Elton, Edwin J. and Martin J. Gruber, "Optimum Centralized Portfolio Construction with Decentralized Portfolio Management," *The Journal of Financial and Quantitative Analysis*, 2004, *39* (3), 481–494.

Fischer, Paul and Steven Huddart, "Optimal contracting with endogenous social norms," *American Economic Review*, 2008, *98* (4), 1459–1475.

Fligstein, Neil and Adam Goldstein, "The anatomy of the mortgage securitization crisis," *Institute for Research on Labor and Employment, Working Paper Series*, feb 2010.

_ **and Alexander Roehrkasse**, "All of the incentives were wrong: opportunism and the financial crisis," in "The American Sociological Association Annual Meeting" New York, NY 2013.

Foster, Dean P. and Peyton H. Young, "Gaming performance fees by portfolio managers," *The Quarterly Journal of Economics*, 2010, *125* (4), 1435–1458.

Hansen, Laura L., "Corporate financial crime: social diagnosis and treatment," *Journal of Financial Crime*, 2009, *16* (1), 28–40.

He, Xuedong, Sang Hu, and Steven Kou, "Separating skilled and unskilled fund managers by contract design," *Working paper*, 2015.

Holmstrom, Bengt and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 1987, *55* (2), 303–28.

_ **and** _ , "Regulating trade among agents," *Journal of Institutional and Theoretical Economics*, 1990, *146*, 85–105.

_ **and** _ , "Multitask principal-agent analyses: incentive contracts, asset ownership, and job design," *Journal of Law, Economics, & Organization*, 1991, *7*, 24–52.

Huck, Steffen, Dorothea Kübler, and Jörgen Weibull, "Social norms and economic incentives in firms," *Journal of Economic Behavior & Organization*, 2012, *83* (2), 173–185.

Kandel, Eugene and Edward P. Lazear, "Peer pressure and partnerships," *Journal of Political Economy*, 1992, *100* (4), 801–817.

Keynes, John M., *The general theory of employment, interest and money*, Palgrave Macmillan, 1936.

Lehrer, Ehud, "Any inspection is manipulable," *Econometrica*, 2001, *69* (5), 1333–1347.

Lettau, Martin and Stijn Van Nieuwerburgh, "Reconciling the return predictability evidence," *Review of Financial Studies*, 2008, *21* (4), 1607–1652.

Lewellen, Jonathan, "Predicting returns with financial ratios," *Journal of Financial Economics*, nov 2004, *74* (2), 209–235.

Myerson, Roger B., "Optimal auction design," *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

Olszewski, Wojciech and Alvaro Sandroni, "Manipulability of future-independent tests," *Econometrica*, 2008, *76* (6), 1437–1466.

Sandroni, Alvaro, "The reproducible properties of correct forecasts," *International Journal of Game Theory*, 2003, *32* (1), 151–159.

_ **, Rann Smorodinsky, and Rakesh V. Vohra**, "Calibration with many checking rules," *Mathematics of Operations Research*, 2003, *28* (1), 141–153.

Sappington, David E. M, "Incentives in Principal-Agent Relationships," *Journal of Economic Perspectives*, may 1991, *5* (2), 45–66.

Scharfstein, David S. and Jeremy C. Stein, "Herd behavior and investment," *American Economic Review*, 1990, *80* (3), 465–479.

Sharpe, William F., "Decentralized Investment Management," *The Journal of Finance*, may 1981, *36* (2), 217–234.

Shmaya, Eran, "Many inspections are manipulable," *Theoretical Economics*, 2008, *3* (3), 367–382.

Simpson, Sally S., "Making Sense of White Collar Crime: Theory and Research," *The Ohio State Journal of Criminal Law*, 2011, *8*, 481–502.

Stracca, Livio, "Delegated portfolio management: a survey of the theoretical literature," *Journal of Economic Surveys*, dec 2006, *20* (5), 823–848.

Torous, Walter, Rossen Valkanov, and Shu Yan, "On Predicting Stock Returns with Nearly Integrated Explanatory Variables," *The Journal of Business*, 2004, *77* (4), 937–966.

Treynor, Jack L and Fischer Black, "How to use security analysis to improve portfolio selection," *The Journal of Business*, 1973, *46* (1), 66–86.

Van Binsbergen, Jules H., Michael W. Brandt, and Ralph S. J. Koijen, "Optimal Decentralized Investment Management," *The Journal of Finance*, aug 2008, *63* (4), 1849–1895.

6 Appendix

Lemma 2. *There exists an $M > 0$ such that for every $(X, q, \alpha) \in Q_1 \times Q_2 \times [0,1]$ and every $m \geqslant M$,*

$$\boldsymbol{E}\left[|X - q|\mathbf{1}_{\{|(1-\alpha)X+\alpha q|>m\}}\right] < \frac{\epsilon}{2}.$$

Proof. By the optimality of the actions in Q_1 and the sub-optimality of the actions in Q_2, we know that $\mathbf{E}[X - q] > \epsilon$ for every $X \in Q_1$ and every $q \in Q_2$. In addition, for every $\alpha \in [0,1]$, one can choose a sufficiently large $M_{X,q,\alpha} > 0$ such that for every $m \geqslant M_{X,q,\alpha}$,

$$\mathbf{E}\left[|X - q|\mathbf{1}_{\{|(1-\alpha)X+\alpha q|>m\}}\right] < \frac{\epsilon}{2}. \tag{8}$$

This follows from the fact that for every $(X, q, \alpha) \in Q_1 \times Q_2 \times [0,1]$, the set of random variables $\{|X-q|)\mathbf{1}_{\{|(1-\alpha)X+\alpha q|>m\}}\}_{m\in\mathbb{N}}$ is a sequence of real-valued measurable functions that are weakly dominated by an integrable function $|X - q|$. That is,

$$|X - q|\mathbf{1}_{\{|(1-\alpha)X+\alpha q|>m\}} \leqslant |X - q|$$

for every $m \in \mathbb{N}$. The sequence converges pointwise to 0 as $m \to \infty$. Hence, by the dominated convergence theorem,

$$\mathbf{E}\left[|X - q|\mathbf{1}_{\{|(1-\alpha)X+\alpha q-Y|>m\}}\right] \to 0 \quad \text{as } m \to \infty.$$

Since Ineq. (8) is strict, there exists an open set $B_{X,q,\alpha} \subseteq Q^2 \times \mathbb{R}$ containing (X, q, α), such that this inequality holds for every $(X', q', \alpha') \in B_{X,q,\alpha}$ and every $m \geqslant M_{X,q,\alpha}$.

The collection of open sets $\{B_{X,q,\alpha}\}_{(X,q,\alpha)\in Q_1\times Q_2\times[0,1]}$ is an open cover of the compact set $Q_1 \times Q_2 \times [0,1]$, hence a finite subcover B exists. Fix a positive number $M = \max_{B_{X,q,\alpha}\in B} M_{X,q,\alpha}$ and note that (8) holds for every $(X, q, \alpha) \in Q_1 \times Q_2 \times [0,1]$ and every $m \geqslant M$. ■

Theorem 2. *For every set of pure actions A, the* Truncated Reward Scheme *f is optimal.*

Proof. Fix a a strategy $\sigma_1 \in Q\backslash Q_1$ and $\sigma_2, \ldots, s_k \in Q$. There exist $X \in Q_1$ and $q \in Q_2$ such that $\sigma_1 = (1-\alpha)X + \alpha q$, where $\alpha > 0$. Without loss of generality, we relate only to Player 1 and prove that $\mathbf{E}[f_1(X, \sigma_2, \ldots, s_k)] > \mathbf{E}[f_1(\sigma_1, \sigma_2, \ldots, s_k)]$. In words, for every profile of strategies $\sigma = (\sigma_1, \ldots, \sigma_k)$, Player 1 can increase his expected payoff by passing to a diversified action that includes only optimal actions.

By the linearity of the sum in f_1, it suffices to prove that

$$\mathbf{E}[\phi(X))] > \mathbf{E}[\phi((1-\alpha)X + \alpha q)], \tag{9}$$

and every diversified action $\sigma_1 \notin Q_1$, that includes sub-optimal actions q, is dominated by some diversified action $X \in Q_1$.

Assume to the contrary that (9) does not hold, i.e. that

$$\mathbf{E}[\phi(X)] \leqslant \mathbf{E}[\phi((1-\alpha)X + \alpha q)]. \tag{10}$$

Consider the real-valued function $\psi(x) = x - \phi(x)$, and note that $\phi(x) = x - \psi(x)$. Then, Ineq. (10) is recast as

$$\mathbf{E}[X - \psi(X)] \leqslant \mathbf{E}[(1-\alpha)X + \alpha q - \psi((1-\alpha)X + \alpha q)]$$

or, equivalently,

$$\mathbf{E}[X] - \mathbf{E}[(1-\alpha)X + \alpha q] \leqslant \mathbf{E}[\psi(X)] - \mathbf{E}[\psi((1-\alpha)X + \alpha q)]. \tag{11}$$

Since X is a convex combination of optimal actions and q is a convex combination of sub-optimal actions, it follows from (3) that

$$\begin{aligned} \mathbf{E}[X] - \mathbf{E}[(1-\alpha)X + \alpha q] &= \mathbf{E}[X - (1-\alpha)X - \alpha q] \\ &= \alpha\mathbf{E}[X - q] \\ &> \alpha\epsilon. \end{aligned}$$

Combining the last inequality with Ineq. (11) we obtain

$$\begin{aligned} \mathbf{E}[\psi(X) - \psi(X - \alpha(X-q))] &= \mathbf{E}[\psi(X)] - \mathbf{E}[\psi((1-\alpha)X + \alpha q)] \\ &\geqslant \mathbf{E}[X] - \mathbf{E}[(1-\alpha)X + \alpha q] \\ &> \alpha\epsilon. \end{aligned}$$

Denote $\gamma = \psi(X) - \psi(X - \alpha(X-q))$. We contradict the last inequality by showing that $\mathbf{E}[\gamma] < \alpha\epsilon$.

Consider the intervals $I_1 = (-\infty, -M)$, $I_2 = [-M, M]$, and $I_3 = (M, \infty)$. One can write ψ explicitly as

$$\psi(x) = \begin{cases} x + M, & \text{if } x \in I_1, \\ 0, & \text{if } x \in I_2, \\ x - M, & \text{if } x \in I_3. \end{cases}$$

Note that $\psi(x) < 0$ iff $x \in I_1$ and $\psi(x) > 0$ iff $x \in I_3$.

Overall, there are 9 cases we need to consider where $X \in I_i$ and $X - \alpha(X-q) \in I_j$, for every $i, j = 1, 2, 3$ (denote these events by A_{ij}):

Event A_{11}**.** If $i = j = 1$, then

$$\gamma = X + M - [X - \alpha(X-q) + M] = \alpha(X-q).$$

Event A_{33}**.** If $i = j = 3$, then

$$\gamma = X - M - [X - \alpha(X-q) - M] = \alpha(X-q).$$

Event A_{22}**.** If $i = j = 2$, then $\gamma = 0 - 0 = 0$.

Event A_{12}**.** If $i = 1$ and $j = 2$, then $\gamma = X + M < 0$. The inequality $\gamma < 0$ also holds in events A_{23} and A_{13}.

Event A_{32}**.** If $i = 3$ and $j = 2$, then $X - \alpha(X-q) < M$, or equivalently, $X - M < \alpha(X-q)$. This means that $\gamma = \psi(X) = X - M < \alpha(X-q)$.

Event A_{31}**.** If $i = 3$ and $j = 1$, then

$$\begin{aligned} \gamma &= X - M - (X - \alpha(X-q) + M) \\ &= \alpha(X-q) - 2M \\ &< \alpha(X-q). \end{aligned}$$

Event A_{21}**.** If $i = 2$ and $j = 1$, then $-M < X < M$, which implies that $-X - M < 0$. Thus,

$$\begin{aligned} \gamma &= 0 - (X - \alpha(X-q) + M) \\ &= -X - M + \alpha(X-q) \\ &< \alpha(X-q). \end{aligned}$$

This covers all nine possible cases. To conclude, we showed that in A_{i1} when $i = 1, 2$, and in A_{3j} when $j = 1, 2, 3$, the inequality $\gamma < \alpha(X-q)$ holds, and in all other events $\gamma < 0$. Note that $\bigcup_{j=1}^{3} A_{3j} = \{X > M\} \subseteq \{|X| > M\}$ and

$$\begin{aligned} A_{11} \cup A_{21} &= \{X \leqslant M,\ X - \alpha(X-q) < -M\} \\ &= \{X \leqslant M,\ (1-\alpha)X + \alpha q < -M\} \\ &\subseteq \{|(1-\alpha)X + \alpha q| > M\}. \end{aligned}$$

Therefore,

$$
\begin{aligned}
\mathbf{E}[\gamma] &= \sum_{i,j=1}^{3} \mathbf{E}\left(\gamma \mathbf{1}_{A_{ij}}\right) \\
&< \sum_{i=1}^{2} \mathbf{E}\left([\alpha(X-q)]\mathbf{1}_{A_{i1}}\right) + \sum_{j=1}^{3} \mathbf{E}\left([\alpha(X-q)]\mathbf{1}_{A_{3j}}\right) \\
&= \alpha\mathbf{E}[(X-q)\mathbf{1}_{\{X>M\}}] + \alpha\mathbf{E}\left[(X-q)\mathbf{1}_{\{X\leqslant M,\ X-\alpha(X-q)<-M\}}\right] \\
&\leqslant \alpha\mathbf{E}[|X-q|\mathbf{1}_{\{X>M\}}] + \alpha\mathbf{E}\left[|X-q|\mathbf{1}_{\{X\leqslant M,\ X-\alpha(X-q)<-M\}}\right] &(12)\\
&\leqslant \alpha\mathbf{E}[|X-q|\mathbf{1}_{\{|X|>M\}}] + \alpha\mathbf{E}\left[|X-q|\mathbf{1}_{\{|(1-\alpha)X+\alpha q|>M\}}\right] &(13)\\
&< \alpha\frac{\epsilon}{2} + \alpha\frac{\epsilon}{2} = \alpha\epsilon. &(14)
\end{aligned}
$$

Here, Ineq. (12) follows from the abosulte values, Ineq. (13) follows from increasing the subset over which the expected values are taken, and Ineq. (14) follows from (8). A contradiction.

We proved that for every player, every optimal strategy (pure or diversified) $X \in Q_1$ dominates every sub-optimal strategy $\sigma_1 = (1-\alpha)X + \alpha q \in Q \backslash Q_1$. Hence, by eliminating sub-optimal strategies, the players will play only optimal strategies.

Let Δ_{Q_1} be the probability simplex over the respective pure actions in Q_1. That is, every diversified optimal action of a player i could be represented by a probability vector in Δ_{Q_1}. By the fixed point theorem on a convex compact set Δ_{Q_1}, we know that an equilibrium exists, and the result follows. ∎

Theorem 3. *Fix $M > 0$ and assume that for every subset A of strictly-bounded actions (i.e., $\Pr(|X_i| > M) = 1$ for every action X_i), the investment game G_f has a dominant-strategy optimal equilibrium. Then, $f_i(r)$ is linear in r_i for every r_{-i}.*

Proof. Without loss of generality, fix $i = 1$ and take $r_{-i} = (r_2, \ldots, r_k) \in (-M, M)^{k-1}$ such that $r_i \geqslant r_{i+1}$ for every $2 \leqslant i \leqslant k-1$. Define $g(t) = f_1(t, r_{-i})$. We need to prove that $g(t)$ is linear.

Fix $r_1 > r_2$ and consider $-M < y < x < M$ and $\lambda \in (0,1)$ such that $r_1 = \lambda x + (1-\lambda)y$. We start by proving that $g(\lambda x + (1-\lambda)y) = \lambda g(x) + (1-\lambda)g(y)$.

Define the constant random variables $X_j = r_j$ for every $1 \leqslant j \leqslant k$. Take $\epsilon > 0$ such that $\lambda + \epsilon < 1$ and define Z_+ such that

$$
Z_+ = \begin{cases} x, & \text{w.p. } \lambda + \epsilon, \\ y, & \text{w.p. } 1 - \lambda - \epsilon. \end{cases}
$$

Clearly, $\mathbf{E}[Z_+] > \mathbf{E}[X_j]$ for every $1 \leqslant j \leqslant k$. Let $A_+ = \{X_1, X_2, \ldots, X_k, Z_+\}$ be a set of strictly-bounded actions. Since an optimal dominant-strategy equilibrium exists, it follows that Z_+ is a dominant strategy of player 1. Thus,

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= f_1(\lambda x + (1-\lambda)y, r_{-i}) \\ &= \mathbf{E}[f_1(X_1, X_2, \ldots, X_k)] \\ &\leqslant \mathbf{E}[f_1(Z_+, X_2, \ldots, X_k)] \\ &= (\lambda+\epsilon) f_1(x, r_{-i}) + (1-\lambda-\epsilon) f_1(y, r_{-i}) \\ &= (\lambda+\epsilon) g(x) + (1-\lambda-\epsilon) g(y). \end{aligned}$$

Taking the limit when ϵ tends to 0, gives the inequality $g(\lambda x + (1-\lambda)y) \leqslant \lambda g(x) + (1-\lambda)g(y)$.

Now, take $\epsilon > 0$ such that $\lambda - \epsilon > 0$ and define Z_- such that

$$Z_- = \begin{cases} x, & \text{w.p. } \lambda - \epsilon, \\ y, & \text{w.p. } 1 - \lambda + \epsilon. \end{cases}$$

Let $A_- = \{X_1, X_2, \ldots, X_k, Z_-\}$ be a set of strictly-bounded actions, and similarly to the previous reasoning, we get that X_1 is a dominant strategy, as $\mathbf{E}[X_1] = r_1 > \mathbf{E}[Z_-]$. Therefore,

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= f_1(\lambda x + (1-\lambda)y, r_{-i}) \\ &= \mathbf{E}[f_1(X_1, X_2, \ldots, X_k)] \\ &\geqslant \mathbf{E}[f_1(Z_-, X_2, \ldots, X_k)] \\ &= (\lambda-\epsilon) f_1(x, r_{-i}) + (1-\lambda+\epsilon) f_1(y, r_{-i}) \\ &= (\lambda-\epsilon) g(x) + (1-\lambda+\epsilon) g(y). \end{aligned}$$

Taking the limit where $\epsilon \to 0$, it follows that $g(\lambda x + (1-\lambda)y) \geqslant \lambda g(x) + (1-\lambda)g(y)$. To conclude, we proved that $g(\lambda x + (1-\lambda)y) = \lambda g(x) + (1-\lambda)g(y)$ for the specific case where $\lambda x + (1-\lambda)y = r_1 \geqslant r_2$. Nevertheless, this result holds for every $y < r_1 < x$, and a straightforward examination shows that

$$\frac{g(y) - g(r_1)}{y - r_1} = \frac{g(x) - g(r_1)}{x - r_1} = \frac{g(y) - g(x)}{y - x},$$

which implies linearity, as required. ∎

Theorem 4. *For every set of actions A such that U is bounded and uniquely maximized, the following* General Reward Scheme f *is U-optimal:*

$$f_i(r) = \frac{1}{k} + \frac{1}{2Mk}\left[U(kr_i) - \frac{1}{k-1}\sum_{j\neq i} U(kr_j)\right]\mathbf{1}_{\{\forall i,\ |U(kr_i)|\leqslant M\}}.$$

Proof. Let $q^* \in Q$ be the unique diversified action such that $E[U(kq^*)] > E[U(kq)]$ for every $q \in Q\backslash\{q^*\}$. We prove that for every Player i, for every profile of diversified actions $(\sigma_1, \sigma_2, \ldots, \sigma_k) \in Q^k$ of players $1, \ldots, k$ respectively, and for every strategy $\sigma_i \neq q^*$ of Player i, the inequality

$$\mathbf{E}\left[f_i\left(\sigma_1, \ldots, \sigma_{i-1}, q^*, \sigma_{i+1}, \ldots, \sigma_k\right)\right] > \mathbf{E}\left[f_i\left(\sigma_1, \ldots, \sigma_k\right)\right],$$

holds.

Similarly to the proof of Theorem 1, assume w.l.o.g. that $i = 1$ and

$$\begin{aligned}
\mathbf{E}\left[f_1\left(q^*, \sigma_2, \ldots, \sigma_k\right)\right] &= \mathbf{E}\left[\frac{1}{k} + \frac{1}{2Mk}\left(U(kq^*) - \frac{1}{k-1}\sum_{j=2}^{k} U(k\sigma_j)\right)\right] \\
&= \frac{1}{k} + \frac{1}{2Mk}\left[\mathbf{E}\left[U(kq^*)\right] - \frac{1}{k-1}\sum_{j=2}^{k}\mathbf{E}\left[U(k\sigma_j)\right]\right] \\
&> \frac{1}{k} + \frac{1}{2Mk}\left[\mathbf{E}\left[U(k\sigma_1)\right] - \frac{1}{k-1}\sum_{j=2}^{k}\mathbf{E}\left[U(k\sigma_j)\right]\right] \\
&= \mathbf{E}\left[\frac{1}{k} + \frac{1}{2Mk}\left(U(k\sigma_1) - \frac{1}{k-1}\sum_{j=2}^{k} U(k\sigma_j)\right)\right] \\
&= \mathbf{E}\left[f_1\left(\sigma_1, \sigma_2, \ldots, \sigma_k\right)\right].
\end{aligned}$$

The inequality implies that q^* is a strictly dominant strategy for every player i. Thus, $(q^*, \ldots, q^*) \in Q^k$ is a dominant-strategy equilibrium. Since all the players use the same diversified action q^* in the unique equilibrium σ, Eq. (7) is satisfied and the result follows. ■

Theorem 5. *In the case of two players, there is no non-constant universal reward scheme.*

Proof. Let $x < y < z$. We first prove that $f_1(y,y) \geqslant f_1(x,y)$. Assume to the contrary that $f_1(y,y) < f_1(x,y)$. Let $A = \{X_1, X_2\}$ be a set of two actions X_1 and X_2 with a joint probability distribution

$X_1 \backslash X_2$	x	y
x	0	0
y	1	0

Note that $\mathbf{E}[X_1] > \mathbf{E}[X_2]$. However,

$$\mathbf{E}[f_1(X_2, X_1)] = f_1(x,y) > f_1(y,y) = \mathbf{E}[f_1(X_1, X_1)],$$

implying that (X_1, X_1) is not an equilibrium in G_f, since Player 1 can benefit from deviating to X_2. Thus,

$$f_1(y,y) \geqslant f_1(x,y). \tag{15}$$

For similar reasons,

$$f_2(y,y) \geqslant f_2(y,x). \tag{16}$$

Next we prove that $f_1(x,x) \geqslant f_1(y,x)$. Let p be a number in $(0,1)$ and let $A = \{X_1, X_2\}$ be a set of two actions X_1 and X_2 with a joint probability distribution,

$X_1 \backslash X_2$	x	y
x	0	p
z	$1-p$	0

A direct computation shows that

$$\mathbf{E}[f_1(X_1, X_1)] - \mathbf{E}[f_1(X_2, X_1)] \quad = \quad p(f_1(x,x) - f_1(y,x)) + (1-p)\,(f_1(z,z) - f_1(x,z)).$$

Recall that $f_1(z,z) - f_1(x,z)$ is bounded. If $f_1(x,x) < f_1(y,x)$, then there is p smaller than, but sufficiently close to 1, such that for every z, $\mathbf{E}[f_1(X_1, X_1)] - \mathbf{E}[f_1(X_2, X_1)] < 0$. In other words,

$$\mathbf{E}[f_1(X_1, X_1)] < \mathbf{E}[f_1(X_2, X_1)]. \tag{17}$$

Now one can choose z to be sufficiently large, so that $\mathbf{E}[X_1] > \mathbf{E}[X_2]$. Inequality (17) implies that (X_1, X_1) is not an equilibrium in G_f, since Player 1 can benefit from deviating to X_2. Hence,

$$f_1(x,x) \geqslant f_1(y,x). \tag{18}$$

A similar argument shows that

$$f_2(x, x) \geqslant f_2(x, y). \tag{19}$$

We now sum up inequalities (15), (16), (18), and (19) to obtain, $f_1(y, y) + f_2(y, y) + f_1(x, x) + f_2(x, x) \geqslant f_1(x, y) + f_2(y, x) + f_1(y, x) + f_2(x, y)$. Due to Eq. (2), equality holds. Thus, (15), (16), (18), and (19) are actually equalities. Therefore,

$$f_1(x, x) = f_1(x, y) = f_1(y, x) = f_1(y, y),$$

and the proof is complete. ■

Theorem 6. *If f is a strongly universal reward scheme, then every profile of strategies is an equilibrium.*

Proof. Let f be a strongly universal reward scheme. Clearly, Theorem 5 implies that the result holds for the case of $k = 2$. Fix $k \geqslant 3$. We prove the theorem by showing that for every player i and for every vector of outcomes $r \in \mathbb{R}^k$, the ith coordinate $f_i(r)$ of the reward scheme is non-decreasing and non-increasing in r_i.

Assume to the contrary that there exists a player i, a vector of outcomes $r \in \mathbb{R}^k$, and $w_i \in \mathbb{R}$, such that $f_i(w_i, r_{-i}) > f_i(r_i, r_{-i})$ where $w_i < r_i$. Define the random variable X such that $\Pr(X = x) > 0$ if $x = r_j$ when $1 \leqslant j \leqslant k$. Assume that $\Pr(X = r_i) > \Pr(X = r_j)$ for every $j \neq i$. In addition, define a set of i.i.d. random variables $X_j \sim X$ where $1 \leqslant j \leqslant k$. Define the vector-valued random variable (W, X_{-i}) by

$$\Pr((W, X_{-i}) = x) = \Pr((X_i, X_{-i}) = x), \quad \forall x \neq r,$$

and

$$\Pr((W, X_{-i}) = (w_i, r_{-i})) = \Pr((X_i, X_{-i}) = r).$$

Clearly, (W, X_{-i}) and W are well defined. A direct computation shows that $\mathbf{E}[W] < \mathbf{E}[X]$. However, the vector (X_i, X_{-i}) is not an equilibrium, as Player i can deviate to W and increase his payoff, since $f_i(w_i, r_{-i}) > f_i(r_i, r_{-i})$. Hence, $f_i(\cdot, r_{-i})$ is non-decreasing for every i and every r_{-i}.

Now assume to the contrary that $f_i(r_i, r_{-i})$ is strictly increasing in r_i. That is, there exists a player i, a vector of outcomes $r \in \mathbb{R}^k$, and $y_i \in \mathbb{R}$, such that $f_i(y_i, r_{-i}) > f_i(r_i, r_{-i})$ where $y_i > r_i$.

Let $\bar{z}, \underline{z} \in \mathbb{R}$ be two real numbers such that $\bar{z} > r_j > \underline{z}$ for every $1 \leqslant j \leqslant k$ and let p be a number in $(0,1)$. Define the random variable Y such that, w.p. p, it follows that $\Pr(Y = r_j) > 0$ for every $1 \leqslant j \leqslant k$. Assume that $\Pr(Y = r_i) > \Pr(Y = r_j)$ for every $j \neq i$. In addition, w.p. $1-p$, the random variable Y equals $\bar{z}$. Define a set of i.i.d. random variables $Y_j \sim Y$ where $1 \leqslant j \leqslant k$. Define the vector-valued random variable (Z, Y_{-i}) by

$$\Pr((Z, Y_{-i}) = y) = \Pr((Y_i, Y_{-i}) = y) \;\; \forall y \neq r, y_j \neq \bar{z} \; \forall j,$$

$$\Pr((Z, Y_{-i}) = (y_i, r_{-i})) = \Pr((Y_i, Y_{-i}) = r),$$

and if there exists a coordinate j of $y \in \mathbb{R}^k$ such that $y_j = \bar{z}$, then

$$\Pr((Z, Y_{-i}) = (\underline{z}, y_{-i})) = \Pr((Y_i, Y_{-i}) = y).$$

Clearly, (Z, Y_{-i}) and Z are well defined. Note that

$$\begin{aligned}
\mathbf{E}[f_i(Z, Y_{-i})] &= \mathbf{E}\left[f_i(Y_i, Y_{-i})\mathbf{1}_{\{Y_{-i} \neq r, Y_j \neq \bar{z} \;\, \forall j\}}\right] + f_i(y_i, r_{-i})\Pr((Y_i, Y_{-i}) = r) \\
&+ \sum_{\substack{y \in \mathbb{R}^k: \\ \exists j, y_j = \bar{z}}} f_i(\underline{z}, y_{-i})\Pr((Y_i, Y_{-i}) = y) \\
&> \mathbf{E}\left[f_i(Y_i, Y_{-i})\mathbf{1}_{\{Y_j \neq \bar{z} \;\, \forall j\}}\right] + \sum_{\substack{y \in \mathbb{R}^k: \\ \exists j, y_j = \bar{z}}} f_i(\underline{z}, y_{-i})\Pr((Y_i, Y_{-i}) = y) \quad (20) \\
&= \mathbf{E}\left[f_i(Y_i, Y_{-i})\right] + \sum_{\substack{y \in \mathbb{R}^k: \\ \exists j, y_j = \bar{z}}} (f_i(\underline{z}, y_{-i}) - f_i(y))\Pr((Y_i, Y_{-i}) = y), \quad (21)
\end{aligned}$$

where Ineq. (20) follows from the assumption that $f_i(y_i, r_{-i}) > f_i(r_i, r_{-i})$. The sum in Eq. (21) is bounded, therefore we can choose a p sufficiently close to 1 (but still smaller than 1), such that $\mathbf{E}[f_i(Z, Y_{-i})] > \mathbf{E}\left[f_i(Y_i, Y_{-i})\right]$ for every $\bar{z}$. Taking a sufficiently large $\bar{z}$ and a sufficiently low $\underline{z}$ guarantees that $\mathbf{E}[Y] > \mathbf{E}[Z]$.

In conclusion, the vector (Y_i, Y_{-i}) is not an equilibrium, as Player i can deviate to Z and increase his payoff. A contradiction. Hence, $f_i(\cdot, r_{-i})$ is non-increasing for every i and every r_{-i}. The combination of the two results proves that f_i is independent of the ith coordinate. This implies that the expected payoff of every player i is independent of the player's actions and that every profile of actions is an equilibrium. ■