Continuous-time incentives in hierarchies

Emma Hubert

Received: 6 September 2021 / Accepted: 7 February 2023 / Published online: 15 June 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
This paper studies continuous-time optimal contracting in a hierarchy, generalising the model of Sung (Math. Financ. Econ. 9:195–213, 2015). More precisely, in this hierarchical model, the principal (she) can contract with a manager (he) to incentivise him to act in her best interest, despite only observing the net benefits of the total hierarchy. The manager in turn subcontracts with the agents below him. Both the agents and the manager independently control in continuous time a stochastic process representing their outcome. First, we show through this continuous-time adaptation of Sung’s model that even if the agents only control the drift of their outcome, their manager controls the volatility of their continuation utility by choosing their contract sensitivities. This first illustrative example justifies the use of recent results by Cvitanić et al. (Finance Stoch. 22:1–37, 2018) on optimal contracting for drift and volatility control to carefully study continuous-time incentive problems in hierarchy. Some technical and numerical comparisons are provided to highlight the differences with Sung’s model. Then, in a second more theoretical part, we provide the methodology to tackle a more general hierarchy model. The solution is based on the theory of second-order backward stochastic differential equations (2BSDEs), and extends the results in (Cvitanić et al. in Finance Stoch. 22:1–37, 2018) to a multitude of agents with non-trivial interactions, especially concerning volatility control.

Keywords Principal–agent problem · Moral hazard · Hierarchical contracting · Second-order BSDEs

Mathematics Subject Classification (2020) 91A65 · 91B41 · 60H30 · 93E20

JEL Classification C73 · D82 · D86

The author gratefully acknowledges the support of the ANR project PACMAN ANR-16-CE05-0027, the FACE Foundation – Thomas Jefferson Fund, as well as the support of University Gustave Eiffel through a mobility grant.

✉️ E. Hubert
eh3988@princeton.edu

1 Department of Operations Research and Financial Engineering (ORFE), Princeton University, Princeton, NJ 08544, USA
1 Introduction

1.1 A little bit of context

In an organisation, a hierarchy usually consists of a power entity at the top with subsequent levels of power underneath. This particular structure, which is a dominant mode in our contemporary society, raises many questions on its efficiency, its cost, its optimal size, etc. To answer these questions, an abundant literature has emerged in the last century in a wide variety of fields, from philosophy to mathematics, through social and management sciences. The first mathematical model for the study of the optimal hierarchy seems to be the work of Williamson [59], but as he mentioned, this question, which presents a serious dilemma for business theory, was originally introduced by Knight in 1933 (see Knight [37, preface to the re-issue] for a recent edition). Many authors have followed this trend, including Calvo and Wellisz [11, 12] and Keren and Levhari [36], as well as Qian [52] who takes into account the notion of incentives.

In the meantime, mathematical models on incentive theory became widespread in the 1970s, especially through the work of Mirrlees [46], and were applied a few years later to hierarchical organisations by Stiglitz [55] and Mirrlees [47]. Incentive theory is strongly related to contract theory and principal–agent problems, and is associated with a vast literature that cannot be mentioned here for the sake of conciseness. In the case of a hierarchy, we are dealing with a succession of interlinked principal–agent problems, or in other words, a sequence of nested Stackelberg equilibria. The interest of this mathematical formalism lies in the modelling of information asymmetries within a hierarchy, in particular moral hazard, as in the works by Laffont [39], Yang [60], Macho-Stadler and Pérez-Castrillo [44], Itoh [32] and Jost and Lammers [34]. However, it should be noted that the above-mentioned works are discrete-time models.

1.2 Moving to continuous time

In the late 1980s, the literature on contract theory expanded to include models in continuous time. The first, and seminal, paper of this class is by Holmström and Milgrom [28]. This work was then extended, and main contributors in these regards are Schättler and Sung [54], Sannikov [53], Biais et al. [8] as well as Cvitanić and Zhang [18, Part III]. The most recent breakthrough is due to Cvitanić et al. [15] who have developed a general theory that allows to address a wide spectrum of principal–agent problems. With their approach, the problem faced by the principal becomes a standard optimal control problem. Most importantly, this method enables to tackle volatility control problems. It has subsequently been extended and applied in many different situations. We can mention in a non-exhaustive way the applications to finance by the same authors in [14] and by Cvitanić and Xing [17], the works of Aid et al. [1] and Alasseur et al. [2] for applications related to the energy sector, as well as other various extensions, e.g. the works of Hernández Santibañez and Mastrolia [27] and Hu et al. [29]. Recently, principal–agent problems in continuous time have been enhanced to models with several agents, through the works of Koo et al. [38],...
Élie and Possamaï [23] and Baldacci et al. [4], and even to a continuum of agents with mean-field interactions by Élie et al. [22], Carmona and Wang [13] and Élie et al. [21].

This extension to a large number of agents is a significant step towards the application of continuous-time contract theory to hierarchies. Nevertheless, this type of models seems for the moment to be countable on one hand. First, Miller and Yang [45] consider a hierarchy of \( n + 1 \) players, each with a principal–agent relationship. Using the approach of Evans et al. [24], they identify conditions under which a dynamic programming construction of an optimal contract can be reduced to only a one-dimensional state space and one-dimensional control set, independent of the size of the hierarchy. However, the approach in [24] to characterise optimal contracts is less general than the one in Cvitanić et al. [15] on which we rely in the present paper. In particular, it does not allow volatility control, which seems inescapable in our framework. Then Li and Yu [41] develop a method using FBSDEs to characterise the equilibrium of a generalised Stackelberg game with multi-level hierarchy in a linear–quadratic setting. Finally, Keppo et al. [35] have developed (concurrently with the present paper) a hierarchical principal–agent problem similar to the two aforementioned models; but the mathematical approach is more related to the one we consider here, since it is based on [15], to take into account that the partner controls the volatility of the output. Nevertheless, in these three hierarchical models in continuous time, it is assumed that the entities of the hierarchy observe and control the same output process.

1.3 Sung’s model

The framework of the present work is inspired by the model developed by Sung [58], where a manager is hired by a principal to subcontract with \( n \) agents and each worker (manager and agents) controls his own output process. This model includes a bi-level moral hazard. First, the manager does not observe the effort of the agents, but only the resulting outputs. Second, the principal observes only the net benefit of the hierarchy, defined by the difference between the sum of the outputs of all workers (later denoted by \( X^i \) for the \( i \)th worker) and the sum of the contracts paid to the agents (\( \xi^i \) for agent \( i \)), i.e.,

\[
\zeta = \sum_{i=0}^{n} X^i - \sum_{i=1}^{n} \xi^i. \tag{1.1}
\]

Instead of studying a continuous-time version of the model, Sung states that “for ease of exposition and without loss of generality, we formulate a discrete-time model which is analogous to its continuous-time counterpart” [58, p. 2]. Extending the reasoning of Holmström and Milgrom [28], he therefore restricts the study to linear contracts, in the sense that they are linear with respect to the outcome, and states that “this assumption is without loss of generality, as long as our results are interpreted in the context of continuous-time models” [58, p. 3].

However, while the restriction to linear contracts can be justified in Sung’s framework for the first Stackelberg equilibrium, this is no longer the case for the contract
offered by the principal to the manager: although the agents control only the drift of their outcomes, the manager controls both the drift and the volatility of the net benefit $\zeta$ defined above. Indeed, as the manager chooses the contract sensitivity for the agents, he impacts the volatility of each contract $\xi_i$ and thus the volatility of $\zeta$. Therefore, according to Cvitanić et al. [15], the type of contracts considered by Sung is sub-optimal: in continuous time, it is not sufficient to limit oneself to linear contracts (in the sense of [28]) when the volatility of the state variable is controlled. In fact, the optimal form of contracts should contain an additional part indexed on the quadratic variation of the net benefit $\zeta$, in order to incentivise the manager to design appropriate contracts for the agents. However, in the one-period model of Sung, this controlled quadratic variation cannot be estimated (unlike in continuous time), which leads to a fundamental gap between these two frameworks.

1.4 Main contributions

This paper consists of two parts. We first present the continuous-time version of Sung’s model [58] in order to highlight the aforementioned differences, and to motivate the rigorous study of continuous-time general hierarchical problems in the second part. More precisely, the continuous-time version of Sung’s model is introduced and solved in Sect. 2. This opening example highlights the differences between the one-period model and its continuous-time equivalent, especially regarding the volatility control and the form of the contracts. A discussion on these differences is provided in Sect. 3, both from a technical and an economic point of view, and illustrated by numerical results. In particular, this example leads to the conclusion that in order to rigorously study a continuous-time hierarchy problem, it is not possible to consider the associated one-period model with linear contracts. In our opinion, these conclusions justify the use of the theory of second-order backward stochastic differential equations (2BSDEs) to tackle problems of moral hazard within a hierarchy.

Motivated by the previous results, the second part of this paper provides a systematic method to solve a more general class of continuous-time incentive problems in a hierarchy. In particular, in Sect. 4, we consider a more comprehensive structure of hierarchy with $m$ managers, who in turn subcontract with the agents in their teams. Moreover, the workers now control both the drift and the volatility of their output. Finally, we consider general preferences, allowing us to recover the exponential utility functions (CARA) of Sung’s model [58], but also other cases such as risk-neutral preferences. The search for the optimal contract in such a framework triggers the use of the theory of 2BSDEs as in Cvitanić et al. [15]. However, the results of [15] do not apply directly here as we have to address the multi-agent case. In particular, at the managers’ level, since their stochastic optimisation problems are coupled, the managers’ continuation utilities satisfy a multidimensional 2BSDE, which is an extension of the pair of 2BSDEs considered by Possamaï et al. [51, Eqs. (3.22) and (3.23)] in a zero-sum game with two interacting players. Therefore, the present paper is the first to deal with the principal–multi-agent case with volatility control, and especially with non-trivial interactions between the agents concerning this volatility control. This explains the relevance of defining and solving the problem thoroughly. However, the focus of Sect. 4 is primarily methodological, and we therefore report some technical results and important proofs in the Appendix.
Throughout this study, we determine the optimal form of continuous-time incentives for a particular hierarchical structure, which can be extended in a straightforward way to a larger-scale hierarchy. Although theoretical, the results we obtain give intuitions based on solid conceptual considerations to know which levers could be activated to incentivise workers within a hierarchy. In particular, as explained further in Sect. 3.3, the indexation of the contract on the quadratic variation of net profits for the managers argues in favour of remunerating them through stock options, known as employee stock option (ESO), in addition to regular shares of stock, as highlighted by Lazear [40] for example. These results can be applied to problems of incentives in firms with a hierarchical structure, but also and above all as soon as work is delegated to an external entity. For example, these multi-layered incentive problems can be used to model the relationships between a firm and its subsidiaries/subcontractors, or between insurance/reinsurance companies and policyholders.

2 An opening example: Sung’s model in continuous time

In order to illustrate the differences between discrete- and continuous-time hierarchy models, and thus justify the motivation for studying a general continuous-time model in Sect. 4, we first present a simple hierarchical contracting problem, inspired by Sung [58]. In this model, the principal contracts with one manager who in turn subcontracts with \( n \) agents, as illustrated in Fig. 1. Despite its simplicity, this illuminating example shows the need to take volatility control into account to determine optimal incentives. As the main purpose of this section is to illustrate the relevance of a continuous-time study, the reasoning will for now remain informal; we refer to Sects. 4 and 5 for a rigorous model setting and the exact methodology to solve it.

For ease of comparison between discrete- and continuous-time models, we consider in this section the continuous-time equivalent of Sung’s one-period model [58]. Between 0 and some time \( T > 0 \), denoting the maturity fixed in the contract, the firm has \( n + 1 \) tasks which have to be carried out by \( n + 1 \) workers. The outputs of the tasks are represented by \( n + 1 \) stochastic processes, denoted by \( X_i \) for \( i \in \{0, \ldots, n\} \), with dynamics

\[
\text{d}X_i^t = \alpha_i^t \text{d}t + \sigma_i^t \text{d}W_t^i, \quad \sigma_i^t > 0, t \in [0, T], \text{ and } X_i^0 = 0.
\]

More precisely, the \( i \)th worker carries out the task with outcome \( X_i \) by choosing a costly effort \( \alpha_i \in \mathcal{A}_i \) taking values in \( \mathbb{R} \), where \( \mathcal{A}_i \) is the set of admissible control processes. For simplicity, we assume that \( W_i \) for \( i \in \{0, \ldots, n\} \) are independent.

![Fig. 1 Hierarchy in Sung’s model [58]](image-url)
Brownian motions and that the efforts of a worker only impact his own project, which means that all projects are independent and no workers collude. We also consider the quadratic cost of effort

\[ c^i(a) = a^2 / (2k^i), \quad k^i > 0, \quad \text{for} \; a \in \mathbb{R}, i \in \{0, \ldots, n\}. \]

The benefit of each worker \( i \) is represented by a CARA utility function with risk aversion coefficient \( R^i > 0 \). For convenience, we later use the notation \( \tilde{R}^i := k^i + R^i |\sigma^i|^2 \) for \( i \in \{0, \ldots, n\} \). A worker is also characterised by his reservation utility, i.e., the minimum level of utility that must be guaranteed to that worker; otherwise he will refuse to perform the task. The principal is risk-neutral and seeks to maximise the expected difference between the sum of the outcomes and the sum of the compensations paid to the workers.

In a **direct contracting case** (DC), the principal can directly contract with the agents, without the help of a manager. This traditional principal–multi-agent problem can be solved in a straightforward way, since only the drift of the output is controlled, by extending the results obtained by Holmström and Milgrom [28] to a multitude of agents (see Koo et al. [38] for a rigorous result, or the more general model of Êlie and Possamaï [23]). This case is also discussed by Sung [58], but the main point of [58] is to study the case where the principal contracts with a manager, who in turn subcontracts with the agents.

In the **hierarchical contracting case** (HC) considered in [58], the principal cannot directly contract with the workers. She hires a manager (the worker indexed by \( i = 0 \)) who

(i) carries out his own task with output \( X^0 \) by choosing an effort process \( \alpha^0 \in \mathcal{A}^0 \);
(ii) hires \( n \) agents to carry out the \( n \) remaining tasks; each agent \( i \) handles the task with outcome \( X^i \), by choosing his effort level \( \alpha^i \in \mathcal{A}^i \), for \( i \in \{1, \ldots, n\} \);
(iii) reports to the principal the net benefit \( \zeta \), i.e., the difference between the sum of the outcomes and the sum of the compensations to be paid to the agents, as defined in (1.1).

We show below that in the continuous-time framework, we can improve the results in [58] by considering a more general form of contracts. As already mentioned, this gap is strongly related to the definition of \( \zeta \): by choosing the agents’ contract sensitivity, the manager actually controls the volatility of \( \zeta \), which necessitates to introduce a more sophisticated form of contracts, as demonstrated in Cvitanić et al. [15].

### 2.1 A continuous-time principal–manager–agents problem

In this section, we define the continuous-time equivalent of the agents’, manager’s and principal’s optimisation problems considered in Sung [58].

Each agent \( i \in \{1, \ldots, n\} \) controls his own output \( X^i \) with dynamics (2.1) by choosing an admissible effort \( \alpha^i \in \mathcal{A}^i \). Given a contract \( \xi^i_T \) offered by his manager, his value function is defined by

\[
V^i_0(\xi^i_T) := \sup_{\alpha^i \in \mathcal{A}^i} J^i_0(\xi^i_T, \alpha^i), \quad \text{where} \; J^i_0(\xi^i_T, \alpha^i) := \mathbb{E} \left[ -e^{-R^i(\xi^i_T - \int_0^T c^i(\alpha^i_t) \, dt)} \right].
\]
As in Sung [58], we define the reservation utility of the $i$th worker by his utility without contract, i.e., here $V_0^i(0) = -1$. We assume that the manager only observes the outcome processes $X_i$ for $i \in \{1, \ldots, n\}$ and not directly the efforts of the agents, which implies the first level of moral hazard. In order to follow Sung’s model as closely as possible, we also assume that the compensation for the $i$th agent can only be indexed on his own performance, i.e., his outcome process $X_i$, and denote the set of admissible contracts by $C^i$.

The manager controls his own output $X_0$ through an effort $\alpha_0^0 \in A^0$. He also designs the compensations for the agents, namely the collection of contracts

$$\xi_A^T \in C^A := \{ (\xi_i^T)_{i=1}^n \text{ such that } \xi_i^T \in C^i \text{ for all } i \in \{1, \ldots, n\} \}.$$ 

Although we consider that the manager designs the contracts for the agents, all compensations, whether for the agents or the manager, are paid by the principal. Given a contract $\xi_0^T$ designed by the principal, the manager’s value function is defined by

$$V_0^0(\xi_0^T) := \sup_{(\alpha_0^0, \xi_A^T) \in A^0 \times C^A} J_0(\xi_0^T, \alpha_0^0, \xi_A^T),$$

with

$$J_0(\xi_0^T, \alpha_0^0, \xi_A^T) := \mathbb{E}\left[-e^{-R_0(\xi_0^T - \int_0^T c^0(\alpha_t)dt)}\right].$$

As in [58], the second level of moral hazard is related to the fact that the manager only reports to the principal the net benefit $\zeta$. Therefore, the principal does not separately observe all the outcomes of the workers, or their efforts, or even the precise compensations paid to each agent. This assumption is natural in this hierarchical framework, since the main idea is that the principal delegates control to the manager. Therefore, the manager carefully observes the outcomes of each agent to determine their individual payment, but only summarises to the principal the global result of his working team, hence this net benefit $\zeta$. In a continuous-time framework, $\zeta$ is the process defined by

$$\zeta_t = \sum_{i=0}^n X_i^T - \sum_{i=1}^n \xi_i^T, \quad t \in [0, T]. \quad (2.2)$$

where for all $i \in \{1, \ldots, n\}$, $(\xi_i^T)_{t \in [0, T]}$ corresponds to the certainty equivalent of the $i$th agent’s continuation utility, which is defined below in (2.3). The fact that the principal observes $\zeta$ at least at the terminal time $T$ is very natural, since it is simply the sum of the total profit minus the payments to be made to agents. Nevertheless, its definition in continuous time requires further justification. The certainty equivalent of the $i$th agent’s continuation utility at time $t \in [0, T]$, namely $\xi_i^T$, can in fact be seen, as mentioned by He [26], as the agent’s deferred compensation: it is the monetary amount that needs to be set aside at time $t$ in order to provide the agent with the corresponding terminal payment $\xi_i^T$ at the end of the contracting period. In other words, the net profit at time $t$ is the difference between the total cash flow received until $t$ and the deferred compensation fund for the agents.

Finally, we consider a risk-neutral principal whose objective is to maximise the total net benefit of the hierarchy by choosing the optimal contract $\xi_0^T$ for the manager.
As mentioned above, the principal only observes the process $\zeta$ defined by (2.2), and therefore the contract $\xi_t^0$ for the manager can only be indexed on $\zeta$. The corresponding set of admissible contracts is denoted by $C^0$. In particular, we shall see later in this example that by designing the compensation for the agents, namely the collection $\xi_T^A$, the manager impacts the volatility of the net benefit $\zeta$. More precisely, the volatility of each $\xi^i$, and therefore of $\zeta$, depends on the contract sensitivity chosen by the manager for the $i$th agent. This is precisely why we have to consider more general contracts, indexed on the quadratic variation of $\zeta$, so that the principal can encourage the manager to offer decent contracts to the agents.

Mathematically speaking, the principal’s problem is given by

$$V_0^P := \sup_{\xi_t^0 \in C^0} J_0^P(\xi^0_T),$$

with $J_0^P(\xi^0_T) := \mathbb{E}\left[\sum_{i=0}^{n} X_T^i - \sum_{i=1}^{n} \xi_T^i - \xi^0_T\right] = \mathbb{E}[\zeta_T - \xi^0_T]$.

**Remark 2.1** In Sung [58], the principal only observes the terminal total benefit, i.e., $\zeta_T$ for $T = 1$. This assumption is obviously quite natural for a one-period framework. When moving to continuous time, it is standard to assume that the whole output process, between time 0 and time $T$ is observable, and therefore that the contract $\xi_T^0$ can depend on the paths of the output process until time $T$. We strongly believe that Sung has a continuous-time model of this form in mind when he says that “for ease of exposition and without loss of generality, we formulate a discrete-time model which is analogous to its continuous-time counterpart”, since he then refers to the works by Schättler and Sung [54], Sannikov [53] and Cvitanić et al. [16]. The other possibility would be to assume that only the final value of the outputs is observable and contractible, i.e., $X_T^i$ for $i \in \{1, \ldots, n\}$ and $\zeta_T$; but in our opinion, this would go against the very notion of a continuous-time model, which is conventionally considered as the limiting case of a discrete-time formulation with many periods. In some continuous-time models, such as Holmström and Milgrom [28], the optimal contract ultimately boils down to a linear indexation on the final value of the output, but this result does not hold in more general frameworks.

### 2.2 Solving the sequence of Stackelberg equilibria

In order to solve the above principal–manager–agents problem, one can follow the general theory developed in Cvitanić et al. [15] for each Stackelberg equilibrium, starting with the manager–agent problem.

#### 2.2.1 Revealing contract for an agent

Consider $i \in \{1, \ldots, n\}$ to focus on the contract for the $i$th agent. Recall that his contract $\xi_T^i$ is assumed to be a measurable function of the paths until $T$ of his output $X^i_T$. By applying classical results of contract theory (see e.g. the aforementioned papers by Holmström and Milgrom [28], Schättler and Sung [54], Sannikov [53]),
the optimal form of contract is the terminal value $\xi_i^T$ of the certainty equivalent of the continuation utility, defined for some $\xi_0^i \in \mathbb{R}$ by

$$\xi_i^t = \xi_0^i - \int_0^t \mathcal{H}^i(Z_t^i)\,ds + \int_0^t Z_t^i\,dX_t^i + \frac{1}{2} R^i \int_0^t |Z_s^i|^2 d\langle X^i \rangle_s, \quad t \in [0, T]. \tag{2.3}$$

where $Z^i_t$ is a payment rate chosen by the manager and $\mathcal{H}^i(z) := \sup_{a \in \mathbb{R}} (az - c^i(a))$ for all $z \in \mathbb{R}$ is the Hamiltonian of the $i$th agent. This form of contract is exactly the continuous-time equivalent of the linear contract considered by Sung in [58, Eq. (7)]. Note that the compensation includes a fixed part $\xi_0^i \in \mathbb{R}$, which is chosen so as to satisfy the $i$th agent’s participation constraint. This form of contract leads us to define by $\mathcal{V}^i$ the collection of all $\mathbb{R}$-valued processes $Z^i_t$, predictable with respect to the filtration generated by $X^i$, and satisfying appropriate integrability conditions. We also denote $\mathcal{V}^0 := \prod_{i=1}^n \mathcal{V}^i$. Given a contract of the form (2.3), we can easily solve the agent’s problem, mainly by maximising the above Hamiltonian. The following proposition is thus a direct consequence of the considered form of contracts.

**Proposition 2.2** Fix $i \in \{1, \ldots, n\}$. Let $(\xi_0^i, Z^i) \in \mathbb{R} \times \mathcal{V}^i$ and consider the associated contract $\xi_i^T$ defined as the terminal value of (2.3). Given this contract, the optimal effort $\alpha_i^{\text{HC}}$ of the $i$th agent is given by $\alpha_i^{\text{HC}} := k^i Z_t^i$ for $t \in [0, T]$. Under this effort, the dynamics of $X^i$ and $\xi^i$, respectively, satisfy for all $t \in [0, T]$ that

$$dX_t^i = k^i Z_t^i \,dt + \sigma^i \,dW_t^i \quad \text{and} \quad d\xi_t^i = \frac{1}{2} \tilde{R}^i |Z_t^i|^2 \,dt + Z_t^i \sigma^i \,dW_t^i.$$  

The previous result is classical in the literature (see Koo et al. [37, Propositions 3.1 and 3.2] for a rigorous result, or the more general result by Élie and Possamaï [23, Theorem 4.1]), and its proof is therefore omitted here.

**2.2.2 Towards volatility control**

In addition to his effort $\alpha^0 \in A^0$, the manager chooses the collection of contracts $\xi_T^A \in \mathcal{C}^A$ for the agents. As mentioned above, instead of studying all possible contracts $\xi_T^i \in \mathcal{C}^i$ for the $i$th agent, it is sufficient to restrict the study to contracts of the form (2.3). Therefore, the manager’s problem boils down to a more standard optimal control problem: to design the compensation for the $i$th agent, the manager only has to choose the process $Z^i \in \mathcal{V}^i$.

**Remark 2.3** Recall that the manager designs the contracts for the agents, but that all compensations are paid by the principal. For this reason, we assume that the principal chooses, for all $i \in \{1, \ldots, n\}$, the constant $\xi_0^i$ in the $i$th agent’s contract, noticing that it must be fixed in such a way that his participation constraint is satisfied. At first, it might be more natural to consider that the manager decides the constant $\xi_0^i$ in the contract of the $i$th agent. However, as the manager does not pay the remuneration, he is in fact indifferent to the value of $\xi_0^i$ as long as the agent’s participation constraint is satisfied, while it is optimal for the principal to set this constant in order to actually saturate the agent’s participation constraint.
Since the manager reports to the principal only the benefit $\xi$ in continuous time, his compensation offered by the principal can only be a measurable function of $\xi$. However, even if the agents only control the drift of their outcomes, the manager controls both the drift and the volatility of $\xi$, as we can see from its dynamics under the optimal efforts of the agents which is

$$d\xi_t = \left(\alpha_0^0 + \sum_{i=1}^{n} \left( k^i Z^i_t - \frac{1}{2} \tilde{R}^i |Z^i_t|^2 \right) \right) dt + \sigma^0 dW^0_t + \sum_{i=1}^{n} \sigma^i (1 - Z^i_t) dW^i_t$$

for $t \in [0, T]$. Indeed, by choosing optimally the payment rate in each agent’s contract ($Z^i$ for all $i \in \{1, \ldots, n\}$), the manager controls the volatility of the certainty equivalent of the agents’ continuation utilities $\xi^i$ (through the term $Z^i_t \sigma^i dW^i_t$) and thus the volatility of $\xi$. Therefore, we must consider a more extensive class of contracts than the one used by Sung [58]. Indeed, in continuous time, it is not sufficient to limit oneself to linear contracts (in the sense of Holmström and Milgrom [28]) when the volatility of the state variable is controlled, as demonstrated by Cvitanić et al. [15]. This is where our model and Sung’s diverge: he considers that the one-period model is equivalent, and therefore continues to restrict the study to contracts that are linear with respect to the outcome.

### 2.2.3 Revealing contract for the manager

Taking into account the previous discussion, it is necessary to use the recent results in [15] on optimal incentives for drift and volatility control (see also Cvitanić et al. [14] for an application with exponential utilities, as well as the work by Lin et al. [43] for an extension to a random time horizon). With this in mind, let $V$ be the set of all $(z, \gamma) \in \mathbb{R}^2$ such that $\tilde{R}^i z - |\sigma^i|^2 \gamma > 0$ for all $i \in \{1, \ldots, n\}$. Similarly to $V^0$, we define by $V$ the collection of all processes $(Z, \Gamma)$ taking values in $V$ which are predictable with respect to the filtration generated by $\xi$ and satisfy appropriate integrability conditions. The optimal form of contract offered by the principal to the manager is then given by the terminal value $\xi^0_T$ of the process

$$\xi^0_t = \xi^0_0 - \int_0^t \mathcal{H}^0(Z_s, \Gamma_s) ds + \int_0^T Z_s d\xi_s + \frac{1}{2} \int_0^T (\Gamma_s + R^0 |Z_s|^2) d\langle \xi \rangle_s$$

for $t \in [0, T]$, characterised by $(\xi^0_0, Z, \Gamma) \in V$ chosen by the principal and where the manager’s Hamiltonian $\mathcal{H}^0$ is defined for all $(z, \gamma) \in V$ by

$$\mathcal{H}^0(z, \gamma) := \frac{\gamma |\sigma^0|^2}{2} + \sup_{a \in \mathbb{R}} (az - c^0(a)) + \sum_{i=1}^{n} \sup_{z^i \in \mathbb{R}} \left( z^i \left( k^i z^i - \frac{1}{2} \tilde{R}^i |z^i|^2 \right) + \frac{\gamma}{2} |\sigma^i (1 - z^i)|^2 \right).$$

The main difference between the optimal form (2.3) of contracts for the agents and this form for the manager is the indexation on the quadratic variation of the contractible variable through a parameter (process) $\Gamma$ chosen by the principal. This new
indexation parameter comes from the fact that the manager controls the volatility of ζ. Indeed, by choosing the contract sensitivity \( Z^i \) for each agent, the manager impacts the volatility of \( \xi^i \) and thus the volatility of ζ. Therefore, in order to incentivise the manager to design the best contract for the agents, i.e., to choose the optimal contract sensitivity \( Z^i \), the principal should index the contract \( \xi^0_T \) on the quadratic variation of ζ. Through the restriction to linear contracts, Sung actually sets \( \Gamma = -R^0|Z|^2 \) (see Sect. 3.1 for more details). Then, as for the agent’s problem, using the form (2.4) of contracts, we can establish the following result whose proof is straightforward and therefore omitted here.

**Proposition 2.4** Let \((Z, \Gamma) \in \mathcal{V}\) and consider the associated contract \( \xi^0_T \) defined as the terminal value of (2.4). Given this contract, the optimal effort and the optimal control on the \( i \)th agent’s compensation \((i \in \{1, \ldots, n\})\) chosen by the manager are respectively given by \( \alpha^0_{i, HC} := k^0Z_t \) and \( Z^i_{t, HC} := z^i_{HC}(Z_t, \Gamma_t) \) for \( t \in [0, T] \), where

\[
z^i_{HC}(z, \gamma) := \frac{k^iz - |\sigma^i|^2\gamma}{\tilde{R}^i_z - |\sigma^i|^2\gamma}, \quad (z, \gamma) \in \mathcal{V}.
\]  

(2.5)

For these optimal controls, the dynamics of \( \xi \) and \( \xi^0 \) are respectively given by

\[
d\xi_t = \left( k^0Z_t + \sum_{i=1}^n \left( k^iZ^i_{t, HC} - \tilde{R}^i_z Z^i_{t, HC} \right) \right) dt + dV_t,
\]

\[
d\xi^0_t = \frac{1}{2}|Z_t|^2 \left( \tilde{R}^0 + R^0 \sum_{i=1}^n |\sigma^i (1 - Z^i_{t, HC})|^2 \right) dt + Z_t dV_t,
\]

where \( dV_t := \sigma^0dW^t_0 + \sum_{i=1}^n \sigma^i(1 - Z^i_{t, HC})dW^i_t, \quad t \in [0, T] \).

**2.2.4 Solving the principal’s problem**

Since the optimal form of contract is given by (2.4), the principal only has to choose the parameters \((Z, \Gamma) \in \mathcal{V}\) to design the contract for the manager, and his optimisation problem boils down to a more standard control problem. Recall that we assume that the principal also chooses all the constants \( \xi^0_i \) in each agent’s contract, as well as the constant \( \xi^0_0 \) in the manager’s contract, in addition to the contract sensitivities \((Z, \Gamma) \in \mathcal{V}\). The main results are summarised in the following proposition, the proof of which follows the usual arguments and can be found in Hubert [31, Sect. 2.5.1, proof of Proposition 2.1.6].

**Proposition 2.5** (i) The optimal payment rates for the manager are given by the constant processes \( Z^* := z^{HC} \) and \( \Gamma^* := -R^0(z^{HC})^3 \), with \( z^{HC} \) a maximiser of

\[
\tilde{V}_0^P := \sup_{z > 0} \left( k^0z - \frac{1}{2} \tilde{R}^0|z|^2 + \sum_{i=1}^n h^i_{HC}(z, -R^0z^3) \right),
\]

(2.6)
where $\tilde{R}_0 := k^0 + R^0|\sigma|^2$ and for all $i \in \{1, \ldots, n\}$ and $(z, \gamma) \in \mathcal{V}$,

$$
\bar{h}_i,HC(z, \gamma) := k^i z_i,HC(z, \gamma) - \frac{1}{2} \tilde{R}_i |z_i,HC(z, \gamma)|^2 - \frac{1}{2} R^0|\sigma| z_i(1 - z_i,HC(z, \gamma))|^2.
$$

Using (2.6), the value function of the principal is therefore equal to $V^P_0 = T \bar{V}^P_0$.

(ii) The optimal contract offered by the principal to the manager is given by

$$
\xi^0_T = -\mathcal{H}^0(z^HC, -R^0(z^HC)^3) T + z^HC \zeta_T + \frac{1}{2} R^0|z^HC|^2(1 - z^HC) \langle \zeta \rangle_T,
$$

(2.7)

where $\mathcal{H}^0$ is the manager’s Hamiltonian. In particular, the optimal choice for $\xi^0_0$ is such that he obtains exactly his reservation utility equal to $-1$, i.e., $\xi^0_0 = 0$.

(iii) For $i \in \{1, \ldots, n\}$, the optimal contract offered by the principal to the $i$th agent is

$$
\xi^i_T = -\mathcal{H}^i(z^{i,HC}, \ast(z^HC)) T + z^{i,HC,\ast}(z^HC) X^i_T + \frac{1}{2} R^i|z^{i,HC,\ast}(z^HC)|^2 \langle X^i \rangle_T,
$$

where $\mathcal{H}^i$ is the $i$th agent’s Hamiltonian and $z^{i,HC,\ast}(z) := z^{i,HC}(z, -R^0z^3)$ for all $z > 0$, recalling that $z^{i,HC}$ is defined by (2.5) in Proposition 2.4. In particular, as for the manager, the optimal choice of $\xi^0_0$ is 0.

Remark 2.6 In the optimal contract for the $i$th agent, given by point (iii) of Proposition 2.5, the quadratic variation of $X^i$ at time $T$, i.e., $\langle X^i \rangle_T$, can be replaced by its (constant) value $T|\sigma|^2$. However, this simplification is not possible for $\langle \zeta \rangle_T$ in the contract for the manager. Indeed, this term plays a significant incentive role, even if at the end – under the optimal effort of the manager – the quadratic variation is also constant. More precisely, while the indexation on $\zeta_T$ incentivises the manager to increase the drift, the term $\langle \zeta \rangle_T$ encourages him to make an effort on the volatility. Offering him a contract with a constant value of the quadratic variation instead of the quadratic variation itself does not give him an incentive to choose the optimal volatility, just as indexing the contract on the optimal value of $\zeta_T$ does not give him incentives to increase the drift. In other words, the principal offers the optimal contract $\xi^0_T$ given by (2.7), indexed on the final results $\zeta_T$ and $\langle \zeta \rangle_T$, in order to incentivise the manager to apply the optimal effort and to design the best contract for the agent. Then under the optimal efforts, one can compute the effective payment the manager receives. This difference between the optimal contract – which indicates how the remuneration is indexed to performance, in order to provide incentives to perform – and the effective payment is fundamental.

3 Discussion on the benefits of continuous time

Now that we have completely solved the continuous-time equivalent of the one-period model considered by Sung [58], we can highlight the differences in terms of results, both from a theoretical and an economic point of view.
3.1 Non-optimality of linear contracts in continuous time

The approach previously described requires that the optimal contracts for the manager must be indexed on the quadratic variation of the net profit $\zeta$ through the parameter $\Gamma^*: = -R^0(\zeta^T H C)^3$. However, in [58], Sung restricts the analysis to linear contracts: although he remarks that decisions on middle managerial contracts are affecting the volatility of the net profit of the firm, he chooses to view them as a case similar to unobservable project choice decisions modelled in [57]. More precisely, he states that “the top manager turns out to choose not only the mean of the outcome of his own effort but, in effect, the volatility of the total profit of the firm as he chooses middle managerial contracts. Thus, our problem turns out to be similar to the unobservable project choice problem in [57]”. In [57], Sung studies a principal–agent problem in continuous time where the volatility can be controlled. He distinguishes two cases:

(a) One case is where the variance is observed, but since the Brownian motion is only one-dimensional, there is no moral hazard on the volatility’s effort any more. Indeed, the variance is equal to the square of the volatility effort, and since the variance is observed, the effort can be computed by the principal. Therefore, she directly controls the agents’ effort on the volatility, and the model degenerates to the first-best case for the volatility.

(b) One case is where the variance is not observed, and thus the principal cannot index the contract on the quadratic variation of the outcome, which obliges her to consider only linear contracts.

In [58], as Sung considers that the variance is not observed, the principal cannot offer a contract to the manager indexed on it. In our extended class of contracts defined by (2.4), this is equivalent to forcing $\Gamma^* + R^0|Z^*|^2 = 0$.

Nevertheless, in continuous time, it seems natural to consider that the principal observes the quadratic variation $\langle \zeta \rangle$ of the total benefit and can therefore contract on it. Indeed, by observing $\zeta$ in continuous time, she can estimate the quadratic variation through the sum of the squared increments. Moreover, a result of Bichteler [9] (see Neufeld and Nutz [48, Proposition 6.6] for a modern presentation) states that this quadratic variation, even controlled, can be defined independently of the probability associated to the effort. Therefore, contrary to (b) above, the contract can be indexed on the quadratic variation. In addition, since the process $\zeta$ is naturally driven by $n + 1$ independent Brownian motions, the principal does not perfectly observe the controls $Z^I$ of the manager, but only a functional of them. This prevents the volatility control part from degenerating into the first-best case, contrary to (a).

In conclusion, in the case of volatility control, it is not possible to consider the one-period model by limiting the study to linear contracts and expect to obtain the same results as in continuous time. In our opinion, this justifies the full study of the continuous-time version of hierarchical models, and thus the use of the recent theory on 2BSDEs, from a theoretical point of view. This motivates us to describe in Sect. 4 a more general problem of optimal incentives in a hierarchical organisation, and provide in Sect. 5 a methodology to solve this problem.

3.2 Numerical illustrations

To highlight the differences between the two models, and especially the benefits of the continuous-time approach, we present in this section some numerical simulations,
using the parameters detailed in [58, Sect. 5], in particular in the case of identical workers, i.e., \( k^i = 1000, \ R^i = 50 \) and \( \sigma^i = 1 \) for all \( i \in \{0, \ldots, n\} \). We represent in Fig. 2 the optimal pay-for-performance sensitivities, respectively, for the agents and for the manager in three cases:

(i) the direct contracting (DC) case (blue line);
(ii) with contracts indexed on the quadratic variation (green curve);
(iii) without indexation on the quadratic variation, as in [58] (orange curve).

The value of the principal per worker, i.e., \( V^P_0 / (n + 1) \), is represented in Fig. 3 (left), also for these three situations. All curves are plotted with respect to the number of workers, starting from 2 (i.e., \( n = 1 \)) to consider at least two agents in the DC case, or one agent and one manager in the HC case. Pay-for-performance sensitivity (PPS) is a common proxy for the strength of incentives (see Gryglewicz et al. [25]). In our framework, this sensitivity is directly related to the efforts of the workers. Indeed, for all \( i \in \{1, \ldots, n\} \), the \( i \)th agent’s optimal effort is given by \( \alpha^{i,HC} := k^i Z^i \), where
$Z^i$ is precisely the PPS for the $i$th agent’s contract. A similar relation holds for the manager’s effort and his PPS.

As already outlined in [58], in this hierarchical contracting framework, the manager can benefit from the results of agents’ efforts and transfers his own compensation risk to them through their contracts. As a consequence, agents are induced to work harder than implied in the direct contracting case (see Fig. 2, left). To counterbalance this undesirable risk-shifting motivation of the manager, the principal sets his contract sensitivity to a level lower than that of the contract in the DC case. Consequently, the manager makes less effort than what would be required in the DC situation (see Fig. 2, right). In addition, the bigger the company, the more the manager shifts the risk onto the agents. Consequently, the larger the number of workers, the lower the sensitivity of the manager’s incentive contract.

Nevertheless, our sophisticated contracts allow an improvement of the results: the PPSs we obtain, both for the manager and for the agents, are closer to the PPS in the DC case, compared to those obtained in [58]. This results in less effort on the agents’ side (see Fig. 2, left), but more effort on the manager’s side (see Fig. 2, right). The relative gain in terms of the manager’s effort increases with the number of workers and reaches for example 60% for 30 workers in the company, which means that the manager exerts 60% more effort thanks to these contracts.

Finally, we represent in Fig. 3 the value function of the principal per worker (left) and the associated relative gain (right). With the indexation on the quadratic variation, the principal’s value is obviously higher than without. Even if the relative gain per worker seems small when the number of workers is large, this gain must be multiplied by the number of workers. Moreover, one should consider the benefit of these contracts not only from the principal’s point of view, but also from a global managerial perspective: by developing these contracts, the principal better monitors the manager’s efforts and therefore regulates his risk-shifting motivation, which results in improved conditions for the agents and a better division of work and risk among workers. In particular, it allows to reduce the effort gap between the agents and their manager.

Remark 3.1 The principal’s value per worker is higher when there are two workers instead of three, while it is then increasing with the number of workers. Indeed, when only one agent is supervised by the manager, the principal has access to the indexation parameter $Z^1$ (up to a sign) by observing the quadratic variation of $\zeta$. Therefore in this particular case, there is ‘less’ moral hazard on volatility control, explaining the higher value of the principal.

In this hierarchical model, the principal’s value is smaller than the value obtained in a direct contracting framework. Indeed, the cash flow is assumed to be additive among all workers, and the hierarchy here introduces another layer of moral hazard, which makes this organisational structure less beneficial to the principal. However, in some cases, the hierarchical organisation may produce better results than direct contracting, for example when there is complementarity between the manager’s and the agents’ effort. In fact, the present model can be extended in a straightforward way to add an ability parameter for the manager, in order to justify his position in
the hierarchy, and thus obtain better results than in the DC case. It is also possible to study different types of reporting from the manager to the principal, other than the net benefit (see Hubert [31, Sect. 2.4]).

3.3 Economic implications

To summarise, the previous numerical results share the same features as those outlined in [58]: since the manager subcontracts with the agents, he can benefit to some extent from the results of agents’ efforts and transfers his own compensation risk to them. As a consequence, agents are induced to work harder than implied in the direct contracting case, while the manager makes less effort. Sung concludes that the results obtained with this model on the low managerial effort can serve as an explanation of the empirical finding of Jensen and Murphy [33] that the average CEO contract sensitivity of large firms is lower than that of small firms.

Nevertheless, our sophisticated contract allows an improvement of the results: it mitigates the undesirable risk-shifting motivation of the risk-averse manager. In other words, the manager still benefits from the results of agents’ efforts and transfers a part of his own compensation risk to them, but less than with linear contracts. Consequently, agents are still induced to work harder than in the DC case, but significantly less than in [58]. In addition, the reduction of the agents’ efforts is compensated by an increase in the manager’s effort, and as a result, the burden is more evenly distributed between the agents and their manager. More precisely, using a contract indexed on the quadratic variation for the manager allows the principal to better monitor his performance, and especially the contract sensitivity he chooses for his agents.

The main economic or managerial question might be how to implement these contracts in practice. Indeed, while the indexation on the net benefit is standard to implement, by granting a share of the company’s stock to the manager, the indexation on its quadratic variation is less conventional. Nevertheless, as thoughtfully described by Baldacci and Possamaï [3, Sect. 3.2.2] in the case of similar contracts but applied to green bonds, the quadratic variation part can actually be replicated using vanilla options. Therefore, our sophisticated contract with the indexation on the quadratic variation of net profits argues in favour of remunerating managers through stock options, known as employee stock options (ESOs), in addition to regular shares of stock, as mentioned for example by Lazear [40]. Indeed, in our framework, $\xi$ can be related to the market value of the considered company, and thus $\langle \xi \rangle$ can be replicated through vanilla options on the stock.

4 A general hierarchy model

For the general model, we consider the hierarchy represented in Fig. 4: the principal contracts with $m$ managers, and each manager $j \in \{1, \ldots, m\}$ in turn subcontracts with $n_j$ agents, indexed by $(j, i)$ for $i \in \{1, \ldots, n_j\}$. As before, the term workers will refer to the actors in the hierarchy who are in charge of managing a project, i.e., both agents and managers. The total number of workers is thus given by $w := m + \sum_{j=1}^{m} n_j$. We fix throughout the general model $d \in \mathbb{N}$, which represents
the dimension of the noise which affects each project. The \((j, i)\)th agent manages the project with output \(X^{j,i}\), while the \(j\)th manager is in charge of the project with output \(X^{j,0}\). We assume here for simplicity that the outputs are one-dimensional and uncorrelated, meaning that tasks to be performed have been clearly separated. Moreover, each worker in the hierarchy can only impact his own project. In fact, interactions between workers naturally occur at the level of managers, and a way to handle this will be explained when solving the managers’ problem in Sect. 5.2. In particular, this principal–multi-agent model is the first to deal with non-trivial interactions between agents, especially regarding volatility control; hence the relevance of defining and solving the problem thoroughly. However, the focus of this section is primarily methodological, and we therefore report some important but technical results and proofs in the Appendix (further details can be found in [31, Chap. 3]).

**Notations.** Let \(\mathbb{N} := \mathbb{Z}_0 \setminus \{0\}\) be the set of positive integers and recall that \(T > 0\) denotes some maturity fixed in the contract. For any \((\ell, c) \in \mathbb{N}^2\), \(\mathbb{M}^{\ell,c}\) denotes the space of \(\ell \times c\) matrices with real entries. When \(\ell = c\), we set \(\mathbb{M}^{\ell} := \mathbb{M}^{\ell,\ell}\). For any \(d \in \mathbb{N}\), let \(C([0, T]; \mathbb{R}^d)\) denote the set of continuous functions from \([0, T]\) to \(\mathbb{R}^d\), endowed with the topology of uniform convergence on the compact \([0, T]\). For a probability space of the form \(\Omega := C([0, T]; \mathbb{R}^k) \times \tilde{\Omega}\) and an associated filtration \(\mathbb{F}\), we have to consider processes \(\psi : [0, T] \times C([0, T]; \mathbb{R}^k) \to E\), taking values in some Polish space \(E\), which are \(\mathbb{F}\)-optional, i.e., \(\mathcal{O}(\mathbb{F})\)-measurable, where the optional \(\sigma\)-field \(\mathcal{O}(\mathbb{F})\) is generated by \(\mathbb{F}\)-adapted right-continuous processes. In particular, such a \(\psi\) is non-anticipative in the sense that \(\psi(t, x) = \psi(t, x_{\leq t})\) for all \(t \in [0, T]\) and \(x \in C([0, T]; \mathbb{R}^k)\).

For \(j \in \{1, \ldots, m\}\), \(i \in \{0, \ldots, n_j\}\) and \(x^{j,i} \in S^{j,i}\) for some set \(S^{j,i}\), we use for vectors the notations

\[
\begin{align*}
x^j &:= (x^{j,i})_{i=0}^{n_j}, & x &:= (x^j)_{j=1}^m, & x^M &:= (x^{j,0})_{j=1}^m, \\
x^{-j} &:= (x^\ell)_{\ell \in \{1,\ldots,m\} \setminus \{j\}}.
\end{align*}
\]
and the corresponding sets

\[ S^j := \prod_{i=0}^{n_j} S^{j,i}, \quad S := \prod_{j=1}^{m} \prod_{i=0}^{n_j} S^{j,i}, \quad S^M := \prod_{j=1}^{m} S^{j,0}, \]

\[ S^{-j} := \prod_{\ell=1, \ell \neq j}^{m} S^j. \quad (4.2) \]

In a similar way, we also define

- \( x^A \in S^A \), the vector obtained by suppressing the elements \( x^M \) of \( x \);
- \( x^{-}(j,i) \in S^{-}(j,i) \), the vector obtained by suppressing the element \( x^{j,i} \) of \( x^A \);
- \( x^{j\setminus 0} := (x^{j,i})_{i \in \{1, \ldots, n_j\}} \in S^{j\setminus 0} \) and \( x^{-j\setminus 0} := (x^{\ell\setminus 0})_{\ell \in \{1, \ldots, m\} \setminus \{j\}} \in S^{-j\setminus 0} \).

We also use for sums the notations

\[ \bar{x}^j := \sum_{i=0}^{n_j} x^{j,i} \in \overline{S}^j, \quad \bar{x}^{-j} := (\bar{x}^{\ell})_{\ell \in \{1, \ldots, m\} \setminus \{j\}} \in \overline{S}^{-j}. \quad (4.3) \]

4.1 Moral hazard and reporting in the hierarchy

The most important aspect that we wish to address in this paper is the loss of information by moving up the hierarchy. To model this, we assume that each manager \( j \) only reports the results of his team to the principal through a (possibly multidimensional) variable \( \zeta^j \), in the spirit of Sung’s model detailed previously.

4.1.1 More general contracts

In the previous part of this paper, to provide an appropriate comparison with Sung’s model [58], we chose to follow [58] as closely as possible, in particular by restricting the contract for the \( i \)th agent to be indexed only on his own \( X^i \). Under this restriction, the payment rate \( Z^i \) chosen by the manager should not depend on anything other than \( X^i \); in particular, it cannot depend more generally on \( \zeta \) that contains information on other workers. In this particular framework, we obtain at the end that the optimal payment rate is deterministic. However, in general, the optimal control \( Z^i \) of the manager on the \( i \)th agent’s contract may be a function of \( X^i \) and thus cannot be computed by the principal, since she only observes the global result \( \zeta \). Moreover, in full generality, the aforementioned optimal control \( Z^i \) is a function of two processes \( Z \) and \( \Gamma \) chosen by the principal, and thus assumed to be predictable with respect to what she observes. Indeed, since the manager’s contract is restricted to be a measurable function of \( \zeta \), the payment rates \( Z \) and \( \Gamma \) indexing this contract should also be functions of \( \zeta \). We thus obtain a contradiction: from the agent’s point of view, \( Z^i \) can only be a function of \( X^i \), but from the principal’s point of view, \( Z^i \) can only be a function of \( \zeta \).

Nevertheless, in this particular example, all optimal efforts and controls turn out to be deterministic (even constant), and we can thus assume that the contract for the \( i \)th agent is only indexed on his own output \( X^i \). However, in a more general framework such as the one studied in this second part, we can no longer restrict the
study to such contracts. More precisely, we are led to consider that an agent’s contract can be indexed on the other agents’ outputs. The manager’s controls will thus be predictable with respect to the filtration generated by $\zeta$ and can then be computed by the principal. The precise framework is detailed below.

4.1.2 Different level of monitoring

We study a case of loss of information by proceeding up the hierarchy, modelled by the fact that each manager $j$ reports to the principal only the variable $\zeta^j$ representing the performance of his work team. Indeed, if we consider for example that the principal represents the company’s shareholders as in [58], it is logical to assume that she is not aware of the precise results of each team led by a manager and is probably only interested in the profits and costs, or even the net profits/benefit of each team. We thus assume that the principal only observes a vector $\zeta := (\zeta^j)_{j \in \{1, \ldots, m\}}$. Under this assumption, the admissible contracts for the managers are measurable with respect to the filtration generated by $\zeta$. In other words, the contract for the $j$th manager depends on the reported result $\zeta^j$ of his working team, but also on the results reported by other teams.

Therefore, each manager receives a contract indexed on $\zeta$, and we should make some assumptions ensuring that it is the only state variable of his control problem, in order to avoid the more challenging case where the manager’s problem depends on another process unobservable by the principal. Under these assumptions (see in particular Assumption 5.9 below), even if a manager observes independently the outcomes of his agents, his optimal controls are adapted to the filtration generated by $\zeta$, and thus computable at the optimum by the principal. Nevertheless, the managers can use the detailed information they have to index on it the contracts for their agents. Indeed, if a manager indexes his agents’ contracts only on $\zeta$, he does not benefit from the information he knows over that known by the principal, and there is then no loss of information between the manager and the principal. Therefore, it must be in the interest of the $j$th manager to index the compensation for his workers on all the information he has. In the context of a hierarchy in a company, we can assume that a manager is well informed about the results of his agents, and that during meetings between managers, everyone communicates only the overall result of his team. Therefore, using the notations defined by (4.1) and (4.3), we assume that the $j$th manager observes in continuous time

- the output of all workers of his team, including his own, i.e., the $(n_j + 1)$-dimensional process $X^j$, plus
- the $(m - 1)$-dimensional process $X^{-j}$ representing the sum of the results of each of the other teams.

Then the canonical space of each agent should contain every process observable by his manager. Indeed, we choose to focus here on the loss of information by going up the hierarchy; considering in addition a loss of information by going down the hierarchy would seriously complicate the problem and thus require further study. For the sake of clarity and simplicity, we actually consider that the agents observe all the outputs of the workers.
Remark 4.1 In practice, it seems difficult to imagine that a worker would have access to the individual result of an agent from another team, and this justifies our assumption that a manager only observes the sum of the results of each of the other teams. We nevertheless assume, for the sake of simplicity, that the agents observe all the outputs. This consideration makes it possible to define a single canonical space for all agents, regardless of their team, and is actually without loss of generality. Indeed, as each agent can only impact his own project, and given the form of his objective function, the state variables of his problem are only those on which the contract is indexed, i.e., the variables observable by the manager.

In summary, we consider the following framework: the agents, at the bottom of the hierarchy, observe all the outputs, i.e., the process \( X \) taking values in \( \mathbb{R}^w \). As explained above, they do not use all the information contained in \( X \); the important thing is that they at least have access to the information held by their manager. Each manager perfectly observes the outputs of his agents as well as his own. However, he does not have access to the detailed results of the other teams, but only to the sum of the outputs produced per team. Finally, the principal only observes \( m \) variables, namely \( \zeta := (\zeta^j)^m_{j=1} \), each representing the aggregation of a team’s results.

4.1.3 Theoretical formulation for the workers

Let \( j \in \{1, \ldots, m\} \) and \( i \in \{0, \ldots, n_j\} \). Each worker \((j, i)\) takes care of his own task by choosing a pair \( v^{j,i} := (\alpha^{j,i}, \beta^{j,i}) \), where \( \alpha^{j,i} \) and \( \beta^{j,i} \) are respectively \( A^{j,i} \)- and \( B^{j,i} \)-valued for some subsets \( A^{j,i} \) and \( B^{j,i} \) of Polish spaces. In addition, we consider the functions

\[
\lambda^{j,i} : [0, T] \times A^{j,i} \to \mathbb{R}^d, \quad \sigma^{j,i} : [0, T] \times B^{j,i} \to \mathbb{R}^d,
\]

assumed to be bounded. More precisely, the scalar product between the two functions represents the drift of the outcome of the \((j, i)\)th worker, while \( \sigma^{j,i} \) represents its volatility. We set for simplicity \( U^{j,i} := A^{j,i} \times B^{j,i} \) and call \( U \) the Cartesian product of the sets \( U^{j,i} \), following the notations defined by (4.2). To easily write the dynamics of the column vector \( X \) composed by the collection of all the \( X^{j,i} \), we denote by \( \Lambda : [0, T] \times U \to \mathbb{R}^{dw} \) and \( \Sigma : [0, T] \times U \to \mathbb{M}^{dw,w} \) the functions corresponding to the drift and the volatility of \( X \). More precisely, these functions are defined by

\[
\Lambda(t, u) := (\sigma^{j,i}(t, b^{j,i}) \cdot \lambda^{j,i}(t, a^{j,i}))_{j,i}, \quad t \in [0, T], u = (a, b) \in U, \quad (4.4)
\]

in the sense that \( \Lambda \) is a column vector with \( \Lambda^{j,i}(t, u) := \sigma^{j,i}(t, b^{j,i}) \cdot \lambda^{j,i}(t, a^{j,i}) \), and

\[
\Sigma(t, b) := \bigoplus_{j=1}^{m} \bigoplus_{i=0}^{n_j} \sigma^{j,i}(t, b^{j,i}), \quad (t, b) \in [0, T] \times B, \quad (4.5)
\]

where \( \oplus \) symbolises direct sum of matrices (vectors in this case).
To be consistent with the weak formulation of control problems, we need to define the canonical space $\Omega$ for the workers. For the sake of conciseness, the rigorous formulation is postponed to Appendix A.1; we only summarise here the important definitions. First, this canonical space includes in particular $X$, which represents the collection of the $w$ one-dimensional outcomes $X_{j,i}$ controlled by the workers. Each $X_{j,i}$ is affected by a $d$-dimensional idiosyncratic noise $W_{j,i}$, and $W$ is the collection of the $w$ noises $W_{j,i}$, using the notations defined in (4.1). The canonical space is equipped with the canonical filtration $\mathcal{F} := (\mathcal{F}_t)_{t \in [0,T]}$. Then by Definition A.1, we can construct on $(\Omega, \mathcal{F}_T)$ a subset $\mathcal{P}$ of appropriate probability measures, in the sense that all $\mathbb{P} \in \mathcal{P}$ are associated to an $\mathcal{F}$-predictable effort process $\nu^\mathbb{P} := (\alpha_{j,i}^\mathbb{P}, \beta_{j,i}^\mathbb{P})_{j,i}$, and such that by Lemma A.2, we have for $X$ the representation

$$X_t = x_0 + \int_0^t \Lambda(s, \nu_s^\mathbb{P}) ds + \int_0^t \Sigma(s, \beta_s^\mathbb{P})^\top dW_s, \quad t \in [0, T], \mathbb{P}\text{-a.s.}$$

More precisely, for any $j \in \{1, \ldots, m\}, i \in \{0, \ldots, n_j\}$,

$$X_{j,i}^t = x_{j,i}^0 + \int_0^t \sigma_{j,i}(s, \beta_{j,i}^\mathbb{P}) \cdot (\lambda_{j,i}(s, \alpha_{j,i}^\mathbb{P}) ds + dW_{j,i}^s), \quad t \in [0, T], \mathbb{P}\text{-a.s.}$$

The set $\mathcal{P}$ therefore represents the set of admissible controls of all workers.

### 4.2 The principal–managers–agents hierarchy

The hierarchy is modelled by a series of interlinked principal–agent problems. The principal offers a contract $\xi^j$ to the $j$th manager for $j \in \{1, \ldots, m\}$, and then this manager must design the compensation $\xi^{j,i}$ of each agent $(j, i)$ for $i \in \{1, \ldots, n_j\}$. Finally, as in the illustrative model developed in the first part of the paper, we assume that the principal chooses the continuation utilities of the agents at time $t = 0$ (in addition to those of the managers), i.e., the expected value finally obtained by the agents at time $T$, denoted by $Y_0^{j,i} \in \mathbb{R}$ for the $(j, i)$th agent.

As usual, the two Stackelberg games are solved from bottom to top; but at first, it seems more appropriate to think about this hierarchy from top to bottom. As mentioned above, the principal chooses the initial values of the agents’ continuation utility. This leads us to fix, until the principal’s problem, $Y_0^{j,i} \in \mathbb{R}$ for $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$, summarised by $Y_0^\Lambda \in \mathbb{R}^{w \times m}$. The principal also offers to each manager a contract indexed on the variable $\zeta$ she observes, which leads to the first Stackelberg game. Given his contract and the choices of other managers, the $j$th manager chooses

(i) an effort $\nu^j,0$ and the associated probability,

(ii) a compensation scheme for the agents he manages, i.e., $(\xi^{j,i})_{i=1}^{n_j}$.

The agents determine their optimal efforts in response to the choices of the managers, in particular to their contracts. To write their optimisation problems, it is therefore necessary to fix a contract and the efforts chosen by the managers. In addition, since an agent receives a contract which depends on his colleagues’ output, his optimal response must also be defined in relation to the efforts of the other agents.
Therefore, to consider the stochastic control problem of a particular agent \((j, i)\) for \(j \in \{1, \ldots, m\}\) and \(i \in \{1, \ldots, n_j\}\), we should fix a contract \(\xi^{j,i}\), the efforts of the managers represented by some probability measure \(\mathbb{P}^M \in \mathcal{P}^M\), as well as the other agents’ efforts, also summarised by a probability measure \(\mathbb{P}^{-(j,i)} \in \mathcal{P}^{-(j,i)}\). Appendix A.2 details the construction of the subsets \(\mathcal{P}^M\) and \(\mathcal{P}^{-(j,i)}\) of probability measures representing, respectively, the set of admissible controls for the group of managers and for all other agents (apart from the \((j, i)\)th agent). Then given two probability measures \(\mathbb{P}^M \in \mathcal{P}^M\) chosen by the group of managers and \(\mathbb{P}^{-(j,i)} \in \mathcal{P}^{-(j,i)}\) chosen by other agents, we can define the set of admissible responses of the \((j, i)\)th agent to others’ actions, denoted by \(\mathcal{P}^{j,i}(\mathbb{P}^{-(j,i)}, \mathbb{P}^M)\) (see Definition A.3).

### 4.2.1 A Nash equilibrium between the agents

To consider a particular agent, we fix throughout this section \(j \in \{1, \ldots, m\}\) and \(i \in \{1, \ldots, n_j\}\). Following the discussion in Sect. 4.1.2, we assume that the \(j\)th manager observes the output \(X^j\) of the agents under his supervision, but also the sum \(\bar{X}^{-j}\) of the outputs of each of the other teams. Formally, \(\xi^{j,i}\) must be a measurable functional of the paths of \(X^j\) and \(\bar{X}^{-j}\), i.e.,

\[
\xi^{j,i} : C([0, T]; \mathbb{R}^{nj+1}) \times C([0, T]; \mathbb{R}^{m-1}) \to \mathbb{R},
\]

\[
(X^j, \bar{X}^{-j}) \mapsto \xi^{j,i}(X^j, \bar{X}^{-j}).
\] (4.7)

One can notice that the contracts for the agents of team \(j\) are measurable with respect to \(\mathcal{G}_j\), defined as the natural filtration generated by \(X^j\) and \(\bar{X}^{-j}\), i.e., all agents in a team receive a contract indexed on the same outputs. The set of admissible contracts for the \((j, i)\)th agent is denoted by \(\mathcal{C}^{j,i}\), and we refer to Definition 4.4 below for a rigorous description.

Given this contract and two probability measures \((\mathbb{P}^{-(j,i)}, \mathbb{P}^M) \in \mathcal{P}^{-(j,i)} \times \mathcal{P}^M\) chosen by other workers, we introduce the \((j, i)\)th agent’s objective function as

\[
J^{j,i}(\mathbb{P}, \xi^{j,i}) := \mathbb{E}^\mathbb{P} \left[ K_{0,T}^{j,i,\mathbb{P}} g^{j,i}(X^{j,i} \wedge T, \xi^{j,i}) - \int_0^T K_{0,s}^{j,i,\mathbb{P}} c^{j,i}(s, X^{j,i}, v^{j,i,i}_{s, \mathbb{P}}) ds \right]
\] (4.8)

for \(\mathbb{P} \in \mathcal{P}^{j,i}(\mathcal{P}^{-(j,i)}, \mathbb{P}^M)\), where

(i) \(g^{j,i} : C([0, T]; \mathbb{R}) \times \mathbb{R} \to \mathbb{R}\) is a utility function assumed to be Borel-measurable in each argument and such that for any \(x \in C([0, T]; \mathbb{R})\), the map \(\xi \mapsto g^{j,i}(x, \xi)\) is invertible with inverse \(\overline{g}^{j,i}\);

(ii) \(c^{j,i} : [0, T] \times C([0, T]; \mathbb{R}) \times U^{j,i} \to \mathbb{R}\) is a cost function assumed to be Borel-measurable in each argument and such that for any \(u \in U^{j,i}\), the map

\[
(t, x) \mapsto c^{j,i}(t, x, u)
\]

is \(\mathbb{F}\)-optional, and there exists some \(p > 1\) such that

\[
\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ \int_0^T |c^{j,i}(t, X^{j,i}, v^{j,i,i}_{t, \mathbb{P}})|^p dt \right] < \infty;
\]
(iii) the discount factor is for \( 0 \leq t \leq s \leq T \) and \( \mathbb{P} \in \mathcal{P}^{j,i}(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}) \) given by
\[
K_{t,s}^{j,i,\mathbb{P}} := \exp \left( - \int_t^s \int_U k^{j,i}(v, X^{j,i}, u^{j,i}) \Pi(dv, du) \right),
\]
where the function \( k^{j,i} : [0, T] \times C([0, T]; \mathbb{R}) \times U^{j,i} \to \mathbb{R} \) is assumed to be bounded and Borel-measurable in each argument and such that for any \( u \in U^{j,i} \), the map \((t, x) \mapsto k^{j,i}(t, x, u)\) is \( \mathbb{F} \)-optional.

Note that the function \( g^{j,i} \) introduced above in (i) is referred to as a utility function, although it may not satisfy the usual concavity assumption to be a utility function in the economic sense. This slight abuse of terminology is justified as in most principal–agent problems (as in the model described in Sect. 2), this function is specified as a CARA utility function. Nevertheless, one can consider in this framework more general functions, commonly referred to as reward functions in control theory.

The above setting and assumptions are relatively standard in contract theory (see [15]). In addition, the map \( g^{j,i} \) is assumed to be invertible so that one can recover the contract \( \xi^{j,i} \) from the continuation utility of the agent. More precisely, our first goal is to obtain the dynamics of the continuation utility of the \((j,i)\)th agent, which is denoted by \((Y^{j,i}_t)_{t \in [0, T]} \) and satisfies at the end of the contracting period the equality \( Y^{j,i}_T = g^{j,i}(X^{j,i}, \xi^{j,i}) \). The contract is thus given by \( \xi^{j,i} = g^{j,i}(X^{j,i}, Y^{j,i}_T) \). However, we are forced to make an additional hypothesis, and the reader is referred to Remark 5.2 below for the motivation.

**Assumption 4.2** There exist functions \( c^{j,i}_x, k^{j,i}_x : [0, T] \times C([0, T]; \mathbb{R}) \to \mathbb{R} \), as well as \( c^{j,i}_u, k^{j,i}_u : [0, T] \times U^{j,i} \to \mathbb{R} \), with the property that
\[
c^{j,i}(t, x, u) = c^{j,i}_x(t, x) + c^{j,i}_u(t, u), \quad k^{j,i}(t, x, u) = k^{j,i}_x(t, x) + k^{j,i}_u(t, u)
\]
for all \((t, x, u) \in [0, T] \times C([0, T]; \mathbb{R}) \times U^{j,i}\).

Given a compensation \( \xi^{j,i} \) and probabilities \((\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}) \in \mathcal{P}^{-(j,i)} \times \mathcal{P}^{M}\) chosen by other workers, the optimisation problem faced by the \((j,i)\)th agent is given by
\[
V^{j,i}_0(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}, \xi^{j,i}) := \sup_{\mathbb{P} \in \mathcal{P}^{j,i}(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M})} J^{j,i}(\mathbb{P}, \xi^{j,i}). \tag{4.9}
\]
For \( V^{j,i}_0(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}, \xi^{j,i}) \) to make sense, we require minimal integrability on the contracts and thus impose that there is some \( p > 1 \) such that
\[
\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ |g^{j,i}(X^{j,i}, \xi^{j,i})|^p \right] < \infty. \tag{4.10}
\]

**Definition 4.3** A probability measure \( \mathbb{P} \in \mathcal{P}^{j,i}(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}) \) is an optimal response to the probabilities \( \mathbb{P}^{-(j,i)} \) and \( \mathbb{P}^{M} \) chosen by others and to a contract \( \xi^{j,i} \in \mathcal{C}^{j,i} \) if \( V^{j,i}_0(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}, \xi^{j,i}) = J^{j,i}(\mathbb{P}, \xi^{j,i}) \). We denote by \( \mathcal{P}^{j,i,\ast}(\mathbb{P}^{-(j,i)}, \mathbb{P}^{M}, \xi^{j,i}) \) the collection of all such optimal probability measures.
Using the notation (4.1), we define by $\xi^A$ the collection of contracts for all agents, i.e., $\xi_{j,i} \in C_{j,i}$ for $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$, and the corresponding set is denoted by $C^A$. Given a collection $\xi^A \in C^A$ of contracts for the agents and the managers’ choice $\mathbb{P}^M \in \mathcal{P}^M$, a Nash equilibrium $\mathbb{P}^*$ between the agents can thus be defined as an optimal response for any agent (see Definition A.4), and we write $\mathbb{P}^* \in \mathcal{P}^* \times (\mathbb{P}^M, \xi^A)$.

To simplify the scope of our study and avoid unnecessary complexities, we subsequently require that all eligible contracts for agents induce a unique Nash equilibrium between agents. Indeed, if we do not assume uniqueness, we have to represent the preferences of agents and managers between different Nash equilibria (see [23, Sect. 4.1.1] for an example).

**Definition 4.4** Fix $\mathbb{P}^M \in \mathcal{P}^M$ and $Y^A_0 \in \mathbb{R}^{w-m}$, the collection of $Y_{j,i}^0 \in \mathbb{R}$ for all $j \in \{1, \ldots, m\}$, $i \in \{1, \ldots, n_j\}$. A contract of the form (4.7) satisfying Condition (4.10) is called admissible. The corresponding class is denoted by $C^A_{j,i}$. Moreover, a collection $\xi^A$ of contracts for the agents is admissible, and we write $\xi^A \in C^A$, if

(i) the Nash equilibrium $\mathbb{P}^*$ between agents is unique, i.e., $\mathcal{P}^* \times (\mathbb{P}^M, \xi^A) = \{\mathbb{P}^*\}$;

(ii) for all $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$, we have $\xi_{j,i} \in C^A_{j,i}$ and the $(j,i)$th agent’s value function at the Nash equilibrium is equal to $Y_{j,i}^0$.

In addition, similarly to classical principal–agent problems, we assume that the $(j,i)$th agent has a reservation utility level $\rho_{j,i} \in \mathbb{R}$ below which he refuses to work. Mathematically speaking, this means that a contract should satisfy the inequality

$$V_{j,i}^0(\mathbb{P}^-(j,i), \mathbb{P}^M, \xi_{j,i}) \geq \rho_{j,i}.$$  \hspace{1cm} (4.11)

Note that if the collection $\xi^A$ of contracts for the agents is admissible in the sense of Definition 4.4, then there exists a unique Nash equilibrium $\mathbb{P}^*$ between agents which satisfies in addition $V_{j,i}^0(\mathbb{P}^-(j,i), \mathbb{P}^M, \xi_{j,i}) = Y_{j,i}^0$. Therefore, in order to ensure that the participation constraint (4.11) of the $(j,i)$th agent is satisfied, the principal only has to choose $Y_{j,i}^0 \geq \rho_{j,i}$ in the end.

### 4.2.2 A Nash equilibrium between the managers

**Throughout the following**, we fix $j \in \{1, \ldots, m\}$ to informally define the $j$th manager’s optimisation problem. Recall that the $j$th manager is in charge of a task that generates an output $X_{j,0}$. His effort to improve his output $X_{j,0}$ is defined in an informal way by a pair $v_{j,0} := (\alpha_{j,0}, \beta_{j,0}) \in U_{j,0}$ taking values in $U_{j,0} := A_{j,0} \times B_{j,0}$.

We suppose that the $j$th manager reports in continuous time to the principal the variable $\zeta^j$, assumed to be $h$-dimensional for some $h \in \mathbb{N}$ and measuring the global result of his entire working team (including himself). Therefore, $\zeta^j$ depends on the outcomes of his team, namely $X_j$ (the outcomes of the agents he manages and his own) and the collection $\bar{\xi}_{j,0} := (\xi_{j,i}^0)_{i=1}^n$ of compensations to be paid to the agents. Therefore, the principal only knows the result of each team, represented by the process $\zeta := (\zeta_{j})_{j=1}^m$, and thus the contract she can design for the $j$th manager depends
exclusively on the path of $\zeta$, i.e.,

$$\xi^{j,0} : C([0, T]; \mathbb{R}^{hm}) \ni \zeta \mapsto \xi^{j,0}(\zeta) \in \mathbb{R}. \tag{4.12}$$

In other words, we consider contracts for managers which are $G_T$-measurable, where $G := (\mathcal{G}_t)_{t \in [0, T]}$ is the natural filtration of $\zeta$. The set of admissible contracts for the $j$th manager is denoted by $C^{j,0}$ (see Definition 4.7 for a rigorous description).

We define the characteristics of the $j$th manager to consist of

(i) a utility function $g^{j,0} : C([0, T]; \mathbb{R}^h) \times \mathbb{R} \to \mathbb{R}$, assumed to be Borel-measurable in each argument and such that for any $x \in C([0, T]; \mathbb{R}^h)$, the map $y \mapsto g^{j,0}(x, y)$ is invertible with inverse $g^{j,0};$

(ii) a cost function $c^{j,0} : [0, T] \times C([0, T]; \mathbb{R}^h) \times U^{j,0} \to \mathbb{R}$, assumed to be Borel-measurable in each argument and such that the map $(t, x) \mapsto c^{j,0}(t, x, u)$ is $G$-optional for any $u \in U^{j,0}$ and satisfies for some $p > 1$ that

$$\sup_{P \in \mathcal{P}} \mathbb{E}^P \left[ \int_0^T |c^{j,0}(t, \zeta^j, \nu^j, P)|^p dt \right] < \infty; \tag{13}$$

(iii) a discount factor $k^{j,0} : [0, T] \times C([0, T]; \mathbb{R}^h) \times U^{j,0} \to \mathbb{R}$, assumed to be bounded and Borel-measurable in each argument and such that for any $u \in U^{j,0}$, the map $(t, x) \mapsto k^{j,0}(t, x, u)$ is $G$-optional, with associated quantity, for $P \in \mathcal{P},$

$$K^{j,0, P}_{t,s} := \exp \left( -\int_t^s \int_U k^{j,0}(v, \zeta^j, u^{j,0}) \Pi (dv, du) \right) \quad \text{for } 0 \leq t \leq s \leq T;$$

(iv) a reservation utility level $\rho^{j,0} \in \mathbb{R}$ below which he refuses to work.

Given a probability $P \in \mathcal{P}$ on the canonical space, as well as a contract $\xi^{j,0}$ designed for him by the principal, the objective of the $j$th manager is given by

$$J^{j,0}(P, \xi^{j,0}) := \mathbb{E}^P \left[ K^{j,0, P}_{0,T} g^{j,0}(\xi^j, \xi^{j,0}) - \int_0^T K^{j,0, P}_{0,t} c^{j,0}(t, \xi^j, \nu^{j,0, P}) dt \right]. \tag{14}$$

Remark 4.5 Note that the manager’s cost and utility functions $g^{j,0}, c^{j,0}$ and $k^{j,0}$ depend on $\zeta^j$ and not on the global output $X$. Indeed, these functions can only depend on the variable observed by the principal. Otherwise, if they depend on a variable she does not observe, she cannot compute the managers’ Hamiltonians even for their optimal efforts. This would raise major issues, not yet addressed in the literature (at least in continuous time), which would require a full study before it can be considered in our case. Nevertheless, it is worth noticing that some works attempt to address similar problems, such as the paper by Huang et al. [30].

The $j$th manager must optimise the specific criterion defined by (4.14), given the contract he receives, but also given the choices of the other managers. We thus fix a contract $\xi^{j,0}$ of the form (4.12) as well as the decisions of other managers, i.e., for all $\ell \in \{1, \ldots, m\} \setminus \{j\},$
(i) the effort $v^{\ell,0} \in U^{\ell,0}$ of the $\ell$th manager;
(ii) the collection of contracts $\xi^{\ell\setminus 0} \in C^{\ell\setminus 0}$ offered by the $\ell$th manager to his $n_\ell$ agents.

Given this, the $j$th manager must thus choose an optimal control $v^{j,0} \in U^{j,0}$ as well as a contract $\xi^{j,i} \in C^{j,i}$ for each agent $(j,i)$ under his supervision. We summarise the controls of the managers by a tuple $\tilde{\chi} \in \mathcal{X}$ so that

$$\tilde{\chi}_j \equiv (v^{j,0}, (\xi^{j,i})_{i=1}^{n_j}) \in \mathcal{X} := U^{j,0} \times \prod_{i=1}^{n_j} C^{j,i}$$

is the control of the $j$th manager for $j \in \{1, \ldots, m\}$. Informally, the previous probability $\mathbb{P}^{M} \in \mathcal{P}^M$ results from the effort choice of all managers, namely $\tilde{\chi} \in \mathcal{X}$.

As mentioned in Definition 4.4, we require that eligible contracts for the agents induce a unique Nash equilibrium between them, i.e., $\mathcal{P}^{A,*}(\mathbb{P}^A, \xi^A) = \{\mathbb{P}^*(\tilde{\chi})\}$, where the notation is made to highlight the dependence of the probability $\mathbb{P}^*$ on the control $\tilde{\chi}$. The $j$th manager’s optimisation problem can then be written informally as

$$V^{j,0}(\xi^{j,0}, \tilde{\chi}^{-j}) := \sup_{\tilde{\chi}_j \in \mathcal{X}_j} J^{j,0}(\mathbb{P}^*(\tilde{\chi}), \xi^{j,0}). \quad (4.15)$$

Similarly to the agents’ problem, we require minimal integrability on the contracts,

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[|g^{j,0}(\xi^j, \xi^{j,0})|^p] < \infty \quad \text{for some } p > 1. \quad (4.16)$$

**Definition 4.6** Given a collection $\xi^M := (\xi^{j,0})_{j=1}^m \in C^M$ of contracts, a Nash equilibrium between the managers is a tuple $\tilde{\chi} \in \mathcal{X}$ of controls such that

$$V^{j,0}(\xi^{j,0}, \tilde{\chi}^{-j}) = J^{j,0}(\mathbb{P}^*(\tilde{\chi}), \xi^{j,0}) \quad \text{for all } j \in \{1, \ldots, m\}.$$  

We denote by $\mathcal{P}^{M,*}(\xi^M)$ the set of Nash equilibria given $\xi^M \in C^M$.

**Definition 4.7** A contract of the form (4.12) satisfying Condition (4.16) is called admissible. The corresponding class is denoted by $C^{j,0}$. The product of the sets $C^{j,0}$ for $j \in \{1, \ldots, m\}$ such that the resulting collection $\xi^M$ of contracts for the managers induces a unique Nash equilibrium between them is denoted by $C^M$.

4.2.3 A principal at the top

It remains to define, still informally at this point, the principal’s problem. Contrary to the managers, her problem is more classical as she does not directly control any
process; she just designs the collection of contracts $\xi^M$ for the $m$ managers. Her criterion is defined by

$$J^P(\xi^M) := E^{P}(\xi^M)[K^P_{0,T} g^P(\zeta, \xi^M)],$$

where $P(\xi^M)$ can be seen informally as the probability resulting from the optimal controls of all managers and agents under her supervision given a collection of contracts $\xi^M$, and

(i) $g^P : \mathbb{R}^{hm} \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a given utility function, nonincreasing and concave in the second argument;

(ii) the discount factor $K^P_{s,t} := e^{\int_s^t k^P(r, \zeta)dr}$ for $0 \leq s \leq t \leq T$ is defined by means of a bounded function $k^P : [0, T] \times C([0, T] ; \mathbb{R}^h) \rightarrow \mathbb{R}$ which is $\mathcal{G}$-optional, recalling that $\mathcal{G} := (\mathcal{G}_t)_{t \in [0,T]}$ is the natural filtration of $\zeta$.

The principal must therefore optimise the specific criterion defined by (4.17) by choosing the collection of contracts $\xi^M \in \mathcal{C}^M$ as well as the initial value of the agents’ continuation utility $Y^A_0 \in \mathbb{R}^{w-m}$. Since we assume uniqueness of the Nash equilibrium between managers for admissible contracts, the principal’s problem is

$$V^P_0 := \sup_{Y^A_0} \sup_{\xi^M \in \mathcal{C}^M} J^P(\xi^M).$$

(4.18)

Note that even if $Y^A_0$ does not appear directly in the principal’s optimisation problem, its choice has an impact on the optimal controls of the workers as well as on the value of $\zeta$. However, to lighten the notation, we choose not to emphasise this dependency.

5 Reduction to a standard stochastic control problem

In a classical way, the two Stackelberg games are solved from bottom to top. First, we look for the optimal response of an agent to arbitrary choices of others. Next, the Nash equilibrium between the agents can be solved, under fixed contracts and efforts of the managers. Then, knowing the optimal response of the agents, each manager chooses his efforts and the remuneration for his agents in order to optimise his criterion, given the choices of other managers and the contracts designed by the principal. Since the manager’s choices depend on those of other managers, it is also necessary to find a Nash equilibrium between managers. Finally, given the optimal response of each manager to a contract, the principal can design the optimal contract for each manager in order to optimise her own criterion.

Extending the recent results in Ćvitanić et al. [15], it is relatively straightforward to solve the $(j, i)$th agent’s problem. Indeed, by first restricting the study to revealing contracts indexed on the dynamics of the outputs $X^j$ and $\overline{X}^{-j}$ and their quadratic variations, the agent’s optimal efforts are given by the maximisers of his Hamiltonian. Adapting the reasoning developed in [15], we can then show that the restriction to contracts of this form is without loss of generality from the manager’s point of view.

The main result of the present paper is that this reasoning can be extended to the Stackelberg game between the managers and the principal, even with non-trivial
interactions between the managers through drift and volatility control. Therefore, following the same intuition, we limit our study to contracts for the managers indexed on a well-chosen state variable, namely $\zeta$, and its quadratic variation, through a tuple of parameters $\mathcal{Z} \in \mathcal{V}$ chosen by the principal. We prove that the problem of the managers is relatively simple to solve for this particular class of contracts, and that this restriction is in fact without loss of generality for the principal. More precisely, we establish in Theorem 5.16 that at the end of the day, the principal’s problem defined by (4.18) boils down to the standard control problem

$$V^P_0 = \sup_{Y \geq \rho} \sup_{\mathcal{Z} \in \mathcal{V}} \mathbb{E}^{P^*(\mathcal{Z})} [K^P_{0,T} g^P(\zeta, \xi^M)],$$

where

(i) the inequality $Y_0 \geq \rho$ must be understood componentwise as $Y_{0,j} \geq \rho_{j,i}$ for all $j \in \{1,\ldots,m\}$, $i \in \{0,\ldots,n_j\}$, and thus ensures the participation of all workers;

(ii) $P^*(\mathcal{Z})$ is the unique Nash equilibrium between the workers, given the control $\mathcal{Z} \in \mathcal{V}$ chosen by the principal;

(iii) $\xi^M$ is the collection of revealing contracts for the managers, characterised by the choice of $Y^M_0 \in \mathbb{R}^m$ and $\mathcal{Z} \in \mathcal{V}$.

5.1 Contracting with the agents

This subsection is devoted to solving the problem of a particular agent. With this in mind, we fix throughout the following $j \in \{1,\ldots,m\}$ and $i \in \{0,\ldots,n_j\}$, $Y_{0,j} \in \mathbb{R}$, as well as the probabilities $P^{-(j,i)}$ and $P^M$ chosen by other workers, and thus the associated efforts $\tilde{\nu} := (\nu^{-(j,i)}, \nu^M) \in \mathcal{U}^{-(j,i)} \times \mathcal{U}^M$. We have assumed previously that given his manager’s observation, an admissible contract $\xi^{j,i} \in \mathcal{C}^{j,i}$ for the $(j,i)$th agent is restricted to functions of the form (4.7), and more precisely satisfies Definition 4.4. Therefore, in view of his objective function (4.8), we can already point out that the state variables of his optimisation problem (4.9) are $X^j$ and $X^\perp_j$. By considering the dynamic version $Y^{j,i}_t$ of his value function, we should have

$$Y^{j,i}_0 = V^{j,i}_0 (P^{-(j,i)}, P^M, \xi^{j,i}), \quad Y^{j,i}_T = g^{j,i} (X^{j,i}_T, \xi^{j,i}).$$

From this, we notice that we have an explicit relationship between the compensation $\xi^{j,i}$ and the terminal value function $Y^{j,i}_T$. Given the probabilities chosen by other workers and the associated efforts, we first write the Hamiltonian of the considered agent $(j,i)$ in Sect. 5.1.1. Intuitively, this Hamiltonian appears by simply applying Itô’s formula to the dynamic function of the consumer and by considering the associated Hamilton–Jacobi–Bellman (HJB for short) equation. The next step is to derive a class of so-called revealing contracts, thus extending to a many-agent framework the results of Cvitanić et al. [15] who considered general moral hazard problems with one agent, or alternatively extending to volatility control the results of Élie and Possamaï [23] where the agents controlled only the drift of the output process $X$. Similarly, the class of revealing contracts can be obtained by formally applying Itô’s formula. Finally, Theorem 5.8 states that the restriction to revealing contracts is without loss of generality.
5.1.1 Agent’s Hamiltonian

Recalling that the agent’s problem is defined by (4.9), and since his contract is restricted to functions of the form (4.7), the state variables of his problem are $X_j$ and $X_{-j}$. Nevertheless, the agent only controls the process $X_{j,i}$, while the dynamics of the other state variables are fixed through the probabilities $P^{-j,i}$ and $P^M$. Adapting the reasoning in [15], the agent’s Hamiltonian is a sum of two components:

(i) One is the classical Hamiltonian part as in [15], given by the supremum on the agent’s effort $u := (a, b) \in U_{j,i}$ of the function

$$h^{j,i}(t, x, y, z, \gamma, u) := \Lambda^{j,i}(t, u)z + \frac{\gamma}{2} \|\sigma^{j,i}(t, b)\|^2 - c^{j,i}(t, x, u) - k^{j,i}(t, x, u)y,$$

for $(t, x, y, z, \gamma) \in [0, T] \times C([0, T]; \mathbb{R}) \times \mathbb{R}^3$ and $\Lambda^{j,i}(t, u) := \sigma^{j,i}(t, b) \cdot \lambda^{j,i}(t, a)$.

(ii) The second part is related to the indexation of the contract on the outputs of other workers and thus depends on their effort $\hat{\nu} := (\nu^{-(j,i)}, \nu^M) \in U^{-(j,i)} \times U^M$, fixed by $P^{-j,i}$ and $P^M$. It is defined, for $t \in [0, T]$ and $(z, \tilde{z}) \in \mathbb{R}^{nj+1} \times \mathbb{R}^{m-1}$, by

$$H^{j,i}(t, z, \tilde{z}, \hat{\nu}) := z \cdot \left( \sum_{\ell=0}^{nj} \Lambda^{j,\ell}(t, \tilde{\nu}^{j,\ell}) \right) + \tilde{z} \cdot \left( \sum_{\ell=0}^{nm} \Lambda^{k,\ell}(t, \tilde{\nu}^{k,\ell}) \right).$$

Remark 5.1 Without the assumption on the independence of drift and volatility with respect to the outputs $X$, the second part of $H^{j,i}$, defined by (5.1), would depend on the outputs of the other teams, which are not supposed to be observable by the manager of the $j$th team. As a result, the manager would not be able to compute the Hamiltonian of his agents, which would lead to the more challenging case already mentioned in Remark 4.5 where the agents’ problem depends on a process unobservable by the manager.

The Hamiltonian $\mathcal{H}^{j,i}$ of the agent $(j, i)$ is then defined, along $\hat{\nu} \in U^{-(j,i)} \times U^M$ and for any $(t, x, y, z, \tilde{z}, \gamma) \in [0, T] \times C([0, T]; \mathbb{R}) \times \mathbb{R} \times \mathbb{R}^{nj+1} \times \mathbb{R}^{m-1} \times \mathbb{R}$, as

$$\mathcal{H}^{j,i}(t, x, y, z, \tilde{z}, \gamma, \hat{\nu}) := \sup_{u \in U_{j,i}} h^{j,i}(t, x, y, z, \gamma, u) + H^{j,i}(t, z, \tilde{z}, \gamma, \hat{\nu}).$$

Using Assumption 4.2, we can already notice that a maximiser of the Hamiltonian, if it exists, can be written as a function $u^{j,i,:} : [0, T] \times \mathbb{R}^3 \rightarrow A^{j,i} \times B^{j,i}$, i.e.,

$$u^{j,i,:}(t, y, z^i, \gamma) := (a^{j,i,:}, b^{j,i,:})(t, y, z^i, \gamma).$$

This maximiser, which is proved later to be the optimal effort of the $(j, i)$th agent, only depends on time, the variable $y$ (the agent’s continuation utility) and the parameters $z^i$ and $\gamma$ (the indexation of the contract on his output $X^{j,i}$ and its quadratic variation, respectively). In particular, this maximiser does not depend on the effort $\hat{\nu}$ of the other workers, nor on the indexation of the agent’s contract on the outputs of the others, represented by the parameter $\tilde{z}^{-j}$. Therefore, an agent in fact optimises his efforts independently of others.
Remark 5.2 Without Assumption 4.2, the maximiser of the Hamiltonian would also be a function of $X^{j,i}$. In this case, even if the agent’s supervisor could still compute his optimal effort, the manager of another team could not, which would be an issue similar to the one mentioned in Remark 5.1.

Similarly, we can define the Hamiltonian of any agent in the same way $\mathcal{H}^{j,i}$ is defined for the $(j,i)$th agent. Therefore, the function (5.3) is defined for any $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$. Apart from the $(j,i)$th agent, we denote by $\nu^{-(j,i),*}$ the collection of the efforts of other agents such that their Hamiltonians are maximised. To simplify the reasoning from now on, we make the following assumption, which ensures a unique optimal effort for each agent and therefore a unique Nash equilibrium.

Assumption 5.3 For $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$, there exists a unique Borel-measurable map $u^{j,i,*} : [0, T] \times \mathbb{R}^3 \to U^{j,i}$, defined by (5.3), maximising the Hamiltonian $\mathcal{H}^{j,i}$ given by (5.2).

5.1.2 A relevant form of contracts leading to a Nash equilibrium

We define the relevant subset of contracts, similarly as Cvitanić et al. [15, Definition 3.2] but extended to a multi-agent framework, in the spirit of Élie and Possamaï [23], although with volatility control in addition.

Let $\mathcal{V}^j := \mathbb{R}^{n_j+1} \times \mathbb{R}^{m-1} \times \mathbb{R}$. For any $\mathcal{G}^j$-predictable processes $Z := (Z, \tilde{Z}, \Gamma)$ valued in $\mathcal{V}^j$ and any $Y^{j,i}_0 \in \mathbb{R}$, we introduce the process $Y^{j,i}$, for $t \in [0, T]$, by

$$
Y^{j,i}_t := Y^{j,i}_0 - \int_0^t \mathcal{H}^{j,i}(r, Y^{j,i}_r, \nu^{(j,i)}_r, \nu^{M}_r, Y^{j,i}_r, Z_r, \tilde{Z}_r) \, dr + \int_0^t Z_r \cdot dX^{j,i}_r + \int_0^t \tilde{Z}_r \cdot d\overline{X}^{-j}_r + \frac{1}{2} \int_0^t \Gamma_r \, d\langle X^{j,i} \rangle_r \quad \mathbb{P}\text{-a.s. for all } \mathbb{P} \in \mathcal{P}^{j,i}(\mathbb{P}^{-(j,i)}, \mathbb{P}^M), \quad (5.4)
$$

recalling that $\mathcal{H}^{j,i}$ is defined by (5.2) and where $\hat{\nu}^* := (\nu^{-(j,i),*}, \nu^{M})$ is the collection of optimal efforts of the other agents, given by their Hamiltonian maximisers through (5.3), and for fixed efforts of the managers. This process $Y^{j,i}$ represents the continuation utility of the $(j,i)$th agent, given the actions of others.

Remark 5.4 Note that the process $Y^{j,i}$ defined by (5.4) is the solution of an ODE with random coefficients. Since the Hamiltonian $\mathcal{H}^{j,i}$ defined by (5.2) is Lipschitz-continuous in the variable $y$ (due to the discount factor $k^{j,i}$ being bounded), $Y^{j,i}$ is well defined as the unique solution of (5.4).

Remark 5.5 The process defined from (5.4) only includes an indexation on $X^{j}, \overline{X}^{-j}$ and the quadratic variation of $X^{j,i}$. At first, one should expect that this process would also be indexed on the quadratic variations and covariations of both processes $X^{j}$ and $\overline{X}^{-j}$. However, the processes are independent and so the covariations are equal to 0. Then, since the $(j,i)$th agent only controls the volatility of $X^{j,i}$, the additional terms simplify with the corresponding terms in the standard Hamiltonian.
Definition 5.6 Let $Y_{0,j,i} \in \mathbb{R}$. We denote by $\mathcal{V}^{j,i}$ the set of $\mathcal{V}^j$-valued $\mathbb{G}^j$-predictable processes $Z$ such that $Y_{j,i}$ defined by (5.4) satisfies the integrability condition
\[
\sup_{P \in \mathcal{P}} \mathbb{E}^P \left[ \sup_{t \in [0,T]} |Y_{j,i}^t|^p \right] < \infty \quad \text{for some } p > 1.
\]

For any $Z \in \mathcal{V}^{j,i}$, we call random variables of the form $\xi_{j,i} = g_{j,i}(X_{j,i}, Y_{j,i}^T)$ revealing contracts for the $(j, i)$th agent, and denote the corresponding set by $\Xi_{1j,i}$.

By considering revealing contracts, we are able to compute the optimal effort of each agent, which was given informally by (5.3): intuitively, maximising each agent’s Hamiltonian is sufficient to obtain his optimal effort. Since the agent’s optimal effort does not depend on the efforts of the others, the computation of the Nash equilibrium is then straightforward. Still informally, Assumption 5.3 is in force to ensure existence and uniqueness of the Nash equilibrium, thus avoiding additional technical considerations at the level of the managers’ problem. The result is rigorously presented in the following proposition, for a fixed probability $P^M$ chosen by the managers. Although intuitive, its formal proof is based on the theory of 2BSDEs and follows the reasoning developed for example by Élie et al. [21, Theorem 3.4] or in Appendix B.3, where the proof of a similar result at the managers’ level is reported.

Proposition 5.7 For any $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$, let $Y_{0,j,i} \in \mathbb{R}$ and $Z_{j,i} := (Z_{j,i}^j, \tilde{Z}_{j,i}^j, \Gamma_{j,i}^j) \in \mathcal{V}^{j,i}$. Define $Y_{j,i}$ and the associated contract $\xi_{j,i} \in \Xi_{1j,i}$ as in Definition 5.6. Denote by $\xi^A \in \Xi^A$ the resulting collection of agents’ contracts. Then $\xi^A \in C^A$, and the unique Nash equilibrium $P^* \in \mathcal{P}^A, \star(P^M, \xi^A)$ is characterised as follows:

(i) The optimal effort of the $(j, i)$th agent is given by the unique maximiser of his Hamiltonian, defined by (5.3), i.e.,
\[
\nu_{j,i}^t := u_{j,i}^t, \star(t, Y_{j,i}^t, (Z_{j,i}^t)^j, \Gamma_{j,i}^t) \quad \mathbb{P}^*\text{-a.s., for each } t \in [0,T].
\]

(ii) $Y_{0,j,i} = V_{0,j,i}(P^\star(j,i), \star, P^M, \xi_{j,i})$.

The next step is to prove that the restriction of our study to revealing contracts is not restrictive from the managers’ point of view. This is precisely the purpose of the following section.

5.1.3 Optimality of the revealing contracts

Let us fix $j \in \{1, \ldots, m\}$ in order to focus on the $j$th manager’s problem. Given an admissible contract $\bar{\xi}_{j,0} \in C_{j,0}$ in the sense of Definition 4.7 as well as the decisions of other managers summarised by the control $\bar{\chi}_{-j}$, we recall that the $j$th manager’s optimisation problem is defined by (4.15). Following the general approach in [15], we can prove that there is no loss of generality for the manager to restrict to contracts in $\Xi^{A,0}$ in the sense of Definition 5.6, instead of contracts in $C_{j,0}$. More precisely, to incentivise the $(j, i)$th agent under his supervision, it is sufficient to offer him
a revealing contract \( \xi^{j,i} \in \Xi^{j,i} \), parametrised by a process \( Z^{j,i} \in \mathcal{V}^{j,i} \), instead of considering all admissible contracts in \( C^{j,i} \). Therefore, the \( j \)th manager has to choose

(i) his own effort \( v^{j,0} \in U^{j,0} \);

(ii) a triple of payment rates for each agent under his supervision, i.e., \( Z^{j,i} \in \mathcal{V}^{j,i} \) for all \( i \in \{1, \ldots, n_j\} \).

His control can thus be summarised by a process \( \chi^j := (v^{j,0}, (Z^{j,i})_{i=1}^{n_j}) \in \mathcal{X}^j \), and extending this reasoning, the control of all managers is denoted by \( \chi \), defined by

\[
\chi := (\chi^j)_{j=1}^m \in \mathcal{X}, \quad \text{where} \quad \mathcal{X} := \prod_{j=1}^m \mathcal{X}^j \text{ and } \mathcal{X}^j := U^{j,0} \times \prod_{i=1}^{n_j} \mathcal{V}^{j,i}.
\]

Note that for all \( j \in \{1, \ldots, m\} \), the process \( \chi^j \) takes values in \( \mathcal{X}^j := U^{j,0} \times (\mathcal{V}^{j})^{n_j} \). The process \( \chi \) thus takes values in \( \mathcal{X} \), naturally defined as the Cartesian product of all the \( \mathcal{X}^j \). With this in hand, we can now turn to the main theorem of this first Stackelberg game. Since its proof is similar to the one of the corresponding result for the next Stackelberg equilibrium, namely Theorem 5.16, we omit it here.

**Theorem 5.8** Consider a collection of admissible contracts \( \xi^M := (\xi^{j,0})_{j=1}^m \in \mathcal{C}^M \) for the managers in the sense of Definition 4.7. Then we have the equality

\[
V_0^{j,0}(\xi^{j,0}, \mathcal{X}^j) = \sup_{\chi^j \in \mathcal{X}^j} J_0^{j,0}(P^\star(\chi), \xi^{j,0}) \quad \text{for all } j \in \{1, \ldots, m\}, \tag{5.5}
\]

where \( P^\star(\chi) \) is the unique Nash equilibrium between the agents, given the control \( \chi \in \mathcal{X}^j \) of the managers.

To summarise the agents’ problem, Proposition 5.7 first solves the Nash equilibrium for a probability \( P^M \) and a collection of revealing contracts \( \xi^A \in \Xi^A \) chosen by the managers. Then Theorem 5.8 states that the restriction to revealing contracts is without loss of generality. Using the previous results and notations, we can write the value function of each agent at the equilibrium as

\[
V_0^{j,i,*}(\chi) := V_0^{j,i}(P^\star(j,i), \mathcal{X}^j, \xi^{j,0}) \quad \text{for } j \in \{1, \ldots, m\}, i \in \{1, \ldots, n_j\}. \tag{5.6}
\]

### 5.2 Contracting with the managers

We now fix \( j \in \{1, \ldots, m\} \) to focus our attention on the \( j \)th manager’s problem. Recall that the \( j \)th manager controls his own project with outcome \( X^{j,0} \) as well as the compensations for his \( n_j \) agents. Theorem 5.8 ensures that it is sufficient to restrict the admissible contract space for the \( (j,i) \)th agent to \( \Xi^{j,i} \), and thus limit the \( j \)th manager’s optimisation problem to choosing an optimal effort \( v^{j,0} \in U^{j,0} \) as well as \( n_j \) triples \( Z^{j,i} := (Z^{j,i}, \tilde{Z}^{j,i}, F^{j,i}) \in \mathcal{V}^{j,i} \) for \( i \in \{1, \ldots, n_j\} \), to set up the contracts of the agents under his supervision. Therefore, the manager’s goal is to optimally choose his control process \( \chi^j := (v^{j,0}, (Z^{j,i})_{i=1}^{n_j}) \in \mathcal{X}^j \) defined above, given the results of his working team \( \xi^j \) and his compensation \( \xi^{j,0} \) chosen by the principal.
5.2.1 Main assumption

Recall that the contract $\xi^j_0$ for the manager can only be indexed on $\zeta := (\zeta^j)_{j=1}^m$ as defined in (4.12), since it is the only variable observable by the principal. Nevertheless, each $\zeta^j$ measures the global result of the entire $j$th working team (including the manager) and can therefore depend on the outcomes of the team, namely $X^j$, and the collection of payments $\xi^{j,i}$ for $i \in \{1, \ldots, n_j\}$ to be made to the agents, denoted by $\xi^{j,0}$. For example, we could consider that $\zeta^j := \sum_{i=0}^{n_j} X^{j,i} - \sum_{i=1}^{n_j} \xi^{j,i}$ as it was the case in the illustrative model of Sung [58]. More generally here, we can assume that for all $t \in [0, T]$, we have $\zeta^j = f^j(t, X^j, \xi^{j,0})$ for some function $f^j : [0, T] \times \mathbb{R}_+^{n_j+1} \times \mathbb{R} \rightarrow \mathbb{R}^h$.

In full generality and without any particular assumptions on the function $f^j$, there is no reason the dynamics of $\zeta^j$ should not depend on the collections of all outputs and continuation utilities, which would therefore constitute the state variables of the $j$th manager. Unfortunately, this would lead us once again to the case where the principal does not observe all the state variables of the managers’ problem. As already explained in Remarks 4.5 and 5.1, this would raise major issues requiring a comprehensive study before being addressed in our framework. We are thus forced to make the following major assumption on the shape of the induced dynamics for $\zeta$.

Assumption 5.9 There exist two functions $\Lambda_M : [0, T] \times C([0, T]; \mathbb{R}^{mh}) \times \mathcal{X} \rightarrow \mathbb{R}^{dw}$ and $\Sigma_M : [0, T] \times C([0, T]; \mathbb{R}^{mh}) \times \mathcal{X} \rightarrow \mathbb{M}^{hm,dw}$ which are bounded, satisfy that $\Lambda_M(\cdot, z)$ and $\Sigma_M(\cdot, z)$ are $\mathcal{G}$-optional for any $z \in \mathcal{X}$ and are such that $\zeta$ is a weak solution to the SDE

$$d\zeta_t := \Sigma_M(t, \zeta, \chi_t)(\Lambda_M(t, \zeta, \chi_t)dt + dW_t) \quad \text{for all } t \in [0, T].$$

(5.7)

Although restrictive, this assumption nevertheless allows the study of interesting frameworks that we may have in mind, including the context described by Sung [58]. For more details on when this hypothesis is valid, the reader is referred to [31, Sect. 3.6.3]. Throughout the following, we assume that Assumption 5.9 holds.

Even under Assumption 5.9, the dynamics of $\zeta$ is controlled by all managers in a non-trivial way. Indeed, although each agent controls his own outcome – the $(j,i)$th agent only impacts the dynamic of $X^{j,i}$ –, the $j$th manager does not only control the variable $\zeta^j$, but also the other components of the vector $\zeta$. This is due to the fact that the component $\zeta^k$ (for $k \neq j$) depends on the collection of contracts $\xi^k\setminus 0$, which are indexed in particular on $X^{-k}$ and thus depend in particular on the optimal effort of the $(j,i)$th agent given by

$$\nu^{j,i,*}_t := u^{j,i,*}(t, Y^{j,i}_t, (Z^{j,i}_t)^i, (\Gamma^{j,i}_t)^i) \quad \mathbb{P}^*-\text{a.s.}, \text{ for each } t \in [0, T].$$

Since the pair $(Z^{j,i}, \Gamma^{j,i})$ is chosen by the $j$th manager, he somehow controls the volatility of $\zeta^k$ for all $k \neq j$. Note that he also controls the volatility through his own effort $\nu^{j,0}$. This leads us to a new principal–agent problem with interacting agents (here managers), as for example in the work of Élie and Possamaï [23] (see also Élie et al. [22] for the case of an infinite number of interacting agents), but with volatility control in addition, and especially with non-trivial interactions between managers through this volatility control.
5.2.2 Weak formulation for the managers’ problems

Before addressing the optimisation problem of the manager, we should make the following remark. Through the equality (5.5), the manager’s value function is similar to that of an agent in a classical principal–agent problem since given a contract \( \xi^{j,0} \), the manager chooses his optimal controls. However, the state variable \( \zeta^j \), as a function of \( \xi^{j,0}, X^j \) and \( X^{-j} \), seems to be considered partially in a strong and partially in a weak formulation. Indeed, \( \xi^{j,0} \) is considered in a strong formulation (indexed by the control \( \chi^j \in X^j \)), while the vector of outputs \( X \) is considered in a weak formulation (the control \( \chi^j \in X^{-j} \) only impacts the distribution of \( X \) through \( \mathbb{P}^* \)). It makes little sense to consider a control problem of this form directly, and we should adopt the weak formulation to state the problem of each manager, since this is the one which makes sense for the agents’ problem.

The appropriate weak formulation, detailed in Appendix A.3, is inspired by the work of Possamaï et al. [51, Sect. 6.1], where a zero-sum game is considered between two players controlling both the drift and the volatility of the same output process. Therefore, it only requires to extend their formulation to a nonzero-sum game with \( m \) interacting players. Similarly as for the agents’ problem, we first define a subset \( \mathcal{P}^M \) of probability measures \( \mathbb{P} \) on a new canonical space \( \Omega^M \) such that each \( \mathbb{P} \in \mathcal{P}^M \) is associated to an \( \mathbb{F}^M \)-predictable control process \( \chi^P \), and such that the representation (5.7) for the dynamics of \( \zeta \) holds for the control \( \chi := \chi^P \). However, this representation only gives access to an admissible set of controls in terms of probability measures for all managers, namely \( \mathcal{P}^M \), and thus we still need to define the control of a particular manager in response to others. With this in mind, we fix \( j \in \{1, \ldots, m\} \) as well as the controls \( \chi^{-j} \in X^{-j} \) chosen by other managers. We can then define the relevant subset \( \mathcal{P}^j(\chi^{-j}) \) of probability measures \( \mathbb{P} \), associated to an \( \mathbb{F}^j \)-predictable control process \( \chi^j^P \), to obtain the representation (5.7) for the dynamics of \( \zeta \), but controlled by \( \chi := \chi^j^P \otimes_j \chi^{-j} \), where the notation \( \otimes_j \) means that \( (u \otimes_j \chi_t^{-j})^j = u \) and \( (u \otimes_j \chi_t^{-j})^k = \chi_t^k \) for \( k \neq j \), for all \( u \in X^{-j} \).

Therefore, given \( \chi^{-j} \in X^{-j} \) chosen by other managers, the \( j \)th manager must choose an optimal probability measure \( \mathbb{P} \in \mathcal{P}^j(\chi^{-j}) \), which leads to consider for his optimisation problem (4.15) the weak formulation

\[
V_{0}^{j,0}(\xi^{j,0}, \chi^{-j}) := \sup_{\mathbb{P} \in \mathcal{P}^j(\chi^{-j})} J_{0}^{j,0}(\mathbb{P}, \xi^{j,0}),
\]

recalling that \( J_{0}^{j,0} \) is defined by (4.14). We can then adapt Definition 4.6 to define a Nash equilibrium between the managers in weak formulation (see Definition A.6), and denote by \( \mathcal{P}^M,^* (\xi^M) \) the set of Nash equilibria.

5.2.3 Relevant form of contracts for the managers

Recall that Assumption 5.9 is enforced to ensure that \( \zeta \) is the only state variable of the \( j \)th manager’s optimisation problem. Following the line developed in Sect. 5.1.1 for the agents, the Hamiltonian of the \( j \)th manager is defined by

\[
\mathcal{H}^{j}(t, x, y, z, \gamma, \chi^{-j}) := \sup_{u \in X^j} h^{j}(t, x, y, z, \gamma, \chi^{-j}, u)
\]
for \((t, x, y) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}, (z, \gamma) \in \mathbb{R}^{hm} \times \mathbb{M}^{hm}\) and \(\chi^{-j} \in X^{-j}\) chosen by other managers, where
\[
 h^j(t, x, y, z, \gamma, \chi^{-j}, u) := - (c^{j,0} + yk^{j,0})(t, x_j, u) + (\Sigma M \Lambda M)(t, x, u \otimes_j \chi^{-j}) \cdot z + \frac{1}{2} \text{Tr}((\Sigma M^{\top}M)(t, x, u \otimes_j \chi^{-j} \gamma) ), \quad u \in X^j.
\]

Similarly, we can define the Hamiltonian of any manager in the same way \(H^j\) is defined for the \(j\)th manager.

**Assumption 5.10** For all \(j \in \{1, \ldots, m\}\), there exists a Borel-measurable function \(u^j : [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R} \times \mathbb{R}^{hm} \times \mathbb{M}^{hm} \times X^{-j} \rightarrow X^j\) satisfying
\[
 H^j(t, x, y, z, \gamma, \chi^{-j}) = h^j(t, x, y, z, \gamma, \chi^{-j}, u^j(t, x, y, z, \gamma, \chi^{-j}))
\]
for all \((t, x, y, z, \gamma) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R} \times \mathbb{R}^{hm} \times \mathbb{M}^{hm}\) and given the actions of other managers \(\chi^{-j} \in X^{-j}\).

Under the previous assumption, a maximiser of the \(j\)th manager’s Hamiltonian exists. Notice that in full generality, this may depend on time, the paths of the state variable \(\zeta\), the parameter \(y\) (the manager’s continuation utility) and the parameters \(z\) and \(\gamma\) (the indexation of the contract on the reporting \(\zeta\) and its quadratic variation). More importantly, contrary to the optimal efforts of agents which were independent of the efforts of other agents, the maximiser of the Hamiltonian here depends on the efforts of other managers, i.e., on \(\chi^{-j}\). Similarly, these efforts \(\chi^{-j}\) are defined through the maximiser of other managers’ Hamiltonians and thus depend on \((t, x, y^{-j}, z^{-j}, \gamma^{-j})\), but also on \(\chi^j\). This leads to consider the fixed point of a multidimensional Hamiltonian, where each component is the Hamiltonian of a manager.

**Assumption 5.11** There exists a unique Borel-measurable map
\[
u^* : [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times \mathbb{M}^{hm,m} \times (\mathbb{M}^{hm})^m \rightarrow \prod_{j=1}^m X^j
\]
such that the \(j\)th component \(u^{j,*}\) takes values in \(X^j\) and satisfies
\[
 H^j(t, x, \eta^j, u^{-j,*}(t, x, \eta)) = h^j(t, x, \eta^j, u^{-j,*}(t, x, \eta), u^{j,*}(t, x, \eta))
\]
for all pairs \((t, x) \in [0, T] \times C([0, T]; \mathbb{R}^{hm})\) and for all tuples \(\eta := (y^j, z^j, \gamma^j)_{j=1}^m\) in \(\mathbb{R}^m \times \mathbb{M}^{hm,m} \times (\mathbb{M}^{hm})^m\), where \(z^j \in \mathbb{R}^{hm}\) is the \(j\)th column of \(z\) and \(\gamma^j \in \mathbb{M}^{hm}\), and this for all \(j \in \{1, \ldots, m\}\).

**Remark 5.12** Assuming the existence of such a function is classical in multi-agent problems to ensure existence of an equilibrium between players (see for example...
Thanks to Assumption 5.11, we can define the $j$th manager’s Hamiltonian under optimal efforts of all managers, for $(t, x, \eta) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times \mathbb{V}$, where $\mathbb{V} := \mathcal{M}^{hm,m} \times (\mathcal{M}^{hm})^m$, and taking values in $\mathbb{R}$ as

$$
\mathcal{H}^{j,*}(t, x, \eta) := \mathcal{H}^j(t, x, \eta^j, u^{-j,*}(t, x, \eta)). \tag{5.10}
$$

We can now define the relevant subset of contracts, similarly as for the agents, except that the contract must be indexed on $\zeta$ and the managers’ Hamiltonians are coupled. For any $\mathbb{V}$-valued $\mathcal{G}$-predictable processes $(Z, \Gamma)$ and $Y^M_0 := (Y^j_0)_{j=1}^m \in \mathbb{R}^m$, consider the multidimensional process $\mathcal{Y}^M$ such that each component $\mathcal{Y}^j$ satisfies

$$
\mathcal{Y}^j_t := Y^j_0 - \int_0^t \mathcal{H}^{j,*}(r, \zeta, \mathcal{Y}^M_r, Z_r, \Gamma_r)dr + \int_0^t Z^j_r \cdot d\zeta_r + \frac{1}{2} \int_0^t \text{Tr}(\Gamma_r^j \hat{\Sigma}_r)dr, \tag{5.11}
$$

for $t \in [0, T]$. For $j \in \{1, \ldots, m\}$, $\mathcal{Y}^j$ represents the continuation utility of the $j$th manager, given the actions of others. As mentioned in Remark 5.4 for the agents’ level, each component $\mathcal{Y}^j$ of the process $\mathcal{Y}^M$ is defined by (5.11) as a solution to an ODE with random coefficients. The following assumption is made to ensure that this ODE is well defined, so that the solution exists and is unique.

**Assumption 5.13** The multidimensional Hamiltonian $\mathcal{H}^*$ whose components are defined by (5.10) is uniformly Lipschitz-continuous with respect to $y \in \mathbb{R}^m$.

One can notice that if for all $j \in \{1, \ldots, m\}$, the discount factor $k^{j,0}$ is not controlled, meaning that $k^{j,0} : [0, T] \times C([0, T]; \mathbb{R}^p) \to \mathbb{R}$, then the previous assumption is not necessary. Indeed, in this case, the optimal effort of the managers defined through the function $u^*$ is independent of the variable $y \in \mathbb{R}^m$, thus also implying the independence of the Hamiltonian $\mathcal{H}^*$. Similarly, in the case of CARA utility functions, a change of variable allows to eliminate the dependence of the control on the variable $y$. Therefore, in these two cases, Assumption 5.13 is trivially satisfied, ensuring the well-definedness of the process $\mathcal{Y}^M$.

**Definition 5.14** Let $Y^M_0 \in \mathbb{R}^m$. We denote by $\mathcal{V}$ the set of $\mathbb{V}$-valued $\mathcal{G}$-predictable processes $Z$ such that for all $j \in \{1, \ldots, m\}$, each component $\mathcal{Y}^j$ of the $m$-dimensional process $\mathcal{Y}^M$ defined by (5.11) satisfies

$$
\sup_{P \in \mathcal{P}^m} \mathbb{E}^P \left[ \sup_{t \in [0, T]} |\mathcal{Y}^j_t|^p \right] < \infty \quad \text{for some } p > 1.
$$

\({\copyright}\) Springer
Let \( \overline{g}^M : C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \to \mathbb{R}^m \) be such that for all \( j \in \{1, \ldots, m\} \), the \( j \)th component is given by the map \( \overline{g}^{j,0} : C([0, T]; \mathbb{R}^h) \times \mathbb{R} \to \mathbb{R} \) defined in the managers’ problem. For \( \mathcal{Z} \in \mathcal{V} \) and \( Y^0_M := (Y^j_0)_{j=1}^m \in \mathbb{R}^m \), we consider the \( m \)-dimensional random variable \( \xi^M := \overline{g}^M(\zeta, Y^M_T) \) and denote the corresponding set by \( \Xi^M \). We say that \( \xi^M \in \Xi^M \) is a collection of \textit{revealing contracts} for the managers. In particular, for all \( j \in \{1, \ldots, m\} \), its \( j \)th component satisfies \( \xi^j = \overline{g}^{j,0}(\zeta^j, Y^j_T) \), and the corresponding set is denoted by \( \Xi^j \).

### 5.2.4 Nash equilibrium and optimality of revealing contracts

The following proposition characterises the Nash equilibrium for a collection of revealing contracts \( \xi^M \in \Xi^M \) chosen by the principal. Contrary to the agents’ level, the managers’ optimal efforts are now coupled. The formal proof is based on 2BSDE theory and is reported in Appendix B.3.

**Proposition 5.15** Let \( Y^0_M := (Y^j_0)_{j=1}^m \in \mathbb{R}^m \) and \( \mathcal{Z} := (Z, \Gamma) \in \mathcal{V} \). By Definition 5.14, consider the \( m \)-dimensional process \( Y^M \) and the associated collection of contracts \( \xi^M := (\xi^j)_{j=1}^m \in \Xi^M \). Then \( \xi^M \in \mathcal{C}^M \) in the sense of Definition 4.7, and there exists a unique Nash equilibrium in the sense of Definition A.6, i.e., a control \( \chi^* \in \mathcal{X} \) associated to a probability measure \( \mathbb{P}^* \). This Nash equilibrium is characterised as follows:

(i) For all \( j \in \{1, \ldots, m\} \), the optimal effort of the \( j \)th manager is given by the \( j \)th component of the unique map defined in Assumption 5.11, i.e.,

\[
\chi^{i,*}_t := u^{i,*}(t, \zeta, Y^M_t, Z_t, \Gamma_t) \quad \mathbb{P}^*\text{-a.s., for each } t \in [0, T].
\]

(ii) \( Y^j_0 = V^{j,0}_0(\xi^j, \chi^{-j,*}) \).

At the equilibrium, we can write the value function of the \( j \)th manager as

\[
V^{j,0,*}_0(\xi^M) := V^{j,0}_0(\xi^j, \chi^{-j,*}) \quad \text{for all } j \in \{1, \ldots, m\}. \tag{5.12}
\]

Finally, the following result ensures that the specialisation of our study to revealing contracts is not restrictive from the principal’s point of view. This result echoes Theorem 5.8 for the manager–agent problem; its proof is postponed to Appendix B.3.

**Theorem 5.16** Recalling that \( V^P \) is defined by (4.18), we have the equality

\[
V^P = \sup_{Y_0 \geq \rho} \overline{V}^P(Y_0), \quad \text{where } \overline{V}^P(Y_0) := \sup_{\mathcal{Z} \in \mathcal{V}} \mathbb{E}^P[\mathcal{K}^P_{0,T}g^P(\zeta, \xi^M)], \tag{5.13}
\]

where

(i) the inequality \( Y_0 \geq \rho \) must be understood componentwise, i.e., \( Y^{j,i}_0 \geq \rho^{j,i} \) for all \( j \in \{1, \ldots, m\} \) and \( i \in \{0, \ldots, n_j\} \), and ensures the participation of all workers;

(ii) \( \mathbb{P}^*(\mathcal{Z}) \) is the unique Nash equilibrium between the workers, given the control \( \mathcal{Z} \in \mathcal{V} \) chosen by the principal;

(iii) \( \xi^M \) is the collection of revealing contracts for the managers, characterised by the choice of \( Y^0_M \in \mathbb{R}^m \) and \( \mathcal{Z} \in \mathcal{V} \).
5.3 Principal’s problem

Following the previous reasoning, ζ and the continuation utilities \( \mathcal{Y}^M \) of the managers, which are the only variables observable by the principal, are also clearly the only state variables of her problem.

5.3.1 The final standard control problem

Before addressing the principal’s problem, we need to write the dynamics of \( \zeta \) and \( \mathcal{Y}^M \) under the managers’ optimal efforts. With this in mind and recalling the definition of the map \( u^* \) in Assumption 5.11, we define two functions

\[
\Lambda^*_M, \Sigma^*_M : [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times \mathcal{M}^{hm,m} \times (\mathcal{M}^{hm})^m \rightarrow \mathbb{R}^{dw}, \mathcal{M}^{hm,dw}
\]

that for \((t, x, y) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \) and \( v := (z, \gamma) \) in \( \mathcal{M}^{hm,m} \times (\mathcal{M}^{hm})^m \) satisfy

\[
\Lambda^*_M(t, x, y, v) := \Lambda_M(t, x, u^*(t, x, y, v)),
\]

\[
\Sigma^*_M(t, x, y, v) := \Sigma_M(t, x, u^*(t, x, y, v)).
\]

We also define \( c^{M,*}, k^{M,*} : [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times \mathcal{M}^{hm,m} \times (\mathcal{M}^{hm})^m \rightarrow \mathbb{R}^m \) such that for all \( j \in \{1, \ldots, m\} \), the \( j \)th components of \( c^{j,*} \) and \( k^{j,*} \) satisfy

\[
c^{j,*}(t, x, y, v) := c^{j,0}(t, x^j, u^{j,*}(t, x, y, v)),
\]

\[
k^{j,*}(t, x, y, v) := k^{j,0}(t, x^j, u^{j,*}(t, x, y, v)),
\]

respectively, for \((t, x, y, v) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times \mathcal{V} \) with \( x := (x^j)_{j=1}^m \) and for \( x^j \in C([0, T]; \mathbb{R}^h) \).

With these notations, \( \zeta \) is a weak solution to the SDE, for \( t \in [0, T] \),

\[
\zeta_t = \zeta_0 + \int_0^t \Sigma^*_M(s, \zeta, \mathcal{Y}^M_s, \mathcal{Z}_s)(\Lambda^*_M(s, \zeta, \mathcal{Y}^M_s, \mathcal{Z}_s)ds + d\mathcal{W}_s)
\]

(5.14)

under \( \mathbb{P}^*(\mathcal{Z}) \), for some \( \mathcal{Z} := (Z, \Gamma) \in \mathcal{V} \). Then we can compute the value at the equilibrium of the multidimensional Hamiltonian \( \mathcal{H}^* \) defined by (5.10) and use the dynamics (5.14) of \( \zeta \) to write the SDE satisfied by \( \mathcal{Y}^M \). More precisely, starting from (5.11), we can state that each component \( \mathcal{Y}^j \) of \( \mathcal{Y}^M \) satisfies, for \( t \in [0, T] \),

\[
\mathcal{Y}^j_t = Y^j_0 + \int_0^t (c^{j,*} + \mathcal{Y}^j_t k^{j,*})(s, \zeta, \mathcal{Y}^M_s, \mathcal{Z}_s)ds + \int_0^t (Z^j_s)^\top \Sigma^*_M(s, \zeta, \mathcal{Y}^M_s, \mathcal{Z}_s)d\mathcal{W}_s.
\]

Note that the column vector process \( \mathcal{Y}^M \) taking values in \( \mathbb{R}^m \) thus satisfies the multidimensional SDE, for \( t \in [0, T] \),

\[
d\mathcal{Y}^M_t = (c^{M,*} + \mathcal{Y}^M_t k^{M,*})(t, \zeta, \mathcal{Y}^M_t, \mathcal{Z}_t)dt + Z^\top_t \Sigma^*_M(t, \zeta, \mathcal{Y}^M_t, \mathcal{Z}_t)d\mathcal{W}_t.
\]

(5.15)
Recall that the principal’s problem is initially defined by (4.18) and then simplified by Theorem 5.16. As before (and thus omitted here), we should rigorously define the weak formulation of the principal’s problem by constructing a subset $P \in P$ of probability measures such that each $P \in P$ is associated to an $\mathbb{F}^P$-predictable control process $Z^P \in V$ and the representations (5.14) and (5.15) hold respectively for $\zeta$ and $\mathcal{Y}^M$. With these notations, the principal’s problem can finally be written as a standard stochastic control problem, namely

$$V^P(Y_0) = \sup_{P \in P} \mathbb{E}^P[\mathcal{K}_{0,T} g^P(\zeta, \mathcal{Y}^M)],$$

and thus $V^P_0 = \sup_{Y_0 \geq \rho} V^P(Y_0)$.

5.3.2 On solving the principal’s problem

We can define the principal’s Hamiltonian, denoted by $\mathcal{H}^P$, and are then led to consider, for $(t, x, y) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m$, the HJB equation

$$\begin{cases}
-\partial_t V(t, x, y) - \mathcal{H}^P(t, x, y, V(t, x, y), \nabla V(t, x, y), \nabla^2 V(t, x, y)) = 0, \\
V(T, x, y) = g^P(x, \mathcal{Y}^M(x, y)).
\end{cases} \quad (5.16)$$

Given this equation, it is clear that the principal’s problem $V^P$ boils down to a more standard control problem. Nevertheless, (5.16) is path-dependent since the Hamiltonian at time $t \in [0, T]$ depends on the paths of the variable $\zeta$ up to $t$. Therefore, in this general case, solving the principal’s problem for $V^P(Y_0)$ is equivalent to solving a path-dependent partial differential equation (PDE) under appropriate conditions for the solution. We refer to the works by Ekren et al. [19, 20] for more details on this type of problem and for the notion of viscosity solutions. Intuitively, an optimal control $Z^\star$ corresponds to a maximiser of the Hamiltonian. The final step is then to find an optimal $Y_0 \geq \rho$ in order to maximise the previously obtained value function.

If we consider a Markovian framework in the sense that the function $g^P$ only depends on $\zeta_T$ and the Hamiltonian at time $t \in [0, T]$ only depends on the current value $\zeta_t$, then solving the principal’s problem for $V^P$ boils down to solving a more standard PDE. In this case, following the line of Cvitanić et al. [15, Theorem 3.9], we could write a verification result for the problem for $V^P$. In particular, assume that there exist a function $V : [0, T] \times \mathbb{R}^{hm} \times \mathbb{R}^m \to \mathbb{R}$ which is a solution to the HJB equation (5.16), and a function $v^\star : [0, T] \times \mathbb{R}^{hm} \times \mathbb{R}^m \to \mathbb{V}$ attaining the supremum in the principal’s Hamiltonian. Then under appropriate conditions on these two functions, we should obtain that $V^P(Y_0) = V(0, \zeta_0, Y^M_0)$ and that the process $Z^\star$ defined for all $t \in [0, T]$ by $Z^\star_t := v^\star(t, \zeta, \mathcal{Y}^M_t)$ is an optimal control for the principal. As mentioned above, the final step is to optimise on the initial value of the workers’ continuation utility.

Finally, the main aspect to notice regarding the principal’s problem is that thanks to the optimal form of contracts for the managers, and in particular by Theorem 5.16, the dimension of this problem does not explode. More precisely, if the principal supervises $m$ managers, then her problem has $2m$ state variables, potentially multidimensional but of a dimension independent of the number of managers. Indeed, on the
one hand, each manager $j$ communicates his results through a variable $\zeta^j$ of fixed dimension $h$ which constitutes a state variable for the principal. On the other hand, thanks to the elegant reasoning of Sannikov [53], later developed in Cvitanić et al. [15], considering in addition the continuation utility of the said manager is sufficient to solve the principal’s problem. Therefore, the well-known method used to solve a classical contracting problem can be extended to a hierarchical structure and preserves the same main features, namely that the principal’s problem boils down to a more classical control problem with two state variables per agent under her direct supervision.

6 Conclusion

In the first part of this paper, we have introduced and solved the continuous-time version of Sung’s model developed in [58]. This opening example highlights the differences between the one-period model and its continuous-time equivalent, in particular concerning the form of the contracts. More precisely, when studying the continuous-time model, we are allowed to consider an extended class of contracts for the managers, indexed in particular on the quadratic variation of the net benefit $\zeta$ observed by the principal, because of volatility control. Therefore, in order to rigorously study a continuous-time hierarchy problem, it is not possible to consider the associated discrete-time model with linear contracts, which justifies and even requires the use of the theory of 2BSDEs to deal with problems of moral hazard within a hierarchy. The second part of this paper focuses on a more general model and provides a systematic method to solve any hierarchy problems of this sort, by a method that can be extended in a straightforward way to a larger-scale hierarchy.

In the general model, we assumed that the agents (at the bottom of the hierarchy) do not interact with each other, in the sense that each of them controls his own output and that these outputs are uncorrelated. We could actually assume that they interact, in the same way as the managers finally do. Furthermore, we could also consider instead of a finite number of agents a continuum with mean-field interaction. There is no reason why this issue could not be addressed by applying the results of Élie et al. [21] in our framework. However, several assumptions are necessary to complete our study, notably on the shape of the dynamics of the state variables. Even if they are satisfied in the most common and interesting examples, one might want to weaken those assumptions. Most of these hypotheses prevent the case where a principal does not observe one of the state variables of her agent’s problem. Therefore, to hope for an extension of our model, it would be necessary to solve this issue, which is for now scarcely addressed in the literature in continuous time (see however the work of Huang et al. [30] for a particular example).

Appendix A: Weak formulations

This section details the rigorous weak formulations for the workers. The one for the principal is more standard (see e.g. Élie et al. [21]) and thus omitted here.
A.1 Initial canonical space

To model the output processes of the workers and their controls, we consider the canonical space

$$
\Omega := C([0, T]; \mathbb{R}^w) \times C([0, T]; \mathbb{R}^{d_w}) \times U,
$$

where $U$ is the collection of all finite positive Borel measures on $[0, T] \times U$ whose projection on $[0, T]$ is the Lebesgue measure. In other words, every $q \in U$ can be dis-integrated as $q(dt, du) = q_t(du)dt$ for $t \in [0, T]$ and an appropriate Borel-measurable kernel $q_t$. The weak formulation requires to consider a subset of $U$, namely the set $U_0$ of all $q \in U$ such that the kernel $q_t$ is of the form $\delta_{\phi_t}(du)$ for some Borel function $\phi$, where as usual $\delta_{\phi_t}$ is the Dirac mass at $\phi_t$. This space supports a canonical process $(X, W, \Pi)$ defined for $(t, x, \omega, q) \in [0, T] \times \Omega$ by

$$
X_t(x, \omega, q) := x(t), \quad W_t(x, \omega, q) := \omega(t), \quad \Pi(x, \omega, q) := q.
$$

Less formally, $X$ represents the collection of the $w$ one-dimensional outcomes $X^{j,i}$ controlled by the workers. Each $X^{j,i}$ is affected by a $d$-dimensional idiosyncratic noise $W^{j,i}$, and $W$ is the collection of the $w$ noises $W^{j,i}$, using the notations defined in (4.1). Then the canonical filtration $\mathcal{F}$ defined by

$$
\mathcal{F}_t := \sigma\left(\left\{ (X_s, W_s, \Delta_s(\varphi)) : (s, \varphi) \in [0, t] \times C_b([0, T] \times U; \mathbb{R}) \right\}\right), \quad t \in [0, T],
$$

where $C_b([0, T] \times U; \mathbb{R})$ is the set of all bounded continuous functions from $[0, T] \times U$ to $\mathbb{R}$ and

$$
\Delta_s(\varphi) := \int_0^s \int_U \varphi(r, u) \Pi(dr, du) \quad \text{for } (s, \varphi) \in [0, T] \times C_b([0, T] \times U; \mathbb{R}).
$$

Finally, let $C^2_b(\mathbb{R}^k; \mathbb{R})$ for any $k \in \mathbb{N}$ denote the set of bounded twice continuously differentiable functions from $\mathbb{R}^k$ to $\mathbb{R}$ whose first and second derivatives are also bounded. For any $(t, \psi) \in [0, T] \times C^2_b(\mathbb{R}^w \times \mathbb{R}^{d_w}, \mathbb{R})$, we set

$$
M^A_t(\psi) := \psi(X_t, W_t) - \int_0^t \int_U \left( \tilde{\Lambda}(s, u) \cdot \nabla \psi(X_s, W_s) + \frac{1}{2} \text{Tr}(\nabla^2 \psi(X_s, W_s) \tilde{\Sigma}(s, u) \tilde{\Sigma}(s, u)^\top) \right) \Pi(ds, du),
$$

where $\nabla^2 \psi$ denotes the Hessian matrix of $\psi$ and $\tilde{\Lambda}, \tilde{\Sigma}$ are respectively the drift vector and the diffusion matrix of the $(w + d)\cdot$-dimensional vector process $(X, W)^\top$, i.e.,

$$
\tilde{\Lambda}(s, u) := \left( \Lambda(s, u) \quad 0_{w d} \right) \in \mathbb{R}^{w + w d},
$$

$$
\tilde{\Sigma}(s, u) := \begin{pmatrix} 0_{w, w} & \Sigma(s, b)^\top \end{pmatrix} \in \mathbb{M}^{w + w d}
$$

for $s \in [0, T]$, $u := (a, b) \in U$, where $\Lambda, \Sigma$ are defined by (4.4), (4.5).
We fix some initial conditions, namely $x_0 \in \mathbb{R}^w$ representing the initial value of $X$, and let $\mathcal{M}$ be the set of all probability measures on $(\Omega, \mathcal{F}_T)$.

**Definition A.1** The subset $\mathcal{P} \subseteq \mathcal{M}$ is composed of all $\mathbb{P}$ such that

1. $M^A(\psi)$ is an $(\mathbb{F}, \mathbb{P})$-local martingale on $[0, T]$ for all $\psi \in C^2_p(\mathbb{R}^w \times \mathbb{R}^d, \mathbb{R})$;
2. there exists some $w_0 \in \mathbb{R}^d$ such that $\mathbb{P}[(X_0, W_0) = (x_0, w_0)] = 1$;
3. $\mathbb{P}[\Pi \in \mathbb{U}_0] = 1$.

Definition A.1 does not give us access directly to the dynamics of $X$. It is, however, a classical result that, enlarging the canonical space if necessary, one can construct Brownian motions allowing to write the dynamics of $X$; see for instance Stroock and Varadhan [56, Theorem 4.5.2]. Here, since we enlarged the canonical space right from the start to account for the idiosyncratic noises, any further enlargement is not required. Indeed, arguing as in the proof of Lin et al. [42, Lemma 2.2], we can prove the following.

**Lemma A.2** For all $\mathbb{P} \in \mathcal{P}$, we have $\Pi(ds, du) = \delta_{\nu_\mathbb{P}}(du)ds$ $\mathbb{P}$-a.s. for some $\mathbb{F}$-predictable process $\nu_\mathbb{P} := (\alpha_{j,i,\mathbb{P}}, \beta_{j,i,\mathbb{P}})_{j,i}$, and we get the representation (4.6) for $X$.

### A.2 Admissible response of an agent

Let now $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n_j\}$ and consider the $(j, i)$th agent. In order to define his admissible response given the choices of other workers (managers and agents), we should define two subsets of the canonical space $\Omega$. Using the notations (4.2), we consider the sets $U^M$ and $U^{-(j,i)}$ from the collection of sets $U_{j,i}$. In the same manner we defined $\mathbb{U}$ and the corresponding set $\mathbb{U}_0$ above, we let

1. $U^{-(j,i)}$ be the collection of finite positive Borel measures on $[0, T] \times U^{-(j,i)}$ whose projection on $[0, T]$ is the Lebesgue measure, with the associated subset $\mathbb{U}_0^{-(j,i)}$;
2. $U^M$ be the collection of finite positive Borel measures on $[0, T] \times U^M$ whose projection on $[0, T]$ is the Lebesgue measure, with the associated subset $\mathbb{U}_0^M$.

Informally, the set $\mathbb{U}_0$ allows us to define the set of admissible efforts of all the workers, the set $\mathbb{U}_0^{-(j,i)}$ is used for the efforts of other agents apart from the $(j, i)$th, and $\mathbb{U}_0^M$ for the efforts of the managers. In the same way as we defined $\Omega$ and an appropriate subset $\mathcal{P}$ of probability measures in Definition A.1, we now consider the two canonical spaces

1. $(\Omega^{-(j,i)}, \mathcal{F}_T^{-(j,i)})$ for other agents (apart from the $(j, i)$th agent), where
   \[\Omega^{-(j,i)} := C([0, T]; \mathbb{R}^{w-m-1}) \times C([0, T]; \mathbb{R}^{d(w-m-1)}) \times U^{-(j,i)},\]
   with canonical process $(X^{-(j,i)}, W^{-(j,i)}, \Pi^{-(j,i)})$ and canonical filtration $\mathbb{F}^{-(j,i)}$, on which we define an appropriate subset $\mathcal{P}^{-(j,i)}$ of probability measures;
2. $(\Omega^M, \mathcal{F}_T^M)$, the canonical space for the managers, where
   \[\Omega^M := C([0, T]; \mathbb{R}^m) \times C([0, T]; \mathbb{R}^{dm}) \times U^M,\]
with canonical process \((X^M, W^M, \Pi^M)\) and filtration \(\mathbb{F}^M\), on which we define an appropriate subset \(\mathcal{P}^M\) of probability measures.

Informally, the canonical space \(\Omega^{-(j,i)}\) contains the same information as \(\Omega\) except that the components concerning the \((j, i)\)th agent and the managers are removed, and thus \(\Omega^{-(j,i)}\) is a subset of \(\Omega\). Similarly, the space \(\Omega^M\) contains the information of \(\Omega\) concerning the managers. We can now define for the \((j, i)\)th agent the set of his admissible response to others’ actions.

**Definition A.3** Consider two probability measures \(P^M \in \mathcal{P}^M\) and \(P^{-(j,i)} \in \mathcal{P}^{-(j,i)}\), chosen by the managers and the other agents, respectively. The set of admissible responses of the \((j, i)\)th agent, denoted by \(P^{j,i}(P^{-(j,i)}, P^M)\), is given by all probability measures \(P \in \mathcal{P}\) on \((\Omega, F_T)\) such that

(i) the restriction of \(P\) to \((\Omega^{-(j,i)}, F^{-(j,i)}_T)\) is \(P^{-(j,i)}\);

(ii) the restriction of \(P\) to \((\Omega^M, F^M_T)\) is \(P^M\).

Then, using Definition 4.3, we can properly define a Nash equilibrium between agents.

**Definition A.4** Fix a probability measure \(P^M \in \mathcal{P}^M\) chosen by the managers and a collection \(\xi^A \in C^A\) of contracts for the agents. We denote by \(\mathcal{P}^{A,\ast}(P^M, \xi^A)\) the set of Nash equilibria between agents, i.e., the set of probability measures \(P \in \mathcal{P}\) such that for any \(j \in \{1, \ldots, m\}\) and any \(i \in \{1, \ldots, n_j\}\), \(P \in \mathcal{P}^{j,i,\ast}(\mathcal{P}^{-(j,i)}, P^M, \xi^{j,i})\), where \(\mathcal{P}^{-(j,i)}\) is the restriction of \(P\) to \(\Omega^{-(j,i)}\).

### A.3 Canonical space and weak formulation for the managers

Under Assumption 5.9 and given the form (4.12) for the managers’ contracts, \(\zeta\) is clearly the only state variable of the managers’ problems. We are thus led to consider the canonical space

\[
\Omega^M := C([0, T]; \mathbb{R}^{hm}) \times C([0, T]; \mathbb{R}^{wd}) \times \mathbb{X},
\]

where \(\mathbb{X}\) and its subset \(\mathbb{X}_0\) used below are defined in a similar way as \(U\) and \(U_0\) for the initial canonical space in Appendix A.1. The canonical process is denoted by \((\zeta, W, \Pi^M)\), where for any \((t, \omega, \sigma, q) \in [0, T] \times \Omega\),

\[
\zeta_t(\omega, \sigma, q) := \omega_t, \quad W_t(\omega, \sigma, q) := \sigma_t, \quad \Pi^M(\omega, \sigma, q) := q.
\]

The associated canonical filtration is defined by \(\mathbb{F}^M := (\mathcal{F}^M_t)_{t \in [0, T]}\) with

\[
\mathcal{F}^M_t := \sigma \left( \left( \zeta_s, \int_0^s \int_X \varphi(r, u) \Pi^M(dr, du) : (s, \varphi) \in [0, t] \times C_b([0, T] \times \mathbb{X}; \mathbb{R}) \right) \right)
\]
for \( t \in [0, T] \). Then for any \((t, \psi) \in [0, T] \times C^2_b(\mathbb{R}^{hm} \times \mathbb{R}^{dw}, \mathbb{R})\), we set
\[
M_t^M(\psi) := \psi(\xi_t, W_t) - \int_0^t \int_X \left( \tilde{A}_M(s, \xi, u) \cdot \nabla \psi(\xi_s, W_s) + \frac{1}{2} \text{Tr}(\nabla^2 \psi(\xi_s, W_s)(\tilde{\Sigma}_M \tilde{\Sigma}_M^T)(s, \xi, u)) \right) \Pi^M(ds, du),
\]
where \( \tilde{A}_M \) and \( \tilde{\Sigma}_M \) are respectively the drift vector and the diffusion matrix of the \((hm + dw)\)-dimensional vector process \((\xi, W)\)\(\uparrow\), i.e.,
\[
\tilde{A}_M(s, x, u) := \begin{pmatrix} (\Sigma_M \Lambda_M)(s, x, u) \\ 0_{wd} \end{pmatrix}, \quad \tilde{\Sigma}_M(s, x, u) := \begin{pmatrix} 0_{hm, hm} & \Sigma_M(s, x, u) \\ 0_{wd,hm} & 1_{wd} \end{pmatrix},
\]
for \((s, x, u) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}\), with \( \Lambda_M, \Sigma_M \) defined in Assumption 5.9.

**Remark A.5** One may notice that the set \( X \) above is different from the set \( \mathbb{U}_M \) defined in the previous section, while both are related to the managers’ controls. The set \( \mathbb{U}_M \) is related to the managers’ controls impacting the dynamic of \( X \), i.e., their control \( \nu^M := (\alpha^M, \beta^M) \). These controls are first fixed, together with the contracts, in order to solve the agents’ problems. Once the agents’ problems are solved, we can address the managers’ problems. Initially, each manager chooses a contract (in a strong formulation) and his effort (in a weak formulation). As it makes little sense to consider this mixed formulation, the purpose of this section is to rigorously write the weak formulation of the manager’s problem. This is why we define at this stage a new set \( X \) related to all the controls of the managers – not only \( \nu^M \) impacting the dynamics of \( X \), but also the sensitivity parameters \( Z \) and \( \Gamma \) in the agents’ contracts. This collection of controls actually corresponds to \( \chi \), impacting the dynamics of the new state variable \( \xi \).

We fix an initial condition for the process \( \xi \), namely \( \varrho_0 \in \mathbb{R}^{hm} \). Similarly as in Definition A.1 for the initial control problem, we define the subset \( \mathcal{P}_M \) of probability measures \( \mathbb{P} \) on \((\Omega^M, \mathcal{F}_T^M)\) satisfying the following conditions:

(i) For all \( \psi \in C^2_b(\mathbb{R}^{hm} \times \mathbb{R}^{dw}, \mathbb{R}) \), the process \( M^M(\psi) \) is an \((\mathbb{P}_M, \mathbb{P})\)-local martingale on \([0, T] \);

(ii) \( \mathbb{P}[\{\xi_0, W_0\} = (\varrho_0, w_0)] = 1 \);

(iii) \( \mathbb{P}[\Pi^M \in X_0] = 1 \).

Similarly to Lemma A.2, we know that for all \( \mathbb{P} \in \mathcal{P}_M \), there exists an \( \mathbb{P} \)-predictable control process \( \chi^\mathbb{P} \) such that \( \Pi^M(ds, du) = \delta_{\chi^\mathbb{P}}(du)ds \) \( \mathbb{P} \)-a.s., and we thus obtain the representation \((5.7)\) for the dynamics of \( \xi \), but controlled by \( \chi := \chi^\mathbb{P} \). However, this representation only gives access to an admissible set of controls in terms of probability measures for all managers, namely \( \mathcal{P}_M \).

To properly define the choice of a particular manager in response to others, we need to define his own canonical space. With this in mind, we fix throughout the following \( j \in \{1, \ldots, m\} \) as well as the controls \( \chi^{-j} \in \mathcal{X}^{-j} \) chosen by other managers.
The canonical space for the \( j \)th manager is defined by

\[
\Omega_j := C([0, T]; \mathbb{R}^{hm}) \times C([0, T]; \mathbb{R}^{wd}) \times X_j,
\]

where \( X_j \) and its subset \( X_{0j} \) used below are defined in a similar way as \( X \) and \( X_0 \) above. The canonical process is denoted by \((\zeta, W, \Pi_1j)\), and the associated canonical filtration \( F_j := (F_j^t)_{t \in [0, T]} \) is defined as usual. Then, recalling that \( \chi^{-j} \) is fixed throughout this section, we set, for any \((t, \psi) \in [0, T] \times C^2_b(\mathbb{R}^{hm} \times \mathbb{R}^{wd}, \mathbb{R})\),

\[
M_j^t(\psi) := \psi(\zeta_t, W_t) - \int_0^t \int_{X_j} \left( \tilde{\Lambda}_M(s, \zeta, u \otimes_j \chi^{-j}_s) \cdot \nabla \psi(\zeta_s, W_s) \right. \\
+ \frac{1}{2} \text{Tr} \left( \nabla^2 \psi(\zeta_s, W_s)(\tilde{\Sigma}_M \tilde{\Sigma}_M^\top)(s, \zeta, u \otimes_j \chi^{-j}_s) \right) \Pi_j(ds, du),
\]

where \((u \otimes_j \chi^{-j}_t)^i = u_i \) and \((u \otimes_j \chi^{-j}_t)^k = \chi_k^j \) for \( k \neq j \).

We can then define the relevant subset \( \mathcal{P}^j(\chi^{-j}) \) of probability measures \( \mathbb{P} \) on \((\Omega^j, \mathcal{F}_T^j)\), satisfying \( \mathbb{P}[\Pi^j \in X_{0j}] = 1 \) and that \( M_j^t(\psi) \) is an \((\mathcal{F}_t, \mathbb{P})\)-local martingale for all \( \psi \in C^2_b(\mathbb{R}^{hm} \times \mathbb{R}^{wd}, \mathbb{R}) \). Then for all \( \mathbb{P} \in \mathcal{P}^j(\chi^{-j}) \), we obtain that \( \Pi_j(ds, du) = \delta_{\chi^j_P}(du)ds \mathbb{P}\text{-a.s.} \) for some \( \mathbb{P}^j\)-predictable control process \( \chi^j_P \), and we obtain the representation (5.7) for the dynamics of \( \zeta \), but controlled by \( \chi := \chi^j_P \otimes_j \chi^{-j} \).

**Definition A.6** Fix a collection \( \xi^M := (\xi^{j,0})_{j=1}^m \in C^M \) of contracts designed by the principal for the managers. A Nash equilibrium between the managers is a control \( \chi \in \mathcal{X} \) such that there exists \( \mathbb{P}^* \in \mathcal{P}^M \) satisfying \( \mathbb{P}^* \in \mathcal{P}^j(\chi^{-j}) \) for \( j \in \{1, \ldots, m\} \) and such that the supremum in (5.8) is attained for this \( \mathbb{P}^* \). We denote by \( \mathcal{P}^{M, *}(\xi^M) \) the set of Nash equilibria.

**Appendix B: The underlying theory of 2BSDEs**

The theoretical framework developed throughout this paper strongly relies on the recent theory of 2BSDEs, which is thus presented in this appendix. More precisely, Appendix B.1 defines some additional notations and Appendix B.2 clarifies the link between a Nash equilibrium for the managers and the theory of 2BSDEs. Finally, Appendix B.3 regroups the proofs of the results established only for the principal–managers level. The results and proofs for the manager–agent level are similar and thus omitted in this paper for the sake of brevity, but they can be found in [31, Sect. 3.7.1].

**B.1 Additional notations**

Throughout this section, let \( \mathbb{K} := (K_t)_{t \in [0, T]} \) be an arbitrary filtration and \( \mathcal{P} \) any set of probability measures on \((\Omega, \mathcal{F}_T)\).
B.1.1 Filtrations

We denote by $K^+_t := (K^+_t)_{t \in [0,T]}$ the right limit of $K$, i.e., $K^+_t := \bigcap_{s > t} K_s$ for all $t \in [0,T)$ and $K^+_T := K_T$. For any $\mathbb{P} \in \mathcal{P}$, we denote by $K^+_t := (K^+_t)_{t \in [0,T]}$ the completed filtration, where $K^+_t$ is the completion (in $\mathcal{F}_T$) of the $\sigma$-field $K_t$ under $\mathbb{P}$. Denote by $K^{\mathbb{P}}$ the right limit of $K^+_t$, so that $K^{\mathbb{P}}$ satisfies the usual conditions. In addition, the filtrations $K^{\mathbb{P}}_t := (K^{\mathbb{P}}_t)_{t \in [0,T]}$ are defined by $K^{\mathbb{P}}_t := \bigcap_{\mathbb{P}' \in \mathcal{P}} K^{\mathbb{P}'}_t$, $K^{\mathbb{P}}_t + := K^{\mathbb{P}}_t +$ for $t \in [0,T)$, $K^{\mathbb{P}}_T := K^{\mathbb{P}}_T$. Finally, we use the notation, for any $(\mathbb{P}, t) \in \mathcal{P} \times [0,T)$,

$$\mathcal{P}(t, \mathbb{P}, K) := \{ \mathbb{P}' \in \mathcal{P} : \mathbb{P}[E] = \mathbb{P}'[E] \text{ for all } E \in K^+_t \}. \quad (B.1)$$

B.1.2 Canonical spaces and norms

Let $\Sigma$ be an $S^\ell$-valued process, where $S^\ell$ denotes the set of symmetric positive matrices in $M^\ell$ for $\ell \in \mathbb{N}$, and $p > 1$. To define the solution of a 2BSDE in our framework, we consider the following spaces with their associated norms:

(i) $H^p_\ell(K, \mathbb{P}, \Sigma)$ is the space of $K^\mathbb{P}$-progressively measurable $\mathbb{R}^\ell$-valued processes $Z$ satisfying

$$\|Z\|_{H^p_\ell(K, \mathbb{P}, \Sigma)} := \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ \left( \int_0^T Z_t^\top \Sigma_t Z_t dt \right)^{p/2} \right] < \infty;$$

(ii) $D^p(K, \mathbb{P})$ is the space of $K^{\mathbb{P}}$-optional $\mathbb{R}$-valued càdlàg processes $Y$ with

$$\|Y\|_{D^p(K, \mathbb{P})} := \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ \sup_{0 \leq t \leq T} |Y_t|^p \right] < \infty;$$

(iii) $\mathcal{I}^p(K, \mathbb{P})$ is the space of all $K^{\mathbb{P}}$-optional càdlàg and nondecreasing processes $K \in D^p(K, \mathbb{P})$ satisfying $K_0 = 0$. 

B.2 2BSDE representation for a manager

B.2.1 Another representation for the set of measures

The general approach to moral hazard problems in Cvitanić et al. [15] requires to distinguish between the efforts of the agents which give rise to absolutely continuous probability measures in $\mathcal{P}^M$, namely those for which only the drift changes, and those for which the volatility control changes, while keeping fixed the quadratic variation of $\zeta$. The goal of this subsection is to provide the appropriate formulation in our setting.

Definition B.1 We define as $\mathcal{P}^M$ the set of probabilities $\mathcal{P}$ on $(\Omega^M, \mathcal{F}^M_T)$ such that
(i) the canonical vector process \((\zeta, W)^\top\) is an \((\mathcal{F}_\mathbb{M}, \mathbb{P})\)-local martingale for which there exists an \(\mathbb{F}_\mathbb{M}\)-predictable and \(\mathcal{X}\)-valued process \(\chi^\mathbb{P}\) such that the \(\mathbb{P}\)-quadratic variation of \((\zeta, W)^\top\) is equal to

\[
\begin{pmatrix}
\Sigma M \Sigma^\top M(t, \zeta, \chi^\mathbb{P}_t) & \Sigma M(t, \zeta, \chi^\mathbb{P}_t)
\end{pmatrix}
\begin{pmatrix}
\Sigma M \Sigma^\top M(t, \zeta, \chi^\mathbb{P}_t)
I_{\text{wd}}
\end{pmatrix},
t \in [0, T], \mathbb{P}\text{-a.s.};
\]

(ii) \(\mathbb{P}[\Pi \in \mathbb{H}_0] = 1\).

We thus know that for all \(\mathbb{P} \in \mathcal{P}^\mathbb{M}\), we have for \(\zeta\) the representation

\[
\zeta_t = \zeta_0 + \int_0^t \Sigma M(s, \zeta, \chi^\mathbb{P}_s) dW_s,
t \in [0, T], \mathbb{P}\text{-a.s.}
\]

We can also define a pathwise version of the \(\mathbb{F}_\mathbb{M}\)-predictable quadratic variation \((\zeta)\), allowing us to define the nonnegative symmetric matrix \(\hat{\Sigma}_t\) for all \(t \in [0, T]\) by

\[
\hat{\Sigma}_t := \lim \sup_{n \to \infty} n((\langle \zeta \rangle_t - \langle \zeta \rangle_{t-1/n})).
\]  

Since \(\hat{\Sigma}_t\) takes values in \(\mathbb{S}_{\text{hm}}\), we can naturally define its square root \(\hat{\Sigma}_t^{1/2}\).

**Definition B.2** Let \(\mathbb{P} \in \mathcal{P}^\mathbb{M}\) and consider the process \(\chi^\mathbb{P}\) associated to \(\mathbb{P}\) in the sense of Definition B.1 (i). For any \(\mathcal{X}\)-valued and \(\mathbb{F}_\mathbb{M}\)-predictable process \(\chi\) such that

\[
\Sigma M \Sigma^\top M(t, \zeta, \chi^\mathbb{P}_t) = \Sigma M(t, \zeta, \chi^\mathbb{P}_t),
t \in [0, T], \mathbb{P}\text{-a.s.,}
\]

we define the equivalent measure \(\mathbb{P}^\chi\) via its Radon–Nikodým density on \(\mathcal{F}_T^\mathbb{M}\) by

\[
\frac{d\mathbb{P}^\chi}{d\mathbb{P}} := \exp\left(\int_0^T \Lambda M(s, \zeta, \chi_s) \cdot dW_s - \frac{1}{2} \int_0^T |\Lambda M(s, \zeta, \chi_s)|^2 ds\right).
\]

Notice that \(\mathbb{P}^\chi\) is well defined since \(\Lambda M\) is bounded. It is then immediate to check that the set \(\mathcal{P}^\mathbb{M}\) coincides exactly with the set of all probability measures of the form \(\mathbb{P}^\chi\) which satisfy in addition that \(\mathbb{P}^\chi \circ (\zeta_0, W_0)^{-1} = \delta(\varrho, \iota)\) for some \(\iota \in \mathbb{R}^\text{wd}\).

Following the reasoning developed in Sect. 5.2, it is necessary to characterise the admissible sets for the actions of a considered manager in response to other managers’ choices. In particular, this leads to the definition in Appendix A.3 of \(\mathcal{P}^j(\chi^{-j})\) in addition to the definition of \(\mathcal{P}^\mathbb{M}\) on the whole canonical space, when actions of other managers are fixed through \(\chi^{-j} \in \mathcal{X}^{-j}\). We are therefore led to consider the set \(\mathcal{P}^j(\chi^{-j})\) corresponding to \(\mathcal{P}^j(\chi^{-j})\) in the same way that we just constructed \(\mathbb{P}^\mathbb{M}\) corresponding to \(\mathcal{P}^\mathbb{M}\) in Definition B.2.
B.2.2 Best-reaction function of a manager

In this subsection, for a given admissible contract $\xi_j \in C^{j,0}$ in the sense of Definition 4.7 and for the choices $\chi^{-j}$ of other managers, we wish to relate the best-reaction function $V^{j,0}_0(\xi^{j,0}, \chi^{-j})$ of the $j$th manager, defined by (5.8), to an appropriate 2BSDE. With this in mind, we fix throughout the following $j \in \{1, \ldots, m\}$ as well as the effort of other managers summarised by $\chi^{-j} \in \mathcal{X}^{-j}$ in order to focus on the $j$th manager. For simplicity, we now set $\mathbb{P} := \mathbb{P}_j(\chi^{-j})$.

First, we define for any $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^{hm})$ and $S \in \mathcal{S}_j^l(x, \chi^{-j})$

$$S_j^l(x, \chi^{-j}) := \{ \Sigma \Sigma^\top t(x, u \otimes_j \chi_i^{-j}) \in \mathbb{S}^{hm}_{+} : u \in \mathcal{X}^l \},$$

$$\tilde{X}_j^l(x, \chi^{-j}, S) := \{ u \in \mathcal{X}^l : \Sigma \Sigma^\top t(x, u \otimes_j \chi_i^{-j}) = S \}.$$

With these notations, we can isolate the partial maximisation with respect to the squared diffusion in the Hamiltonian of the $j$th manager. Indeed, we can define a map $F_j$ for all $(t, x, y, z, \chi_{-j}, S) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R} \times \mathbb{R}^{hm} \times \mathcal{X}_{-j}^{j} \times \mathbb{S}^{hm}_{+}$ by

$$F_j(t, x, y, z, \chi_{-j}, S) := \sup_{u \in \tilde{X}_j^l(x, \chi^{-j}, S)} \tilde{h}_j(t, x, y, z, \chi_{-j}, u), \quad (B.3)$$

where in addition, for $u \in \tilde{X}_j^l(x, \chi^{-j}, S)$,

$$\tilde{h}_j(t, x, y, z, \chi_{-j}, u) := -(c^{j,0} + yk^{j,0})(t, x^j, u) + (\Sigma \Lambda \Sigma^\top t)(t, x, u \otimes_j \chi_i^{-j}) \cdot z.$$

We thus obtain that the Hamiltonian $\mathcal{H}_j$ of the $j$th manager, defined by (5.9), satisfies for all $(t, x, y, z, \gamma, \chi_{-j}) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R} \times \mathbb{R}^{hm} \times \mathcal{M}^{hm} \times \mathcal{X}_{-j}^{j}$ that

$$\mathcal{H}_j(t, x, y, z, \gamma, \chi_{-j}) = \sup_{S \in \mathcal{S}_j^l(x, \chi^{-j})} \left( F_j(t, x, y, z, \chi_{-j}, S) + \frac{1}{2} \text{Tr}(S \gamma) \right).$$

Given an admissible contract $\xi_j \in C^{j,0}$ and using the definition of $F_j$ in (B.3), we are led to consider the 2BSDE, indexed by $j$,

$$\mathcal{Y}_t = g^{j,0}(\xi_j^l, \xi_j^j) + \int_t^T F_j(s, \xi_j^l, Z_s, \chi^{-j}, \tilde{\Sigma}_s) ds$$

$$- \int_t^T Z_s \cdot d\xi_s + \int_t^T dK_s. \quad (B.4)$$

The following definition adapts the classic notion of 2BSDEs to our framework, using the notations defined in Appendix B.1.1. In particular, recall here that $\mathcal{P} := \overline{\mathbb{P}}_j(\chi^{-j})$ and that $\mathcal{G}$ is the natural filtration generated by $\zeta$.

**Definition B.3** We say that $(\mathcal{Y}, Z, K)$ is a solution to (B.4) if (B.4) holds $\mathcal{P}$-q.s. and if for some $k > 1$, $\mathcal{Y} \in \mathbb{D}^k(\mathcal{G}, \mathcal{P})$, $Z \in \mathbb{H}^k_{hm}(\mathcal{G}, \mathcal{P}, \tilde{\Sigma})$, $K \in \mathbb{I}^k(\mathcal{G}, \mathcal{P})$, where $K$ satisfies
in addition the minimality condition

\[ 0 = \mathbb{P}^{-}\text{ess inf}_{\mathbb{P}' \in \mathcal{P}(t, \mathbb{P}, G)} \mathbb{E}^{\mathbb{P}'}[K_T - K_t | \mathcal{G}^+_t], \quad 0 \leq t \leq T, \mathbb{P}\text{-a.s. for all } \mathbb{P} \in \mathcal{Q}, \]

recalling that \( \mathcal{P}(t, \mathbb{P}, G) \) is defined by (B.1) and \( \mathcal{G}^+ \) is the right limit of the completion of \( G \) under \( \mathbb{P} \).

The following result relates the solution to the above 2BSDE to the best-reaction function of the \( j \)th manager.

**Proposition B.4** Fix \( \chi^{-j} \in \mathcal{X}^{-j} \) as well as \( \xi^j \in \mathcal{C}^{j,0} \). Let \( (\mathcal{Y}, Z, K) \) be a solution to (B.4). We have

\[ V^{j,0}_0(\xi^j, \chi^{-j}) = \sup_{\mathbb{P} \in \mathcal{P}^j(\chi^{-j})} \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_0]. \]

Conversely, the (dynamic) value function \( (V^{j,0}_t(\xi^j, \chi^{-j}))_{t \in [0, T]} \) always provides the first component \( \mathcal{Y} \) of a solution to (B.4). Moreover, any optimal effort \( \tilde{\chi}^{j,*} \) and the optimal measure \( \mathbb{P}^\tilde{\chi} \) must be such that \( \mathbb{P}^\tilde{\chi} \) is defined from \( \mathbb{P} \) and \( \tilde{\chi} \) by Definition B.2, where \( \tilde{\chi} := \chi^{j,*} \otimes_j \chi^{-j} \).

**Proof** The proof is classical and follows the lines of [15, Propositions 4.5 and 4.6]. We thus only mention here why the assumptions required to apply the results of Possamaï et al. [50] are satisfied in our framework.

First of all, recall that by the definitions of \( k^{j,0} \) in Sect. 4.2.2 and \( \Sigma_M, \Lambda_M \) in Assumption 5.9, these functions are assumed to be bounded. As in [15, Proposition 4.5], it follows that for all \( (t, x, y, z) \in [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R} \times \mathbb{R}^{hm}, S \in S^j_t(x, \chi^{-j}) \) and \( (y', z') \in \mathbb{R} \times \mathbb{R}^{hm} \), we have

\[ |F^j(t, x, y, z, \chi^{-j}, S) - F^j(t, x, y', z', \chi^{-j}, S)| \leq |k^{j,0}|_{\infty} |y - y'| + |\Lambda_M|_{\infty} |S|^{1/2}(z - z'), \]

ensuring that \( F^j \) is Lipschitz-continuous in \( y \) and in \( S^{1/2}z \) as required in [50, Assumption 2.1 (i)]. Moreover, by Definition 4.7 of the set \( \mathcal{C}^{j,0} \) of admissible contracts, the terminal condition \( g^{j,0}(\xi^j, \mathcal{X}) \) satisfies (4.16). Using in addition the integrability condition (4.13) for \( \mathcal{C}^{j,0} \), it then follows that the terminal condition \( g^{j,0}(\xi^j, \mathcal{X}) \) and \( F^j \) satisfy the integrability properties in [50, Assumption 1.1 (ii)]. Indeed, we have

\[ \sup_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[|g^{j,0}(\xi^j, \mathcal{X})|^p] < \infty. \]
by Condition (4.16), and by (4.13), for some $p > 1$,
\[
\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ \int_t^T |F_j(s, \zeta, 0, 0, \chi^{-j}, \Sigma_s)|^p ds \right] \leq \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P} \left[ \int_0^T |c^{i,0}(s, \zeta, \chi^{i,\mathbb{P}})|^p ds \right]
\]
is also finite for all $t \in [0, T]$. In addition, [50, Assumption 3.1] is satisfied thanks to the integrability condition (4.13) for $c^{i,0}$, as explained in [15, Proposition 4.5]. Next, [50, Assumption 1.1 (iii)–(v)] are satisfied by the set of measures $\mathcal{Q}^j := \mathbb{P}^j(\chi^{-j})$; see for instance Nutz and van Handel [49]. Finally, the set $\mathcal{Q}^j$ is saturated in the sense of [50, Definition 5.1]; see [50, Remark 5.1].

\[\Box\]

**B.2.3 Characterisation of the Nash equilibrium between managers**

With Proposition B.4 in hand, we can now characterise Nash equilibria between the managers by a collection of coupled 2BSDEs, reminiscent of the multidimensional BSDE obtained in the setting of Élie and Possamaï [23] where only the drift of the process was controlled.

**Theorem B.5** Let $\xi^M \in C^M$ be the collection of contracts for the managers, meaning that the $j$th manager receives a contract $\xi^j \in C^{j,0}$. The unique Nash equilibrium $\mathbb{P} \in \mathcal{P}^M(\xi^M)$ in the sense of Definition A.6 is characterised by $\mathbb{P} = \mathbb{P}^{\tilde{\chi}^*}$, where $\tilde{\chi}^*$ is such that for any $j \in \{1, \ldots, m\}$,

\[
K_j^t = 0,
\]

\[
\tilde{\chi}^{j,*} = \arg \max_{u \in \tilde{X}_t^j(\tau, \tilde{\chi}^{-j,*}, \Sigma_t)} \tilde{h}^j(t, \tau, Y_j^t, Z_j^t, \tilde{\chi}^{-j,*}, u), \quad \forall t \in [0, T], \mathbb{P}^{\tilde{\chi}^*}\text{-a.s.,}
\]

where $(Y_j, Z_j, K_j)$ is a solution to (B.4) in the sense of Definition B.3 on $\mathbb{P}^{\tilde{\chi}^*}(\tilde{\chi}^{-j,*})$.

**Proof** As in the statement of the theorem, we fix a collection $\xi^M \in C^M$ of contracts for the managers. Recall that by definition of the set $C^M$, the collection of contracts $\xi^M$ leads to a unique Nash equilibrium between the managers (see Definition 4.7). By Proposition B.4, we have a characterisation of the best-reaction function of the $j$th manager to an arbitrary tuple of controls $\chi^{-j}$ chosen by the other managers. A Nash equilibrium $\mathbb{P}^*$ associated to an optimal effort $\tilde{\chi}^* := (\tilde{\chi}^{j,*})_{j=1}^m$ then necessitates only that for all $j \in \{1, \ldots, m\}$, $\mathbb{P}^*$ is the best-reaction function of the $j$th manager to $\tilde{\chi}^{-j,*}$. In other words, the probability $\mathbb{P}^*$ and the associated effort $\tilde{\chi}^*$ have to satisfy (B.5) for all $j \in \{1, \ldots, m\}$, which ends the proof. \[\Box\]

Theorem B.5 leads us to consider, if it exists, a map $\tilde{u}$ taking values in $\mathbb{R}^m$ and satisfying for each coordinate $j \in \{1, \ldots, m\}$ that

\[
\tilde{u}^j(t, x, y, z, S) = \arg \max_{u \in \tilde{X}_t^j(x, \tilde{\chi}^{-j}(x, y, z, S), S)} \tilde{h}^j(t, x, y, z, S, u)
\]
Continuous-time incentives in hierarchies

for \((t, x) \in [0, T] \times C([0, T]; \mathbb{R}^{hm})\), \(y = (y^j_{m})_{j=1} \in \mathbb{R}^m\), \(z = (z^j_{m})_{j=1} \in (\mathbb{R}^{hm})^m\), \(S \in \mathbb{S}^{hm}_+\). We then define \(F^* : [0, T] \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^m \times (\mathbb{R}^{hm})^m \times \mathbb{S}^{hm}_+ \to \mathbb{R}^m\) such that each component satisfies

\[
F_j^*(t, x, y, z, S) := F^j(t, x, y_j, z_j, \tilde{u}^j(t, x, y, z, S)), \quad j \in \{1, \ldots, m\}.
\]

We can then consider a triple \((\mathcal{Y}, Z, K)\) which is a solution of a multidimensional 2BSDE in the sense that each component \((\mathcal{Y}^j, Z^j, K^j)\) is a solution to the 2BSDE

\[
\mathcal{Y}^j_t = g^{j, 0}(\xi^j_t, \xi^j_t) + \int_t^T F_j^*(s, \xi, \mathcal{Y}^j_s, Z^j_s, \tilde{\Sigma}_s)s - \int_t^T Z^j_s \cdot d\xi_s + \int_t^T dK^j_s
\]

\(\mathcal{P}^j(\tilde{\chi}^{j,*})\)-q.s., where \(\tilde{\chi}^j\) is defined component by component for all \(t \in [0, T]\) and \(j \in \{1, \ldots, m\}\) by \(\tilde{\chi}^j_t := \tilde{u}^j(t, \xi, \mathcal{Y}^j_t, Z^j_t, \tilde{\Sigma}_t)\).

This multidimensional 2BSDE is an extension of the pair of 2BSDEs considered by Possamaï et al. \[51, \text{Eqs. (3.22), (3.23)}\] in their framework of a zero-sum game with two interacting players. One can also relate this multidimensional 2BSDE to the mean-field and McKean–Vlasov 2BSDEs obtained respectively in Élie et al. \[21\] and by Barrasso and Touzi \[5\] in a framework with a continuum of agents with mean-field interactions.

### B.3 Technical proofs for the managers’ problem

Thanks to the previously established results, we now have everything in hand to prove Proposition 5.15 and Theorem 5.16, which will complete our study.

**Proof of Proposition 5.15** Let \(Y^0_0 := (Y^0_j)_{j=1}^m\) and \(Z := (Z, \Gamma) \in \mathcal{Y}\). By Definition 5.14, consider the \(m\)-dimensional process \(\mathcal{Y}^m := (\mathcal{Y}^j)_{j=1}^m\) as well as the associated collection of contracts \(\xi^m := (\xi^j)_{j=1}^m \in \Xi^m\). Note that each \(\xi^j\) naturally satisfies the properties to be admissible in the sense of Definition 4.7, and that it suffices to prove uniqueness of the Nash equilibrium to ensure that \(\xi^m \in C^M\).

**Existence of a Nash equilibrium.** We first fix \(j \in \{1, \ldots, m\}\) and assume that the other managers apart from the \(j\)th are playing according to \(\chi^{j,*} \) defined by (i), i.e.,

\[
\chi^{-j,*} = (\chi^{\ell,*})_{\ell=1, \ell \neq j},
\]

\[
\chi^{\ell,*}_t = u^{\ell,*}(t, \xi, \mathcal{Y}, Z, \Gamma), \quad t \in [0, T], \ell \neq j. \tag{B.6}
\]

We look for the best admissible response of the \(j\)th manager with respect to the effort of the others, i.e., a probability \(\mathbb{P} \in \mathcal{P}^j(\chi^{j,*})\) that maximises his utility. Informally, we want to prove that \(\chi^{j,*}\) is also given by (B.6). By our assumption on the contract, namely that \(\xi^j \in \Xi^j\), we have in particular that the continuation utility \(\mathcal{Y}^j\) of the \(j\)th manager satisfies (5.11). Recalling the definition of \(\mathcal{H}^{j,*}\) in (5.10), it is easy to see that \(\mathcal{Y}^j\) satisfies for all \(t \in [0, T]\) that

\[
\mathcal{Y}^j_t = Y^0_j - \int_0^t \mathcal{H}^j(r, \xi, \mathcal{Y}^j_r, Z^j_r, \Gamma_r, \chi^{-j,*}_r)dr + \int_0^t Z^j_r \cdot d\xi_r + \frac{1}{2} \int_0^t \text{Tr}(\Gamma^j_r \tilde{\Sigma}_r)dr.
\]
Recalling that we have denoted by $\hat{\Sigma}$ the pathwise version of the quadratic variation $\langle \xi \rangle$ (see (B.2)) and using the definition of $F^j$ in (B.3), we let, for all $r \in [0, T]$,
\[
dK^j_r := \left( H^j (r, \xi, \gamma^j_r, Z^j_r, \chi^j_r, \hat{\Sigma}_r) - \frac{1}{2} \text{Tr}(\Gamma^j_r \hat{\Sigma}_r) - F^j (r, \xi, \gamma^j_r, Z^j_r, \chi^j_r, \hat{\Sigma}_r) \right) \text{d}r.
\]
Plugging this into the above expression for $\gamma^j_t$, we obtain
\[
\gamma^j_t = \gamma^j_0 - \int_0^t F^j (r, \xi, \gamma^j_r, Z^j_r, \chi^j_r, \hat{\Sigma}_r) \text{d}r + \int_0^t Z^j_r \cdot \text{d}\xi_r - \int_0^t dK^j_r.
\]
Finally, recalling that the contract satisfies $\xi^j = \bar{g}^j(\xi^j, \gamma^j_t)$, we get $\gamma^j_T = g^j(\xi^j, \xi^j)$ and can rewrite the previous equation in backward form as
\[
\gamma^j_t = g^j(\xi^j, \xi^j) + \int_t^T F^j (r, \xi, \gamma^j_r, Z^j_r, \chi^j_r, \hat{\Sigma}_r) \text{d}r - \int_t^T Z^j_r \cdot \text{d}\xi_r + \int_t^T dK^j_r,
\]
which exactly corresponds to (B.4) under $\mathbb{P} := \mathbb{P}^j(\chi^j, \xi^j)$. By the definition of $F^j$, we can directly check that $K^j$ is always a nondecreasing process, which vanishes on the support of any probability measure corresponding to the efforts $\chi^j, \xi^j$ defined in the statement of the proposition. To ensure that $(\gamma^j, Z^j, K^j)$ solves (B.4), it therefore remains to check that all the integrability requirements in Definition B.3 are satisfied. The one for $\gamma^j$ is immediate by the definition of the set $V$. The required integrability for $(Z^j, K^j)$ follows from Bouchard et al. [10, Theorem 2.1 and Proposition 2.1].

**Uniqueness.** We have obtained that the candidate provided in the statement of the proposition is indeed an equilibrium. To prove uniqueness, let $\chi^{j, *}$ be the arbitrary efforts of other managers. In this case, the continuation utility of the $j$th manager, given a contract $\xi^j \in \Xi^j$, does not satisfies (B.4) since the other agents’ efforts are not optimal. However, $\Xi^j \subseteq C^{j, 0}$ and by Proposition B.4, an optimal effort $\chi^{j, *}$ is a maximiser of the map $F^j$ and coincides with a maximiser of his Hamiltonian $H^j$ given by (5.9). Recall that the existence of such a maximiser is ensured by Assumption 5.10, but uniqueness is not assumed. However, this reasoning is valid for all managers, implying that $\chi^*$ satisfies for all $j \in \{1, \ldots, m\}$ that $\chi^{j, *}$ is a maximiser of the map $u^*$ in Assumption 5.11. Then by the assumption on the uniqueness of the map $u^*$ guaranteeing the maximisation of each manager’s Hamiltonian simultaneously, this induces a unique equilibrium in terms of efforts, given by $\chi^*$.

**Proof of Theorem 5.16** The main point is to prove that the restriction to revealing contracts in the sense of Definition 5.14 is without loss of generality. This proof relies on arguments similar to the ones developed initially in Cvitanić et al. [15, Theorem 3.6]. In a nutshell, we consider an arbitrary collection $\xi^M \in C^M$ of contracts in the sense that the $j$th manager receives the contract $\xi^j \in C^{j, 0}$. Starting from this admissible collection of contracts, the goal is to prove that for all managers, we can define an
Continuous-time incentives in hierarchies

approximation \( \xi^{j,\varepsilon} \) of his contract \( \xi^j \) such that the new collection of contracts \( \xi^\varepsilon \) is admissible and gives the same Nash equilibrium. Moreover, we have to verify that for all \( j \in \{1, \ldots, m\} \), we have \( \xi^{j,\varepsilon} = \xi^j \) at the Nash equilibrium, and that the associated continuation utility \( Y^\varepsilon \) satisfies the representation (5.11), in addition to the required integrability conditions, to ensure that \( \xi^\varepsilon \) is a collection of revealing contracts in the sense of Definition 5.14.

First, using Proposition B.4 and Theorem B.5, we know that for a collection of contracts \( \xi^M \in C^M \), there exists a unique equilibrium \( P^* \in \mathcal{P}^M,*(\xi^M) \), associated to an optimal effort \( \tilde{\chi}^* \in \mathcal{X}^* \), satisfying for any \( j \in \{1, \ldots, m\} \) that

\[
\tag{B.7}
K^j = 0,
\]

\[
\tilde{\chi}_t^* = \arg \max_{u \in \tilde{\mathcal{X}}_t^j (\zeta, \tilde{\chi}^{j-}, \Sigma_t)} \tilde{\mathcal{Y}}_t^j (t, \zeta, \tilde{\chi}_t^j, Z_j^i, \tilde{\chi}^{j-}, u), \quad t \in [0, T], \mathbb{P}^*\text{-a.s.},
\]

where \( (\mathcal{Y}^j, Z_j, K^j) \) is a solution to (B.4) in the sense of Definition B.3.

Given this arbitrary but admissible collection \( \xi^M \) of contracts, the idea is to use the above solution \( (\mathcal{Y}^j, Z_j, K^j) \) to construct an approximation \( \xi^{j,\varepsilon} \) of the contract \( \xi^j \). Similarly to the agents’ problem, let us fix some \( \varepsilon > 0 \) and define for all \( j \in \{1, \ldots, m\} \) an absolutely continuous approximation of \( K^j \) by

\[
K^{j,\varepsilon}_t := \frac{1}{\varepsilon} \int_{(t-\varepsilon)^+}^t K^j_s \, ds, \quad t \in [0, T].
\]

The process \( K^{j,\varepsilon} \) naturally inherits some properties of \( K^j \). More precisely, given the effort \( \chi^{j-} \) of other managers and recalling the notation \( \mathcal{P} := \mathcal{P}^j (\chi^{j-}) \), we have that \( K^{j,\varepsilon} \) is \( \mathcal{G}^\mathcal{P} \)-predictable, nondecreasing \( \mathbb{P}^*\text{-q.s.} \) and

\[
K^{j,\varepsilon} = 0 \quad \mathbb{P}^*\text{-a.s.} \text{ for all } \mathbb{P}^* \in \mathcal{P}^M,*(\xi^M). \tag{B.8}
\]

Recalling that the optimal efforts \( \tilde{\chi}^{j-,*} \) of other managers are defined \( \omega \) by \( \omega \) through (B.7), we next define the \( m \)-dimensional process \( \mathcal{Y}^\varepsilon \) such that each component \( \mathcal{Y}^{j,\varepsilon} \) satisfies \( \mathcal{P}^j (\tilde{\chi}^{j-,*}) \)-q.s. for \( t \in [0, T] \) the SDE

\[
\mathcal{Y}^{j,\varepsilon}_t := Y^j_0 - \int_0^t F^j(s, \zeta, \mathcal{Y}^{j,\varepsilon}_s, Z_j^i, \tilde{\chi}^{j-,*}_s, \tilde{\Sigma}_s) \, ds + \int_0^t Z_j^i \cdot d\zeta_s - \int_0^t dK^{j,\varepsilon}_s. \tag{B.9}
\]

We first verify that for all \( j \in \{1, \ldots, m\} \), the triple \( (\mathcal{Y}^{j,\varepsilon}, Z_j^i, K^{j,\varepsilon}) \) solves (B.4) under \( \mathcal{P} := \mathcal{P}^j (\chi^{j-,*}) \) with terminal condition \( \mathcal{Y}^{j,\varepsilon}_T \) and generator \( F^j \). With this in mind, fix \( j \in \{1, \ldots, m\} \) as well as the other components \( (\mathcal{Y}^{j-,*}, Z_j^{j-}, K^{j-,*}) \). First, by (B.8), \( K^{j,\varepsilon} \) clearly satisfies the required minimality condition. Then \( K^{j,\varepsilon} \leq K^j \) inherits the integrability of \( K^j \), and moreover we can verify that

\[
sup_{\mathcal{P} \in \mathcal{P}^M} \mathbb{E}^\mathcal{P} [|\mathcal{Y}^{j,\varepsilon}_T|^p] < \infty, \text{ similarly to the equivalent proof for the manager–agents problem. Therefore, by Possamaï et al. [50, Theorem 4.4], we have the estimate}
\]

\[
\|\mathcal{Y}^{j,\varepsilon}\|_{\mathcal{H}^p(G, \mathfrak{N})} + \|Z^j_i\|_{\mathcal{H}^p^G(G, \mathfrak{N}, \Sigma)} < \infty \quad \text{for } \tilde{p} \in (1, p). \tag{B.10}
\]
We finally observe that a probability measure $\mathbb{P}$ satisfies $K^j = 0 \mathbb{P}$-a.s. if and only if it satisfies $K^e = 0 \mathbb{P}$-a.s. An approximation $\xi^{j,e}$ of the admissible contract $\xi^j$ can thus be defined, $\omega$ by $\omega$, from the terminal value of $\mathcal{Y}^{j,e}$ by $\xi^{j,e} := \mathcal{G}^{j,0}(\xi^j, \mathcal{Y}^{j,e})$, recalling that $\mathcal{G}^{j,0}$ corresponds to the inverse of $g^{j,0}$ with respect to the second variable. In other words, the approximation $\xi^{j,e}$ satisfies $\mathcal{Y}^{j,e} = g^{j,0}(\xi^j, \xi^{j,e})$.

To prove that the contract $\xi^{j,e}$ is revealing, meaning that it belongs to $\Sigma^{j,0}$, we should make the parameter $\Gamma$ appear in the representation (B.9). With this in mind, notice that for any $(t, \omega, x, y, z) \in [0, T] \times \Omega^M \times C([0, T]; \mathbb{R}^{hm}) \times \mathbb{R}^{hm} \times \mathbb{R}^{m+1}$, the map

$$
\gamma \mapsto \mathcal{H}^j(t, x, y, z, \gamma, \tilde{\chi}^{-j,*}) - \frac{1}{2} \text{Tr}(\gamma S(\omega)) - F^j(t, x, y, z, \tilde{\chi}^{-j,*}, S(\omega))
$$

is surjective onto $(0, \infty)$. Indeed, it is nonnegative (by the definition of $\mathcal{H}^j$ and $F^j$), convex, continuous on the interior of its domain, and coercive by the boundedness of the functions $\Lambda_M, \Sigma_M, k^{j,0}$ and $c^{j,0}$. Let $\tilde{K}^{j,e}$ denote the density of the absolutely continuous process $K^{j,e}$ with respect to Lebesgue measure. Applying a classical measurable selection argument (the maps appearing here are continuous and we can use the results from Beneš [6, 7]), we can deduce the existence of a $\mathcal{G}$-predictable process $\Gamma^{j,e}$ such that for $s \in [0, T]$,

$$
\dot{K}^{j,e}_s = \mathcal{H}^j(s, \zeta, \mathcal{Y}^{j,e}_s, Z^j_s, \Gamma^{j,e}_s, \tilde{\chi}^{-j,*}_s) - \frac{1}{2} \text{Tr}(\Gamma^{j,e}_s \tilde{\Sigma}_s) - F^j(s, \zeta, \mathcal{Y}^{j,e}_s, Z^j_s, \tilde{\chi}^{-j,*}_s, \tilde{\Sigma}_s).
$$

Indeed, if $\dot{K}^{j,e}_s > 0$, the existence of $\Gamma^{j,e}_s$ is clear from (B.11), and if $\dot{K}^{j,e}_s = 0$, then $\Gamma^{j,e}_s$ can be chosen arbitrarily. Substituting in (B.9), we obtain for all $t \in [0, T]$ for $\mathcal{Y}^{j,e}$ the representation

$$
\mathcal{Y}^{j,e}_t = Y^j_0 - \int_0^t \left( \mathcal{H}^j(r, \zeta, \mathcal{Y}^{j,e}_s, Z^j_s, \Gamma^{j,e}_s, \tilde{\chi}^{-j,*}_s)ds - Z^j_s \cdot d\zeta_s - \frac{1}{2} \text{Tr}(\Gamma^{j,e}_s \tilde{\Sigma}_s)ds \right).
$$

This shows that the continuation utility $\mathcal{Y}^{j,e}$ has the required dynamics (5.11) since at the equilibrium, the effort $\tilde{\chi}^* = \chi^*$ is unique. The fact that the contract $\xi^{j,e}$ induced by $Y^{j,e}$ belongs to $\Sigma^j$ then stems from (B.10). Moreover, notice that the admissible contract $\xi^j$ and its approximation $\xi^{j,e}$ coincide at the equilibrium in the sense that $\xi^{j,e} = \xi^j \mathbb{P}$-a.s. This reasoning is true for all $j \in \{1, \ldots, m\}$, and we have therefore constructed a well-suited approximation of the collection $\xi^M$ of contracts, belonging to $\Sigma^M$. Using Propositions B.4 and 5.15, we can then conclude as in the proof of [15, Theorem 3.6] since both collections of contracts lead to the same unique Nash equilibrium.

Finally, the equality (5.13) is now trivial. Indeed, by Definition 5.14, choosing a collection $\xi^M \in \Sigma^M$ of contracts is strictly equivalent to choosing both a pair of payment rates $\mathcal{Z} := (Z, \Gamma) \in \mathcal{V}$ to index the contract of each manager respectively on $d\zeta$ and $d(\zeta)$, and a constant $Y^0_0 := (Y^j_0)_{j=1}^m \in \mathbb{R}^m$. However, for all $j \in \{1, \ldots, m\}$, the constant $Y^j_0 \in \mathbb{R}$ must be chosen so that the participation constraint for the $j$th
Continuous-time incentives in hierarchies 659

manager is satisfied. Moreover, the principal also chooses the initial value $Y_{0}^{A} \in \mathbb{R}$ of the agents’ continuation utility such that Condition (4.11) is satisfied. Using Propositions 5.7 and 5.15 for the agents and the managers, respectively, these conditions are satisfied if and only if for all $j \in \{1, \ldots, m\}$ and $i \in \{0, \ldots, n_j\}$,

$$Y_{j,i}^i = V_{j,i,\ast}^i(\chi^{\ast}) \geq \rho_{j,i}^i \quad \text{and} \quad Y_{0}^j = V_{0}^{j,0,\ast}(\xi^M) \geq \rho_{j,0}^j,$$

recalling that $V_{0}^{j,i,\ast}$ and $V_{0}^{j,0,\ast}$ are defined by (5.6) and (5.12), respectively. This justifies the equality (5.13) and ends the proof. \hfill \Box

Acknowledgements The author is thankful to Dylan Possamaï (Department of Mathematics, ETH Zürich) for his insightful ideas and useful advice. Finally, the author is also thankful to anonymous referees for their careful and patient reading of the manuscript and for valuable comments.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Aïd, R., Possamaï, D., Touzi, N.: Optimal electricity demand response contracting with responsiveness incentives. Math. Oper. Res. 47, 2112–2137 (2022)
2. Alasseur, C., Farhat, H., Saguan, M.: A principal–agent approach to capacity remuneration mechanisms. Int. J. Theor. Appl. Finance 23, 2050052 (2020)
3. Baldacci, B., Possamaï, D.: Governmental incentives for green bonds investment. Math. Financ. Econ. 16, 539–585 (2022)
4. Baldacci, B., Possamaï, D., Rosenbaum, M.: Optimal make-take fees in a multi market-maker environment. SIAM J. Financ. Math. 12, 446–486 (2021)
5. Barrasso, A., Touzi, N.: Controlled diffusion mean field games with common noise and McKean–Vlasov second order backward SDEs. Theory Probab. Appl. 66, 613–639 (2022)
6. Beneš, V.E.: Existence of optimal strategies based on specified information, for a class of stochastic decision problems. SIAM J. Control 8, 179–188 (1970)
7. Beneš, V.E.: Existence of optimal stochastic control laws. SIAM J. Control 9, 446–472 (1971)
8. Biasi, B., Mariotti, T., Rochet, J.C., Villeneuve, S.: Large risks, limited liability, and dynamic moral hazard. Econometrica 78, 73–118 (2010)
9. Bichteler, K.: Stochastic integration and $L^p$-theory of semimartingales. Ann. Probab. 9, 49–89 (1981)
10. Bouchard, B., Possamaï, D., Tan, X., Zhou, C.: A unified approach to a priori estimates for supersolutions of BSDEs in general filtrations. Ann. Inst. Henri Poincaré Probab. Stat. 54, 154–172 (2018)
11. Calvo, G.A., Wellisz, S.: Supervision, loss of control, and the optimum size of the firm. J. Polit. Econ. 86, 943–952 (1978)
12. Calvo, G.A., Wellisz, S.: Hierarchy, ability, and income distribution. J. Polit. Econ. 87, 991–1010 (1979)
13. Carmona, R., Wang, P.: Finite-state contract theory with a principal and a field of agents. Manag. Sci. 67, 4725–4741 (2021)
14. Cvitanić, J., Possamaï, D., Touzi, N.: Moral hazard in dynamic risk management. Manag. Sci. 63, 3328–3346 (2017)
15. Cvitanić, J., Possamaï, D., Touzi, N.: Dynamic programming approach to principal–agent problems. Finance Stoch. 22, 1–37 (2018)
16. Cvitanić, J., Wan, X., Zhang, J.: Optimal compensation with hidden action and lump-sum payment in a continuous-time model. Appl. Math. Optim. 59, 99–146 (2009)
17. Cvitanić, J., Xing, H.: Asset pricing under optimal contracts. J. Econ. Theory 173, 142–180 (2018)
18. Cvitanić, J., Zhang, J.: Contract Theory in Continuous-Time Models. Springer, Berlin (2012)
19. Ekren, I., Touzi, N., Zhang, J.: Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I. Ann. Probab. **44**, 1212–1253 (2016)
20. Ekren, I., Touzi, N., Zhang, J.: Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part II. Ann. Probab. **44**, 2507–2553 (2016)
21. Élie, R., Hubert, E., Mastrolia, T., Possamaï, D.: Mean-field moral hazard for optimal energy demand response management. Math. Finance **31**, 399–473 (2021)
22. Élie, R., Mastrolia, T., Possamaï, D.: A tale of a principal and many, many agents. Math. Oper. Res. **44**, 440–467 (2018)
23. Élie, R., Possamaï, D.: Contracting theory with competitive interacting agents. SIAM J. Control Optim. **57**, 1157–1188 (2019)
24. Evans, L.C., Miller, C.W., Yang, I.: Concavity and optimality conditions for continuous-time principal–agent problems. Technical report, University of California, Berkeley and Massachusetts Institute of Technology (2015). https://math.berkeley.edu/~miller/papers/EvansMillerYang2015.pdf
25. Gryglewicz, S., Hartman-Glaser, B., Zheng, G.: Growth options, incentives, and pay for performance: theory and evidence. Manag. Sci. **66**, 1248–1277 (2020)
26. He, Z.: A model of dynamic compensation and capital structure. J. Financ. Econ. **100**, 351–366 (2011)
27. Hernández Santibáñez, N., Mastrolia, T.: Contract theory in a VUCA world. SIAM J. Control Optim. **57**, 3072–3100 (2019)
28. Holmström, B., Milgrom, P.: Aggregate and linearity in the provision of intertemporal incentives. Econometrica **55**, 303–328 (1987)
29. Hu, K., Ren, Z., Touzi, N.: Continuous-time principal–agent problem in degenerate systems. Preprint (2019). https://arxiv.org/abs/1910.10527
30. Huang, Y., Ju, N., Xing, H.: Optimal contracting with unobservable managerial hedging. Preprint (2017). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3079655
31. Hubert, E.: Interactions and incentives: between contract theory and mean-field games. PhD thesis, Université Gustave Eiffel (2020). https://theses.hal.science/tel-03325128
32. Itoh, H.: Job design and incentives in hierarchies with team production. Hitotsubashi J. Commer. Manag. **36**, 1–17 (2001)
33. Jensen, M.C., Murphy, K.J.: Performance pay and top-management incentives. J. Polit. Econ. **98**, 225–264 (1990)
34. Jost, P.J., Lammers, F.: Organization of project evaluation and implementation under moral hazard. J. Inst. Theor. Econ. **166**, 212–238 (2010)
35. Keppo, J., Touzi, N., Ruiting, Z.: Dynamic contracting in asset management under investor–partner–manager relationship. Preprint (2021). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3761499
36. Keren, M., Levhari, D.: The optimum span of control in a pure hierarchy. Manag. Sci. **25**, 1162–1172 (1979)
37. Knight, F.H.: Risk, Uncertainty and Profit. Courier Corporation, Chelmsford (2012)
38. Koo, H.K., Shim, G., Sung, J.: Optimal multi–agent performance measures for team contracts. Math. Finance **18**, 649–667 (2008)
39. Laffont, J.J.: Analysis of hidden gaming in a three–level hierarchy. J. Law Econ. Organ. **6**, 301–324 (1990)
40. Lazez, E.P.: Compensation and incentives in the workplace. J. Econ. Perspect. **32**(3), 195–214 (2018)
41. Li, N., Yu, Z.: Forward–backward stochastic differential equations and linear–quadratic generalized Stackelberg games. SIAM J. Control Optim. **56**, 4148–4180 (2018)
42. Lin, Y., Ren, Z., Touzi, N., Yang, J.: Second-order backward SDE with random terminal time. Electron. J. Probab. **25**, 1–43 (2020)
43. Lin, Y., Ren, Z., Touzi, N., Yang, J.: Random horizon principal–agent problems. SIAM J. Control Optim. **60**, 355–384 (2022)
44. Macho-Stadler, I., Pérez-Castrillo, J.D.: Centralized and decentralized contracts in a moral hazard environment. J. Ind. Econ. **46**, 489–510 (1998)
45. Miller, C.W., Yang, I.: Optimal dynamic contracts for a large-scale principal–agent hierarchy: a concavity-preserving approach. Preprint (2015). https://arxiv.org/abs/1506.05497
46. Mirrlees, J.A.: An exploration in the theory of optimum income taxation. Rev. Econ. Stud. **38**, 175–208 (1971)
47. Mirrlees, J.A.: The optimal structure of incentives and authority within an organization. Bell J. Econ. **7**, 105–131 (1976)
48. Neufeld, A., Nutz, M.: Measurability of semimartingale characteristics with respect to the probability law. Stoch. Process. Appl. **124**, 3819–3845 (2014)
49. Nutz, M., van Handel, R.: Constructing sublinear expectations on path space. Stoch. Process. Appl. 123, 3100–3121 (2013)
50. Possamaï, D., Tan, X., Zhou, C.: Stochastic control for a class of nonlinear kernels and applications. Ann. Probab. 46, 551–603 (2018)
51. Possamaï, D., Touzi, N., Zhang, J.: Zero-sum path-dependent stochastic differential games in weak formulation. Ann. Appl. Probab. 30, 1415–1457 (2020)
52. Qian, Y.: Incentives and loss of control in an optimal hierarchy. Rev. Econ. Stud. 61, 527–544 (1994)
53. Samikov, Y.: A continuous-time version of the principal–agent problem. Rev. Econ. Stud. 75, 957–984 (2008)
54. Schättler, H., Sung, J.: The first-order approach to the continuous-time principal–agent problem with exponential utility. J. Econ. Theory 61, 331–371 (1993)
55. Stiglitz, J.E.: Incentives, risk, and information: notes towards a theory of hierarchy. Bell J. Econ. 6, 552–579 (1975)
56. Stroock, D., Varadhan, S.: Multidimensional Diffusion Processes. Springer, Berlin (1997)
57. Sung, J.: Linearity with project selection and controllable diffusion rate in continuous-time principal–agent problems. Rand J. Econ. 26, 720–743 (1995)
58. Sung, J.: Pay for performance under hierarchical contracting. Math. Financ. Econ. 9, 195–213 (2015)
59. Williamson, O.E.: Hierarchical control and optimum firm size. J. Polit. Econ. 75, 123–138 (1967)
60. Yang, Y.L.: Degree of supervision, moral hazard, and hierarchical control. J. Econ. Behav. Organ. 26, 93–105 (1995)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.