

## Simple Contracts under Observable and Hidden Actions

Bo Chen · Yu Chen · David Rietzke

Received: October 17, 2018 / Accepted: March 4, 2019

**Abstract** We consider a general framework for multitask moral hazard problems with observable and hidden actions. Ideally, the principal in our framework can design optimal contracts that depend on both observable (and verifiable) actions and realized outcomes. Given a mild assumption on the existence of a punishment scheme, we identify a general equivalence result, dubbed the “forcing principle,” which states that every optimal contract in our framework is strategically equivalent to a simple forcing contract, which only specifies an outcome-contingent reward scheme and an action profile, and the agent receives the outcome-contingent reward only if he follows the recommended observable actions (and is otherwise punished severely). The forcing principle has useful implications: It confers analytic advantage for the existence and computation of optimal contracts in our setting. It also highlights and makes explicit the importance of the existence of the punishment scheme in characterizing first-best benchmarks in moral hazard problems.

**Keywords** First-Best Benchmark · Forcing Contract · Forcing Principle, · Moral Hazard · Observable Actions

**JEL Classifications** C61 C62 D86

---

Yu Chen passed away unexpectedly in February of 2019, just weeks before the paper was accepted. His friendship and talents will be missed by his coauthors. We thank our Referee and Editor for very constructive comments. We also gratefully acknowledge helpful comments and discussion from Michael Alexeev, Sascha Baghestanian, Robert Becker, Filomena Garcia, Junichiro Ishida, Jianpei Li, Frank Page, Daniela Puzzello, Eric Rasmussen, Satoru Takahashi, Ning Yang, and Jinghua Yao. Yu Chen was supported by the National Natural Science Foundation of China (Grant No.71673133). The usual disclaimer applies.

Bo Chen  
Department of Economics, Southern Methodist University, USA,  
E-mail: bochen@smu.edu

Yu Chen  
Department of Economics, University of Graz, Austria

David Rietzke  
Department of Economics, Lancaster University, UK  
emaild.rietzke@lancaster.ac.uk

## 1 Introduction

In many real-world contracting problems, an agent's decision will involve various dimensions, some of which may be more easily observed than others. Holmström and Milgrom (1991) provide the first formal analysis of principal-agent problems with multidimensional tasks.<sup>1</sup> An important insight of this study (and subsequent literature on multitasking) is that the optimal design of incentive contracts depends crucially on the precision with which actions can be measured. While this literature has produced many useful insights, much of the work is somewhat limited in scope, and the literature has tended to focus on the particular case of linear contracts, an agent with exponential utility, and signals that are jointly normally distributed.<sup>2</sup>

We study an optimal contracting model with moral hazard and multidimensional tasks, but diverge from previous multitasking studies in two dimensions.<sup>3</sup> First, we focus on an environment in which the agent takes both observable and unobservable actions. Such scenarios are not difficult to imagine. For example, although the time a worker spends at work may be observable, the mental or physical effort he devotes to his job may not be easily measured. In a relationship between a venture capitalist and an entrepreneur, while the entrepreneur's investment in capital is easily verifiable, the entrepreneur's daily effort to run the business may be too costly to monitor. Finally, in the context of insurance provision, the insurer may be able to observe the number of doctor's visits per year, but the daily preventive care taken by the insured party is difficult to observe. Although other studies have allowed for both observable and unobservable actions, we explicitly take both into account, and explore—in a general framework—how the agent's observable actions enter optimally into contracts in a somewhat simple manner.

Second, we impose minimal structure: We only require the outcome, action, and reward spaces to be metric spaces, and each of these sets can be multi-dimensional, potentially non-compact, finite or infinite, countable or uncountable. Moreover, contracts in our framework are only assumed to be Borel measurable, although other contracting constraints can also be considered and included in our framework. Finally, only Borel measurability is required for the principal and agents' utility functions, allowing for a wide range of preferences for the contracting parties.

Specifically, we consider an abstract moral hazard setting in which a principal hires an agent for a project with multiple tasks. The agent chooses both observable and hidden actions, resulting in stochastic outcomes in which both observable actions and realized outcomes are verifiable. With so little structure in our baseline model, one might hypothesize that an optimal contract (should one exist) would employ com-

---

<sup>1</sup> A precursor of the model in Holmström and Milgrom (1991) can be found in Holmström and Milgrom (1987) on linear contracts as optimal compensation schemes in dynamic environments.

<sup>2</sup> This is the so-called "linear-exponential-normal (LEN)" model in the literature.

<sup>3</sup> We equate multitasks with multidimensional actions in our analysis, although there could be literal differences in different contexts.

plex reward schemes that link the agent’s payment non-trivially to both observable actions and outcomes.<sup>4</sup>

To the contrary, we establish a general *forcing principle*, which demonstrates the relatively simple manner by which observable actions should optimally be taken into account by the principal. Under a relatively mild assumption which ensures the existence of a severe punishment available to the principal, we first show that there is no loss of generality for the principal to restrict attention to a particular set of contracts (i.e., punishing contracts) by which the agent receives outcome-contingent payment *only if* he has chosen the observable actions recommended by the principal, and is severely punished (i.e., obtaining a payoff no more than the agent’s outside option) otherwise. Under such forcing contracts, both the principal and the agent obtain the same expected utilities as in any other optimal contract in the setting. Using this result as a building block, we then establish the forcing principle: The solution to the original contracting problem can be obtained by solving an equivalent, but simplified forcing contracting problem in which the principal “chooses” the observable action directly, and can restrict attention to outcome-contingent reward schemes in which reward variations are independent of the observable actions. We also highlight the relevance of the punishment condition by demonstrating that the equivalence cannot be preserved if the condition is violated (see Example 3).

Our simple contracts (where the optimal contracts depend on observable actions in a simple fashion) are closely related to the concept of forcing contracts in many applications of contract theory. The use of these simple contracts dates to the early literature on moral hazard, and this structure is frequently mentioned in the contracting literature, especially in the characterization of first-best benchmarks (see Section 4.1).<sup>5</sup> It is worth noting that the equivalence in forcing principle, while perhaps intuitive, is far from trivial to establish, particularly so in our general moral hazard setting with little structure: Since both hidden and observable actions enter the agent’s payoff function in a general way, any feasible forcing contract must satisfy incentive compatibility with respect to *both* hidden and observable actions simultaneously, together with a corresponding individual rationality condition which is also affected by both types of actions. Hence, the equivalence result suggested by our forcing principle cannot be regarded as an immediate extension of forcing contracts in first-best benchmarks where there are only observable actions and individual rationality conditions.<sup>6</sup> The crux and the difficulty in establishing the equivalence result, as we will show, lies in carefully constructing a series of auxiliary (and intermediate) constrained maximization problems and then demonstrating that all these maximization problems are strategically equivalent.

Moreover, as far as we are aware, the usefulness and robustness of such contracts has *not* been explored (at least not in a general framework). Our forcing principle provides a single sufficient assumption on the existence of a suitable punishment scheme

<sup>4</sup> For example, the famous Gantt’s task and bonus plan in industrial management specifies a worker’s reward scheme that varies in both the worker’s output and working hours in a complex way, which are called *hybrid* reward schemes hereafter. See Example 1.

<sup>5</sup> See Mirrlees (1974) and Grossman and Hart (1983), among many others.

<sup>6</sup> It is clear though that a typical forcing contract in a first-best benchmark is only a special case of our general forcing principle.

available to the principal, under which any feasible contract can be equivalently written as a simple forcing contract. Crucially, we do not impose the usual technical assumptions (i.e., continuity, compactness, etc.) to establish the forcing principle. Our result thus demonstrates remarkable robustness of the optimality of simple forcing contracts in contracting problems with partially observable actions, in the sense that such simple contracts are not sensitive to the details of the model. Indeed, the power of our result comes from the general conditions under which it holds, implying its applicability to a wide range of settings.

Our forcing principle has useful practical implications in general settings of applications with both observable and unobservable actions. In particular, it implies that there is no need for the principal to design and implement complex hybrid contracts with rewards that depend intricately on both observable actions and outcomes, as long as the observable action is contractible. In fact, it suffices for the principal to consider a simplified contract where the agent's reward depends mainly on the output of the project.

The forcing principle also has useful theoretical applications. The first application of the forcing principle has to do with the characterization of first-best benchmarks in moral hazard problems with only observable and verifiable actions, which is a special case of our framework (without hidden actions). The standard approach in the literature is to characterize the first-best benchmark using a forcing contract (similar to our simple forcing contracts) in which the principal forces an action on the agent without any incentive compatibility constraint. The forcing principle, when applied to such a setting, implies that the standard approach of considering a forcing contract for the first-best benchmark is valid *if* a suitable and severe punishment scheme is available to the principal. When this assumption is violated, however, we demonstrate via an example that the forcing contract employed in the standard approach may fail to identify the first-best benchmark. Our result hence highlights the importance of the assumption on the existence of a severe punishment scheme in the characterization of first-best benchmarks, which has not been emphasized explicitly in the literature.<sup>7</sup>

Another useful application of the forcing principle is that it greatly simplifies the contracting problem and enables us to establish the existence of optimal *deterministic* contracts in our setting under mild topological conditions. Our approach here is most closely related to Page (1987), who establishes the existence of optimal deterministic contracts in a pure moral hazard setting. Page's approach is fairly general, allowing in particular for an infinite dimensional contract space and for the principal and the agent to hold different beliefs about the random outcomes generated by the agent's actions. Nevertheless, his approach cannot be directly applied to our setting with both observable and hidden actions.<sup>8</sup> We establish that under the forcing principle, one avoids the complications of a hybrid reward scheme, and Page's approach can then be adapted to establish existence in our setting.

---

<sup>7</sup> While this is perhaps a minor issue in that the existence of such a punishment is realistic and is either implicitly or explicitly recognized in most of the moral hazard literature, we think that it is still sensible and useful to raise the issue explicitly.

<sup>8</sup> Technically, it is difficult to establish sequential continuity of the two parties' expected utility functions via Delbaen's Lemma, crucial in Page's approach, when the observable action component also enters the reward scheme. We discuss this in detail in Section 4.2.

Finally, our forcing principle can simplify the analytical computation of optimal contracts in our moral hazard setting. The principal can now search for optimal reward schemes in a strictly smaller set, and the incentive constraint for the agent is simplified to the choice of only hidden actions. Technically, a standard first-order approach (FOA) cannot be directly applied to find optimal contracts in the original general contracting problem; however, given the forcing principle, the FOA approach can be applied in the equivalent but simpler contracting problem. We discuss these simplifications at greater length in Section 4.3 (and Example 4), in which we also compute the explicit optimal contract for a simple moral hazard problem with the help of our forcing principle.

### 1.1 Related Literature

As stated, our paper is related to the literature on multitask principal-agent problems that originated with Holmström and Milgrom (1991). One particular result in this literature is that the power of incentives for some tasks depends on the principal's ability to monitor various aspects of the agent's performance. High-powered incentive contracts can reduce performance, because the agent can shift his effort from poorly measured activities to better measured and more highly compensated ones. While our paper also suggests that observed actions be used solely as thresholds rather than active components of the agent's payments, our results are obtained in an abstract setting and are motivated by consideration of *simple contracts*.

Our paper is also related to a strand of literature that analyzes optimal contracting issues in some specific multitask moral hazard problems with both hidden and observable actions (e.g., Laux, 2001; Zhao, 2008; Chen, 2010, 2012). These studies consider similar settings, in which a cost-minimizing principal hires risk-neutral agent(s) protected by limited liability to work on multiple tasks and the agents' effort choices in some tasks are observable. However, they analyze different issues such as task clustering and job design (Laux, 2001); choosing between input monitoring and output monitoring (Zhao, 2008; Chen, 2010); and all-or-nothing payment structures (Chen, 2012). While these studies are important precursors and have characterized optimal contracts as forcing contracts similar to the ones in this study, we consider a more general setting and a different objective, by which we aim to establish a general forcing principle and put this principle to work. Nevertheless, in the context of this literature, our paper makes an important contribution by providing a foundation for focusing on forcing contracts rather than on more complicated reward schemes in these settings, which, given our forcing principle, is without loss of generality.

Finally, the existence of optimal contracts is an important issue in moral hazard problems, and the forcing principle also enables us to establish the existence of an optimal contract in a general moral hazard setting with both hidden and observable actions. As noted above, Page (1987) provides the first methodological contribution on the existence of deterministic optimal contracts in a (pure) moral hazard model

with a general topological structure.<sup>9</sup> More recently, similar to our existence result, Ke and Xu (2017) establish existence of optimal *deterministic* contracts in a general pure moral hazard model. Their focus, however, is on a setting in which utility functions are concave in reward (money), and the agent's utility is quasi-separable between reward and (hidden) action (as in Grossman and Hart, 1983), whereas we do not necessarily impose such conditions for our existence result.<sup>10</sup> Our existence analysis hence is more closely related to that in Page (1987), but our setting differs from his by including observable actions. Specifically, our analysis takes advantage of the forcing principle and follows Page's approach to establish the existence of an optimal deterministic contract in general moral hazard settings with both unobservable and observable actions.

## 2 The Model

We consider a principal-agent relationship in which a principal (she) hires an agent (he) to make a one-time choice of multi-dimensional actions on a project. The timing of the contracting game is standard: The principal proposes a take-it-or-leave-it contract to the agent. The agent then decides whether to accept the contract. If he accepts the contract, the agent chooses a vector of actions and is then rewarded, according to the contract, upon the outcomes generated by his actions.

The agent chooses two types of actions, observable and unobservable. We assume that the agent's observable actions are contractible and his unobservable actions are not. Let  $A = A_0 \times A_1$  be the set of all possible action profiles available to the agent where  $A_0$  is a metric space containing all possible observable actions and  $A_1$  is a metric space containing all possible unobservable actions available to the agent.<sup>11</sup> A typical action profile of the agent is denoted as  $(a_0, a_1) \in A$  with an observable component  $a_0 \in A_0$  and an unobservable component  $a_1 \in A_1$ .

After the agent chooses his actions  $(a_0, a_1)$ , nature moves by stochastically generating an *outcome*  $\omega$  based on  $(a_0, a_1)$ , where  $\omega \in \Omega$  and  $\Omega$  is a metric space of all possible outcomes. Here an outcome  $\omega$  can be multi-dimensional and may include both monetary components (e.g., sales revenues generated by a salesperson) and non-monetary components (e.g., service evaluations of the salesman by clients). Let  $\mathcal{B}(\Omega)$  denote the Borel  $\sigma$ -field over the outcome space  $\Omega$ . For each action profile  $(a_0, a_1)$ , let  $P(\cdot; a_0, a_1) : \Omega \rightarrow [0, 1]$  be a probability measure defined on  $(\Omega, \mathcal{B}(\Omega))$ . Observe that  $P(\cdot; a_0, a_1)$  is  $(a_0, a_1)$ -dependent, and it is common knowledge between the principal and the agent that the observable outcomes will be generated according to the (same) probability measure  $P(\cdot; a_0, a_1)$ .

<sup>9</sup> Page (1991, 1992b) extend this methodology to general models with both moral hazard and adverse selection. Page (1992a) presents an existence result for Bayesian incentive compatible mechanisms for Stackelberg games with incomplete information.

<sup>10</sup> Recently, Kadan et al. (2017) also establish an existence result for principal-agent problems that encompass pure moral hazard, pure adverse selection, and problems with both. However, the existence result in Kadan et al. (2017) is mainly for optimal *randomized* mechanisms.

<sup>11</sup> Here and in the sequel, the *observability* of a given object implies both verifiability by a third party and measurability of the object with respect to the corresponding distribution function.

We next define the principal's strategies. A strategy of the principal is a *reward scheme* contingent on outcome  $\omega$  and action  $a_0$ , which are both observable and contractible. Let  $D$  be a metric space of all possible end-of-period *rewards* for the agent, with typical element  $d \in D$ . Like outcome  $\omega$ , a reward  $d$  can be multi-dimensional and contain both monetary (e.g., salaries) and non-monetary (e.g., promotions) components. Define a function set

$$\mathcal{M} := \{s : \Omega \rightarrow D \mid s \text{ is a Borel Measurable function}\},$$

with element  $s$ .  $\mathcal{M}$  is hence the set of pure outcome-contingent reward schemes.

Let  $\mathcal{K} \subseteq \mathcal{M}$  represent all available, pure outcome-contingent reward schemes under certain primitive restrictions on contracting practices.<sup>12</sup> Given that observable actions are contractible, it is natural for the principal to also include observable actions in her contract. We define the set of all possible *hybrid* reward schemes to be

$$\mathcal{R}_{\mathcal{K}} := \{f : \Omega \times A_0 \rightarrow D \mid f \text{ is Borel measurable, and } f(\cdot, a_0) \in \mathcal{K} \forall a_0 \in A_0\},$$

where its typical element is  $f$ . By definition, a hybrid reward scheme  $f$  typically depends on both outcomes and observable actions.

The principal benefits from  $\omega$  and pays  $f(\omega, a_0)$  to the agent. Denote the principal's vNM utility function as  $u : \Omega \times D \times A \rightarrow \mathbb{R}$ , and the agent's vNM utility function defined over rewards and actions as  $v : D \times A \rightarrow \mathbb{R}$ . Both  $u$  and  $v$  are Borel measurable functions. The expected utility of each party can be written as:

$$\text{Principal: } U(f, a_0, a_1) = \int_{\Omega} u(\omega, f(\omega, a_0), a_0, a_1) P(d\omega; a_0, a_1),$$

$$\text{Agent: } V(f, a_0, a_1) = \int_{\Omega} v(f(\omega, a_0), a_0, a_1) P(d\omega; a_0, a_1).$$

To allow for general preferences, we have implicitly built the reward and the effort cost in the utility functions  $u$  and  $v$ . Notice that we only impose Borel measurability on the two parties' utility functions. In particular, this general setting of utility allows for arbitrary interdependencies among the agent's various actions, arbitrary risk attitudes, and non-separable and discontinuous utility functions for the principal and the agent.

The principal's objective is to design an optimal (possibly) *hybrid* contract  $(f, a_0, a_1)$  consisting of a reward scheme  $f$  and action recommendations  $(a_0, a_1)$ . The principal's problem **(P1)** is formulated as follows:<sup>13</sup>

$$\begin{aligned} & \max_{(f, a_0, a_1) \in \mathcal{R}_{\mathcal{K}} \times A} U(f, a_0, a_1) \\ \text{(P1)} \quad & \text{s.t. (IR): } V(f, a_0, a_1) \geq r \\ & \text{(IC): } (a_0, a_1) \in \arg \max_{(a'_0, a'_1) \in A} V(f, a'_0, a'_1) \end{aligned}$$

<sup>12</sup> To be specific,  $\mathcal{K}$  denotes the set of all practically available pure outcome-contingent reward schemes, given current market/technology conditions and legal customs and the principal's abilities for computation and accounting, etc., as in Page (1987). For instance,  $\mathcal{K}$  can be the set of bounded and monotone contracts or a set of contracts that are restricted to be linear in outcomes.

<sup>13</sup> In **(P1)** if  $A_0$  is a singleton, then the problem reduces to a pure moral hazard model, while if  $A_1$  is a singleton, then **(P1)** corresponds to the first-best contracting problem with only observable (and verifiable) actions. We use the notation "max" rather than "sup" in the principal's and agent's problems in **(P1)**. The existence of solutions in **(P1)** will be established with further assumptions in Section 4.

As usual, the (IR) constraint guarantees the agent's expected utility from the contract to be at least his reservation utility  $r$ , and the (IC) constraint requires that the recommended action profile be optimal for the agent. Importantly, notice that in the (IC) constraint, though  $a_0$  is verifiable, there is no *forcing* in **(P1)**, since the principal does not directly "choose" the recommended  $a_0$  for the agent.<sup>14</sup>

**Definition 1** A contract  $(f, a_0, a_1)$  is **feasible** if  $f \in \mathcal{R}_{\mathcal{X}}$  and  $f$  satisfies (IR) and (IC).

To illustrate the general principal-agent problem, consider a concrete example:

*Example 1* Consider an employer-employee relationship. The employee chooses observable working hours  $a_0 \in [0, 24]$  and hidden effort  $a_1 \in [0, 1]$ . An action profile  $(a_0, a_1)$  stochastically generates a monetary outcome  $z \in Z$  via a distribution function  $G(z; a_0, a_1)$  with density  $g(z; a_0, a_1)$ . The employer offers a monetary reward  $d \in D$  according to a scheme  $f : Z \times A_0 \rightarrow D$ . The employee's payoff is  $v(d, a_0, a_1) = \varphi(d) - C(a_0, a_1)$ , with  $\varphi(d)$  being the utility from the reward  $d$  and  $C(a_0, a_1)$  the cost from  $(a_0, a_1)$ . The principal's utility is  $u(z - d)$ . The resulting principal-agent problem is

$$\begin{aligned} & \max_{(f, a_0, a_1) \in \mathcal{R}_{\mathcal{X}} \times A} \int_Z u(z - f(z, a_0)) g(z; a_0, a_1) dz \\ \text{s.t. } & \int_Z \varphi(f(z, a_0)) g(z; a_0, a_1) dz - C(a_0, a_1) \geq r \\ & (a_0, a_1) \in \arg \max_{(a'_0, a'_1) \in A} \int_Z \varphi(f(z, a'_0)) g(z; a'_0, a'_1) dz - C(a'_0, a'_1) \end{aligned}$$

Overall, we have imposed little structure in our model in order to obtain a general moral hazard framework: The spaces  $A_0, A_1, \Omega$  and  $D$  can all have multiple dimensions, with some being discrete and some continuous, and the two parties' utility functions are allowed to be non-separable in the corresponding arguments. As we shall see shortly, such a framework is more or less sufficient for our forcing principle to hold so that a simple forcing contract solves the principal's optimization problem **(P1)**.

### 3 The Forcing Principle

The celebrated informativeness principle of Holmström (1979) posits that additional information that is incrementally informative about an agent's action should be included in the optimal contract. It stands to reason that optimal contracts in our setting should involve complex reward schemes that depend non-trivially on both the outcomes and the observable actions as addressed in **(P1)**.<sup>15</sup> A case in point here is Gantt's famous task and bonus plan in industrial management (Periasamy, 2010, p. 390):

<sup>14</sup> For future comparison, such *forcing* will take place in **(P2)** in Section 3.

<sup>15</sup> It is known, however, that the informativeness principle may not hold in a multitask principal-agent models (see, e.g., Holmström and Milgrom, 1991)



*Example 2* Gantt's task and bonus plan involves the careful design of an incentive system using time rate, differential piece rate and bonus. To illustrate, the principal observes both the time worked  $a_0$  and output  $\omega$  (but not how hard the agent works  $a_1$ ). Gantt's plan specifies a wage scheme  $f(\omega, a_0)$  with parameters  $\alpha$ ,  $\beta$ , and  $\bar{\omega}$ :

$$f(\omega, a_0) = \begin{cases} \alpha a_0 & \text{if } \omega < \bar{\omega} \\ (\alpha + 20\%)a_0 & \text{if } \omega = \bar{\omega} \\ \beta \omega & \text{if } \omega > \bar{\omega} \end{cases}$$

i.e., a time rate  $\alpha$  is applied when output  $\omega$  is below a standard  $\bar{\omega}$ , a bonus time rate ( $\alpha + 20\%$ ) is employed when  $\omega$  hits the standard, and a higher piece rate  $\beta$  is employed otherwise. This task and bonus plan was identified as quite successful by Gantt (1919).<sup>16</sup>

Given our general framework and the complexity of reward schemes, however, the characterization of optimal contracts in **(P1)** is rather unmanageable. An alternative and simpler contracting approach for the principal is to downplay (but not neglect) the role of observable actions and focus instead on the main *performance indicator* of the agent, i.e., on the outcome  $\omega$ .<sup>17</sup> Specifically, the principal now imposes an explicit choice of  $a_0$  and searches for a pure *outcome-contingent* reward scheme, rather than a complex hybrid reward scheme. Such a contracting procedure is formally described as

$$\begin{aligned} \text{(P2)} \quad & \max_{(s, a_0, a_1) \in \mathcal{K} \times A} \int_{\Omega} u(\omega, s(\omega), a_0, a_1) P(d\omega; a_0, a_1) \\ \text{s.t.} \quad & \text{(IR): } \int_{\Omega} v(s(\omega), a_0, a_1) P(d\omega; a_0, a_1) \geq r \\ & \text{(IC): } a_1 \in \arg \max_{\hat{a}_1 \in A_1} \int_{\Omega} v(s(\omega), a_0, \hat{a}_1) P(d\omega; a_0, \hat{a}_1) \end{aligned}$$

An optimal contract solving **(P2)** is called a *simple forcing contract*. Such a formulation of forcing contracts is analogous to the standard treatment of observable and verifiable actions in the literature: Since observable actions are verifiable, the choice of  $a_0$  can be forced upon the agent, and it is as if the principal now directly chooses  $a_0$  for the agent. As a result, the choice of  $a_0$  does not enter the **(IC)** constraint.

The contracting problem **(P2)** is simpler than **(P1)**: The **(IC)** constraint in **(P2)** only involves  $a_1$  and is hence simpler for both the principal and the agent. In addition, the principal chooses a pure outcome-contingent reward scheme in  $\mathcal{K}$  (rather than a more complex hybrid reward scheme in  $\mathcal{R}_{\mathcal{K}}$ ). The simpler formulation in **(P2)** then motivates the natural question of whether (and if yes, when) optimal solutions for **(P2)** are also optimal contracts for **(P1)**, and vice versa. We demonstrate in the rest of this section that this is indeed the case for the two contracting problems **(P1)** and **(P2)** as long as a severe punishment scheme is available for the principal (Assumption 1). As a result, simple forcing contracts can be optimal for the original contracting problem **(P1)**.

<sup>16</sup> "For Bethlehem Steel, the average monthly output of the shop from March 1, 1900, to March 1, 1901, was 1,173,000 pounds, and from March 1, 1901, to August 1, 1901 (having implemented the task and bonus system) was 2,069,000 pounds. The shop had 700 men in it and we were paying on the bonus plan only about 80 workmen out of the entire 700." (Chapter VII, Gantt 1919).

<sup>17</sup> This intuitive approach is perhaps consistent with many business practices and academic research emphasizing the importance of **key performance indicators** for successes in various organizations.

The forcing contracts solving **(P2)** (and **(P1)** under Assumption 1) are *simple* in the sense that compared to optimal hybrid contracts, such forcing contracts are easier to characterize in **(P2)** and the reward schemes for these contracts only need to be outcome-contingent.<sup>18</sup> In addition, again compared to hybrid contracts, these forcing contracts simplify the decision making for both the principal and the agent in the setting.

Importantly, notice that optimal forcing contracts from **(P2)** are *not* necessarily always (weakly) worse than optimal contracts from **(P1)**: Although the principal can choose a more general hybrid reward scheme in **(P1)**, the two contracting problems are not *nested*. In particular, while the principal restricts attention to outcome-contingent reward schemes in **(P2)**, the (IC) constraint in **(P2)** is less stringent in that it does not require the agent to optimally choose  $a_0$ . As a result, the set of *feasible contracts* (cf. Definition 1) in **(P2)** is not necessarily strictly smaller than that in **(P1)**. Indeed, we show later in Example 3 (Section 4.1) that a forcing contract for **(P2)** can actually lead to a strictly better outcome for the principal compared to an optimal hybrid contract for **(P1)**.

To formally investigate the relationship between the contracting problems **(P1)** and **(P2)**, we start with the following assumption:

**Assumption 1 (Punishment)** There exists a punishment scheme  $t \in \mathcal{H}$  such that for all  $(a_0, a_1) \in A$ ,  $\int_{\Omega} v(t(\omega), a_0, a_1) P(d\omega; a_0, a_1) \leq r$ .

Assumption 1 says that the principal has the option of using an (outcome-contingent) punishment scheme that offers the agent no more than his reservation utility, no matter what actions he takes. Thus, facing the reward scheme  $t$ , the agent will have no incentive to participate given his outside option, so the punishment scheme  $t$  is *severe*. The existence of such a punishment scheme is a relatively mild assumption in contracting scenarios. For example, as long as the principal has the option of offering a sufficiently low reward scheme, i.e., if  $f(\omega, a_0) = c$  for all  $(\omega, a_0)$  is an element in  $\mathcal{R}_{\mathcal{H}}$  such that the agent's utility  $v(c, a_0, a_1) \leq r$  for all  $(a_0, a_1)$ , then Assumption 1 will be satisfied.<sup>19</sup> In addition, given that our reward (metric) space  $D$  can be multi-dimensional with various components (which are not necessarily in  $\mathbb{R}$ ), the punishment  $t(\omega)$  in Assumption 1 can be monetary and/or non-pecuniary. Moreover, since  $t(\omega)$  is only defined from the agent's point of view, the punishment may or may not benefit the principal.<sup>20</sup>

We first present an intermediate step toward the equivalence between **(P1)** and **(P2)**. Specifically, we show that every reward scheme in a feasible hybrid contract can be equivalently transformed into a simple reward scheme with an outcome-contingent component and a punishment component, which leads to the same expected utility to both the principal and the agent (Proposition 1).<sup>21</sup>

<sup>18</sup> The outcome-contingent portion of the reward scheme however is similar to an optimal contract solving a canonical principal-agent model in Holmström (1979), which may not take a simple form.

<sup>19</sup> If on the other hand the agent is protected by limited liability or minimum-wage legislation from receiving harsh putative wages, then Assumption 1 will be violated. See Section 4.1.

<sup>20</sup> We thank our Referee for bringing the additional implications of  $t(\omega)$  to our attention.

<sup>21</sup> All proofs in the paper are relegated to an Appendix.

**Proposition 1** *Suppose Assumption 1 holds. For any feasible contract  $k^* = (f^*, a_0^*, a_1^*)$ , consider another contract  $\tilde{k} = (\tilde{f}, a_0^*, a_1^*)$  such that  $\tilde{f}(\omega, a_0) = \begin{cases} s(\omega), & \text{if } a_0 = a_0^* \\ t(\omega), & \text{otherwise} \end{cases}$ , where  $s(\omega) \equiv f^*(\omega, a_0^*)$ . Then the contract  $\tilde{k}$  is also feasible, and the principal and the agent receive the same expected utilities under  $k^*$  and  $\tilde{k}$ .*

One perhaps useful angle from which to view Proposition 1 is that it has a similar flavor as the *revelation principle* in the context of mechanism design: Proposition 1 implies that for problem **(P1)**, it is without loss of generality to consider a proper subset of hybrid contracts with typical element  $\tilde{k}$  rather than a general hybrid contract in  $\mathcal{R}_H$ . The scheme  $\tilde{f}$  in the contract  $\tilde{k}$  employs a simple punishment to induce the agent to choose the principal's recommended observable actions. The observable actions merely serve a role similar to a threshold in the reward scheme, and the main variations in the reward scheme are contingent only on the outcomes to incentivize the agent to choose the recommended unobservable actions. The contract  $\tilde{k}$  in Proposition 1 can be regarded as a *forcing contract with punishment*. It is intuitive and perhaps consistent with some contracts in practice. For example, to enforce an eight-hour workday requirement (the recommended observable action  $a_0^*$ ) for a worker, a manager can lower his wage or fire him, if the worker fails to meet the requirement.

Given Proposition 1, we can restrict the principal's choice of reward schemes to the class of pure outcome-contingent reward schemes, augmented by the punishment scheme  $t(\omega)$  in Assumption 1. Using this as a crucial building block, our main result, Theorem 1 below, formally presents the *forcing principle*, which states that the original contracting problem **(P1)** can be equivalently transformed into the simple forcing contracting problem **(P2)** under Assumption 1. Such an equivalence comes into play implicitly via the forcing contract with punishment established in Proposition 1.

**Theorem 1 (Forcing Principle)** *Under Assumption 1, **(P1)** is strategically equivalent to **(P2)**, i.e.,*

1. *Given an optimal solution  $(f^*, a_0^*, a_1^*)$  to **(P1)**, there is an optimal solution  $(s^\dagger, a_0^\dagger, a_1^\dagger)$  to **(P2)**, where  $(a_0^\dagger, a_1^\dagger) = (a_0^*, a_1^*)$ ,  $s^\dagger(\omega) = f^*(\omega, a_0^*)$  for all  $\omega \in \Omega$ , and the agent receives reward scheme  $t(\omega)$  in Assumption 1 whenever  $a_0 \neq a_0^\dagger$ .*
2. *Given an optimal solution  $(s^\dagger, a_0^\dagger, a_1^\dagger)$  to **(P2)**, there is an optimal solution  $(f^*, a_0^*, a_1^*)$  to **(P1)**, where  $(a_0^*, a_1^*) = (a_0^\dagger, a_1^\dagger)$  and for all  $\omega \in \Omega$ ,*

$$f^*(\omega, a_0) = \begin{cases} s^\dagger(\omega), & \text{if } a_0 = a_0^\dagger \\ t(\omega), & \text{otherwise} \end{cases} \quad (1)$$

3. *Both the principal and the agent attain the same expected utility in **(P1)** and **(P2)**.*

The forcing principle summarized in Theorem 1 implies that for a general moral hazard problem with observable and hidden actions, the optimal contract can be found by solving a simple problem in which the principal chooses  $a_0$  directly, so long as the agent can be severely punished when deviating from the recommended observable action.<sup>22</sup> Moreover, both the principal and the agent will obtain the same expected

<sup>22</sup> In such simple contracts, the principle can simply state some necessary or minimal requirements on  $a_0$  for the agent to get any compensation (for example, minimal working hours, project completion deadlines, exclusive sales territories, etc).

utilities under **(P1)** and **(P2)**, and hence neither one has any incentive to deviate from the simple contracting procedure **(P2)**.

Notice that the equivalence result in Theorem 1, while perhaps intuitive, is **not** straightforward to establish, particularly so in our general moral hazard setting with little structure. The maximization problems **(P1)** and **(P2)** involves qualitatively different domains of contracts, different constraints, and hence different sets of feasible contracts. The crux and the difficulty in establishing the equivalence result, as shown in the proof in Appendix, lies in carefully constructing a series of auxiliary (and intermediate) constrained maximization problems and then demonstrating that all these maximization problems are strategically equivalent.

The optimality of the simple forcing contract in **(P2)** is of practical significance. In implementing the simple forcing contract, the principal only needs to detect whether the agent has deviated from the recommended observable action. In particular, there is no need for the principal to specify different punishments for different deviations in the observable actions—in practice, measuring such deviations precisely can be quite costly. Thus, the informational demands of these simple contracts are significantly less than what is required for more general contracts.

Finally, it is worth noting that Assumption 1 is crucial for Theorem 1: The equivalence between the two contracting problems **(P1)** and **(P2)** breaks down when Assumption 1 is violated—we present an explicit example in Section 4.1 (Example 3) to show that without Assumption 1, the principal may receive strictly higher payoffs from the simple problem **(P2)**. As stated previously, since the sets of feasible contracts in **(P1)** and **(P2)** are not *nested*, without imposing further structure, the general relationship between the two contracting problems **(P1)** and **(P2)** is difficult to characterize absent of Assumption 1.

## 4 Applications of the Forcing Principle

The forcing principle is powerful. In particular, it does not require stringent conditions or mathematical structures on the primitives other than Assumption 1 on the existence of a suitable punishment, which renders the principle applicable in many applications. We present three specific applications in this section to demonstrate the theoretical advantages conferred by the forcing principle.

### 4.1 First-Best Benchmark in Pure Moral Hazard

Our first application of the forcing principle is that it can shed light on the characterization of the first-best benchmark in pure moral hazard problems, which arises frequently in the literature. Our general framework in Section 2 reduces to a perfect information setting *without* moral hazard when the set of unobservable actions  $A_1$  is empty. Such a setting (with only observable and verifiable actions) is typically employed to characterize the first-best benchmark for the underlying principal-agent problem.

A conventional treatment in the literature when characterizing first-best benchmarks is to take the actions out of the hands of the agent and let the principal dictate

what action the agent should choose.<sup>23</sup> The optimal contract in a first-best benchmark typically takes the form of a forcing contract and is derived from the following optimization problem, again denoted as **(P2)**, in which we assume  $A_1$  is empty (or a singleton) for **(P2)** in Section 3:

$$\begin{aligned} \text{(P2)} \quad & \max_{(s, a_0) \in \mathcal{X} \times A_0} \int_{\Omega} u(\omega, s(\omega), a_0) P(d\omega; a_0) \\ & \text{s.t. (IR): } \int_{\Omega} v(s(\omega), a_0) P(d\omega; a_0) \geq r \end{aligned}$$

Applying Theorem 1 to this setting directly leads to the following corollary:

**Corollary 1** *Suppose Assumption 1 holds and that  $A_1$  is empty (or a singleton). Then problem **(P1)** can be equivalently reduced to **(P2)**. The first-best benchmark takes the form of a simple forcing contract where the principal directly chooses an action for the agent.*

Although it is directly implied by Theorem 1, Corollary 1 validates the standard approach to characterizing first-best benchmarks in the moral hazard literature. Importantly, though, the equivalence result in Corollary 1 holds only under the existence of a severe punishment scheme  $t(\omega)$  (Assumption 1), and such an assumption to our knowledge has not been explicitly pointed out in the literature. We next present a simple example to show that Corollary 1 (and also Theorem 1) depends critically on the availability of a suitable punishment scheme.

*Example 3* Consider a principal-agent problem with  $\Omega = \{0, 1\}$ ,  $A_0 = \{1, 2\}$ ,  $A_1 = \{a_1\}$ , and  $D \subset \mathbb{R}$ , i.e., there are two outcomes, “success” ( $\omega = 1$ ) and “failure” ( $\omega = 0$ ), and two observable actions, with an outside option of zero. Moral hazard is degenerate here, given that  $A_1$  is a singleton. Let  $p(a_0)$  be the probability of success and  $p(2) > p(1) > 0$ . Let  $f_{\omega}(a_0)$  be the transfer to the agent after  $\omega \in \{0, 1\}$  and action  $a_0$  from the agent. Given a utility function  $v(d, a_0) = d - a_0$ , the agent’s expected utility is

$$V(f_{\omega}, a_0) = p(a_0)f_1(a_0) + (1 - p(a_0))f_0(a_0) - a_0.$$

Given a utility function  $u(\omega, d) = \alpha\omega - d$  with  $\alpha > 0$ , the principal’s expected utility is

$$U(f_{\omega}, a_0) = p(a_0)(\alpha - f_1(a_0)) - (1 - p(a_0))f_0(a_0).$$

The principal’s revenue is  $\alpha$  from a successful project (and 0 from a failed project).

Let the agent’s outside option be 0. Suppose the principal faces a minimum-wage constraint requiring  $f_{\omega}(a_0) \geq 2$  for each  $\omega \in \{0, 1\}$  and each  $a_0 \in \{1, 2\}$ , i.e., the agent’s minimal wage is 2 as long as he participates. Hence, Assumption 1 is violated here.

First consider problem **(P2)**, in which the principal chooses an outcome-contingent reward scheme  $\{s_{\omega}\} = \{s_0, s_1\}$ :

$$\begin{aligned} & \max_{(s_1, s_2, a_0)} \{p(a_0)(\alpha - s_1) - (1 - p(a_0))s_0\} \\ & \text{s.t. (IR): } p(a_0)s_1 + (1 - p(a_0))s_0 - a_0 \geq 0 \\ & \quad \text{(MW): } s_{\omega} \geq 2 \text{ for } \omega \in \{0, 1\} \end{aligned}$$

<sup>23</sup> See, for instance, the characterization of first-best benchmarks in standard textbooks (Laffont and Martimort, 2009; Bolton and Dewatripont, 2005)

Since the agent is “forced” to choose a specific  $a_0$ , the *(IC)* constraint can be ignored. Given the minimum-wage constraint *(MW)*, one can verify that the optimal contract for **(P2)** is  $(a_0^* = 2, s_0^* = s_1^* = 2)$  whenever  $\alpha > 0$  and  $p(2) > p(1)$ , and the principal’s expected utility is  $U(s_\omega^*, a_0^*) = \alpha p(2) - 2$ .<sup>24</sup>

Next consider the principal’s problem **(P1)** formulated as follows:

$$\begin{aligned} & \max_{(f_1, f_2, a_0)} \{p(a_0)(\alpha - f_1(a_0)) - (1 - p(a_0))f_0(a_0)\} \\ & \text{(IR): } p(a_0)f_1(a_0) + (1 - p(a_0))f_0(a_0) - a_0 \geq 0 \\ & \text{s.t. (IC): } a_0 \in \arg \max_{a'_0 \in \{1, 2\}} p(a'_0)f_1(a'_0) + (1 - p(a'_0))f_0(a'_0) - a'_0 \\ & \text{(MW): } f_\omega(a_0) \geq 2 \text{ for } \omega \in \{0, 1\}, a_0 \in \{1, 2\} \end{aligned}$$

Notice that the *(MW)* constraint implies that the agent can ensure himself an expected utility of at least 1 by choosing  $a_0 = 1$ , i.e., *(IR)* always holds. We discuss two cases:

*Case 1* :  $\alpha \geq \frac{1}{p(2)-p(1)}$ .

Given the *(MW)* constraint, the optimal payment scheme can be derived as

$$f_\omega^*(a_0) = \begin{cases} 2 + \frac{1}{p(2)}, & \text{if } \omega = 1 \text{ and } a_0 = 2, \\ 2, & \text{otherwise.} \end{cases}$$

And under  $f_\omega^*(a_0)$ , it is optimal for the principal to induce  $a_0^* = 2$  given that  $\alpha \geq \frac{1}{p(2)-p(1)}$  and the principal’s expected utility from the optimal contract  $(a_0^* = 2, f_\omega^*(a_0))$  is  $U(f_\omega^*, a_0^*) = \alpha p(2) - 3$ .

*Case 2* :  $0 < \alpha < \frac{1}{p(2)-p(1)}$ .

The optimal payment scheme is  $f_\omega^{**}(a_0) = 2$  for all  $\omega \in \{0, 1\}$  and  $a_0 \in \{1, 2\}$ .

Under  $f_\omega^{**}(a_0)$ , it is optimal for the principal to induce  $a_0^* = 1$  given that  $\alpha \in \left(0, \frac{1}{p(2)-p(1)}\right)$  and the principal’s expected utility from the optimal contract  $(a_0^{**} = 1, f_\omega^{**}(a_0))$  is  $U(f_\omega^{**}, a_0^{**}) = \alpha p(1) - 2$ .

In both cases, problem **(P1)** is **not** equivalent to problem **(P2)**. Moreover, one can verify that from a social welfare point of view it is efficient to induce  $a_0 = 1$  if  $\alpha \in \left(0, \frac{1}{p(2)-p(1)}\right)$ , while inducing  $a_0 = 2$  is efficient when  $\alpha \geq \frac{1}{p(2)-p(1)}$ .<sup>25</sup> Hence, the optimal contract for **(P1)** always leads to the “first-best” outcome (maximized total surplus), which is **not** true for the optimal contract solving **(P2)**. On the other hand, the optimal contract for **(P2)** is strictly better for the principal, i.e.,  $U(s_\omega^*, a_0^*) > \max\{U(f_\omega^*, a_0^*), U(f_\omega^{**}, a_0^{**})\}$ . Hence, regardless of whether the first-best benchmark is defined from the principal’s point of view or from the angle of social surplus, the two problems **(P1)** and **(P2)** lead to different outcomes in this example.

<sup>24</sup> Throughout Example 3, we assume that the principal is always willing to hire the agent given the minimum wage constraint—this can be guaranteed by adjusting the level of the minimum wage without qualitatively affecting the structure of the optimal contracts.

<sup>25</sup> The socially optimal action in the setting is identified by comparing the net surpluses generated by  $a_0 = 2$  and  $a_0 = 1$ , which are  $(\alpha p(2) - 2)$  and  $(\alpha p(1) - 1)$ , respectively.

Example 3 shows that Assumption 1 is an indispensable condition for the problems **(P1)** and **(P2)** to be strategically equivalent. The problem here is that the minimum-wage constraint protects the agent from being severely punished and an equivalent forcing contract to implement the solution to **(P1)**.

Another implication of Example 3 is that it also highlights the important role of Assumption 1 in characterizing first-best benchmarks for moral hazard problems. In particular, the standard approach of using a forcing contract to represent the first-best contract when the agent's actions are observable and verifiable may not be valid without Assumption 1.

While Assumption 1 is a mild assumption, a principal in a real-life contracting problem may be restricted from using arbitrarily large sticks in punishing poor performance from an agent—for example, limited liability and minimum-wage legislation can both prevent the agent from being punished too harshly in practice.<sup>26</sup> Example 3 shows that for such settings, it can be problematic to use a forcing contract solving **(P2)** to represent the corresponding first-best benchmark in that such a forcing contract may not characterize the *actual* first-best benchmark: In Example 3 it is optimal from a social welfare point of view to implement a less costly action  $a_0 = 1$  when the project value  $\alpha$  is small, but a forcing contract from **(P2)** ignores the relevant (IC) constraint and implements  $a_0 = 2$  for the principal. For general moral hazard problems where Assumption 1 is violated, using a forcing contract from **(P2)** may set up a *false* first-best benchmark, leading to erroneous efficiency evaluations of an underlying second-best contract.

Granted that the first-best benchmark in the literature is sometimes only a hypothetical situation and in others the existence of such a punishment is (implicitly or explicitly) assumed and satisfied, the problematic issue in Example 3 does not pose much of a problem to the moral hazard literature. Nevertheless, we believe that it is worthwhile to point out the importance of Assumption 1 explicitly.

#### 4.2 Existence of Optimal Contracts

Our second application of the forcing principle concerns the existence of an optimal solution to the original contracting problem **(P1)**. To be specific, the forcing principle enables us to restrict attention to a simple class of forcing contracts and therefore to establish the existence of optimal contracts under mild topological conditions in our setting.

Our approach to existence is similar to Page's (1987) general topological approach, which establishes the existence of optimal contracts in a pure moral hazard setting under certain topological assumptions, provided that the set of admissible reward schemes—called contracts in Page's setting—satisfies a sequential compactness property. The main idea of Page's approach is to identify conditions that guarantee certain sequential compactness and non-emptiness of the set of feasible contracts—recall Definition 1—and (sequential) continuity of the principal's expected utility,

<sup>26</sup> A number of studies in the previous literature have studied moral hazard problems where limited liability or a lower/minimum bound on the agent's payment is imposed (so that Assumption 1 is possibly violated). See, for example, Sappington (1983); Innes (1990); Kim (1997); Dewatripont et al. (2003), etc.

so that a solution to the principal's problem always exists, akin to the Weierstrass extreme value theorem.<sup>27</sup>

We first establish sequential continuity of the two parties' expected utility functions (by Delbaen's Lemma; Delbaen, 1974; Page, 1987) and sequential closedness of the set of feasible contracts. These results then imply that the expected utility of the principal admits a finite supremum over the set of feasible contracts. Finally, we show that there is a sequence of feasible contracts converging to this finite supremum with its limit being an optimal solution to **(P2)**.

Notice that Page's (1987) approach *cannot* be applied directly under a hybrid reward scheme. The difficulty lies in establishing the sequential continuity of the two parties' expected utility functions using Delbaen's Lemma, given that observable actions also enter a hybrid reward scheme. To be specific, consider the sequential continuity of the agent's expected utility. An application of Delbaen's Lemma requires the condition that for a sequence of reward schemes and actions  $\{(f_n, a_n)\}_n$  in  $\mathcal{R} \times \mathcal{X} \times A$  with  $f_n$  converging to  $f^0 \in \mathcal{X}$  pointwise on  $\Omega$  and  $(a_{0,n}, a_{1,n})$  converging to  $(a_0^0, a_1^0) \in A$  under the metric on  $A$ ,  $v(f_n(a_{0,n}, \omega), a_{0,n}, a_{1,n})$  converges to  $v(f^0(a_0^0, \omega), a_0^0, a_1^0)$  for each  $\omega \in \Omega$ . However, given that  $a_0$  is also an argument of  $f$ , the continuity of the agent's utility function  $v$  is insufficient to guarantee such convergence. Delbaen's Lemma hence cannot be directly applied to establish the sequential continuity of the expected utility functions. Consequently, it is also difficult to establish sequential closedness of the set of feasible contracts.

Our forcing principle helps avoid such complications created by hybrid reward schemes. In particular, the structure of reward schemes in **(P2)** resembles that in Page's 1987 moral hazard model, which renders Delbaen's Lemma applicable, and an approach similar to Page's can then be employed to establish existence for **(P2)**. By Theorem 1, the existence of a solution to **(P2)** then implies the existence of a solution to **(P1)**.

As in Page (1987), we start with a few assumptions:

**Assumption 2 (Action)** The action spaces  $A_0$  and  $A_1$  are compact metric spaces.

**Assumption 3 (Outcome)** The outcome space  $\Omega$  is a compact subset of a Euclidean space.

**Assumption 4 (Sequential Continuity of  $P$ )** For each closed  $E \subseteq \Omega$  and each sequence  $\{(a_{0,n}, a_{1,n})\} \rightarrow (a_0, a_1)$  in  $A$ ,  $P(E; a_{0,n}, a_{1,n}) \rightarrow P(E; a_0, a_1)$ .

**Assumption 5 (Reward)** The reward space  $D$  is a closed interval in  $\mathbb{R}$ .

**Assumption 6 (Continuity of Utility)**  $u$  is continuous on  $\Omega \times D \times A$  and  $v$  is continuous on  $D \times A$ .

**Assumption 7 (Sequential Compactness of  $\mathcal{H}$ )** The outcome-contingent reward scheme constraint set  $\mathcal{H}$  is a sequentially compact subset of  $\mathcal{M}$  under the topology of pointwise convergence.

<sup>27</sup> Notice that we will consider sequential topological properties for the reward-scheme sets. Since such reward-scheme sets are sets of functions, they may not always be metrizable from the outset. Consequently, sequential topological properties will be more meaningful for the proof of existence (for instance, sequential compactness is not equivalent to compactness for sets of functions).



We now provide a more compact representation of the principal's optimization problem in **(P2)** in order to establish our existence result. First consider the agent's incentive issue in **(P2)**. Define a set-valued mapping  $A^* : \mathcal{K} \times A_0 \rightarrow A_1$  as

$$A^*(s, a_0) = \arg \max_{a_1 \in A_1} V(s, a_0, a_1),$$

i.e.,  $A^*(s, a_0)$  denotes the set of optimal unobservable actions  $a_1$  for the agent given  $(a_0, s)$  specified by the principal. In addition, denote the agent's optimal level of expected utility under the contract  $(s, a_0)$  as

$$V^*(s, a_0) = \max_{a_1 \in A_1} V(s, a_0, a_1).$$

Next consider the agent's participation in **(P2)**. Define the *participation-guarantee set*  $\mathcal{L}(r)$  as the set of incentive-compatible contracts that guarantee the agent's participation:

$$\mathcal{L}(r) := \{(s, a_0) \in \mathcal{K} \times A_0 : V^*(s, a_0) \geq r\}.$$

We assume that the reservation utility  $r$  is sufficiently low so that  $\mathcal{L}(r)$  is non-empty:

**Assumption 8 (Reservation Utility)** The reservation utility  $r$  is such that  $\mathcal{L}(r) \neq \emptyset$ .

Finally, the constraint set of **(P2)** is then the graph of  $A^*$  restricted to  $\mathcal{L}(r)$ :

$$Gr(A^*) := \{(s, a_0, a_1) \in \mathcal{K} \times A : (s, a_0) \in \mathcal{L}(r) \text{ and } a_1 \in A^*(s, a_0)\}.$$

Given the above notation, we can compactly rewrite **(P2)** as follows:

$$\max_{(s, a_0, a_1) \in Gr(A^*)} U(s, a_0, a_1)$$

where  $Gr(A^*) = \{(s, a_0, a_1) \in \mathcal{K} \times A : (s, a_0) \in \mathcal{L}(r) \text{ and } a_1 \in A^*(s, a_0)\}$ .

The existence of optimal solutions to the constrained optimization problem **(P2)** and also **(P1)** can then be established as follows:

**Proposition 2 (Existence of Optimal Contract)** *Under Assumptions 1-8, there exists an optimal contract  $(s^*, a_0^*, a_1^*)$  that solves the principal-agent problem **(P2)**. Moreover, there also exists an optimal solution to the original principal-agent problem **(P1)**.*

### 4.3 Computation of Optimal Contracts

The forcing principle can also confer advantages when analytically deriving optimal contracts for our moral hazard contracting problems, since the optimal contracts now take a simpler functional form. This is particularly relevant when applying the commonly employed first-order approach (FOA): When using the FOA for moral hazard with observable actions, one must assume, somewhat ad hoc, differentiability of  $f$  (and further calculations may still be difficult given that  $f$  is a function of both  $\omega$  and  $a_0$ ). Instead, by solving the simplified problem **(P2)** for the reward scheme  $s(\omega)$ , one can avoid such problems when applying the FOA. Example 4 provides an illustration:

*Example 4* A principal hires an agent for a project with binary outcomes, success or failure. The agent chooses  $(a_0, a_1) \in A = [0, 1]^2$  where  $a_0$  is the observable time input and  $a_1$  the unobservable effort intensity. The principal obtains  $\pi$  if the project is successful (with probability  $P(a_0, a_1)$ ), and 0 if the project fails. The payment scheme consists of a transfer  $r_s : A_0 \rightarrow \mathbb{R}$  after a success and a transfer  $r_f : A_0 \rightarrow \mathbb{R}$  after a failure. The agent's utility is  $v(d) - C(a_0, a_1)$  from transfer  $d$  where  $C(a_0, a_1)$  is the effort cost. Let  $v(0) = 0$  and  $r = 0$ , and hence Assumption 1 holds. The principal's problem is:

$$\begin{aligned} & \max_{(r_s, r_f; a_0, a_1) \in \mathcal{M} \times A} (\pi - r_s(a_0))P(a_0, a_1) - r_f(a_0)(1 - P(a_0, a_1)) \\ \text{s.t. } & (IR): v(r_s(a_0))P(a_0, a_1) + v(r_f(a_0))(1 - P(a_0, a_1)) - C(a_0, a_1) \geq 0 \\ & (IC): (a_0, a_1) \in \arg \max_{(a'_0, a'_1) \in A} v(r_s(a'_0))P(a'_0, a'_1) + v(r_f(a'_0))(1 - P(a'_0, a'_1)) - C(a'_0, a'_1) \end{aligned}$$

The FOA cannot be applied without assuming the differentiability of  $r_s(\cdot)$  and  $r_f(\cdot)$ . Such differentiability, however, is hard to justify (see Example 2) and may not necessarily be true for optimal transfers  $r_s^*(a_0)$  and  $r_f^*(a_0)$ . Nevertheless, our forcing principle implies that an optimal contract can be solved from a simplified problem:

$$\begin{aligned} & \max_{(r_s, r_f; a_0, a_1) \in \mathbb{R} \times \mathbb{R} \times A} (\pi - r_s)P(a_0, a_1) - r_f(1 - P(a_0, a_1)) \\ \text{s.t. } & (IR): v(r_s)P(a_0, a_1) + v(r_f)(1 - P(a_0, a_1)) - C(a_0, a_1) \geq 0 \\ & (IC): a_1 \in \arg \max_{a'_1 \in A_1} v(r_s)P(a_0, a'_1) + v(r_f)(1 - P(a_0, a'_1)) - C(a_0, a'_1) \end{aligned}$$

Here, the principal only solves for two scalars  $r_s, r_f$  and the FOA can be readily applied. An optimal contract that solves the simplified problem  $(r_s^\dagger, r_f^\dagger; a_0^\dagger, a_1^\dagger)$  specifies a recommended effort  $(a_0^\dagger, a_1^\dagger)$ , a transfer of  $r_s^\dagger$  or  $r_f^\dagger$  depending on the outcome when  $a_0^\dagger$  is observed, and a transfer of 0 regardless of the outcome when  $a_0^\dagger$  is not observed. While such a form of optimal contract is obvious here, Theorem 1 establishes the optimality of  $(r_s^\dagger, r_f^\dagger; a_0^\dagger, a_1^\dagger)$  for the original problem formally.

We now use an example to illustrate how the forcing principle enables us to *explicitly* compute the optimal contract for a simple moral hazard problem with both observable and hidden actions. Consider a principal-agent problem in which a risk-neutral research organization/funder (principal) delegates an R&D project to a risk-neutral researcher (agent). The R&D project requires two inputs from the researcher: an investment  $x \in [0, 1]$  and effort  $y \in [0, 1]$ . The researcher's cost function  $C(x, y)$  is separable in investment and effort and takes the form

$$C(x, y) = x^3 + y.$$

The investment  $x$  (corresponding to  $a_0$  in our general setting) is observable and verifiable, while the effort choice  $y$  ( $a_1$  in our general setting) is the researcher's private information. The outcome of the project, given  $(x, y)$ , is binary (success or failure) and the probability of success is  $p(x, y) = x\sqrt{y}$ . The funder receives a revenue of  $W > 0$  if the project succeeds and a zero revenue otherwise—for simplicity, we normalize  $W$  to be 1. We assume that the researcher is protected by limited liability, with an outside option of zero.

The funder's contract specifies a reward scheme, a recommended investment  $x$ , and a recommended effort  $y$ . According to our forcing principle (also see Example 4), it is without loss of generality that the reward scheme consists of a transfer  $r_f \in [0, W]$  if the project fails and a transfer  $r_s \in [0, W]$  if the project succeeds only after the recommended  $x$  is observed (and a transfer of 0 otherwise)—recall the limited liability constraint. Here the transfers can be understood as research grants offered by the funder. The funder's contracting problem is then:

$$\begin{aligned} & \max_{r_s, r_f \in [0, W]; x, y \in [0, 1]} x\sqrt{y}(W - r_s) - (1 - x\sqrt{y})r_f \\ \text{s.t.} \quad & \text{(IR): } x\sqrt{y}r_s + (1 - x\sqrt{y})r_f - x^3 - y \geq 0, \\ & \text{(IC): } y \in \arg \max_{y'} x\sqrt{y'}r_s + (1 - x\sqrt{y'})r_f - x^3 - y'. \end{aligned}$$

Since the two parties are risk neutral, there is no loss of generality to specify  $r_f = 0$ , reducing the contracting problem to

$$\begin{aligned} & \max_{r_s \in [0, W]; x, y \in [0, 1]} x\sqrt{y}(W - r_s) \\ \text{s.t.} \quad & \text{(IR): } x\sqrt{y}r_s - x^3 - y \geq 0 \\ & \text{(IC): } y \in \arg \max_{y'} x\sqrt{y'}r_s - x^3 - y' \end{aligned}$$

We now solve for the optimal contract in two stages. First, using the first-order approach, we derive the *unique-valued* best effort response of the agent, i.e., the optimal effort from (IC), given  $r_s$  and  $x$ , as<sup>28</sup>

$$y^*(r_s, x) = \left( \frac{xr_s}{2} \right)^2.$$

Since  $y^*(r_s, x)$  is uniquely valued, we replace  $y$  with  $y^*(r_s, x)$  and the problem becomes

$$\max_{r_s \in [0, W]; x \in [0, 1]} \frac{x^2 r_s}{2} (W - r_s) \text{ s.t. } r_s^2 \geq 4x.$$

The (IR) condition is then binding, given that the funder's expected utility strictly increases in  $x$  for any  $r_s$ . This allows us to solve for  $r_s$  explicitly.<sup>29</sup> Finally, the closed-form solution to the optimal contract can be derived as (recall that  $W = 1$ ):

$$r_s^* = \frac{5W}{6}, x^* = \frac{r_s^{*2}}{4}, y^* = \frac{r_s^{*6}}{64}.$$

<sup>28</sup> It can be readily verified that fixing  $x$ , the agent's expected utility is strictly concave in effort  $y$ , and the classic first-order approach (FOA) can be applied here. When FOA cannot be applied, one can employ a polynomial approach introduced by Renner and Schmedders (2015) which (approximately) transforms the principal's (bilevel) optimization problem into a simpler nonlinear optimization problem. More recently, Ke and Ryan (2018) also provide a general approach to solve the standard moral hazard problem via a max-min-max formulation when FOA fails.

<sup>29</sup> It can be verified that after substituting the binding (IR) constraint, the objective function ( $r_s^5(W - r_s)$ ) is strictly quasiconcave in the relevant range, and the unique global maximizer can be derived via the first-order condition as  $r_s^* = \frac{5W}{6}$ .

## 5 Conclusion

In a general moral hazard framework with both hidden and observable actions, we have identified a useful equivalence result, the forcing principle, which implies that the principal in our setting can use simple forcing contracts with essentially outcome-contingent rewards to fully replicate the effect of a general hybrid contract, provided that the agent can be punished sufficiently upon deviations from the recommended observable actions. We have also shown that the forcing principle facilitates the existence analysis of optimal deterministic contracts, as well as the computation of optimal contracts for such moral hazard problems. The forcing principle also helps identify a potential caveat in the characterization of first-best benchmarks in the moral hazard literature, if the agent is protected from being punished sufficiently severely in the contract setting.

We have obtained our results in a general model. In particular, our forcing principle is robust in the sense that it does not depend on fine details of the model. The generality of our model also implies that our results may have a broad range of applications where observable actions are non-trivial components in contracting problems with moral hazard. In more concrete contexts, the characterization of optimal contracts or the relationship between two different types of actions could be further studied based on the forcing principle. Moreover, we conjecture that a result similar to the forcing principle still holds in a more general setting with adverse selection. However, the interrelation between moral hazard, adverse selection, and observable actions substantially complicates the relevant analysis, and we leave this problem for future work.<sup>30</sup>

## Appendix: Proofs

### Proof of Proposition 1.

Let  $k^* = (f^*, a_0^*, a_1^*)$  be a feasible contract. By Assumption 1, there is a Borel measurable schedule,  $t : \Omega \rightarrow D$  such that for any  $(a_0, a_1) \in A$ ,  $\int_{\Omega} v(t(\omega), a_0, a_1) P(d\omega; a_0, a_1) \leq r$ . Define  $s : \Omega \rightarrow D$  such that  $s(\omega) = f^*(\omega, a_0^*)$  for all  $\omega$ , and consider the proposed  $\tilde{f}$  with the simple punishment  $t(\omega)$ :

$$\tilde{f}(\omega, a_0) = \begin{cases} s(\omega), & \text{if } a_0 = a_0^* \\ t(\omega), & \text{otherwise} \end{cases}$$

By construction,  $s \in \mathcal{K}$ ,  $\tilde{f} \in \mathcal{R}_{\mathcal{K}}$  and  $s, \tilde{f}$  are both Borel-measurable. Moreover,  $\tilde{f}(\omega, a_0^*) = f^*(\omega, a_0^*)$ . We next show that  $\tilde{k} = (\tilde{f}, a_0^*, a_1^*)$  satisfies both (IR) and (IC).

Under  $\tilde{k}$ , if the agent follows the recommendation  $a_0^*$ , then for all  $\omega$ ,  $\tilde{f}(\omega, a_0^*) = s(\omega) = f^*(\omega, a_0^*)$ . Hence, the agent's expected utility under  $\tilde{k}$  from choosing  $a_0^*$  is

$$\int_{\Omega} v(\tilde{f}(\omega, a_0^*), a_0^*, a_1^*) P(d\omega; a_0^*, a_1^*) = \int_{\Omega} v(f^*(\omega, a_0^*), a_0^*, a_1^*) P(d\omega; a_0^*, a_1^*) \geq r,$$

<sup>30</sup> In a specific R&D funding setting with moral hazard, adverse selection, and observable actions, Rietzke and Chen (2018) discuss the characterization of optimal mechanisms by directly adopting forcing contracts.

where the inequality follows from feasibility of  $k^*$ . Thus,  $\tilde{k}$  satisfies (IR).

To show (IC), let  $(a_0, a_1)$  be an arbitrary action profile for the agent, and consider the reward schedule  $\tilde{f}$ . If  $a_0 \neq a_0^*$  then  $\tilde{f}(\omega, a_0) = t(\omega)$  for all  $\omega$ . Given Assumption 1, the agent indeed has incentive to choose  $a_0^*$ . If  $a_0 = a_0^*$  then  $\tilde{f}(\omega, a_0) = f^*(\omega, a_0^*)$  for all  $\omega$ . Now if the agent chooses  $(a_0^*, a_1)$ , his expected utility is:

$$\begin{aligned} \int_{\Omega} v(f^*(\omega, a_0^*), a_0^*, a_1) P(d\omega; a_0^*, a_1) &\leq \int_{\Omega} v(f^*(\omega, a_0^*), a_0^*, a_1^*) P(d\omega; a_0^*, a_1^*) \\ &= \int_{\Omega} v(\tilde{f}(\omega, a_0^*), a_0^*, a_1^*) P(d\omega; a_0^*, a_1^*) \end{aligned}$$

The above follows from the optimality of  $(a_0^*, a_1^*)$  for the agent under the reward scheme  $f^*$  and that  $\tilde{f}(\omega, a_0^*) = f^*(\omega, a_0^*)$  for all  $\omega$ . Hence, the contract  $\tilde{k}$  satisfies both (IC) and (IR), and is feasible. Moreover, the fact that  $\tilde{f}(\omega, a_0^*) = f^*(\omega, a_0^*)$  for all  $\omega$  implies that the principal and the agent attain the same expected utility under  $\tilde{k}$  and  $k^*$ .  $\square$

### Proof of Theorem 1.

By Proposition 1, for any feasible contract  $(f, a'_0, a'_1)$ , there is  $s \in \mathcal{X}$  with  $s(\omega) \equiv f(\omega, a'_0)$  and a corresponding feasible punishment contract  $(\tilde{f}, a'_0, a'_1)$  with

$$\tilde{f}(\omega, a_0) = \begin{cases} s(\omega), & \text{if } a_0 = a'_0 \\ t(\omega), & \text{otherwise} \end{cases}, \quad (2)$$

and the principal obtains the same expected utility under  $(\tilde{f}, a'_0, a'_1)$  and  $(f, a'_0, a'_1)$ . Since we can always choose a feasible contract  $(f, a'_0, a'_1)$  with a punishment contract  $(\tilde{f}, a'_0, a'_1)$ , the original problem **(P1)** is equivalent to the following **(P̃1)**:

$$\begin{aligned} &\max_{(s, a'_0, a'_1) \in \mathcal{X} \times A} \int_{\Omega} u(\omega, \tilde{f}(\omega, a'_0), a'_0, a'_1) P(d\omega; a'_0, a'_1) \\ (\tilde{\mathbf{P}}1) \quad &\text{s.t. (IR')} : \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1) P(d\omega; a'_0, a'_1) \geq r \\ &\quad \quad \quad \text{(IC')} : (a'_0, a'_1) \in \arg \max_{(\tilde{a}_0, \tilde{a}_1) \in A} \int_{\Omega} v(\tilde{f}(\omega, \tilde{a}_0), \tilde{a}_0, \tilde{a}_1) P(d\omega; \tilde{a}_0, \tilde{a}_1) \end{aligned}$$

Notice that the reward scheme  $\tilde{f}$  in **(P̃1)** takes a specific form that is determined jointly by  $s(\omega)$  and  $a'_0$  (see (2)).

Consider a new problem **(P̂1)** and recall that  $t(\omega)$  is defined in Assumption 1:

$$\begin{aligned} &\max_{(s, a'_0, a'_1) \in \mathcal{X} \times A} \int_{\Omega} u(\omega, s(\omega), a'_0, a'_1) P(d\omega; a'_0, a'_1) \quad \text{s.t.} \\ (\hat{\mathbf{P}}1) \quad &\text{(IR)} : \int_{\Omega} v(s(\omega), a'_0, a'_1) P(d\omega; a'_0, a'_1) \geq r \\ &\quad \quad \quad \text{(IC}_0\text{)} : \int_{\Omega} v(s(\omega), a'_0, a'_1) P(d\omega; a'_0, a'_1) \geq \int_{\Omega} v(t(\omega), a_0, a_1) P(d\omega; a_0, a_1), \forall a_0 \neq a'_0, a_1 \in A_1 \\ &\quad \quad \quad \text{(IC}_1\text{)} : a'_1 \in \arg \max_{a_1 \in A_1} \int_{\Omega} v(s(\omega), a'_0, a_1) P(d\omega; a'_0, a_1) \end{aligned}$$

**Lemma 1.** Given the definition in (2), the set of constraints (IR') and (IC') in **(P̃1)** is equivalent to the set of constraints (IR), (IC<sub>0</sub>), and (IC<sub>1</sub>) in **(P̂1)**, i.e., any  $(s, a'_0, a'_1) \in \mathcal{X} \times A$  satisfying (IR') and (IC') also satisfies (IR), (IC<sub>0</sub>), and (IC<sub>1</sub>), and vice versa.

**Proof of Lemma 1.** First suppose that  $(s, a'_0, a'_1)$  satisfies (IR') and (IC'). Then (IR)

holds for  $(s, a'_0, a'_1)$  given the definition of  $\tilde{f}(\omega, a_0)$  in (2) and (IR'). Next, since  $(s, a'_0, a'_1)$  satisfies (IC'), we have that for all  $(a_0, a_1) \in A$ ,

$$\int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1)P(d\omega; a'_0, a'_1) \geq \int_{\Omega} v(\tilde{f}(\omega, a_0), a_0, a_1)P(d\omega; a_0, a_1). \quad (3)$$

In particular, by the definition of  $\tilde{f}(\omega, a_0)$ , for  $a_0 \neq a'_0$ , (3) implies

$$\begin{aligned} \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1)P(d\omega; a'_0, a'_1) &= \int_{\Omega} v(s(\omega), a'_0, a'_1)P(d\omega; a'_0, a'_1) \\ &\geq \int_{\Omega} v(t(\omega), a_0, a_1)P(d\omega; a_0, a_1), \end{aligned}$$

which is (IC<sub>0</sub>) in  $(\tilde{\mathbf{P}}1)$ . Moreover, again by (3), we have for all  $a_1 \in A_1$ ,

$$\begin{aligned} \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1)P(d\omega; a'_0, a'_1) &\geq \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a_1)P(d\omega; a'_0, a_1) \\ &= \int_{\Omega} v(s(\omega), a'_0, a_1)P(d\omega; a'_0, a_1), \end{aligned}$$

which is (IC<sub>1</sub>) in  $(\tilde{\mathbf{P}}1)$ . Hence,  $(s, a'_0, a'_1)$  satisfies (IR), (IC<sub>0</sub>), and (IC<sub>1</sub>).

Now suppose that  $(s, a'_0, a'_1)$  satisfies (IR), (IC<sub>0</sub>), and (IC<sub>1</sub>). Then similarly (IR') holds for  $(s, a'_0, a'_1)$  given the definition of  $\tilde{f}(\omega, a_0)$  in (2) and (IR). We now show that  $(s, a'_0, a'_1)$  also satisfies (IC'). Suppose not, i.e., given  $s$ , there is  $(a''_0, a''_1) \neq (a'_0, a'_1)$  in  $A$  such that  $(s, a''_0, a''_1)$  satisfies (IC<sub>0</sub>) and (IC<sub>1</sub>) but

$$\int_{\Omega} v(\tilde{f}(\omega, a''_0), a''_0, a''_1)P(d\omega; a''_0, a''_1) > \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1)P(d\omega; a'_0, a'_1). \quad (4)$$

Suppose  $a''_0 \neq a'_0$ . Then by definition of  $\tilde{f}(\omega, a_0)$  and  $t(\omega)$ , we have

$$\int_{\Omega} v(\tilde{f}(\omega, a''_0), a''_0, a''_1)P(d\omega; a''_0, a''_1) = \int_{\Omega} v(t(\omega), a''_0, a''_1)P(d\omega; a''_0, a''_1) \leq r,$$

contradicting (IR) and (4). Hence, we must have  $a''_0 = a'_0$ , i.e., (4) becomes

$$\int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a''_1)P(d\omega; a'_0, a''_1) > \int_{\Omega} v(\tilde{f}(\omega, a'_0), a'_0, a'_1)P(d\omega; a'_0, a'_1),$$

which further implies

$$\int_{\Omega} v(s(\omega), a'_0, a''_1)P(d\omega; a'_0, a''_1) > \int_{\Omega} v(s(\omega), a'_0, a'_1)P(d\omega; a'_0, a'_1),$$

which, however, contradicts (IC<sub>1</sub>). Hence,  $(s, a'_0, a'_1)$  satisfies (IR') and (IC'). This proves Lemma 1.

Given Lemma 1 and  $\tilde{f}(\omega, a_0)$  defined in (2), the maximization problems of  $(\tilde{\mathbf{P}}1)$  and  $(\mathbf{P}1)$  are therefore equivalent.<sup>31</sup> Now since in  $(\tilde{\mathbf{P}}1)$ , for any recommended action

<sup>31</sup> In cases in which  $\int_{\Omega} v(t(\omega), a_0, a_1)P(d\omega; a_0, a_1) = r$  for some recommended actions  $(a_0, a_1)$  and the agent is indifferent, we assume that the agent takes the recommended actions  $(a_0, a_1)$ .

$a'_0$ , (IC<sub>0</sub>) holds automatically given (IR) and Assumption 1, we can further rewrite **(P1)** as

$$\begin{aligned} & \max_{(s, a'_0, a'_1) \in \mathcal{H} \times A} \int_{\Omega} u(\omega, s(\omega), a'_0, a'_1) P(d\omega; a'_0, a'_1) \\ \text{s.t. (IR): } & \int_{\Omega} v(s(\omega), a'_0, a'_1) P(d\omega; a'_0, a'_1) \geq r \\ \text{(IC}_1\text{): } & a'_1 \in \arg \max_{a_1 \in A_1} \int_{\Omega} v(s(\omega), a'_0, a_1) P(d\omega; a'_0, a_1) \end{aligned}$$

which is exactly **(P2)**. Since **(P1)** and **(P2)** are equivalent optimization problem under Proposition 1, it follows that bullet points 1, 2, and 3 in Theorem 1 hold as desired.  $\square$

### Proof of Proposition 2.

This proof follows the proof of existence of optimal contracts in Page (1987) by replacing sequences of  $f_n$ 's by sequences of  $(s_n, a_{0,n})$ 's. Delbaen's Lemma (Delbaen, 1974) will imply that  $U$  and  $V$  are sequentially continuous. Moreover, the constraint set of **(P2)**  $Gr(A^*) = \{(s, a_0, a_1) \in \mathcal{H} \times A : (s, a_0) \in \mathcal{L}(r), a_1 \in A^*(s, a_0)\}$  is non-empty and sequentially closed, and  $U^* := \sup_{(s, a_0, a_1) \in Gr(A^*)} U(s, a_0, a_1)$  is finite. Next, since  $U^*$  is a supremum, there is a sequence  $\{(s_n, a_{0,n}, a_{1,n})\}_n$  in  $Gr(A^*)$  such that  $U(s_n, a_{0,n}, a_{1,n}) \rightarrow U^*$ . Given the sequential compactness of  $\mathcal{H}$  and the compactness of  $A$ , there is a subsequence  $\{(s_{n_k}, a_{0,n_k}, a_{1,n_k})\}_k$  in  $Gr(A^*)$  and a triple  $(s^*, a_0^*, a_1^*)$  in  $\mathcal{H} \times A$  such that  $s_{n_k} \rightarrow s^*$  pointwise on  $\Omega$  and  $(a_{0,n_k}, a_{1,n_k}) \rightarrow (a_0^*, a_1^*)$  under the metric on  $A$ . Since  $Gr(A^*)$  is sequentially closed,  $(s^*, a_0^*, a_1^*) \in Gr(A^*)$ . Finally, the sequential continuity of  $U$  implies that  $U(s^*, a_0^*, a_1^*) = U^*$ . Therefore,  $(s^*, a_0^*, a_1^*)$  is the solution to **(P2)**.  $\square$

### References

- Bolton, P. and Dewatripont, M. (2005). *Contract Theory*. MIT Press, Cambridge, Massachusetts.
- Chen, B. (2010). All-or-nothing monitoring: comment. *American Economic Review*, 100(1):625–27.
- Chen, B. (2012). All-or-nothing payments. *Journal of Mathematical Economics*, 48(3):133–142.
- Delbaen, F. (1974). Continuity of the expected utility. In Dreze, J., editor, *Allocation under Uncertainty: Equilibrium and Optimality*, pages 254–256. Macmillan, London.
- Dewatripont, M., Legros, P., and Matthews, S. A. (2003). Moral hazard and capital structure dynamics. *Journal of the European Economic Association*, 1(4):890–930.
- Gantt, H. (1919). *Work, Wages and Profits*. The Engineering Magazine Company, New York, 2nd edition.
- Grossman, S. J. and Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1):7–46.
- Holmström, B. (1979). Moral hazard and observability. *Bell journal of Economics*, 10(1):74–91.
- Holmström, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55(2):303–28.

- Holmström, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law Economics and Organization*, 7:24 – 52.
- Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52(1):45–67.
- Kadan, O., Reny, P. J., and Swinkels, J. M. (2017). Existence of optimal mechanisms in principal-agent problems. *Econometrica*, 85(3):769–823.
- Ke, R. and Ryan, C. T. (2018). A general solution method for moral hazard problems. *Theoretical Economics*, 13(3):1425–1481.
- Ke, R. and Xu, X. (2017). Existence of optimal deterministic contracts in moral hazard problems. *Mimeo*.
- Kim, S. K. (1997). Limited liability and bonus contracts. *Journal of Economics & Management Strategy*, 6(4):899–913.
- Laffont, J. J. and Martimort, D. (2009). *The theory of incentives: the principal-agent model*. Princeton university press.
- Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics*, 32(3):514–527.
- Mirrlees, J. (1974). Notes on welfare economics, information and uncertainty. In *Essays on Economic Behavior under Uncertainty*. North Holland, Amsterdam.
- Page, F. H. (1987). The existence of optimal contracts in the principal-agent model. *Journal of Mathematical Economics*, 16(2):157–167.
- Page, F. H. (1991). Optimal contract mechanisms for principal-agent problems with moral hazard and adverse selection. *Economic Theory*, 1(4):323–338.
- Page, F. H. (1992a). Bayesian incentive compatible mechanisms. *Economic theory*, 2(4):509–524.
- Page, F. H. (1992b). Mechanism design for general screening problems with moral hazard. *Economic theory*, 2(2):265–281.
- Periasamy, P. (2010). *A textbook of financial cost and management accounting*. Himalaya Publishing House.
- Renner, P. and Schmedders, K. (2015). A polynomial optimization approach to principal-agent problems. *Econometrica*, 83(2):729–769.
- Rietzke, D. and Chen, Y. (2018). Push or pull? Performance-pay, incentives, and information. Working Paper: Graz Economics Papers 2018-12.
- Sappington, D. (1983). Limited liability contracts between principal and agent. *Journal of Economic Theory*, 29(1):1–21.
- Zhao, R. R. (2008). All-or-nothing monitoring. *American Economic Review*, 98(4):1619–28.