

# The Power of Information for Intermediate States in Contract Design

Yirui Zhang  
Tsinghua University  
Beijing, China  
zhangyr22@mails.tsinghua.edu.cn

Zhixuan Fang  
Tsinghua University  
Beijing, China  
Shanghai Qi Zhi Institute  
Shanghai, China  
zfang@mail.tsinghua.edu.cn

## ABSTRACT

In the conventional principal-agent problem, a principal delegates a task to an agent and formulates a contract to incentivize the agent's actions on behalf of the principal. However, this framework overlooks the information that is possibly available during the delegation process in some scenarios. To address this limitation, we propose a novel model that incorporates multiple intermediate states to capture such information revealed during the delegation. Furthermore, to evaluate the impact of the information embedded in these intermediate states, we introduce two distinct contracts: the pay-halfway contract, which provides payments based not only on final outcomes but also on intermediate states, and the terminate-halfway contract, which allows the principal to terminate the delegation process upon encountering undesirable intermediate states. This leads to the question of whether and how these contract types can leverage intermediate-state information? In particular, we ask: Can these contract types outperform standard contracts, and if so, when and to what extent? We answer the first question affirmatively and provide several important insights regarding the second, shedding light on the circumstances in which intermediate-state-aware contracts yield substantial advantages.

## KEYWORDS

Contract Design; Information

### ACM Reference Format:

Yirui Zhang and Zhixuan Fang. 2026. The Power of Information for Intermediate States in Contract Design. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 9 pages. <https://doi.org/10.65109/EYFD4743>

## 1 INTRODUCTION

Contract theory is a fundamental area of economics that examines how contractual arrangements are formulated in the presence of information asymmetry. Prior research has typically utilized the principal-agent framework to study the dynamic between the involved parties, where the principal (employer) often lacks complete

information about the agent's (employee's) actions or effort levels. This uncertainty necessitates the creation of incentive-compatible contracts that motivate the agent to act in the principal's best interest while considering the agent's self-interest.

In the classic setup, the principal delegates decision-making authority to the agent and can only observe the final outcome, not the actions taken by the agent. This leads to a potential moral hazard, as the agent may not act in the best interest of the principal. To mitigate this problem, the principal designs incentive-compatible contracts that link compensation to observable outcomes in order to encourage desirable behavior. The agent, seeking to maximize her own benefits, will then take actions based on the promised compensations. Much of the existing literature in this field ([3, 6, 10]) is dedicated to determining the optimal contract arrangement from the principal's perspective.

However, real-world delegation often allows the principal to observe some information during the process. For instance, when assigning a project to an employee, the employer may receive mid-term progress reports, enabling them to monitor the project's advancement. Similarly, in supply chain scenarios, suppliers may provide periodic updates on the production process. Such intermediate information is not captured in the standard principal-agent framework. To address this limitation, we propose a new model within contract theory: the two-stage delegation process. This model incorporates intermediate states to capture information revealed during the delegation. Specifically, the process consists of two stages: (i) The agent takes an initial action and transitions to a specific intermediate state, which embeds the information about the delegation process. (ii) At the intermediate state, the agent takes further action, resulting in an outcome. Importantly, the principal can observe both the intermediate state and the final outcome.

This extended framework provides a more accurate representation of the dynamic nature of real-world contractual environments. Many practical settings fit naturally into our framework, including procurement, crowdsourcing, and adaptive task assignment. For example, in procurement, the evolving stages of a project can be interpreted as intermediate information revealed progressively over the course of the delegation process. Consequently, within this two-stage model that incorporates intermediate states, it becomes both interesting and necessary to analyze how such information revelation influences contractual relationships and outcomes.

Corresponding author: Zhixuan Fang.

This work is supported by Tsinghua University Dushi Program.



This work is licensed under a Creative Commons Attribution International 4.0 License.

*Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). <https://doi.org/10.65109/EYFD4743>

## 1.1 Main Results

In this section, we summarize our main results, focusing on the central question: whether and how information embedded in intermediate states benefits the principal.

To explore this, we introduce two conceptually distinct contractual mechanisms: the pay-halfway contract and the terminate-halfway contract. Unlike standard contracts, which provide compensation solely based on the final outcome, the pay-halfway contract allows for compensation upon the realization of specific intermediate states, while the terminate-halfway contract permits early termination of the delegation process when undesirable intermediate states are observed. We analyze and compare the performance of these contracts against the standard outcome-based contract.

The information embedded in intermediate states comprises two key types: (I) the intrinsic quality of the state, meaning how likely final actions from this state lead to valuable outcomes; and (II) the agent’s initial action choice.

To examine which type(s) of information benefit the principal, and how the proposed contracts leverage them, we consider three specialized two-stage delegation processes that differ in the informational content of their intermediate states:

- In the tree two-stage process, each final outcome is uniquely linked to a single intermediate state, rendering the intermediate state uninformative beyond the final outcome. (No additional information)
- In the stochastic first-stage process, there is only one initial action; although intermediate states do not reveal any information about the agent’s initial choice, they carry predictive value regarding the likelihood of favorable final outcomes. (Type I information only)
- In the deterministic first-stage process, each intermediate state corresponds to a unique initial action, thereby containing information both about the state itself and fully revealing the agent’s initial choice. (Type I and II information)

Our results demonstrate the following:

- In the tree two-stage process, both the pay-halfway and terminate-halfway contracts perform identically to the standard contract due to the lack of informative intermediate states.
- In the stochastic first-stage process, the terminate-halfway contract can significantly outperform the standard contract, while the pay-halfway contract only match the standard contract.
- In the deterministic first-stage process, both contracts can outperform the standard contract and extract welfare from the informative intermediate states.

These findings suggest that the pay-halfway contract primarily utilizes Type II information to infer the agent’s initial action and incentivize desirable behavior. In contrast, the terminate-halfway contract makes use of both Type I and Type II information: it not only deters poor initial actions but also allows the principal to exit early from low-quality intermediate states that are likely to result in unfavorable outcomes.

Finally, we compare the two newly proposed contracts directly. While the terminate-halfway contract can utilize richer information,

our analysis shows that it does not always outperform the pay-halfway contract across all settings.

## 1.2 Related Works

Contract theory, as a fundamental discipline within economics, finds application across diverse contexts. It also boasts a rich historical lineage dating back to seminal works such as [9, 12]. For an introductory exploration of contract theory, interested readers are directed to the comprehensive volumes [3] and the introductory chapters provided by [4], which offer foundational insights.

**Contract Design Without States.** Recent research in contract theory spans a wide array of topics. Without introducing states, studies such as [7], [11], [10], and [5] delve into the effectiveness of diverse contract types. For instance, [7] conduct an analysis of linear contracts, demonstrating their optimality in addressing distributional intricacies. In contrast, [11] and [10] compare various contract types within scenarios characterized by unknown outcome distributions. Additionally, [5] investigate the potency of randomization, focusing on the efficient design of random contract menus. However, it is crucial to emphasize that within the framework of these previous works, there is a notable absence of intermediate states pertaining to the actions of agents. This distinction highlights a gap in the literature that our research aims to address by incorporating intermediate states into contract design.

**Contract Theory in MDPs.** The literature most closely related to our work studies principal-agent problems in Markov Decision Processes (MDPs). Early work by [8] and [4] analyzes dynamic contracts with hidden actions in stochastic environments and characterizes optimal contracts via dynamic programming. [20] consider a related MDP framework with different cost and reward structures and hidden states instead of hidden actions. More recent studies include [17], which examine repeated interactions where the principal conditions decisions on the agent’s strategic state reports, and [15, 16], which focus on MDPs in which the agent can terminate the process but does not control actions. [2] show that Markovian policies are generally suboptimal when agents take costly actions and propose efficient algorithms for computing optimal policies. Finally, [13] study a contractual reinforcement learning setting in which a learning principal influences the agent’s policy through state-contingent payments.

Another research strand focuses on MDP-based contracts under global budget constraints rather than profit-contingent payments. Several studies [1, 14, 18, 19] explore how reward shaping can augment the principal’s profit while considering observable actions. Specifically, [18] formulate a mixed-integer linear program, [14] consider a time-inconsistent biased agent, [19] study the induction of a pre-specified target policy, and [1] focus on instances where polynomial time approximation algorithms can be developed.

Compared with our work and most studies on MDP-based contract models, the core objectives differ fundamentally. MDP-based contract works generally aim to design an optimal contract within a fully dynamic MDP environment. In contrast, our work treats intermediate states as intermediate information and investigates whether and how access to this information benefits the principal relative to scenarios where such information is not leveraged.

The modeling assumptions also differ. We assume that the principal receives reward only upon reaching the final outcome, whereas

MDP-based contract models typically accumulate rewards across states. Moreover, our model adopts a non-recurrent, two-step structure in which the intermediate state represents intermediate information. By contrast, MDP-based contract models may involve recurrent transitions and rely on structural assumptions suited to long-horizon, repeatedly interacting environments.

Intuitively, our model captures a single delegation process in which information is revealed along the way but rewards are realized only at the end. In contrast, most MDP-based contract models reflect successive delegation processes, where the principal commits to contracts and receives rewards at each step. Our framework therefore aligns more naturally with practical delegation settings that involve intermediate information but do not rely on the full machinery of dynamic MDP-based contract models.

## 2 SETUP

In this section, we introduce our new model, the two-stage delegation process. For those interested, the standard contract instance can be found in Appendix.

### 2.1 Model

A two-stage delegation process involves a principal and an agent. The principal assigns a task with  $M$  possible outcomes to the agent, who must take actions to complete it. The process includes  $S$  intermediate states rather than directly producing an outcome. The agent first chooses one of  $N_1$  initial actions, each incurring a cost  $c_i$ . Upon choosing action  $i$ , the agent transitions to an intermediate state  $s$  according to a probability distribution  $\mathcal{F}_i$ , where  $F_{i,s}$  denotes the probability of reaching intermediate state  $s$ . At each intermediate state  $s$ , the agent selects one of  $N_2$  final actions, each with an associated cost  $c_j^s$ . Selecting final action  $j$  at the intermediate state  $s$  leads to an outcome based on a probability distribution  $\mathcal{F}_j^s$ , with  $F_{j,m}^s$  representing the probability of outcome  $m$ . There exists a null initial action, and at every intermediate state, there is also a corresponding null final action, each incurring zero cost, i.e.,  $c_i = 0$  and  $c_j^s = 0$  for the respective null actions. The principal's reward depends solely on the realized outcome: a reward  $r_m$  is received if the outcome  $m$  occurs. The principal observes the intermediate state and the outcome, but not the agent's chosen initial nor the final action. The expected reward of taking the final action  $j$  in intermediate state  $s$  is given by  $R_j^s = \sum_{m=1}^M F_{j,m}^s r_m$ .

An action profile  $\mathbf{a} = (i, j_1, j_2, \dots, j_S)$  specifies that the agent first chooses the initial action  $i$ , and then selects the final action  $j_s$  when transitioning to intermediate state  $s$ . The expected reward of profile  $\mathbf{a}$  is calculated as

$$R_{\mathbf{a}} = \sum_{s=1}^S F_{i,s} R_{j_s}^s.$$

The expected cost is calculated as

$$c_{\mathbf{a}} = c_i + \sum_{s=1}^S F_{i,s} c_{j_s}^s.$$

The expected welfare of profile  $\mathbf{a}$  is defined as  $R_{\mathbf{a}} - c_{\mathbf{a}}$ . The maximal welfare  $\mathcal{W}$  is  $WEL = \max_{\mathbf{a}} \{R_{\mathbf{a}} - c_{\mathbf{a}}\}$ .

### 2.2 Contracts and Agent's Behavior

The principal's objective remains consistent with that of the standard contract instance, aiming to incentivize the agent to take actions advantageous to the principal. The principal achieves this by designing contracts that pledge transfers upon specific occurrences to incentivize agent behavior.

In this section, we will initially outline various contracts. Subsequently, we will delineate the agent's behavior. Finally, we provide illustrative examples to aid understanding and demonstrate the practical application of these contract structures.

**2.2.1 Contracts.** We begin by introducing the standard contract and one of its special cases: the linear contract. Neither of these contracts utilizes information about intermediate states.

**Standard contract.** A standard contract  $\mathbf{t}$  is an  $M$ -dimensional vector, where the  $m$ -th non-negative entry  $t_m$  denotes the transfer made to the agent if outcome  $m$  is realized. We denote the expected profit under the optimal standard contract by *OPT-Standard*.

**Linear Contract.** A linear contract is a specific form of standard contract characterized by a fraction parameter  $\alpha$ . Under this contract, the agent receives a transfer of  $\alpha r_m$  if outcome  $m$  occurs. We denote the expected profit under the optimal linear contract by *OPT-Linear*.

Next, we introduce two novel contract types designed to incorporate intermediate-state information: the pay-halfway contract and the terminate-halfway contract. These two contracts employ seemingly opposite mechanisms: one introduces additional transfers at intermediate states, while the other allows for early termination of the delegation process upon observing certain intermediate states.

**Pay-halfway Contract.** The pay-halfway contract allows the principal to issue non-negative transfers based on intermediate states in addition to final outcomes. It is represented by a  $(S + M)$ -dimensional vector  $(\mathbf{s}, \mathbf{t})$ , where:

- $\mathbf{s} \in \mathbb{R}_{\geq 0}^S$  specifies the transfer  $s_s$  upon reaching intermediate state  $s$ , and
- $\mathbf{t} \in \mathbb{R}_{\geq 0}^M$  specifies the transfer  $t_m$  upon outcome  $m$ .

We denote the expected profit under the optimal pay-halfway contract by *OPT-Pay*.

**Terminate-halfway Contract.** The terminate-halfway contract allows the principal to end the delegation process upon observing certain intermediate states. It is defined by a pair  $(\mathbf{t}, \mathcal{S})$ , where:

- $\mathbf{t} \in \mathbb{R}_{\geq 0}^M$  denotes the transfer  $t_m$  for each final outcome  $m$ , and
- $\mathcal{S} \subseteq \{1, 2, \dots, S\}$  specifies the set of intermediate states at which the process is terminated.

If an intermediate state  $s \in \mathcal{S}$  is reached, the delegation process ends immediately, and both the principal and the agent receive zero payoff. The expected profit under the optimal terminate-halfway contract is denoted by *OPT-Terminate*.

**2.2.2 Agent's Behavior.** We impose the standard assumptions of incentive compatibility (IC) and individual rationality (IR) from contract theory. In our setting, the IR constraint is readily satisfied due to the availability of null actions, which guarantee the agent a nonnegative outside option.

To analyze the IC constraint, note that once the principal commits to a contract, the agent faces a sequential decision problem.

	State 1, <i>fail</i>	State 2, <i>pass</i>
<b>Initial Action 1, <i>effort</i></b> Cost 1.8	0.1	0.9
<b>Initial Action 2, <i>null</i></b> Cost 0	1	0
state 1, <i>fail</i>	<b>Outcome 1, <i>fail</i></b> Reward 0	<b>Outcome 2, <i>success</i></b> Reward 5
<b>Final Action 1, <i>effort</i></b> Cost 2	0.2	0.8
<b>Final Action 2, <i>null</i></b> Cost 0	0.9	0.1
state 2, <i>pass</i>		
<b>Final Action 1, <i>effort</i></b> Cost 1	0	1
<b>Final Action 2, <i>null</i></b> Cost 0	0	1

Table 1: The instance of Example 2.1

The agent selects actions to maximize her expected utility, and the optimal action profile can be characterized by backward induction. The agent first computes the optimal final action and corresponding utility for each intermediate state. Taking these into account, along with possible transfers at intermediate states and the risk of early termination, she then selects the optimal initial action.

Specifically, given a contract: (1)  $t$  (standard), (2)  $(s, t)$  (pay-halfway), or (3)  $(t, S)$  (terminate-halfway), the agent first determines the final optimal action  $j_s^*$  for each intermediate state  $s$  as:

$$j_s^* \in \operatorname{argmax}_j \sum_{m=1}^M F_{j,m}^s t_m - c_j^s.$$

Let  $U^s$  denote the maximum utility the agent can obtain at intermediate state  $s$ :

$$U^s = \max_j \sum_{m=1}^M F_{j,m}^s t_m - c_j^s.$$

Then, the agent selects the initial action  $i^*$  by considering the utility  $U^s$ , the intermediate transfers  $s_s$ , and the risk of termination:

$$i^* \in \operatorname{argmax}_i \sum_{s=1}^S \mathbb{1}(s \notin S) F_{i,s} (U^s + s_s) - c_i.$$

Upon reaching an intermediate state  $s \notin S$ , the agent selects final action  $j_s^*$ . We assume that in case of ties, the agent breaks ties in favor of the principal.<sup>1</sup> Accordingly, the expected payment from the principal to the agent is:

$$T = \sum_{s=1}^S \mathbb{1}(s \notin S) F_{i^*,s} (F_{j_s^*,m}^s t_m + s_s).$$

**2.2.3 Examples of Newly Proposed Contracts.** In this section, we provide examples of the newly proposed contracts to demonstrate their practical use and deepen the understanding of our model.

*Example 2.1 (Pay-halfway Contract).* If a teacher assigns a programming task to a student, her primary concern is the final outcome: success or failure. The teacher receives a reward of  $r = 5$  only

<sup>1</sup>This can be achieved by utilizing arm recommendations.

	State 1, <i>bad</i>	State 2, <i>well</i>
<b>Initial Action 1, <i>effort</i></b> Cost 8	0.01	0.99
<b>Initial Action 2, <i>null</i></b> Cost 0	1	0
state 1, <i>bad</i>	<b>Outcome 1, <i>failure</i></b> Reward 0	<b>Outcome 2, <i>success</i></b> Reward 10
<b>Final Action 1, <i>effort</i></b> Cost 1	0	1
<b>Final Action 2, <i>null</i></b> Cost 0	0	1
state 2, <i>well</i>		
<b>Final Action 1, <i>effort</i></b> Cost 4	0.4	0.6
<b>Final Action 2, <i>null</i></b> Cost 0	0.9	0.1

Table 2: The instance of Example 2.2

if the student successfully completes the program (denoted as "success"). There is a mid-term presentation to evaluate the student's progress, with outcomes either "pass" or "fail". If the student delivers an outstanding presentation, reaching the intermediate state "pass", the likelihood of achieving "success" increases. Otherwise, the program is more likely to fail.

Initially, the student has two choices: to exert effort or not, denoted as "effort" and "null". Exerting effort incurs a cost of 1.8 and gives a high probability ( $p = 0.9$ ) of reaching the "pass" state. Choosing "null" inevitably leads to failing the presentation. After the mid-term presentation, upon reaching an intermediate state, the student again faces two choices: to exert effort or not. If the student passes the presentation, either action leads to a successful final outcome, but "effort" incurs an additional cost of 1. If the student fails the presentation, the "effort" action incurs a cost of 2 and has a 0.8 probability of leading to success, while the "null" action results in only a 0.1 probability of success.

In this scenario, although the teacher doesn't directly benefit from the student's good performance in the mid-term presentation, she may still choose to reward the student for doing well in order to incentivize the student's initial effort. For instance, the teacher might promise a reward of 2 if the student achieves "pass" in the mid-term presentation and 0.1 if the student successfully completes the program. This payment arrangement resembles a pay-halfway contract  $(s, t)$ , where  $s = (0, 2)$ ,  $t = (0, 0.1)$ . Given this "contract", the student employs backward induction. At the "pass" state, she chooses the action "null", resulting in an expected utility of  $U_1 = 0.1$ . At the "fail" state, she also chooses the null action, yielding an expected utility of  $U_2 = 0.01$ . When determining the initial action, if she exerts effort, the expected utility is  $(U_1 + 2)p + U_2(1 - p) - c = 0.091$ . Otherwise, the expected utility is 0.01. Consequently, the student opts to exert effort initially and takes the null action regardless of the state. The expected profit for the principal is  $p(r - 0.1 - 2) + (1 - p)(0.1(r - 0.1)) = 2.659$ .

*Example 2.2 (Terminate-halfway Contract).* When an employer delegates a task to a worker, the employer's primary concern is the

completion of the task. However, there may be a need to estimate the progress of the task. If the worker performs poorly during an interim evaluation, the employer might terminate the delegation process and reassign the task to another worker.

Specifically, there are two potential outcomes: "success" and "failure", along with two intermediate states: "well" and "bad". The employer is rewarded with  $r = 10$  only upon the successful completion of the task. The worker has two initial actions and two final actions: "effort," which entails exertion at a cost of 8 and leads to the "well" state with a high probability of  $p = 0.99$ , and "null," representing initial action with zero cost that leads to the "bad" state. At the "well" state, both final actions result in success. In contrast, at the "bad" state, choosing "effort" results in success with a probability of 0.6 at an additional cost of 4, while selecting "null" results in failure with a probability of 0.9.

In this scenario, the employer may promise to give the worker a payment of 8.2 if the task is completed successfully, i.e., if the "success" outcome is reached. Additionally, the employer will terminate the delegation process if the worker performs poorly at the interim evaluation, i.e., if the "bad" intermediate state is reached. This can be viewed as a terminate-halfway contract represented by  $(t, \mathcal{S})$  where  $t = (0, 8.2)$  and  $\mathcal{S} = \{bad\}$ . Regarding this contract, the worker first computes that at the "well" state, she will take the final action "null," resulting in an expected utility of 8.2. Then, she recognizes that she will get zero utility regardless of the final action she takes at the "bad" state. Consequently, by simple computation, she decides to take the initial action "effort." In this way, the employer earns the expected profit of 1.8.

### 2.3 Specialized Two-stage Delegation Process

Intermediate states primarily convey two types of information:

- (I) the intrinsic quality of the state itself, specifically, the likelihood that subsequent final actions taken from this intermediate state will lead to valuable outcomes; and
- (II) the agent's initial action choice.

In this section, we introduce several specialized two-stage delegation processes that differ in the type and extent of information embedded in intermediate states. Our objective is to examine whether, and which type of, information benefits the principal. More specifically, we aim to assess whether and how the newly proposed contracts can effectively leverage these different types of information.

In the tree two-stage process, each outcome can be reached through only one specific intermediate state. Consequently, the intermediate state appears to offer no additional information beyond what is ultimately revealed by the final outcome.

In the random first-stage process, there is only one initial action, which may lead to multiple intermediate states. These intermediate states do not reveal any additional information about the agent's initial choice (since there is only one). However, the distribution of final actions leading to outcomes can vary across different intermediate states. Some intermediate states may be more likely to result in favorable outcomes than others. Therefore, intermediate states in this setting primarily convey information about the intrinsic quality of the states themselves (Type I information).

Finally, in the deterministic first-stage process, each intermediate state corresponds uniquely to a specific initial action. As a result, observing the intermediate state fully reveals the agent's initial decision and may also convey information about the quality of ensuing outcomes. Hence, intermediate states in this case encapsulate both types of information and are potentially highly informative.

We now proceed to formally define these specialized two-stage delegation processes:

- **Tree Two-stage Process:** In this process, each final outcome can only be realized through a unique intermediate state. Formally, for each outcome  $m$ , there exists at most one intermediate state  $s$  such that  $\sum_{j \in [N_2]} F_{j,m}^s > 0$ .
- **Stochastic First-stage Process:** In this process, there is only one initial action, i.e.,  $N_1 = 1$ .
- **Deterministic First-stage Process:** In this process, each initial action deterministically leads to exactly one intermediate state. Specifically, for every initial action  $i$ , there exists an intermediate state  $s$  such that  $F_{i,s} = 1$ .

## 3 UPPER BOUND RESULTS

In this section, we briefly explore the role and impact of information embedded in intermediate states within our framework, providing several upper-bound results.

We first analyze the extent to which access to intermediate-state information improves the optimal welfare compared to the profit achievable under a standard contract. By establishing upper bounds through comparisons between the welfare and the profit under the optimal standard contract, we derive limits on the maximum profit ratio that our newly proposed contracts can achieve relative to the standard contract.

We show that, in the general two-stage delegation process, the total welfare is at most  $O(SN_1N_2)$  times the profit. We also consider the specialized two-stage delegation process and derive corresponding upper bound results.

**THEOREM 3.1.** *Consider a two-stage delegation process with  $S$  intermediate states,  $N_1$  initial actions, and  $N_2$  final actions, the welfare is no more than  $O(SN_1N_2)$  times the profit obtained by the optimal standard contract, i.e.,*

$$WEL \leq O(SN_1N_2) \cdot OPT\text{-Standard}. \quad (1)$$

**PROOF SKETCH.** Firstly, given that  $OPT\text{-Standard} \geq OPT\text{-Linear}$ , we can constrain the ratio of welfare and profit obtained through the optimal standard contract by the ratio of welfare and profit obtained through the optimal linear contract, i.e.  $\frac{WEL}{OPT\text{-Standard}} \leq \frac{WEL}{OPT\text{-Linear}}$ . Thus, our focus narrows down to analyzing the ratio of welfare and profit achieved through the optimal linear contract, denoted as  $r = \frac{WEL}{OPT\text{-Linear}}$ .

Consequently, we examine the performance of the optimal linear contract, starting with the aspect of breakpoints. A breakpoint  $\alpha$  marks the threshold at which the agent's chosen action undergoes a transition. It can be obtained that the welfare can be bounded in terms of the number of breakpoints and the profit under the optimal standard contract. Specifically, the following inequality holds:

$$WEL \leq nOPT\text{-Standard}$$

where  $n$  is the number of breakpoints.

There are at most  $N_1(N_2)^S$  possible action profiles, so at most that many breakpoints, but this is a loose bound. For each intermediate state  $s$ , there can be at most  $N_2$  breakpoints (one for each possible final action), denoted as  $\alpha_1^s, \alpha_2^s, \dots, \alpha_{N_2}^s$ . Sort these breakpoints in increasing order and relabel them accordingly, resulting in  $\alpha_1, \alpha_2, \dots, \alpha_m$ . Consequently, there are at most  $SN_2$  breakpoints in total for the final actions.

Between any two adjacent breakpoints  $\alpha_i$  and  $\alpha_{i+1}$ , the choice of the agent’s final action remains unchanged, and the expected reward of every initial action remains constant as well. Thus, we can interpret this process as a new standard contract instance, with each intermediate state acting as an outcome. Under this perspective, there are at most  $N_1$  breakpoints between  $\alpha_i$  and  $\alpha_{i+1}$ . Consequently, there are at most  $SN_1N_2$  breakpoints in total.

In conclusion, with  $SN_1N_2$  breakpoints, the ratio of welfare and profit achieved through the optimal linear contract, denoted as  $r$ , is bounded by no more than  $SN_1N_2$ . Thus, we arrive at the desired result.  $\square$

CLAIM 1. Consider a two-stage delegation process with  $S$  intermediate states,  $N_1$  initial actions, and  $N_2$  final actions:

- (1) In a deterministic first-stage process, welfare is bounded by  $WEL \leq O(\min\{S, N_1\} \cdot N_2) \cdot OPT\text{-Standard}$ .
- (2) In a stochastic first-stage process, welfare is bounded by  $WEL \leq O(SN_2) \cdot OPT\text{-Standard}$ .

Notably, total welfare provides only a loose upper bound on the principal’s attainable profit under any contract. Since welfare is defined as the reward minus the agent’s cost, full extraction would leave the agent with negligible utility, which may render such contracts infeasible (e.g., involving negative transfers).

The following claim demonstrates that, in the tree two-stage process, neither of the newly proposed contracts outperforms the standard contract. This result is consistent with the observation that, in the tree two-stage process, the intermediate state provides no additional information beyond what is ultimately revealed by the final outcome.

CLAIM 2. Consider a tree two-stage process with  $S$  intermediate states,  $N_1$  initial actions, and  $N_2$  final actions, the optimal pay-halfway contract and the optimal terminate-halfway can only yield the same profit with the optimal standard contract, i.e.,

$$OPT\text{-Pay} = OPT\text{-Terminate} = OPT\text{-Standard}. \quad (2)$$

## 4 PAY-HALFWAY CONTRACT

In this section, we conduct a detailed analysis of the pay-halfway contract, with a focus on the central questions: *Does the pay-halfway contract leverage information embedded in intermediate states to improve the principal’s profit, and if so, how and to what extent?* Specifically, we compare the performance of the pay-halfway contract with that of the standard outcome-based contract. Our aim is to determine whether the pay-halfway contract can generate strictly higher profit than the standard contract, and to identify the conditions under which such improvement arises, as well as the magnitude of its potential advantage.

We begin by referencing Example 2.1 to affirmatively answer our first question. That is, the pay-halfway contract can indeed

generate higher profits compared to a standard contract. In Example 2.1, while the optimal standard contract  $t = (0, \frac{20}{9})$  yields an approximate profit of 2.53, the pay-halfway contract we provide in the Example  $((s, t),$  where  $s = (0, 2), t = (0, 0.1))$  achieves an expected profit of 2.659. This highlights the effectiveness of the pay-halfway contract.

The remainder of the section is structured as follows. We first present Theorem 4.1, which shows that the pay-halfway contract can substantially reduce the principal’s total payment, thereby illustrating its effectiveness relative to the standard contract. We then analyze a series of specialized delegation processes to investigate the conditions under which the pay-halfway contract outperforms the standard contract, and to what extent.

The following theorem presents an instance that highlights the power of the pay-halfway contract in reducing payments: the ratio of payments required to incentivize the optimal action profile under the standard contract versus the pay-halfway contract can become arbitrarily large.

THEOREM 4.1. There is a two-stage delegation process with  $S$  intermediate states,  $N_1$  initial actions, and  $N_2$  final actions, where the payment ratio between incentivizing the same action profile  $a^*$  by the standard contract and the pay-halfway contract approaches infinity. The action profile  $a^*$  represents the optimal action profile in terms of maximizing welfare.

PROOF SKETCH. We construct an example with three intermediate states and two outcomes. In one state, the final action yields a high reward at zero cost; in another, the high reward is achievable but requires a transfer to cover a cost  $c$ ; in the third state, all final actions lead to zero reward. There are two initial actions: one initial action assigns a higher probability  $p$  to the high-reward zero-cost state and the remaining probability to the zero-reward state, while the other assigns a lower probability  $q < p$  to the high-reward zero-cost state and allocates the rest to the costly state. The welfare-optimal profile chooses the second initial action that involves visiting the costly state.

Under a standard contract, incentives are based only on final outcomes. To incentivize the optimal action profile, the payment for the high-reward outcome must cover the agent’s cost and compensate for incentive constraints arising from the difference in transition probabilities.

In contrast, a pay-halfway contract enables the principal to directly cover the agent’s cost at the critical intermediate state with a much smaller total payment, while still preserving incentive compatibility.

By carefully choosing the transition probabilities, the payment required under the standard contract grows unbounded, while the payment under the pay-halfway contract remains bounded. This results in the ratio of payments approaching infinity, demonstrating the claimed separation.  $\square$

### 4.1 Results in Specialized Delegation Process

We now analyze how the pay-halfway contract leverages the information embedded in intermediate states by comparing its performance to the standard contract across different specialized delegation processes.

We first show that in the stochastic first-stage process, the optimal pay-halfway contract yields the same profit as the optimal standard contract.

**CLAIM 3.** *Consider a stochastic first-stage process with  $S$  intermediate states,  $N_1$  initial actions, and  $N_2$  final actions, the optimal pay-halfway contract yields the same profit with the optimal standard contract, i.e.,*

$$OPT\text{-Pay} = OPT\text{-Standard}. \quad (3)$$

In contrast, in the deterministic first-stage process, the pay-halfway contract can extract nearly the entire welfare and significantly outperform the standard outcome-based contract.

In a deterministic first-stage process, each initial action leads to only one intermediate state. When employing a pay-halfway contract, the agent can simply be incentivized to undertake costly initial actions by offering transfers to the corresponding intermediate state. However, with a standard contract, if we aim to incentivize costly initial actions, we can only raise payments on certain outcomes. Yet, this might inadvertently increase the expected utility for other states, consequently affecting other initial actions as well. Hence, we construct an instance where the pay-halfway contract can specifically incentivize optimal actions, while the standard contract must increase rewards for all initial actions, resulting in inefficient use of payments.

**THEOREM 4.2.** *There is a deterministic first-stage process with  $N_1+1$  intermediate states,  $N_1+1$  initial actions, and  $N_2+1$  final actions, where the profit obtained by the optimal pay-halfway contract is more than  $\Omega(N_1N_2)$  times the profit obtained by the optimal standard contract, i.e.,*

$$OPT\text{-Pay} \geq \Omega(N_1N_2) \cdot OPT\text{-Standard}. \quad (4)$$

**PROOF SKETCH.** Consider two possible outcomes: a zero-reward outcome  $r_1 = 0$ , and a high-reward outcome  $r_2 = r$ , where  $r$  is chosen to be sufficiently large. There are  $N_1 + 1$  initial actions, each deterministically leading to a distinct intermediate state. Among these, a special initial action  $a_{N_1+1}$  incurs a large cost  $c = \sum_{k=1}^{(N_1+1)N_2+1} \lambda^k$  and leads to the state  $s_{N_1+1}$ . All other initial actions incur zero cost and lead to intermediate states  $s_1, \dots, s_{N_1}$ , respectively.

Each intermediate state has  $N_2 + 1$  final actions. In particular, in state  $s_{N_1+1}$ , all final actions incur no additional cost and yield an expected reward of  $R = c + (N_1 + 1)N_2 + 1$ . For each  $s \in [N_1]$  and final action  $a_j^s$  with  $j \in [N_2]$ , the associated cost is  $c_j^s = \sum_{k=1}^{sN_2+j} \lambda^k$ , and the probability of achieving the high-reward outcome is set such that the expected reward is  $R_j^s = c_j^s + sN_2 + j$ .

A pay-halfway contract that pays  $c$  at state  $s_{N_1+1}$  can fully extract the total welfare, which is  $\mathcal{M} = \mathcal{W} = R - c = (N_1 + 1)N_2 + 1$ . In contrast, an optimal standard-contract mechanism can achieve at most the same profit as linear contracts. Due to the exponential nature of the cost structure, the maximum achievable profit under such contracts is bounded by  $O(1)$ .

Therefore, the profit ratio satisfies:  $\frac{\mathcal{M}}{OPT\text{-Standard}} \geq \frac{(N_1+1)N_2+1}{O(1)} = \Omega(N_1N_2)$ .  $\square$

From Claim 1, we observe that the profit under the optimal pay-halfway contract does not exceed  $O(N_1N_2)$  times the profit achievable under the optimal standard contract, as the profit is

bounded above by the welfare. Therefore, the performance of the pay-halfway contract actually matches the tight bound.

**REMARK 1.** *In conclusion, we find that the pay-halfway contract primarily leverages intermediate-state information to infer the agent's initial action choice and thereby incentivize desirable initial behavior. Consequently, this contract structure is particularly effective in deterministic first-stage process, where intermediate states directly reflect the agent's initial actions, but is less suitable for stochastic first-stage process, where such information is absent.*

## 5 TERMINATE-HALFWAY CONTRACT

In this section, we analyze the terminate-halfway contract. Similarly, we investigate whether and how this contract leverages information embedded in intermediate states.

In Example 2.2, the optimal standard contract  $t = (0, 8)$  yields the expected profit of 1.2 which is strictly less than the expected profit of the terminate-halfway contract we mentioned in the example. Thus, the terminate-halfway contract provides a higher expected profit compared to the optimal standard contract, demonstrating the effectiveness of incorporating intermediate state information into the contract design.

We further analyze the aforementioned question by comparing its performance to that of the standard outcome-based contract across various specialized delegation processes.

### 5.1 Results in Specialized Delegation Process

We present several tight counterexamples within specialized delegation processes to illustrate the effectiveness of the terminate-halfway contract. The results simultaneously demonstrate which types of information the terminate-halfway contract can utilize and the extent to which these information types can enhance its performance. We focus on high-level insights, with full technical details provided in the Appendix.

We start with a deterministic first-stage process to show the significant advantage of terminate-halfway contracts.

**THEOREM 5.1.** *There is a deterministic first-stage process with  $N_1 + 1$  intermediate states,  $N_1 + 1$  initial actions, and  $N_2 + 1$  final actions, where the profit obtained by the terminate-halfway contract is more than  $\Omega(N_1N_2)$  times the profit obtained by the optimal standard contract, i.e.,*

$$OPT\text{-Terminate} \geq \Omega(N_1N_2) \cdot OPT\text{-Standard}. \quad (5)$$

**HIGH-LEVEL IDEA.** The terminate-halfway contract permits the delegation process to be terminated if certain unfavorable intermediate states are observed, resulting in an expected utility of zero for the agent. Consequently, this contract can potentially incentivize the agent to select specific initial actions, particularly in deterministic settings. In the deterministic setting, the terminate-halfway contract can motivate the agent to choose a specific initial action by blocking all intermediate states except the one that results from the desired initial action. In contrast, this level of control is not achievable under a standard contract. Therefore, we construct an instance, similar to that in Theorem 4.2, in which any standard contract can obtain at most a  $\frac{1}{\Omega(N_1N_2)}$  fraction of the total welfare, whereas the terminate-halfway contract is able to extract the full welfare.  $\square$

Next, we consider a stochastic first-stage process to further illustrate the superior performance of terminate-halfway contracts.

**THEOREM 5.2.** *There is a stochastic first-stage process with  $2S + 1$  intermediate states, and  $N_2 + 1$  final actions, where the profit obtained by the terminate-halfway contract is more than  $O(SN_2)$  times the profit obtained by the optimal standard contract, i.e.,*

$$OPT\text{-}Terminate \geq \Omega(SN_2) \cdot OPT\text{-}Standard. \quad (6)$$

**HIGH-LEVEL IDEA.** The effectiveness of a terminate-halfway contract stems from the principal’s ability to end the delegation process when adverse intermediate states are observed. This allows the principal to “block” states that would result in negative profits, thereby maximizing overall profit. Consider a scenario with one positive reward outcome and zero-reward outcomes achievable only through optimal final actions in certain states. Incentivizing these actions via the positive outcome can lead to wasteful payments. Furthermore, some states lead exclusively to zero-reward outcomes regardless of the final action. With a terminate-halfway contract, the principal can incentivize the optimal final actions through transfers conditioned on zero-reward outcomes, while blocking states that might result in negative profits. In contrast, a standard contract would either require excessive payments to incentivize desirable actions through the transfer on the high-reward outcome or result in negative profits from undesirable states. Therefore, the terminate-halfway contract enables the principal to extract nearly the full welfare by strategically managing incentives and “blocking” negative outcomes. Conversely, a standard contract forces the principal to choose between incurring wasteful payments or facing negative profits from adverse states.  $\square$

**REMARK 2.** *In summary, the strength of the terminate-halfway contract lies in its ability to reduce excessive payments by allowing the principal to terminate the delegation process upon observing unfavorable signals (i.e., undesirable intermediate states). It can also incentivize the agent to avoid poor initial actions. This contract is particularly effective in the deterministic first-stage process, where intermediate states reveal the agent’s initial action, thereby enabling the principal to encourage desirable initial action. Additionally, it performs well in stochastic first-stage process, where intermediate states may indicate the potential quality of outcomes resulting from final actions, allowing the principal to preemptively terminate low-value trajectories. Therefore, the terminate-halfway contract effectively leverages both types of intermediate-state information: the intrinsic quality of the state and the agent’s initial action.*

## 6 RELATIONSHIP BETWEEN PAY-HALFWAY AND TERMINATE-HALFWAY CONTRACT

In this section, we examine the relationship between the two newly proposed contract types.

Building on previous findings, we observe that the pay-halfway contract primarily leverages intermediate-state information to infer the agent’s initial action and to incentivize desirable initial behavior. Consequently, it tends to perform relatively poorly in the stochastic first-stage process, where intermediate states are less informative about the agent’s initial choice. In contrast, the terminate-halfway

contract appears to make more comprehensive use of intermediate-state information: it not only discourages undesirable initial actions but also allows the principal to avoid low-quality intermediate states that are likely to lead to poor outcomes. These observations naturally lead to the question: Does the terminate-halfway contract exhibit strictly greater effectiveness than the pay-halfway contract? More specifically, does it consistently outperform the pay-halfway contract across all settings?

**THEOREM 6.1.** *There exists an instance where the pay-halfway contract outperforms the terminate-halfway contract.*

**PROOF SKETCH.** We follow the same instance construction as in Theorem 4.1. The key idea is to compare the maximal achievable profit under the pay-halfway contract and the terminate-halfway contract for the optimal action profile.

For the terminate-halfway contract, there are two cases to consider: either some state is blocked, or no state is blocked. If no state is blocked, the contract effectively reduces to a standard contract, which is strictly suboptimal compared to the pay-halfway contract. Conversely, if a state is blocked, the payment necessary to incentivize the optimal action profile yields a strictly lower maximal profit than that of the pay-halfway contract, as the profit in the blocked state remains positive.

By comparing the profit expressions for both contracts, we show that the pay-halfway contract yields strictly higher profit under appropriate parameter settings. This demonstrates the superiority of the pay-halfway contract over the terminate-halfway contract in this setting.  $\square$

This Theorem answers our question by showing that the pay-halfway contract can outperform the terminate-halfway contract in certain cases. One possible explanation is that, although the terminate-halfway contract leverages richer information—both the agent’s initial action and the intrinsic quality of the intermediate state—while the pay-halfway contract mainly relies on the former, the pay-halfway contract may be more effective at incentivizing desirable initial actions.

## 7 CONCLUSION

This paper investigates the role and the impact of information during the delegation process by introducing a novel two-stage contract model with multiple intermediate states. To leverage the information revealed during the process, we propose two conceptually distinct contract types: the pay-halfway contract, which allows transfers based on both final outcomes and selected intermediate states; and the terminate-halfway contract, which permits early termination of the delegation process at certain intermediate states. We conduct a comprehensive analysis of these contracts by comparing their performance to that of the optimal standard contract. This comparison highlights the relative efficiency and effectiveness of the newly proposed contracts that incorporate intermediate information, revealing their respective strengths and limitations. It thereby offers deeper insights into how each contract type exploits information and the practical implications for contract design.

Through this investigation, our work advances the theory of contract design by illustrating how intermediate information can be leveraged to improve performance in delegation settings.

## REFERENCES

- [1] Omer Ben-Porat, Yishay Mansour, Michal Moshkovitz, and Boaz Taitler. 2024. Principal-agent reward shaping in MDPs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 9502–9510.
- [2] Matteo Bollini, Francesco Bacchiocchi, Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. 2024. Contracting with a reinforcement learning agent by playing trick or treat. *arXiv preprint arXiv:2410.13520* (2024).
- [3] Patrick Bolton and Mathias Dewatripont. 2004. *Contract theory*. MIT press.
- [4] Bernard Caillaud and Benjamin E. Hermalin. 2000. Hidden-information agency. *Unpublished manuscript* (2000).
- [5] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. 2022. Designing menus of contracts efficiently: The power of randomization. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 705–735.
- [6] Paul Dütting, Michal Feldman, and Yoav Gal Tzur. 2024. Combinatorial contracts beyond gross substitutes. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 92–108.
- [7] Paul Dütting, Tim Roughgarden, and Inbal Talgam-Cohen. 2019. Simple versus optimal contracts. In *Proceedings of the 2019 ACM Conference on Economics and Computation*. 369–387.
- [8] Drew Fudenberg, Bengt Holmstrom, and Paul Milgrom. 1990. Short-term contracts and long-term agency relationships. *Journal of economic theory* 51, 1 (1990), 1–31.
- [9] Sanford J Grossman and Oliver D Hart. 1992. An analysis of the principal-agent problem. In *Foundations of Insurance Economics: Readings in Economics and Finance*. Springer, 302–340.
- [10] Guru Guruganesh, Jon Schneider, Joshua Wang, and Junyao Zhao. 2023. The power of menus in contract design. In *Proceedings of the 24th ACM Conference on Economics and Computation*. 818–848.
- [11] Guru Guruganesh, Jon Schneider, and Joshua R Wang. 2021. Contracts under moral hazard and adverse selection. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. 563–582.
- [12] Bengt Holmström. 1979. Moral hazard and observability. *The Bell journal of economics* (1979), 74–91.
- [13] Jibang Wu, Siyu Chen, Mengdi Wang, Huazheng Wang, and Haifeng Xu. 2024. Contractual reinforcement learning: Pulling arms with invisible hands. *arXiv preprint arXiv:2407.01458* (2024).
- [14] Guanghui Yu and Chien-Ju Ho. 2022. Environment Design for Biased Decision Makers.. In *IJCAI*. 592–598.
- [15] Hanrui Zhang, Yu Cheng, and Vincent Conitzer. 2022. Efficient algorithms for planning with participation constraints. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 1121–1140.
- [16] Hanrui Zhang, Yu Cheng, and Vincent Conitzer. 2022. Planning with participation constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 5260–5267.
- [17] Hanrui Zhang and Vincent Conitzer. 2021. Automated dynamic mechanism design. *Advances in Neural Information Processing Systems* 34 (2021), 27785–27797.
- [18] Haoqi Zhang and David C Parkes. 2008. Value-Based Policy Teaching with Active Indirect Elicitation.. In *AAAI*, Vol. 8. 208–214.
- [19] Haoqi Zhang, David C Parkes, and Yiling Chen. 2009. Policy teaching through reward function learning. In *Proceedings of the 10th ACM conference on Electronic commerce*. 295–304.
- [20] Hao Zhang and Stefanos Zenios. 2008. A dynamic principal-agent model with hidden information: Sequential optimality through truthful state revelation. *Operations Research* 56, 3 (2008), 681–696.