

Robust Contracts: A Revealed Preference Approach*

Nemanja Antic and George Georgiadis[†]

December 30, 2022

Abstract

We study an agency model in which the principal has outcome data under different incentive schemes and aims to design an optimal contract under minimal assumptions about the way the agent responds to incentives. In particular, the principal knows the agent-optimal actions, which are distributions over outcomes, in response to K “known” contracts but is unaware of other actions available, and importantly, of their costs. The principal seeks a contract that maximizes worst-case profits. The optimal contract is a mixture of the known contracts and the (linear) one that makes the agent residual claimant. Moreover, when $K = 1$, the single known contract maximizes the principal’s profit guarantee, whereas with two known contracts, the optimal mixture puts strictly positive weight on one of the known contracts. Our methodology is straightforward to implement, a point that we demonstrate using data from DellaVigna and Pope’s (2018) experimental study of different incentive schemes.

1 Introduction

Firms and organizations throughout the economy use performance pay to motivate their employees (WorldatWork and Deloitte, 2014). Proper design of incentive schemes however is crucial: when Safelite Autoglass switched from hourly wages to piece rates for their key workers, productivity increased by 44% year-to-year (Lazear, 2000). On the other hand, poorly designed incentives can have dire, sometimes even catastrophic consequences (Jensen, 2002 and Rajan, 2011).

*We are grateful to Alessandro Bonatti, Peter Klibanoff, Michael Powell, to participants at various seminars and conferences for helpful comments, and to David Almog Salama Finkelstein for excellent research assistance.

[†]N. Antic: Northwestern University, nemanja.antic@kellogg.northwestern.edu; G. Georgiadis: Northwestern University, g-georgiadis@kellogg.northwestern.edu

To introduce the main ideas and motivate some of our modeling choices, let us consider an example. Imagine that you run a car dealership and want to design a new incentive plan for your salespeople. To simplify matters, suppose you have settled on rewarding salespeople according to monthly sales, and all that remains to decide is the pay-for-performance relationship. One approach you could take is to adopt industry best practices (see, for example, Zoltners, Sinha and Lorimer, 2006). You could also take guidance from contract theory. The typical approach involves making assumptions about the production environment—the employees’ action set, how actions map into outcomes, and their preferences over money and actions, and then exploiting variation in the offered incentives to recover the unknown parameters (Misra and Nair, 2011 and Georgiadis and Powell, 2022). Some managers may be uncomfortable making such arguably strong assumptions, perhaps due to a lack of information about the production environment. In this paper, we characterize optimal incentives given outcome data from a set of incentive schemes, but otherwise minimal assumptions about the production environment.

In our principal-agent model, events unfold as follows: First, the principal offers a contract, which specifies a non-negative payment to the agent as a function of realized output. Then the agent chooses a costly action—a probability distribution over output—to maximize his expected payoff. Finally, output is drawn according to the chosen distribution, and payoffs are realized. The principal has outcome data under K different contracts which, sidestepping estimation error, enables her to recover the action corresponding to each of these contracts. We assume that the agent best-responds to the offered contract and has quasi-linear preferences over money and actions, but we make no further assumptions about the production environment. The principal does not have prior beliefs about any of the unknown aspects of the environment. Instead, she seeks a contract that maximizes worst-case profit.¹

We begin with the benchmark case in which there is a single known contract (i.e., $K = 1$). In this case, we show that the known contract provides the largest possible profit guarantee in this case. To see why, suppose the agent has only two possible actions to choose from: the known action which is known to be “productive” and a completely unproductive one. If the principal offers a contract that pays more in expectation under the productive action, the principal’s profit clearly decreases relative to the known contract. But, if the principal offers a contract that pays less, then the costs of the productive action may be so high that the agent now chooses the unproductive action, harming the principal again.

¹Our model is conceptually similar to Carroll (2015). The key difference is that the principal does not observe the agent’s cost of choosing *any* action. Instead, she observes the actions chosen in response to the K “known” contracts, which give rise to a set of revealed preference constraints.

We then turn to the case with two known contracts. This case is of particular interest considering that firms are notoriously reluctant to experiment with different incentive schemes, and the majority of studies that exploit variation in incentives feature outcome data from two contracts; see for example Lazear (2018) and the references therein. We show that under certain conditions, a mixture of *one* of the known contracts and the linear one that makes the agent residual claimant maximizes the principal’s profit guarantee and is therefore optimal. Furthermore, if both known contracts are linear, then these conditions are never met, in which case the more profitable of the known contracts is optimal.

With more than three known contracts, we can show that the optimal contract is a convex combination of the known contracts and the linear one that makes the agent residual claimant. In addition, we propose an optimization procedure to obtain the optimal contract numerically.

To demonstrate the applicability of our methodology, we use data from DellaVigna and Pope’s 2018 large-scale experimental study of how different incentive schemes motivate subjects in a real-effort task. For every subset of the seven treatments in which subjects were motivated solely by financial incentives, we take that subset to constitute the set of “known” contracts. We then use the outcome data from each treatment in that subset to compute the empirical distribution function, which corresponds to the agent’s optimal action under that treatment. Next, we assume a value for the principal’s marginal value of output and compute the optimal contract. In each of the 127 subsets and all values of the marginal value of output that we consider, we find that the most profitable of the known contracts provides the largest profit guarantee; i.e., a mixture contract is never optimal. These findings might explain why firms are reluctant to experiment with different incentive schemes: without additional assumptions about the production environment, which managers may be hesitant to make, it is often impossible to find a new contract with a bigger profit guarantee. Thus, it is a case of “better the devil you know”. Our results imply that one would expect to see path-dependence in the contract design process, where otherwise similar firms (or divisions within the same firm) may settle on substantially different incentive plans.

While we make minimal assumptions about the production environment, we do assume that the agent has quasilinear preferences over money and actions. Moreover, because in practice, outcome data is obtained from multiple agents, our framework implicitly assumes that agents are homogeneous.² Naturally, if one were to relax the model by also requiring robustness to unobserved heterogeneity and the agents’ preferences, this would make it

²For example, in the data that we use, each subject participated in a single treatment once. We use the output data from all subjects who participated in each treatment to estimate the corresponding output distribution, thus assuming that all participants chose the same action.

even more likely that the most profitable of the known contracts delivers the largest profit guarantee.

Related Literature. Our paper builds on the literature that studies principal-agent problems under moral hazard pioneered by Holmström (1979) and Mirrlees (1976). In particular, we contribute to the strands of this literature that have sought to relax the knowledge assumptions in the canonical model; see Georgiadis (2022) for an overview. One strand studies models in which the principal is oblivious to one or more parameters and designs a mechanism to elicit this information from the agent (Chade and Swinkels, 2022; Chade et al., 2022; Castro-Pires and Moreira, 2021; Gottlieb and Moreira, 2021).

Another segment of the literature focuses on structural models in which several aspects of the environment are unknown to the principal, who then estimates them using field or experimental data. For example, Misra and Nair (2011) estimates a structural dynamic contracting model and uses it to evaluate counterfactual contracts for the salesforce of a large contact lens manufacturer; see also Chung, Steenburgh and Sudhir (2014) and Chung, Kim and Park (2021) for more elaborate structural models. Georgiadis and Powell (2022) establishes conditions under which the information provided by an A/B test of incentive contracts is sufficient for determining an optimal contract in the classic Holmström model.

The strand which our model is closest to takes the stance that the principal is ambiguity-averse and pursues “robust” contracts that provide the largest profit guarantee. This is as if an adversarial third-party chooses the unknown aspects of the environment after observing the offered contract to minimize the principal’s profit. Carroll (2015) shows that if the principal knows only a subset of the actions available to the agent and their costs, then a linear contract is optimal. Walton and Carroll (2022) provide more general conditions for linear contracts to be optimal. In a setting where the principal also knows a “lower bound” distribution, Antic (2022) shows that optimal contracts are mixtures of debt and equity. See also Carroll (2019) for an overview. Instead, in our model the principal does not know the costs of any of the agent’s actions but can partially infer them from the agent’s revealed preferences. In a concurrent paper, Burkett and Rosenthal (2022) consider, in effect, the same problem.³ While some of our results are similar (e.g., our [Theorem 3](#) and [Corollary 3](#) parallel their Propositions 3 and 5, respectively), they focus on conditions on the underlying data under which the optimal contract is a mixture of *one* of the known contracts and a linear one. Instead, we fully characterize the optimal contract when $K \in \{1, 2\}$, we propose an optimization algorithm to obtain it numerically when $K \geq 3$, and we demonstrate how our methodology can be applied using data from a real-effort experiment.

³See also Rosenthal (2019) for a conceptually similar approach applied to a screening problem.

2 Model

We consider a contractual relationship between a principal (she) and an agent (he) with the following timing: First, the principal designs a contract w , which is an upper-semicontinuous mapping from the set of feasible outputs $\mathcal{X} = [0, \bar{x}]$ to non-negative payments to the agent, $w : \mathcal{X} \rightarrow \mathbb{R}_+$. Then the agent chooses an action F , which is a probability distribution supported on \mathcal{X} , by paying a private cost $C(F)$. Lastly, output $x \sim F$ is drawn and payoffs are realized.

Let $\mathcal{F} \subseteq \Delta(\mathcal{X})$ denote the agent’s action set. The principal does not have full knowledge of this set. However, she knows that it includes a costless action F_0 that generates zero output with certainty, and knows the action that the agent has chosen in response to each of K “known” contracts. We denote these contracts by w_1, \dots, w_K , and the respective actions by F_1, \dots, F_K . Importantly, the principal does not know the cost associated with each of these actions. However, she knows that the agent is rational, and chooses a payoff-maximizing action. An interpretation is that the principal has observational data from having offered each of these contracts enabling her to compute the respective output distribution chosen by the agent.⁴

The agent is risk-neutral, has outside option 0, and is cash constrained. Therefore, any feasible contract must specify non-negative payments. The principal has linear preferences over money and given her knowledge, evaluates each contract according to its worst-case profit. Specifically, this worst-case profit when she offers contract w equals

$$\begin{aligned} \Pi(w) &:= \inf_{\mathcal{F}, C} \int [mx - w(x)] dF(x) \\ \text{s.t. } F &\in \arg \max_{\tilde{F} \in \mathcal{F}} \int w(x) d\tilde{F}(x) - C(\tilde{F}) \end{aligned} \tag{IC}$$

$$F_k \in \arg \max_{\tilde{F} \in \mathcal{F}} \int w_k(x) d\tilde{F}(x) - C(\tilde{F}) \text{ for all } k \in \{1, \dots, K\} \tag{RP}$$

$$\mathcal{F} \supseteq \{F_0, \dots, F_K, F\} \text{ and } C(F) \geq 0 \text{ for all } F \in \mathcal{F} \text{ with } C(F_0) = 0,$$

where m is the principal’s gross profit per unit of output.⁵ This is as-if after the principal offers a contract, an adversarial third-party—*nature*—chooses the agent’s action set and the cost of each action to minimize the principal’s profit subject to (IC), which specifies the agent’s best response to the offered contract, w , and a set of revealed preference constraints

⁴We assume that these contracts are exogenous and that the agent narrowly best-responds to the offered contract without any strategic considerations.

⁵All integrals are evaluated from $x = 0$ to $x = \bar{x}$. We omit these limits for notational simplicity. When convenient, we will also omit the argument of functions and write, for example, w instead of $w(x)$.

given in (RP), which impose that each F_k is a best response to w_k . Naturally the action set \mathcal{F} must include F , as well as the known actions F_0, \dots, F_K , and costs must be nonnegative. The principal's objective is to find a contract that maximizes her worst-case profit:

$$\Pi^* = \sup_{w \geq 0} \Pi(w). \quad (\text{P})$$

A contract is *optimal* if it gives profit guarantee Π^* to the principal.

Finally, we impose three assumptions on the K known contract-action pairs:

- (A.1) At least one of the contracts delivers strictly positive profit to the principal. We adopt the convention that w_1 delivers the largest profit.
- (A.2) Each contract has $w_k(x) \geq 0$ for all $x \in \mathcal{X}$ and $w_k(0) = 0$.
- (A.3) The agent's best responses can be rationalized; i.e.,

$$\int w_k(x) dF_k(x) + \int w_j(x) dF_j(x) - \int w_k(x) dF_j(x) - \int w_j(x) dF_k(x) \geq 0 \quad \forall j, k.$$

Assumption A.1 ensures that the principal does not prefer to walk away. The first part of A.2 states that the known contracts respect limited liability. The second part ensures that none of the known contracts can be trivially improved by a downward shift until the agent's limited liability constraint binds. Assumption A.3 is necessary for the problem to be feasible; if it fails for some j and k , then no action costs can simultaneously rationalize the agent choosing F_j over F_k when contract w_j is offered, and choosing F_k over F_j when w_k is offered.

2.1 The Principal's Problem Simplified

In this subsection, we show that the principal's problem is equivalent to the following simpler, more tractable formulation (due to nature's problem being a linear program):

$$\begin{aligned} \sup_{w_{K+1}} \inf_{F_{K+1}, \mathbf{c}} \int [mx - w_{K+1}(x)] dF_{K+1}(x) & \quad (\text{P}') \\ \text{s.t.} \quad \int w_k(x) dF_k(x) - c_k \geq \int w_k(x) dF_j(x) - c_j & \quad \text{for all } k \text{ and } j \neq k \quad (\text{IC-RP}) \\ w_{K+1}(\cdot) \geq 0, F_{K+1} \in \Delta(\mathcal{X}), \text{ and } \mathbf{c} \in \mathbb{R}_+^{K+1}, & \end{aligned}$$

where $k \in \{1, \dots, K+1\}$ and $j \in \{0, \dots, K+1\}$. In this formulation, for each contract w_{K+1} , instead of choosing the agent's action set and the cost of each action, nature chooses *one* action, F_{K+1} and the vector $\mathbf{c} = \{c_1, \dots, c_{K+1}\}$, where c_k is the cost of action F_k , to minimize the principal's profit subject to a set of incentive compatibility and revealed

preference constraints, which stipulate that F_k is a best response to w_k for each k . Then the principal chooses w_{K+1} to maximize this worst-case profit.

Lemma 1. *A contract w_{K+1} solves (P) if and only if it solves (P').*

Towards a contradiction, suppose that the action set contains at least two actions beyond the known ones. Since the agent can choose at most one of them in response to the offered contract, nature is no worse off by excluding the additional actions from \mathcal{F} . Adding extra actions on the other hand, increases the number of revealed preference constraints, which can only benefit the principal. \square

By contrast, in Carroll (2015) the principal knows a subset of the agent's actions, as well as the costs of those actions. So while the objective function in his model is identical, the costs are known and (IC-RP) applies only to $k = K + 1$, that is, there is an incentive compatibility constraint but no revealed preference constraints.

3 Results

In this section we establish our main results. We start off with the case in which there is *one* known contract (i.e., $K = 1$), and show that continuing to offer the same contract maximizes the principal's worst-case profit. Next we characterize the optimal contract for the case with two known contracts. Finally, we consider the case with an arbitrary number of known contracts.

3.1 A Benchmark: One known contract ($K = 1$)

If the principal knows only the agent's best response to a single contract, then she can do no better than continue to offer that same contract.

Theorem 1. *With one known contract, w_1 , the principal's worst-case profit is maximized when she offers w_1 . That is, any contract which solves (P') is F_1 -a.e. equivalent to w_1 .*

For a sketch of the argument, fix an arbitrary contract $w_2 \neq w_1$. If $\int w_1 dF_1 > \int w_2 dF_1$, then nature can induce the agent to choose the null action F_0 by endowing him with no additional actions (e.g., by setting $F_2 \equiv F_0$) and making action F_1 sufficiently costly. Since F_0 results in non-positive profit, the principal prefers to offer w_1 . If the inequality is reversed, then nature can induce the agent to choose F_1 in response to w (by endowing him with no additional actions), in which case the principal is again better off offering w_1 . Finally, if

the inequality binds, then nature can ensure that the principal earns a vanishingly small worst-case profit, which is strictly smaller than the strictly positive one provided by w_1 .

This result contrasts with much of the robust contracting literature, which shows that linear contracts are optimal (Carroll, 2015; Dai and Toikka, 2022; Walton and Carroll, 2022). The key difference is that this literature assumes that a subset of the agent’s actions and their costs are known, and so the first part of the above sketch breaks down: if $\int (w_1 - w)dF_1$ is sufficiently small, then nature may not be able to induce the agent to choose F_0 .

The revealed preference constraints serve to bound the costs of the known actions. However, if the principal knows the agent’s optimal action in response to only one contract, then this bound, $\int w_1 dF_1$, is too weak to enable her to improve upon the known contract. Tighter bounds may arise in several ways. For example, if the agent is known to earn sufficiently large rents under the known contract, then as the following corollary shows, a linear contract is optimal—in line with most of the robust contracting literature (e.g., Carroll, 2015).

Corollary 1. *Assume there is one known contract, w_1 , and the principal knows that c_1 is no larger than \hat{c} . If \hat{c} is smaller than some threshold \bar{c} , then the linear contract with slope $\sqrt{m\hat{c} / \int x dF_1(x)}$ maximizes the principal’s worst-case profit. Otherwise w_1 is optimal.*

Unlike in the argument for [Theorem 1](#), because the cost of F_1 is bounded, the principal can take away some of the agent’s rents while dissuading him from switching to F_0 . Of course, there are other actions that nature could endow the agent with. Linear contracts ensure that actions which are appealing to the agent are not very harmful to the principal per the standard intuition; see, for example, Carroll (2015).

Another way in which tighter bounds on costs can arise is by having outcome data under multiple contracts, in which case the revealed preference constraints place further restrictions on what costs can be assigned to each action. We turn to this case next.

3.2 Two known contracts ($K = 2$)

In this section we suppose that the principal knows the agent-optimal action in response to each of *two* contracts. This case is empirically relevant because most studies that examine the effects of incentives exploit variation from exactly two incentive schemes; see for example Lazear (2018) and the references therein.

To simplify the exposition we introduce some notation. For each i and j , define

$$v_{ij} := \int w_i(x)dF_j(x) \quad \text{and} \quad \mu_j := \int x dF_j(x)$$

to denote the expected payment under w_i if the agent chooses action F_j , and the expected output under this action, respectively. Next, we define

$$\phi := v_{11} + v_{22} - v_{12} - v_{21}.$$

This quantity is non-negative per Assumption A.3, and it relates to the “wobble room” nature has to hurt the principal by varying c_1 and c_2 while respecting the agent’s revealed preference constraints.⁶ Finally, for each j and $i \neq j$, and conditional on $m\mu_i - v_{ji} \geq \phi$, define the contract

$$w_j^*(x) := \rho_j w_j(x) + (1 - \rho_j)mx, \text{ where } \rho_j := 1 - \sqrt{\phi / (m\mu_i - v_{ji})},$$

which is a mixture of w_j and the linear contract that makes the agent residual claimant, mx . The following theorem shows that under certain conditions, one of these mixture contracts is optimal; otherwise w_1 , the more profitable of the known contracts is optimal.

Theorem 2. *Suppose the principal knows the contract-action pairs (w_1, F_1) and (w_2, F_2) .*

- (i). *If $\sqrt{m\mu_2 - v_{12}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$, then w_1^* is optimal;*
- (ii). *If $\sqrt{m\mu_1 - v_{21}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$, then w_2^* is optimal;*
- (iii). *Otherwise, the more profitable of the known contracts, w_1 , is optimal.*

These conditions are mutually exclusive. The left-hand side of the first and the second condition is the square root of the principal’s profit when she offers w_1^* and w_2^* , respectively, while the right-hand sides are the square root of her profit when she offers w_1 .

To interpret condition (i), suppose that the principal could offer w_1 get the agent to choose F_2 instead of F_1 ; she would benefit if $m\mu_2 - v_{12} > m\mu_1 - v_{11}$. Of course, the principal cannot achieve this aim simply by offering w_1 , because it violates one of the revealed preference constraints. Instead, she must appropriately modify incentives, and ϕ relates to the profit she must give up to do so. The interpretation of condition (ii) is analogous.

Note that conditions (i) and (ii) are easier to satisfy when ϕ is small; in this case however, the optimal contract assigns little weight on mx , so it similar to w_1 or w_2 , respectively.

Which action does the agent choose in response to the optimal contract? If w_1 is optimal, then of course, the agent chooses F_1 . The following corollary characterizes the additional action with which nature endows the agent when the principal optimally offers w_1^* or w_2^* .

⁶The revealed preference constraints (IC-RP) stipulate $v_{11} - c_1 \geq v_{12} - c_2$ and $v_{22} - c_2 \geq v_{21} - c_1$, which can be rewritten as $v_{21} - v_{22} \leq c_1 - c_2 \leq v_{11} - v_{12}$. If $\phi = 0$, then $c_1 - c_2$ is pinned down by these constraints, and the larger ϕ is, the more flexibly nature can choose the costs of the known actions.

Corollary 2. *Suppose the principal optimally offers w_j^* . In response, the agent chooses*

$$F_j^*(x) = \rho_j F_i(x) + (1 - \rho_j) F_0(x),$$

where $i \neq j$, and moreover, the cost of this action is (weakly) smaller than $\rho_j c_i$.

That is, in response to w_1^* , nature endows the agent with an action that is that is a mixture of F_2 and F_0 with respective weights ρ_1 and $1 - \rho_1$ (and analogously for w_2^*). In response, the agent chooses this action at a cost that is no larger than the convex combination of the actions in the mixture.

Linear contracts, which come in various forms including fixed wages, equity, piece rates, and commissions, are among the most common incentive schemes. However, if both known contracts are linear, then conditions (i) and (ii) of [Theorem 2](#) are never satisfied, and so w_1 is optimal.

Corollary 3. *Suppose that both known contracts, w_1 and w_2 , are linear. Then w_1 is optimal.*

Observe that if both known contracts are linear, by [Theorem 2](#), any third contract will also be linear. To see that a third contract cannot improve the principal's profit guarantee, consider nature endowing the agent only with the observed actions, namely, F_0 , F_1 and F_2 . To fix ideas, assume that w_1 offers higher equity, and we have already assumed that it generates higher profit than w_2 . It follows that the new contract offers strictly more equity than w_2 . If the new contract offers more equity than w_1 , then any costs satisfying the revealed preference constraints will result in the agent choosing action F_1 , which is strictly worse for the principal than offering w_1 . For an equity level in-between w_1 and w_2 , nature can choose the costs so that the agent chooses F_2 under the new contract; in this case the new contract is less attractive than w_2 and hence also w_1 .

This result suggests that for an ambiguity-averse principal, experimenting only with linear contracts may be counterproductive; instead, it is valuable to also have outcome data under nonlinear contracts.

3.3 K known contracts

In this section we extend our analysis to an arbitrary number of known contracts. Characterizing the optimal contract in this case is challenging because it involves solving a non-convex optimization program. Nevertheless, we can show that it is a convex combination of the known contracts and $m.x$. Moreover, we propose an optimization procedure to solve for the optimal contract, which we use in our empirical exercise in [Section 4](#).

Consider the following maximization program, which is the dual of (P'):

$$\begin{aligned}
& \sup_{\boldsymbol{\lambda}} \frac{\sum_{j=1}^K \lambda_{K+1,j} \left(m\mu_j + \sum_{k=1}^K \lambda_{k,K+1} v_{kj} \right)}{1 + \sum_{j=0}^K \lambda_{K+1,j}} - \sum_{k=1}^K (\lambda_{k,K+1} + \lambda_{k0}) v_{kk} + \sum_{k=1}^K \sum_{j=1}^K \lambda_{kj} (v_{kj} - v_{kk}) \\
& \text{s.t. } \lambda_{k,K+1} + \lambda_{k0} + \sum_{j=1}^K (\lambda_{kj} - \lambda_{jk}) \geq \lambda_{K+1,k} \quad \text{for all } k \in \{1, \dots, K\} \tag{D'} \\
& \sum_{k=1}^K \lambda_{k,K+1} \leq \sum_{j=0}^K \lambda_{K+1,j} \\
& \boldsymbol{\lambda} \in \mathbb{R}_+^{(K+1) \times (K+2)}
\end{aligned}$$

Each λ_{kj} represents the Lagrange multiplier associated with (IC-RP $_{kj}$), which stipulates that when offered contract w_k , the agent prefers action F_k to F_j . The following theorem shows that, first, every optimal contract is a convex combination of the K known contracts and the (linear) one that makes the agent residual claimant, and second, the principal's problem is equivalent to (D').

Theorem 3. *Given K known contracts, a contract w_{K+1} is optimal if and only if it solves (D'). Moreover, every optimal contract takes the form*

$$w_{K+1}(x) := \sum_{k=1}^K \rho_k w_k(x) + \left(1 - \sum_{k=1}^K \rho_k \right) mx,$$

where $\rho_k = \lambda_{k,K+1} / (1 + \sum_{j=0}^K \lambda_{K+1,j}) \geq 0$ for each k .

Observe that (D') is non-convex owing to the first term in the objective. As a result, standard optimization methods are generally not guaranteed to yield a global maximum. Towards a practical procedure to solve this program, notice that if we fix the multipliers $\{\lambda_{K+1,j}\}_{j=0}^K =: \boldsymbol{\lambda}^{K+1}$, then (D') reduces to a linear program, which can be solved exactly using standard solvers. Denote the objective evaluated at the optimum of this linear program by $\tilde{\Pi}(\boldsymbol{\lambda}^{K+1})$. Then it remains to solve

$$\sup \tilde{\Pi}(\boldsymbol{\lambda}^{K+1}) \text{ subject to } \boldsymbol{\lambda}^{K+1} \in \mathbb{R}_+^{K+1}.$$

While this program is also not convex, its dimension is $K + 1$, whereas (D') has dimension $(K + 1)^2$. Practically, it can be solved relatively swiftly, for example, using a simulated

annealing algorithm.^{7,8}

4 Application

In this section we demonstrate the applicability of our methodology using data from DellaVigna and Pope’s 2018 real-effort experiment conducted on Amazon’s Mechanical Turk. In the experiment, subjects were tasked with repeatedly pressing the ‘a’ and ‘b’ keys in alternating order, and received one *point* for every a/b keystroke pair they managed to complete in a ten-minute period. We focus on the 7 treatments summarized in Table 1, which differ in the monetary incentives offered. Each subject was randomly assigned to a single treatment, received a \$1 participation fee, and performed the task once. In the first treatment, no incentive pay was offered. In treatments 2 to 5, subjects were paid a constant amount per point, whereas in treatments 6 and 7 they received a lump-sum payment (40 and 80 cents, respectively) if they achieved at least 2,000 points.⁹ During the course of the treatment, subjects could see the incentive contract they were on, a countdown clock, as well as a running tally of the points accumulated. The dataset includes the number of points achieved by every subject.

We now describe the exercise that we perform. First, we make an assumption about the principal’s gross profit margin m . Then, for each subset of treatments $\mathcal{W} \subseteq \{\pi^1, \dots, \pi^7\}$, we take it to constitute the set of “known” contracts, and letting K denote the cardinality of \mathcal{W} , we define the K known contracts w_1, \dots, w_K . Next, we use the outcome data from each treatment in this set to compute the corresponding empirical CDF F_k , which we take to be the agents’ best response to w_k .¹⁰ Finally, we compute the optimal contract.

First, we consider all *pairs* of treatments, that is, all sets \mathcal{W} with cardinality 2 (of which

⁷In light of [Theorem 2](#) one might ask whether the optimal contract always puts positive weight on *one* of the known contracts. It turns out that there do exist instances where the optimal contract puts weight on multiple known contracts as [Example 2](#) of [Burkett and Rosenthal \(2022\)](#) demonstrates.

⁸Observe that fixing $\sum_{j=0}^K \lambda_{K+1,j}$, the problem reduces to a quadratic program. In particular, the constraints are linear, and the objective comprises linear terms as well as a quadratic term of the form $\lambda^T Q \lambda$, where λ is the vectorization of the matrix with entries $\lambda_{k,j}$ and Q is a symmetric matrix. Thus an alternative approach is to solve this quadratic program for each $\sum_{j=0}^K \lambda_{K+1,j}$, and then maximize with respect to this sum. In general, Q is not negative-definite, in which case the quadratic program is non-convex (although algorithms that exploit the special structure of quadratic programs to obtain approximate global maxima exist).

⁹To be precise, in treatment 2, they were paid 1 cent per thousand points, and in treatments 3, 4 and 5, they were paid 1, 4, and 10 cents, respectively, for every hundred points. For simplicity, we take x to lie on the unit grid; i.e., $x \in \mathbb{N}$.

¹⁰In doing so, we abstract away from statistical error and we ignore unobserved heterogeneity. We discuss these issues in [Section 5](#). We also define the K contracts in such an order that w_1 generates the largest profit in line with [Assumption A.1](#).

| Incentive Contract | Avg. #points | Std. Dev. | #Subjects |
|---|--------------|-----------|-----------|
| $\pi^1(x) = 0$ | 1521 | 726 | 540 |
| $\pi^2(x) = 0.001x$ | 1883 | 664 | 538 |
| $\pi^3(x) = 0.01x$ | 2029 | 649 | 558 |
| $\pi^4(x) = 0.04x$ | 2132 | 626 | 562 |
| $\pi^5(x) = 0.10x$ | 2175 | 578 | 566 |
| $\pi^6(x) = 40\mathbb{I}_{\{x \geq 2000\}}$ | 2136 | 576 | 545 |
| $\pi^7(x) = 80\mathbb{I}_{\{x \geq 2000\}}$ | 2188 | 530 | 532 |

Table 1: This table describes seven of the experimental treatments in DellaVigna and Pope (2018) that differed in the monetary incentives offered to the subjects. The first column describes the implied incentive contract denominated in cents, where x is the points accumulated. The remaining columns describe, for each treatment, the average number of points accumulated, the standard deviation, and the number of subjects.

there are 21).¹¹ For each pair of treatments and every m between 0.05 and 1 with a grid size of 0.001, we check which of the conditions in Theorem 2 are satisfied to identify the optimal contract. In each of these combinations, the more profitable of the known contracts delivers the largest profit guarantee.¹²

Next, we consider all sets of treatments with cardinality greater than two. To find the optimal contract, we solve (D') using a simulated annealing algorithm.¹³ We repeat this procedure for each set of three or more treatments (of which there are 99) and $m = \{0.05, 0.1, \dots, 1\}$. Again, in every combination, the more profitable of the known contracts is optimal.

5 Discussion

We study an agency model under moral hazard in which the principal faces ambiguity about the actions available to the agent and their costs. The principal has outcome data under K “known” incentive schemes and seeks a contract with the largest profit guarantee. We show that if $K = 1$, then the single known contract is optimal. With two known contracts, a mixture of *one* of the known contracts and the linear contract that makes the agent residual claimant is optimal. If $K \geq 3$, then we show that the optimal contract is a convex

¹¹If $K = 1$, then by Theorem 1 the single known contract provides the largest profit guarantee.

¹²That this is true when both known contracts are linear follows from Corollary 3. Therefore, it suffices to check only the pairs in which one (or both) of the known contracts is π^6 or π^7 . Note that our focusing on $m \geq 0.05$ is to ensure that at least one of the contracts in \mathcal{W} is profitable per assumption A.1.

¹³Simulated annealing is a method for approximating the optimal solution of non-convex optimization programs, where gradient descent algorithms may get “trapped” at a local maximum.

combination of the known ones and the aforementioned linear contract, and propose an algorithm to obtain it numerically. Finally, we demonstrate the applicability of our approach using data from DellaVigna and Pope’s (2018) experimental study of a variety of incentive schemes.

Unobserved heterogeneity. In practice (and in our empirical exercise), one may aggregate outcome data from many agents who are offered the same contract. While our model treats agents as homogeneous, this is likely only a simplifying assumption. In reality, faced with the same contract, different agents may choose different actions and bear different costs. In that case, the empirical distribution function computed using the outcome data from any given contract is a composition of each agent’s (unobserved) action. Such unobserved heterogeneity increases nature’s *leverage*, making it only more likely that one of the known contracts is optimal.

Risk-aversion. We have assumed that the principal knows the agent’s preferences over monetary payments, and in particular, that the agent is risk-neutral. Our model can readily be extended to the case in which the agent’s payoff, for any given wage scheme and action, is $\int u(w(x))dF(x) - C(F)$, where u is a known, strictly increasing, concave function. In this case, the optimal contract is a now nonlinear function of a set of dual multipliers and the known contracts. Of course it is plausible that the principal is oblivious to the agent’s utility function, and requiring that the optimal contract be robust to this type of ambiguity, again, makes it more likely that one of the known contracts is optimal.

Cost restrictions. Our model differs from the robust contracting literature in that the principal does not know the agent’s private cost of taking each action. We place no restrictions on these costs other than those implied by revealed preference. It may be interesting to incorporate restrictions on costs while stopping short of assuming that the cost function is known or can be estimated using outcome data.¹⁴

Estimation error. We have assumed that for each known contract, the principal can identify the agent’s action, that is, the corresponding distribution function over output. In practice of course, outcome data is finite, which gives rise to estimation error. We conjecture that incorporating estimation error would make it more likely that a new contract provides a larger profit “guarantee” (appropriately defined), though a formal analysis is left for future research.

¹⁴For example, one could posit that the cost function $C(F)$ is monotone in first-order stochastic dominance (e.g., Georgiadis, Ravid and Szentes, 2022), or that it comes from parametric family (e.g., f-divergence as in Hébert, 2018) and the principal faces ambiguity over a set of parameters.

References

- Antic, Nemanja.** 2022. “Contracting with unknown technologies.” *Working Paper*.
- Burkett, Justin, and Maxwell Rosenthal.** 2022. “Data-Driven Contract Design.” *Working Paper*.
- Carroll, Gabriel.** 2015. “Robustness and linear contracts.” *American Economic Review*, 105(2): 536–63.
- Carroll, Gabriel.** 2019. “Robustness in mechanism design and contracting.” *Annual Review of Economics*, 11(1): 139–166.
- Castro-Pires, Henrique, and Humberto Moreira.** 2021. “Limited liability and non-responsiveness in agency models.” *Games and Economic Behavior*, 128: 73–103.
- Chade, Hector, and Jeroen Swinkels.** 2022. “Disentangling moral hazard and adverse selection.” Working Paper.
- Chade, Hector, Victoria Marone, Amanda Starc, and Jeroen Swinkels.** 2022. “Multidimensional Screening and Menu Design in Health Insurance Markets.” *NBER Working Paper*.
- Chung, Doug J, Byungyeon Kim, and Byoung G Park.** 2021. “The comprehensive effects of sales force management: A dynamic structural analysis of selection, compensation, and training.” *Management Science*, 67(11): 7046–7074.
- Chung, Doug J, Thomas Steenburgh, and K Sudhir.** 2014. “Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans.” *Marketing Science*, 33(2): 165–187.
- Dai, Tianjiao, and Juuso Toikka.** 2022. “Robust incentives for teams.” *Econometrica*, 90(4): 1583–1613.
- DellaVigna, Stefano, and Devin Pope.** 2018. “What motivates effort? Evidence and expert forecasts.” *The Review of Economic Studies*, 85(2): 1029–1069.
- Georgiadis, George.** 2022. “Contracting with Moral Hazard: A Review of Theory & Empirics.”
- Georgiadis, George, and Michael Powell.** 2022. “A/B contracts.” *American Economic Review*, 112(1): 267–303.

- Georgiadis, George, Doron Ravid, and Balazs Szentes.** 2022. “Flexible Moral Hazard Problems.” *Working Paper*.
- Gottlieb, Daniel, and Humberto Moreira.** 2021. “Simple contracts with adverse selection and moral hazard.” *Theoretical Economics*.
- Hébert, Benjamin.** 2018. “Moral Hazard and the Optimality of Debt.” *The Review of Economic Studies*, 85(4): 2214–2252.
- Holmström, Bengt.** 1979. “Moral hazard and observability.” *The Bell Journal of Economics*, 74–91.
- Jensen, Michael C.** 2002. “Corporate budgeting is broken, let’s fix it.”
- Lazear, Edward P.** 2000. “Performance pay and productivity.” *American Economic Review*, 90(5): 1346–1361.
- Lazear, Edward P.** 2018. “Compensation and incentives in the workplace.” *Journal of Economic Perspectives*, 32(3): 195–214.
- Luenberger, David G.** 1997. *Optimization by vector space methods*. John Wiley & Sons.
- Mirrlees, James A.** 1976. “The Optimal Structure of Incentives and Authority within an Organization.” *The Bell Journal of Economics*, 105–131.
- Misra, Sanjog, and Harikesh S Nair.** 2011. “A structural model of sales-force compensation dynamics: Estimation and field implementation.” *Quantitative Marketing and Economics*, 9(3): 211–257.
- Rajan, Raghuram G.** 2011. “Fault Lines.” In *Fault Lines*. Princeton University Press.
- Rosenthal, Maxwell.** 2019. “A Revealed Preference Approach to Multidimensional Screening.” *Working Paper*.
- Walton, Daniel, and Gabriel Carroll.** 2022. “A General Framework for Robust Contracting Models.” *Working Paper*.
- WorldatWork, and Consulting LLP Deloitte.** 2014. “Incentive Pay Practices Survey: Publicly Traded Companies.”
- Zoltners, Andris A, Prabhakant Sinha, and Sally E Lorimer.** 2006. *The complete guide to sales force incentive compensation: How to design and implement plans that work*. Amacom Books.

A Appendix

A.1 Proof of Theorem 1

Fix an arbitrary contract $w_2 \geq 0$ that satisfies (IC-RP).

Case 1: Suppose $\int w_1(x)dF_1(x) > \int w_2(x)dF_1(x)$. Let nature choose $c_1 = \int w_1(x)dF_1(x)$, $F_2 = F_0$ and $c_2 = 0$. We have

$$\int w_2(x)dF_2(x) - c_2 = w_2(0) = 0 > \int w_2(x)dF_1(x) - c_1,$$

where the first inequality is due to w_2 satisfying the agent's limited liability constraint, and the second inequality follows from the definition of c_1 . That is, F_1 is a best response for the agent when w_1 is offered, and F_2 is a best response for the agent when w_2 is offered. The principal's payoff in this case is $\int [mx - w_2(x)]dF_2(x) = -w_2(0) \leq 0 < \int [mx - w_1(x)]dF_1(x)$, where the second inequality follows from the assumption that the principal earns a positive payoff under w_1 .

Case 2: Suppose $\int w_1(x)dF_1(x) < \int w_2(x)dF_1(x)$. Let nature choose $c_1 = \int w_1(x)dF_1(x)$, $F_2 \equiv F_1$ and $c_2 = c_1$. Then F_1 and F_2 is a best response for the agent when w_1 and w_2 is offered, respectively, and the principal's payoff is

$$\int [mx - w_2(x)]dF_2(x) = \int [mx - w_2(x)]dF_1(x) < \int [mx - w_1(x)]dF_1(x),$$

that is, she is better off offering w_1 .

Case 3: Suppose $\int w_1(x)dF_1(x) = \int w_2(x)dF_1(x)$ and $w_1 \neq w_2$ F_1 -a.e. There exists an $\hat{x} \in \text{supp}(F_1)$ such that $w_1(\hat{x}) < w_2(\hat{x})$.

We distinguish 3 sub-cases. First, suppose that $w_2(0) > 0$. Then let nature pick $F_2 \equiv F_0$, $[\int w_1 dF_1 - w_2(0)]^+ < c_1 < \int w_1 dF_1$, and $c_2 = 0$. It is easy to verify that the agent strictly prefers F_1 in response to w_1 , and strictly prefers F_2 in response to w_2 . However, the principal's payoff when she offers w_2 is $-w_2(x) < 0$. Hence she prefers to offer w_1 , which provides a strictly positive payoff by assumption.

Second, suppose that $w_2(0) = 0$ and $m\hat{x} \leq w_2(\hat{x})$. Then let nature pick $F_2(x) = \mathbb{I}_{x \geq \hat{x}}$, $c_1 = \int w_1 dF_1$, and $w_1(\hat{x}) < c_2 < w_2(\hat{x})$. It is easy to verify that the agent strictly prefers F_1 in response to w_1 , and strictly prefers F_2 in response to w_2 . However, the principal's payoff when she offers w_2 is $m\hat{x} - w_2(\hat{x}) < 0$; hence she prefers to offer w_1 .

Finally suppose that $w_2(0) = 0$ and $m\hat{x} > w_2(\hat{x})$. Pick $0 < \epsilon < \int [mx - w_1(x)]dF_1(x) / [m\hat{x} - w_2(\hat{x})]$, and let nature choose $c_1 = \int w_1 dF_1 - \epsilon[w_1(\hat{x}) + w_2(\hat{x})]/2$, $c_2 = 0$, and $F_2(x) =$

$1 - \epsilon + \epsilon \mathbb{I}_{x \geq \hat{x}}$. That is, F_2 is costless and puts mass $1 - \epsilon$ on $x = 0$ and mass ϵ on \hat{x} . Then

$$\int w_1(x) dF_1(x) - c_1 = \epsilon[w_1(\hat{x}) + w_2(\hat{x})]/2 > \epsilon w_1(\hat{x}) = \int w_1(x) dF_2(x) - c_2, \text{ and}$$

$$\int w_2(x) dF_2(x) - c_2 \geq \epsilon w_2(\hat{x}) > \epsilon[w_1(\hat{x}) + w_2(\hat{x})]/2 = \int w_2(x) dF_1(x) - c_1;$$

i.e., F_1 and F_2 is a best response to w_1 and w_2 , respectively. However, the principal's payoff is $\int [mx - w_2(x)] dF_2(x) = \epsilon[m\hat{x} - w_2(\hat{x})] < \int [mx - w_1(x)] dF_1(x)$, so she strictly prefers to offer w_1 . \square

A.2 Proof of Corollary 1

Assume there is one known contract, and moreover, the principal knows that the cost of F_1 is no larger than $\hat{c} \leq \int w_1(x) dF_1(x)$. Her problem can be expressed as follows:

$$\begin{aligned} \sup_{w_2 \geq 0} \inf_{F_2, c_1, c_2} & \int [mx - w_2(x)] dF_2(x) & (1) \\ \text{s.t.} & \int w_2(x) dF_2(x) - c_2 \geq \int w_2(x) dF_1(x) - c_1 \\ & \int w_1(x) dF_1(x) - c_1 \geq \int w_1(x) dF_2(x) - c_2 \\ & \int w_2(x) dF_2(x) - c_2 \geq \int w_2(x) dF_0(x) \\ & c_1 \leq \hat{c}, c_2 \geq 0, \text{ and } F_2 \in \Delta(\mathcal{X}) \end{aligned}$$

Because $\hat{c} \leq \int w_1 dF_1$ by assumption, the constraint stipulating that the agent prefers F_1 vis-à-vis the null action F_0 when w_1 is offered is redundant. Notice that the right-hand side of the third constraint equals $w_2(0)$, which is nonnegative. We shall guess (and verify later) that this is equal to zero in an optimal solution, and replace this constraint with $\int w_2 dF_2 - c_2 \geq 0$.

Fix a contract w_2 . We have the Lagrangian

$$\begin{aligned}
L(\boldsymbol{\lambda}, w_2) &= \inf \int [mx - (1 + \lambda_1 + \lambda_3)w_2(x) + \lambda_2 w_1(x)] dF_2(x) \\
&\quad + \int [\lambda_1 w_2 - \lambda_2 w_1] dF_1 - (\lambda_1 - \lambda_2)c_1 - (\lambda_2 - \lambda_1 - \lambda_3)c_2 \\
&\text{s.t. } F_2 \in \Delta(\mathcal{X}), 0 \leq c_1 \leq \widehat{c} \text{ and } c_2 \geq 0 \\
&= \min \{mx - (1 + \lambda_1 + \lambda_3)w_2(x) + \lambda_2 w_1(x)\} \\
&\quad + \int (\lambda_1 w_2 - \lambda_2 w_1) dF_1 - [\lambda_1 - \lambda_2]^+ \widehat{c} - \begin{cases} 0 & \text{if } \lambda_2 \leq \lambda_1 + \lambda_3 \\ \infty & \text{if } \lambda_2 > \lambda_1 + \lambda_3, \end{cases}
\end{aligned}$$

where $\lambda_1, \lambda_2, \lambda_3 \geq 0$ are the dual multipliers corresponding to the first, second, third constraint of (1), respectively. To derive the fourth line we used that the integral in the first line is maximized by the degenerate distribution, F_2 , which places all mass at $\hat{x} \in \min_x \{mx - (1 + \lambda_1 + \lambda_3)w_2(x) + \lambda_2 w_1(x)\}$.

Notice that holding w_2 fixed, nature's problem is linear. Hence by the Lagrange Duality Theorem (Luenberger, 1997, Theorem 1, p. 224) strong duality holds, and so the solution to (1) equals

$$\begin{aligned}
&\sup_{w_2 \geq 0, \boldsymbol{\lambda} \geq 0} \min \{mx - (1 + \lambda_1 + \lambda_3)w_2(x) + \lambda_2 w_1(x)\} & (2) \\
&\quad + \int [\lambda_1 w_2(x) - \lambda_2 w_1(x)] dF_1(x) - [\lambda_1 - \lambda_2]^+ \widehat{c} \\
&\text{s.t. } \lambda_2 \leq \lambda_1 + \lambda_3,
\end{aligned}$$

where the constraint follows from the observation that if $\lambda_2 \not\leq \lambda_1 + \lambda_3$, then the objective equals $-\infty$, which cannot be optimal. We shall maximize (2) first with respect to w_2 and then with respect to $\boldsymbol{\lambda}$. Observe that for any given $\boldsymbol{\lambda}$, the optimal contract must be such that the term inside the curly brackets in (2) is constant in x , that is,

$$w_2(x) = \frac{mx + \gamma + \lambda_2 w_1(x)}{1 + \lambda_1 + \lambda_3}$$

for some $\gamma \geq 0$. Since this constant shifts the contract by the same amount for all x , it does not affect the agent's incentives, and so it is optimal to set $\Gamma = 0$. Using this expression we can rewrite (2) as

$$\sup_{\boldsymbol{\lambda} \geq 0} \frac{\lambda_1 m \mu_1 - (1 + \lambda_3) \lambda_2 v_{11}}{1 + \lambda_1 + \lambda_3} - [\lambda_1 - \lambda_2]^+ \widehat{c} \text{ s.t. } \lambda_1 + \lambda_3 \geq \lambda_2 \geq 0,$$

where $\mu_1 := \int x dF_1(x)$ and $v_{11} := \int w_1 dF_1$.

For any fixed $\lambda_1, \lambda_3 \geq 0$, the objective increases in λ_2 at rate

$$\begin{aligned} & -\frac{1 + \lambda_3}{1 + \lambda_1 + \lambda_3} v_{11} + \widehat{c} \text{ if } \lambda_2 \leq \lambda_1, \text{ and} \\ & -\frac{1 + \lambda_3}{1 + \lambda_1 + \lambda_3} v_{11} < 0 \text{ if } \lambda_2 > \lambda_1, \end{aligned}$$

which implies that the objective is maximized by setting

$$\lambda_2 = \begin{cases} 0 & \text{if } (1 + \lambda_1 + \lambda_3)\widehat{c} < (1 + \lambda_3)v_{11}, \text{ and} \\ \lambda_1 & \text{if } (1 + \lambda_1 + \lambda_3)\widehat{c} > (1 + \lambda_3)v_{11}. \end{cases} \quad (3)$$

Because $\lambda_2 \leq \lambda_1$, the constraint in (2) is satisfied for all $\lambda_3 \geq 0$, and since the objective decreases in λ_3 , it is optimal to set $\lambda_3 = 0$.

Using (3) and $\lambda_3 = 0$, we can write the objective solely as a function of λ_1 as

$$\sup_{\lambda_1 \geq 0} \frac{\lambda_1 m \mu_1}{1 + \lambda_1} - \lambda_1 \widehat{c} + \lambda_1 \left[\widehat{c} - \frac{v_{11}}{1 + \lambda_1} \right]^+.$$

This objective increases in λ_1 at rate

$$\begin{aligned} & m\mu_1/(1 + \lambda_1)^2 - \widehat{c} \text{ for } \lambda_1 < v_{11}/\widehat{c} - 1, \text{ and} \\ & (m\mu_1 - v_{11})/(1 + \lambda_1)^2 > 0 \text{ for } \lambda_1 \geq v_{11}/\widehat{c} - 1. \end{aligned}$$

That is, the objective is initially concave, peaking at $\sqrt{m\mu_1/\widehat{c}} - 1$ (provided this is smaller than $v_{11}/\widehat{c} - 1$), it has a kink at $\lambda_1 = v_{11}/\widehat{c} - 1$, and is then increasing. We thus have two candidates for the optimal value of λ_1 : $\sqrt{m\mu_1/\widehat{c}} - 1$ and ∞ . The principal's objective evaluated at the first and second candidate is $(\sqrt{m\mu_1} - \sqrt{\widehat{c}})^2$ and $(m\mu_1 - v_{11})$, respectively, and comparing the two yields that the optimal

$$\lambda_1 = \begin{cases} \sqrt{m\mu_1/\widehat{c}} - 1 & \text{if } \widehat{c} > \bar{c} := 2m\mu_1 - \sqrt{2m\mu_1 - v_{11}}, \text{ and} \\ \infty & \text{otherwise.} \end{cases}$$

In the former case, $\lambda_2 = 0$ and so $w_2(x) \equiv \sqrt{m\widehat{c}/\mu_1} x$ is optimal. In the latter case, $\lambda_2 = \lambda_1$ and so $w_2(x) \equiv w_1(x)$ is optimal. \square

A.3 Proof of Theorem 2

We begin by establishing the following lemma.

Lemma 2. *Suppose the principal knows the pairs (w_1, F_1) and (w_2, F_2) . If a contract other than w_1 is optimal, then it is either w_1^* or w_2^* .*

Proof of Lemma 2. Towards a contradiction, suppose a contract $w_3 \geq 0$ with $w_3(0) = 0$ is optimal.¹⁵ We will show that a contract in the set $\{w_1, w_1^*, w_2^*\}$ gives the principal a higher payoff than w_3 .

By adding and subtracting $(v_{31} - v_{32})$ to ϕ , we can write

$$\phi = [v_{31} + v_{22} - v_{32} - v_{21}] + [v_{32} + v_{11} - v_{31} - v_{12}] \geq 0. \quad (4)$$

Because this sum is nonnegative (per Assumption A.3), at least one of the terms in the square brackets must be nonnegative. Without loss of generality we can label $i, j \in \{1, 2\}$ with $i \neq j$ such that

$$v_{3i} + v_{jj} - v_{3j} - v_{ji} \geq 0. \quad (5)$$

Claim 1. *If the principal's payoff under w_3 is strictly larger than it is under w_1 , then $m\mu_i - v_{3i} > m\mu_1 - v_{11}$ and hence $v_{3i} < v_{ii}$.*

Proof of Claim 1. To see why, suppose towards a contradiction that $v_{3i} \geq v_{ii}$. Let nature choose $F_3 \equiv F_i$, $c_3 = c_i$, and the values of c_i and c_j so that F_3 (or equivalently F_i) is preferred by the agent over F_j when contract w_3 is offered.¹⁶ But then $v_{3i} \geq v_{ii}$ implies that $m\mu_i - v_{3i} \leq m\mu_i - v_{ii} \leq m\mu_1 - v_{11}$, where the last inequality is due to the convention that w_1 delivers a larger payoff to the principal than w_2 . That is, the principal's payoff is weakly larger if she offers w_1 , contradicting the premise that w_3 is optimal. Since we are looking for a contract that increases the principal's payoff vis-a-vis w_1 , we shall henceforth assume that w_3 satisfies $m\mu_i - v_{3i} > m\mu_1 - v_{11}$ and hence $v_{3i} < v_{ii}$. \square

Next, we consider three cases, which distinguish whether $v_{3j} + v_{ii} - v_{3i} - v_{ij}$ is positive or negative.

Case 1(a). Suppose $v_{3j} + v_{ii} - v_{3i} - v_{ij} \geq 0$ and $v_{3j} \geq v_{jj}$

¹⁵The restriction that $w_3(0) = 0$ is without loss of generality per Theorem 3, which shows that for any K , an optimal contract is a mixture of the K known ones, and $m\alpha$. Since each of the known contracts satisfies $w_k(0) = 0$, so does the optimal one.

¹⁶To be specific, nature picks c_i and c_j such that $v_{ii} - c_i \geq v_{ij} - c_j$, $v_{jj} - c_j \geq v_{ji} - c_i$, and $v_{3i} - c_i \geq v_{3j} - c_j$. Such c_i and c_j exist as long as $v_{ii} - v_{ij} \geq v_{ji} - v_{jj}$ and $v_{3i} - v_{3j} \geq v_{ji} - v_{jj}$. These conditions hold since $\phi \geq 0$ and (5), respectively.

In this case, nature will choose $F_3 \equiv F_j$, $c_3 = c_j$, and the values of c_i and c_j such that F_3 (or equivalently F_j) is preferred by the agent over F_i when offered contract w_3 .¹⁷ But then $v_{3j} \geq v_{jj}$ implies that $m\mu_j - v_{3j} \leq m\mu_j - v_{jj}$, contradicting again the premise that w_3 is optimal.

Case 1(b). Suppose $v_{3j} + v_{ii} - v_{3i} - v_{ij} \geq 0$ and $v_{3j} < v_{jj}$

Let nature choose $F_3 \equiv F_0$, $c_3 = 0$, $c_i = v_{3i}$, and $c_j = v_{3j}$. The agent's incentive compatibility and revealed preference constraints are

$$v_{ii} - c_i = v_{ii} - v_{3i} \geq \begin{cases} v_{ij} - v_{3j} = v_{ij} - c_j \\ v_{i3} - c_3 = 0 \end{cases}$$

$$v_{jj} - c_j = v_{jj} - v_{3j} \geq \begin{cases} v_{ji} - v_{3i} = v_{ji} - c_i \\ v_{j3} - c_3 = 0, \end{cases}$$

and $v_{33} \geq v_{3i} - c_i = v_{3j} - c_j = 0$. The first constraint is satisfied per the first condition of case 1(b) and the choice of c_j . The second constraint follows from the first part of [Claim 1](#) and the choice of c_i , the third constraint follows from (5), and the fourth constraint follows from the second condition of case 1(b). Finally, the last constraint is satisfied since $w_3(0) = 0$ by assumption. So when the principal offers contract w_3 , the agent optimally chooses action F_3 (or equivalently F_0), and the principal's payoff is $-w_3(0) = 0 < m\mu_1 - v_{11}$, where the second inequality follows from the assumption that the principal obtains a positive payoff under w_1 . This inequality chain contradicts again the premise that w_3 is optimal.

Case 2. Suppose $v_{3j} + v_{ii} - v_{3i} - v_{ij} < 0$

Let nature choose an action F_3 and a corresponding cost c_3 , both of which remain to be determined, $c_i = v_{3i} - v_{33} + c_3$, and a c_j that also remains to be determined.¹⁸ This choice for c_i ensures that $v_{33} - c_3 = v_{3i} - c_i$; that is, when contract w_3 is offered, the agent weakly prefers F_3 over F_i (in fact he is indifferent). The agent's remaining incentive compatibility

¹⁷To be specific, nature picks c_i and c_j such that $v_{ii} - c_i \geq v_{ij} - c_j$, $v_{jj} - c_j \geq v_{ji} - c_i$, and $v_{3j} - c_j \geq v_{3i} - c_i$. Such c_i and c_j exist as long as $v_{ii} - v_{ij} \geq v_{ji} - v_{jj}$ and $v_{ii} - v_{ij} \geq v_{3i} - v_{3j}$. These conditions hold since $\phi \geq 0$ and per the first condition in case 1(a), respectively.

¹⁸Note that c_3 will have to be greater than $v_{33} - v_{3i}$ to ensure that $c_i \geq 0$.

and revealed preference constraints are

$$\begin{aligned}
v_{ii} - c_i = v_{ii} - v_{3i} + v_{33} - c_3 \geq v_{ij} - c_j &\Leftrightarrow c_j \geq v_{ij} - v_{ii} + v_{3i} - v_{33} + c_3, \\
v_{ii} - v_{3i} + v_{33} - c_3 \geq v_{i3} - c_3 &\Leftrightarrow v_{ii} - v_{3i} + v_{33} - v_{i3} \geq 0, \\
v_{ii} - v_{3i} + v_{33} - c_3 \geq 0 &\Leftrightarrow c_3 \leq v_{ii} + v_{33} - v_{3i}, \\
v_{jj} - c_j \geq v_{ji} - v_{3i} + v_{33} - c_3 &\Leftrightarrow c_j \leq v_{jj} - v_{ji} + v_{3i} - v_{33} + c_3, \\
v_{jj} - c_j \geq v_{j3} - c_3 &\Leftrightarrow c_j \leq v_{jj} - v_{j3} + c_3, \\
v_{jj} - c_j \geq 0 &\Leftrightarrow c_j \leq v_{jj}, \\
v_{33} - c_3 \geq v_{3j} - c_j &\Leftrightarrow c_j \geq v_{3j} - v_{33} + c_3, \text{ and} \\
v_{33} - c_3 \geq v_{30} &\Leftrightarrow c_3 \leq v_{33} - v_{30}.
\end{aligned} \tag{6}$$

For any c_3 , it follows from the first, fourth, fifth and sixth constraint that a c_j satisfying the above constraints exists as long as

$$\begin{aligned}
v_{jj} - v_{ji} + v_{3i} - v_{33} &\geq v_{3j} - v_{33}, \\
v_{jj} - v_{ji} + v_{3i} - v_{33} &\geq v_{ij} - v_{ii} + v_{3i} - v_{33}, \\
v_{jj} - v_{j3} &\geq v_{3j} - v_{33}, \\
v_{jj} - v_{j3} &\geq v_{ij} - v_{ii} + v_{3i} - v_{33}, \\
v_{jj} &\geq v_{3j} - v_{33} + c_3, \text{ and} \\
v_{jj} &\geq v_{ij} - v_{ii} + v_{3i} - v_{33} + c_3.
\end{aligned} \tag{7}$$

The first inequality is satisfied per (5) and the second is satisfied since $\phi \geq 0$. Note also that the right-hand side of the third inequality is smaller than the right-hand side of the fourth inequality, that is, $v_{3j} - v_{33} < v_{ij} - v_{ii} + v_{3i} - v_{33}$ by the condition of case 2. It remains to check that w_3 satisfies (6) and (7). Moreover, c_3 must satisfy

$$v_{33} - v_{3i} \leq c_3 \leq \min \{v_{jj} - v_{3j} + v_{33}, v_{jj} - v_{ij} + v_{ii} - v_{3i} + v_{33}, v_{ii} - v_{3i} + v_{33}, v_{33} - v_{30}\}.$$

Using (5) and that $\phi \geq 0$, $v_{ji} \geq 0$, and $v_{ii} \geq 0$, it is easy to verify that for an appropriate c_3 to exist, it suffices that $v_{3i} \geq v_{30}$.

So far, we have shown that if w_3 increases the principal's payoff (relative to w_1), then (5), (6), (7), $v_{3i} \geq v_{30}$, and the conditions of Claim 1 and Case 2 must hold. Next, we will argue that if there exists a contract that dominates w_1 , then w_1^* or w_2^* maximizes the principal's

payoff. To see this, consider

$$\sup_{w \geq 0} \inf_{F \in \Delta(\mathcal{X})} \int [mx - w(x)] dF(x) \tag{P''}$$

$$\text{s.t. } \int [w(x) - w_j(x)] dF(x) \geq \int [w(x) - w_i(x)] dF_i(x) + \int [w_i(x) - w_j(x)] dF_j(x) \tag{8}$$

$$\int [w(x) - w_i(x)] dF(x) \geq \int [w(x) - w_i(x)] dF_i(x) \tag{9}$$

That is, we consider the principal's max-min problem subject to the constraints (6) and (7), where we have replaced w_3 with the choice variable w . We will check ex-post that the optimal w satisfies [Claim 1](#) and the condition of Case 2, which are necessary for this contract to increase the principal's payoff vis-a-vis w_1 . We will show that whenever the solution to (P'') satisfies these conditions, then it coincides with either w_1^* or w_2^* almost everywhere. Finally, we will remark that this contract also satisfies $\int w(x) dF_i(x) \geq \int w(x) dF_0(x)$, which is the counterpart of $v_{3i} \geq v_{30}$.

Claim 2. *If w satisfies [Claim 1](#), then the right-hand side of (8) is strictly positive and the right-hand side of (9) is strictly negative.*

Proof of Claim 2. The second part of this claim is immediate from the [Claim 1](#), part 1. For the first part of the claim, notice that if the right-hand side of (8) is negative, nature can choose $F = F_0$, in which case the principal's payoff will be no greater than zero. But then the principal would be better off offering contract w_1 , which provides her with a strictly positive payoff by assumption. Since the contract w will be relevant only if it increases the principal's payoff vis-a-vis w_1 , we can henceforth assume that the right-hand side of (8) is strictly positive and the right-hand side of (9) is strictly negative. \square

Let us fix an arbitrary $w \geq 0$ and nonnegative dual multipliers λ and ν . We have the

Lagrangian

$$\begin{aligned}
\mathcal{L}(\lambda, \nu, w) &= \inf_{F \in \Delta(\mathcal{X})} \int [mx - (1 + \lambda + \nu)w(x) + \lambda w_j(x) + \nu w_i(x)] dF(x) \\
&\quad + \lambda \int [w(x) - w_i(x)] dF_i(x) + \lambda \int [w_i(x) - w_j(x)] dF_j(x) \\
&\quad + \nu \int [w(x) - w_i(x)] dF_i(x) \\
&= \min_x \{mx - (1 + \lambda + \nu)w(x) + \lambda w_j(x) + \nu w_i(x)\} + (\lambda + \nu) \int w(x) dF_i(x) \\
&\quad - \lambda \left[\int w_i(x) dF_i(x) - w_i(x) dF_j(x) + w_j(x) dF_j(x) \right] - \nu \int w_i(x) dF_i(x).
\end{aligned}$$

The first integral is minimized by a degenerate distribution F . By the Lagrange Duality Theorem (Luenberger, 1997, Theorem 1, p. 224) strong duality holds, and therefore, the solution to (P^{''}) equals

$$\sup_{w \geq 0} \sup_{\lambda, \nu \geq 0} \mathcal{L}(\lambda, \nu, w). \quad (10)$$

Changing the order of maximization, we fix arbitrary multipliers $\lambda, \nu \geq 0$ and consider $\sup_{w \geq 0} \mathcal{L}(\lambda, \nu, w)$. For each x , a marginal increase in $w(x)$ increases the objective at rate $-(1 + \lambda + \nu) + (\lambda + \nu)dF_i(x) < 0$ if the expression inside the curly brackets is minimized at that particular x , and at rate $(\lambda + \nu)dF_i(x) > 0$ otherwise. Therefore, the expression inside the curly brackets must be constant (in x) and hence the Lagrangian-maximizing contract $w(x)$ satisfies

$$w(x) = \frac{(mx - \gamma) + \lambda w_j(x) + \nu w_i(x)}{1 + \lambda + \nu} \quad (11)$$

for some constant γ . Observe that increasing γ shifts the contract downwards without affecting the agent's incentive constraints, thereby increasing the principal's payoff. Since $w_i(0) = w_j(0) = 0$ by assumption, it is optimal to set $\gamma = 0$, which is the largest value that respects the agent's limited liability constraint.

Substituting the expression for $w(x)$ in (11) into the Lagrangian yields

$$L(\lambda, \nu) := \sup_{w \geq 0} \mathcal{L}(\lambda, \nu, w) = \frac{\lambda + \nu}{1 + \lambda + \nu} (m\mu_i + \lambda v_{ji} + \nu v_{ii}) - \lambda(v_{ii} - v_{ij} + v_{jj}) - \nu v_{ii}.$$

Differentiating $L(\lambda, \nu)$ with respect to each of its arguments yields

$$\frac{dL(\lambda, \nu)}{d\lambda} = \frac{m\mu_i - v_{ji} + \nu(v_{ii} - v_{ji})}{(1 + \lambda + \nu)^2} - \phi \quad \text{and} \quad \frac{dL(\lambda, \nu)}{d\nu} = \frac{m\mu_i + \lambda v_{ji} - (1 + \lambda)v_{ii}}{(1 + \lambda + \nu)^2}. \quad (12)$$

Although the first-order conditions need not be sufficient for a maximum (if the problem is

not concave), they are necessary. We now establish the following claim.

Claim 3. *A contract solves (10) and it (strictly) dominates w_1 only if it is w_1^* or w_2^* .*

Proof of Claim 3. Observe that $dL(\lambda, \nu)/d\nu \leq 0$ if and only if $\lambda > (m\mu_i - v_{ii})/(v_{ii} - v_{ji})$, and moreover if w dominates w_1 , then it must be the case that $v_{ii} - v_{ji} > 0$. To see why the last inequality is true, note that $v_{ji} \leq v_{ii} + v_{ji} - v_{ij} \leq \int w(x)dF_i(x) < v_{ii}$, where the first inequality follows from the fact that $\phi \geq 0$, the second inequality because the right-hand side of (8) is strictly positive (as argued above), and the last inequality follows from Claim 1.

It follows from the first-order conditions in (12) that one of the following pairs (λ, ν) maximizes $L(\lambda, \nu)$:

- i. $\lambda = 0$ and $\nu = \infty$,
- ii. $\lambda = (m\mu_i - v_{ii})/(v_{ii} - v_{ji})$ and $\nu = (v_{ii} - v_{ji})/\phi - (m\mu_i - v_{ii})/(v_{ii} - v_{ji})$, provided $(v_{ii} - v_{ji})^2 > \phi(m\mu_i - v_{ii})$, or
- iii. $\lambda = \sqrt{(m\mu_i - v_{ji})/\phi} - 1$ and $\nu = 0$, provided $\sqrt{(m\mu_i - v_{ji})/\phi} - 1 > (m\mu_i - v_{ii})/(v_{ii} - v_{ji})$.

Under the first pair of multipliers, the corresponding contract is w_i , which of course cannot (strictly) payoff-dominate w_1 . Recall that if w payoff-dominates w_1 , then per Claim 1, it must satisfy $\int w(x)dF_i(x) < v_{ii}$. Substituting the second pair (and $\gamma = 0$) into (11) yields $\int w(x)dF_i(x) = v_{ii}$, which violates the above condition. Next, substituting the third pair of multipliers into (11) yields the contract w_j^* and we have $\int w_j^*(x)dF_i(x) < v_{ii}$ (so that Claim 1 may be satisfied) if and only if $\lambda > (m\mu_i - v_{ii})/(v_{ii} - v_{ji})$. Moreover, this contract (trivially) satisfies $\int w_j^*(x)dF_i(x) \geq \int w_j^*(x)dF_0(x) = 0$, which is the counterpart of $v_{3i} \geq v_{30}$ when we replace w_3 with w_j^* . This completes the proof of Claim 3. \square

To conclude, we have shown that if a contract different from w_1 maximizes the principal's payoff, then this contract is either w_1^* or w_2^* . This completes the proof of Lemma 2. \square

We are now ready to prove the proposition. We have shown that if there exists a contract that payoff-dominates w_1 , then the payoff-maximizing contract is

$$w_j^*(x) := \rho_j w_j(x) + (1 - \rho_j) mx, \text{ where } \rho_j := 1 - \sqrt{\frac{\phi}{m\mu_i - v_{ji}}} \text{ for some } j \in \{1, 2\}.$$

By substituting the optimal multipliers ($\lambda = \sqrt{(m\mu_i - v_{ji})/\phi} - 1$ and $\nu = 0$) into the Lagrangian and using that strong duality holds, we have that the principal's payoff when

she offers w_j^* equals¹⁹

$$\Pi(w_j^*) = (\sqrt{m\mu_i - v_{ji}} - \sqrt{\phi})^2. \quad (13)$$

If she offers w_1 instead, her payoff $\Pi(w_1) = m\mu_1 - v_{11}$. So $\Pi(w_j^*) > \Pi(w_1)$ only if $\sqrt{m\mu_i - v_{ji}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$ as claimed.

It remains to show that the conditions in parts (i) and (ii) of the proposition are mutually exclusive. Towards a contradiction, suppose that $\sqrt{m\mu_1 - v_{21}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$ and $\sqrt{m\mu_2 - v_{12}} - \sqrt{\phi} > \sqrt{m\mu_1 - v_{11}}$ ($\geq \sqrt{m\mu_2 - v_{22}}$). We can rewrite these conditions as

$$\begin{aligned} m\mu_1 - v_{21} &> m\mu_1 - v_{11} + v_{11} + v_{22} - v_{12} - v_{21} + 2\sqrt{\phi(m\mu_1 - v_{11})}, \text{ and} \\ m\mu_2 - v_{12} &> m\mu_2 - v_{22} + v_{11} + v_{22} - v_{12} - v_{21} + 2\sqrt{\phi(m\mu_1 - v_{11})}, \end{aligned}$$

respectively, where we substituted $\phi = v_{11} + v_{22} - v_{12} - v_{21}$. Summing these inequalities yields $\phi + 4\sqrt{\phi(m\mu_1 - v_{11})} < 0$, which is a contradiction since both terms on the left-hand side are positive. \square

A.4 Proof of Corollary 2

Suppose that for some $j \in \{1, 2\}$, condition (j) of [Theorem 2](#) is satisfied. It suffices to show that $\{w_j^*, F_j^*\}$ satisfies (8) and (9) and the principal's objective attains its maximum (which is the square of the expression given in the left-hand side of condition (j) of [Theorem 2](#)). It is straightforward to verify that $\{w_j^*, F_j^*\}$ satisfies (8) with equality, using that $\phi = (1 - \rho_j)^2(m\mu_i - v_{ji})$ by the definition of ρ_j and that $[mx - w_j(x)]dF_0(x) = 0$ by Assumption A.2.

Similarly, by substituting $\{w_j^*, F_j^*\}$ into (8), it is straightforward to show that this constraint is slack using the facts that $\int [w_j^*(x) - w_i(x)]dF_i(x) < 0$ which follows from Claim 2 in the proof of [Theorem 2](#), and that $\int [w_j^*(x) - w_i(x)]dF_0(x) = 0$.

Next, substituting $\{w_j^*, F_j^*\}$ into the principal's objective yields

$$\int [mx - w_j^*(x)]dF_j^*(x) = \left(\sqrt{m\mu_i - v_{ji}} - \sqrt{\phi}\right)^2,$$

which is identical to the left-hand-side of condition (j) in [Theorem 2](#).

Finally, note that incentive compatibility requires that

$$\int w_j^*(x)dF_j^*(x) - C(F_j^*) \geq 0 \text{ and } \int w_j^*(x)dF_j^*(x) - C(F_j^*) \geq \int w_j^*(x)dF_i(x) - c_i.$$

By multiplying both sides of the first constraint by $(1 - \rho_j)$, both sides of the second constraint

¹⁹Recall that whenever w_j^* payoff-dominates w_1 , $m\mu_i - v_{ji} > \phi \geq 0$.

by ρ_j , and adding them, we obtain that $C(F_j^*) \leq \rho_j c_j$ as claimed. \square

A.5 Proof of Corollary 3

Suppose that $w_1(x) = \alpha_1 x$ and $w_2(x) = \alpha_2 x$ for some $\alpha_1, \alpha_2 \in [0, 1]$.²⁰ Then $v_{ik} = \alpha_i \mu_k$ for each $i, k \in \{1, 2\}$ and $\phi = (\alpha_1 - \alpha_2)(\mu_1 - \mu_2)$. Since w_1 payoff-dominates w_2 by convention and $\phi \geq 0$ by Assumption (A3), the parameters $(\alpha_1, \alpha_2, \mu_1, \mu_2)$ must satisfy

$$(m - \alpha_1)\mu_1 \geq (m - \alpha_2)\mu_2 \quad \text{and} \quad \phi = (\alpha_1 - \alpha_2)(\mu_1 - \mu_2) \geq 0. \quad (14)$$

From Theorem 2, w_j^* payoff-dominates w_1 if and only if $\sqrt{\mu_i(m - \alpha_j)} - \sqrt{\phi} > \sqrt{\mu_1(m - \alpha_1)}$. We will show that there does *not* exist a four-tuple $(\alpha_1, \alpha_2, \mu_1, \mu_2)$ such that the last three inequalities are satisfied for any $j \in \{1, 2\}$ and $i \neq j$.

Letting $i = 1$ and $j = 2$, w_2^* payoff-dominates w_1 if and only if

$$\begin{aligned} \sqrt{\mu_1(m - \alpha_2)} &> \sqrt{(\alpha_1 - \alpha_2)(\mu_1 - \mu_2)} + \sqrt{\mu_1(m - \alpha_1)} \\ \Leftrightarrow 1 &> \sqrt{\left(\frac{\alpha_1 - \alpha_2}{m - \alpha_2}\right) \left(1 - \frac{\mu_2}{\mu_1}\right)} + \sqrt{\frac{m - \alpha_1}{m - \alpha_2}} \geq \left|\frac{\alpha_1 - \alpha_2}{m - \alpha_2}\right| + \sqrt{\frac{m - \alpha_1}{m - \alpha_2}}, \end{aligned}$$

where the last inequality follows by rearranging the first part of (14). Letting $x := (\alpha_1 - \alpha_2)/(m - \alpha_2)$ and $y := (m - \alpha_1)/(m - \alpha_2)$, notice that $x + y = 1$, $x \leq 1$, and $y \geq 0$. If $x \geq 0$, then it must be the case that $y \leq 1$ and so $|x| + \sqrt{y} = x + \sqrt{y} \geq x + y = 1$. If instead $x < 0$, then it must be that $y > 1$ and so $|x| + \sqrt{y} \geq 1$. Therefore, the right-hand side of the above display equation is weakly greater than one, implying that w_2^* cannot payoff-dominate w_1 .

Letting $i = 2$ and $j = 1$, w_1^* payoff-dominates w_1 if and only if

$$\begin{aligned} \sqrt{\mu_2(m - \alpha_1)} &> \sqrt{(\alpha_1 - \alpha_2)(\mu_1 - \mu_2)} + \sqrt{\mu_1(m - \alpha_1)} \\ \Leftrightarrow 1 &> \sqrt{\left(\frac{\alpha_1 - \alpha_2}{m - \alpha_1}\right) \left(\frac{\mu_1}{\mu_2} - 1\right)} + \sqrt{\frac{\mu_1}{\mu_2}} \geq \left|\frac{\alpha_1 - \alpha_2}{m - \alpha_1}\right| + \sqrt{\frac{m - \alpha_2}{m - \alpha_1}}, \end{aligned}$$

where the last inequality follows again by rearranging the first part of (14). Letting $\omega := (\alpha_1 - \alpha_2)/(m - \alpha_1)$ and $z = (m - \alpha_2)/(m - \alpha_1)$, notice that $z - \omega = 1$, $\omega \leq 1$, and $z \geq 0$. If $\omega \geq 0$, then it must be that $z = 1 + \omega \geq 1$ and so $|\omega| + \sqrt{z} \geq 1$. Otherwise, it must be that $\omega \in [-1, 0)$ and $z = 1 + \omega \in [0, 1)$, and so $|\omega| + \sqrt{z} \geq |\omega| + z = -\omega + z = 1$. Therefore, the right-hand side of the above display equation is weakly greater than one, implying that w_1^*

²⁰Without loss of generality we can take the intercept to be zero. If the intercept were negative, the contract would violate the agent's limited liability constraint. If it were positive, the principal could trivially raise her payoff by setting the intercept to zero.

cannot payoff-dominate w_1 either.

Since neither w_1^* nor w_2^* dominates w_1 , it follows from [Theorem 2](#) (and in particular [Lemma 2](#)) that w_1 is optimal. \square

A.6 Proof of [Theorem 3](#)

Consider (P') for a given contract w_{K+1} . Fixing the dual multipliers $\lambda_{kj} \geq 0$ for each $k \in \{1, \dots, K+1\}$ and $j \in \{0, \dots, K+1\}$ such that $k \neq j$, the Lagrangian $L(\boldsymbol{\lambda}, w_{K+1})$ equals

$$\begin{aligned}
& \inf_{F_{K+1}, c_{K+1}, c_1, \dots, c_K} \int (mx - w_{K+1}) dF_{K+1} - \sum_{k=1}^{K+1} \sum_{j=0, \neq k}^{K+1} \lambda_{kj} \left(\int w_k dF_k - c_k - \int w_k dF_j + c_j \right) \\
= & \inf_{F_{K+1}, c_{K+1}, c_1, \dots, c_K} \int [mx - (1 + \sum_{j=0}^K \lambda_{K+1,j})w_{K+1} + \sum_{k=1}^K \lambda_{k,K+1}w_k] dF_{K+1} + \sum_{k=1}^{K+1} [\lambda_{k0} + \sum_{j=1, \neq k}^{K+1} (\lambda_{kj} - \lambda_{jk})] c_k \\
& + \sum_{j=0}^K \lambda_{K+1,j} \int w_{K+1} dF_j - \sum_{k=1}^K \lambda_{k,K+1} \int w_k dF_k - \sum_{k=1}^K \sum_{j=0, \neq k}^K \lambda_{kj} w_k (dF_k - dF_j) \\
= & \min_x \{ mx - (1 + \sum_{j=0}^K \lambda_{K+1,j})w_{K+1}(x) + \sum_{k=1}^K \lambda_{k,K+1}w_k(x) \} \\
& + \sum_{j=0}^K \lambda_{K+1,j} \int w_{K+1} dF_j - \sum_{k=1}^K \lambda_{k,K+1} \int w_k dF_k - \sum_{k=1}^K \sum_{j=0, \neq k}^K \lambda_{kj} w_k (dF_k - dF_j) \\
& + \sum_{k=1}^{K+1} \begin{cases} 0 & \text{if } \lambda_{k0} + \sum_{j=1, \neq k}^{K+1} (\lambda_{kj} - \lambda_{jk}) \geq 0 \\ -\infty & \text{otherwise,} \end{cases}
\end{aligned}$$

where the last line follows from the constraint that $c_k \geq 0$ for all $k \geq 1$, and we have used that $c_0 = 0$. Notice that if $\lambda_{k0} + \sum_{j=1, \neq k}^{K+1} (\lambda_{kj} - \lambda_{jk}) < 0$ for any k , then the Lagrangian will be equal to $-\infty$, which cannot be part of an optimal solution. Therefore, we have the constraint

$$\lambda_{k0} + \sum_{j=1, \neq k}^{K+1} (\lambda_{kj} - \lambda_{jk}) \geq 0 \text{ for each } k \in \{1, \dots, K+1\}. \quad (15)$$

By the Lagrange Duality Theorem ([Luenberger, 1997, Theorem 1, p. 224](#)), we have $\Pi(w_{K+1}) = \sup_{\boldsymbol{\lambda} \geq 0} L(\boldsymbol{\lambda}, w_{K+1})$, and so the principal's objective can be rewritten as

$$\sup \{ L(\boldsymbol{\lambda}, w_{K+1}) : \boldsymbol{\lambda} \geq 0 \text{ and } w_{K+1} \geq 0 \}.$$

Without loss of generality, we can change the order of maximization; that is, first we

maximize with respect to the contract w_{K+1} (while holding $\boldsymbol{\lambda}$ fixed), and then we maximize with respect to $\boldsymbol{\lambda}$. For each x , notice that a marginal increase in $w_{K+1}(x)$ raises $L(\boldsymbol{\lambda}, w_{K+1})$ at rate

$$\sum_{j=0}^K \lambda_{K+1,j} dF_j(x) > 0 \text{ if } x \notin \arg \min_{\tilde{x}} \left\{ m\tilde{x} - \left(1 + \sum_{j=0}^K \lambda_{K+1,j} \right) w_{K+1}(\tilde{x}) + \sum_{k=1}^K \lambda_{k,K+1} w_k(\tilde{x}) \right\}$$

$$\sum_{j=0}^K \lambda_{K+1,j} [dF_j(x) - 1] - 1 < 0 \text{ otherwise.}$$

It follows that for given multipliers $\boldsymbol{\lambda}$, the optimal contract is such that the expression in the brackets above is constant, that is,

$$w_{K+1}(x) = \frac{mx - \gamma + \sum_{k=1}^K \lambda_{k,K+1} w_k(x)}{1 + \sum_{j=0}^K \lambda_{K+1,j}}$$

for some constant γ . Notice that raising γ shifts the contract downwards, increasing the principal's payoff by γ without affecting the agent's incentives. It is thus optimal to set it to the smallest value that satisfies the agent's limited liability constraint, which is $\gamma = 0$.

Substituting w_{K+1} into the principal's objective together with (15) yields (D'). Note that the first and second constraint in (D') corresponds to (15) for $k \in \{1, \dots, K\}$ and $k = K + 1$, respectively. We have therefore shown that a contract w_{K+1} is optimal if and only if it solves (D'), and that every optimal contract takes the claimed form. \square